CHEMICAL REVIEWS

pubs.acs.org/CR

Review

Mass Spectrometry-Based Protein Footprinting for Higher-Order Structure Analysis: Fundamentals and Applications

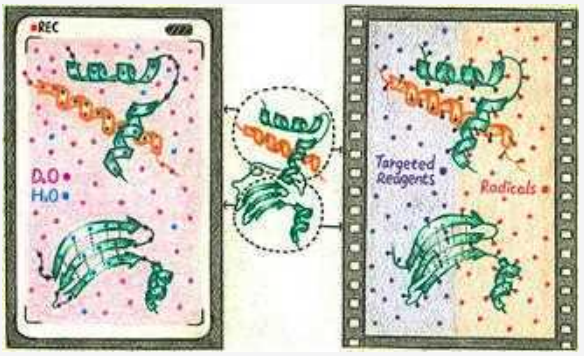
Xiaoran Roger Liu,*,† Mengru Mira Zhang,† and Michael L. Gross*

Cite This: *Chem. Rev.* 2020, 120, 4355−4454

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Proteins adopt different higher-order structures (HOS) to enable their unique biological functions. Understanding the complexities of protein higher-order structures and dynamics requires integrated approaches, where mass spectrometry (MS) is now positioned to play a key role. One of those approaches is protein footprinting. Although the initial demonstration of footprinting was for the HOS determination of protein/nucleic acid binding, the concept was later adapted to MS-based protein HOS analysis, through which different covalent labeling approaches "mark" the solvent accessible surface area (SASA) of proteins to reflect protein HOS. Hydrogen−deuterium exchange (HDX), where deuterium in D_2O replaces hydrogen of the backbone amides, is the most common example of footprinting. Its advantage is that the footprint reflects SASA and hydrogen bonding, whereas one drawback is the labeling is reversible. Another example of footprinting is slow irreversible labeling of functional groups on amino acid side chains by targeted reagents with high specificity, probing structural changes at selected sites. A third footprinting approach is by reactions with fast, irreversible labeling species that are highly reactive and footprint broadly several amino acid residue side chains on the time scale of submilliseconds. All of these covalent labeling approaches combine to constitute a problem-solving toolbox that enables mass spectrometry as a valuable tool for HOS elucidation. As there has been a growing need for MS-based protein footprinting in both academia and industry owing to its high throughput capability, prompt availability, and high spatial resolution, we present a summary of the history, descriptions, principles, mechanisms, and applications of these covalent labeling approaches. Moreover, their applications are highlighted according to the biological questions they can answer. This review is intended as a tutorial for MS-based protein HOS elucidation and as a reference for investigators seeking a MS-based tool to address structural questions in protein science.



CONTENTS

Received: December 16, 2019
Published: April 22, 2020

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

1. INTRODUCTION

Proteins carry out the programmed activities encoded by genes. Although constructed by the polymerization of only 20 distinct amino acids, their numerous biological functions require the high diversity arising from sequence variations and post-translational modifications. As living organisms evolve, proteins acquire specialized abilities that can be organized into different functional classes:[1] enzymes catalyze different intracellular and extracellular reactions, structural proteins provide support and maintain the structural rigidity of the cell, transport proteins facilitate trans-membrane flow of certain materials, regulatory proteins work as sensors and switches to regulate gene expression and protein activities, motor proteins facilitate macroscopic movements, and signaling proteins transduce messages to facilitate the communication of different cellular components. Despite high functionality and diversity, a protein's biological function and corresponding mechanisms of action are determined by their three-dimensional (3D) structure that is encoded in its primary sequence as demonstrated through the famous ribonuclease refolding experiment.[2] The exploration of protein structure−function relationship at the molecular level has developed into an independent subject named structural biology, for which characterization of protein higher-order structure (HOS) is the goal.

1.1. Protein Higher-Order Structures

The understanding of protein HOS is a key topic in biology because the functional mechanisms of proteins are encoded in their 3D structures. Protein structure can be viewed as having four distinct orders (Figure 1).[1,4] Primary structure refers to a

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

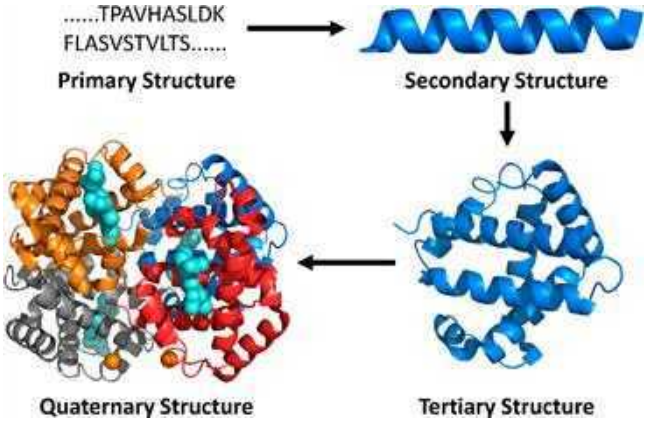


Figure 1. Four orders of protein structure exemplified by human deoxyhemoglobin (PDB 2HHB[3]).

linear combination of amino acids into a polymeric chain, which was first proposed in 1902.[5] Although originally heavily debated, its importance was settled after Sanger and co-workers[6,7] first sequenced insulin. The consensus is that protein **primary sequence** is determined by genetic information encoded in nucleic acids, that genetic information is translated into the order or sequence of the 20 common amino acids in a protein. The principal means of determining primary structure of proteins has become MS-based sequencing or proteomics analysis.

Upon forming linear polymeric chain, proteins fold to yield local structures, including α-helices, β-sheets, β-turns, Ω-loop, etc., with the first two being most common. Such local ordering represents the **secondary structure** of proteins.

Tertiary structures of proteins are their overall 3D structures that result from the folding of the secondary structural components and other unstructured motifs. Although many proteins function individually, there are others that are not biologically active until they interact with other proteins or ligands, and these interactions give rise to **quaternary structure**. An overall 3D structure of a protein complex that contains multiple protein subunits comprises quaternary structure. The term "protein higher-order structure" often refers to the secondary, tertiary, and quaternary structures of a protein. This is the subject of "structural proteomics", and protein footprinting is a key component of the subject.

Besides the covalent peptide bonds that assemble amino acid building blocks into primary sequence, noncovalent interactions stabilize protein HOS: these interactions include hydrogen bonding,[8,9] charge−charge interactions (salt bridge),[10,11] hydrophobic interactions,[12,13] aromatic−aromatic interactions ($\pi-\pi$ stacking),[14] cation−π interactions,[15] and van der Waals forces.[16] There is also a covalent contributor that stabilizes protein HOS, namely disulfide bonds.[17,18] All of these forces work together to overcome the conformational entropy of protein folding (decrease in entropy from random coil to folded protein) and to stabilize a protein in its folded state.[19−21]

1.2. Biophysical Approaches for Characterizing Protein Higher-Order Structure

In 1958, Sir John Kendrew[22] first reported the high-resolution structure of sperm whale myoglobin by X-ray crystallography, for which he shared the 1962 Nobel Prize in Chemistry with Max Perutz[23] for their studies of the structures of globular proteins. Their contributions are a milestone in the field of structural biology, opening the field to pursuits of protein HOS.

After decades of development, X-ray crystallography and nuclear magnetic resonance (NMR) became the "gold standards" for determining protein HOS. The interaction of an ordered protein crystal with X-ray radiation causes the incident beam to diffract. By measuring their diffraction angles and intensities, it is possible to obtain a 3D electron density map of a crystal, from which the protein 3D structure is constructed at atomic resolution.[25] X-ray crystallography provides atomic-level resolution of various protein sizes, making it the most widely accepted spectroscopic approach that provides the most definitive information on protein HOS to date although Cryo-EM will challenge it. Its disadvantage is obvious as well; a crystallized protein sample is a must, and protein crystallization has always been challenging.[26,27] Moreover, X-ray crystallography provides a solid-state structure of proteins, posing concerns about whether proteins alter their structures upon crystallization as compared with their structures in solution or in vivo.[28,29]

NMR resolves protein structure by determining chemical shifts and structural restrains.[30−32] The development of multidimensional NMR (pioneered by Richard Ernst, who was awarded the 1991 Nobel Prize in Chemistry for his contributions to high resolution NMR) allowed protein structures to be determined with near-atomic resolution by employing structural restraints obtained from homonuclear and heteronuclear couplings.[33,34] Solution NMR for determining protein 3D structures was pioneered by Kurt Wüthrich,[35,36] who was later recognized with the 2002 Nobel Prize in Chemistry. Solution NMR determines the liquid-state protein structures, which should resemble native states in the absence of interactions with proteins, ligands, and salts. Although mainly applicable to small proteins, recent work demonstrates that, in special cases, the approach can accommodate protein molecular weights up to 1 MDa.[37]

Another unique feature of solution NMR is its ability to measure protein dynamics in solution, which is not easily accessible at high resolution by other techniques.[38] For those proteins that are not readily soluble, membrane[39] and fibril[40] proteins, for example, solid-state NMR contributes significantly.[41,42] The "dark side" of NMR-based protein structural elucidation is the sample amount, which is generally on the order of milligrams. The data acquisition and analysis are also time-consuming and labor intense. Some proteins cannot maintain their structure for long data acquisitions and signal averaging.

X-ray crystallography and NMR have distinct advantages and drawbacks, but they complement each other.[43,44] As two of the most important pillars in protein HOS elucidation, the combination has contributed over 96% of high resolution protein structures in Protein Data Bank (PDB).[45]

Another emerging technique in high-resolution protein HOS elucidation is cryogenic electron microscopy (Cryo-EM), whose name is self-explanatory (EM at cryogenic temperature).[46,47] Upon snap-freezing the sample at cryogenic temperature, water molecules in the sample remain amorphous, thus minimizing distortion of the protein structure.[48] Damage by the electron beam to the sample is minimized at the cryogenic temperature.[49] The ability to average multiple randomly aligned EM pictures accelerated the development of Cyro-EM.[50] These factors combined to enable structural

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

determination of proteins with high-resolution. An early electron crystallographic study by Henderson and co-workers[51] demonstrates the capability of cryo-EM for determining membrane protein structure with near-atomic resolution (<4 Å). Continued technical advances allow single-particle cryo-EM to characterize noncrystalline samples[52,53] and determine protein structures with atomic resolution (to 1.8 Å).[47,54] Cryo-EM is also capable of characterizing protein complexes[55,56] and membrane proteins.[57] As cryo-EM generally favors proteins with high molecular weights,[58] recent advances made possible the characterization of proteins with sizes <100 kDa.[54] Its requirements of low sample amount, straightforward sample preparation, and high structural resolution make it the "rising star" in structural proteomics.[46,47] With X-ray crystallography and NMR, these three methods are the basis for determining protein structure with the highest resolution.

In addition to the high-resolution approaches, there are others including mass spectrometry (MS) that can characterize the proteins with lower resolution. Circular dichroism (CD),[59,60] ultraviolet (UV) resonance Raman,[61] and Fourier transform infrared spectroscopy (FT-IR)[62,63] can provide a general or overall characterization of protein secondary structure without providing atomic coordinates. UV at 280 nm is most commonly used in fast quantification of protein concentrations but contributes little to structure.[64] Fluorescence and Fürster resonance energy transfer (FRET) specialize in probing protein−protein interactions and protein conformational changes for regions of a protein.[65−67] There are other approaches that characterize the protein from a global perspective, including dynamic light scattering[68,69] and its coupling with size exclusion chromatography,[70] to focus on the hydrodynamic radius of proteins, negative staining electron microscopy,[71,72] and atomic force microscopy, which examine the morphologies of proteins,[73] isothermal titration calorimetry, which discloses the thermodynamic properties of proteins,[74,75] surface plasmon resonance, which measures the kinetic properties of proteins upon interacting with others,[76,77] and others. All of these techniques combine to provide a biophysical toolbox for protein HOS elucidations; they are summarized schematically in Figure 2.

In addition to these biophysical approaches, computer modeling also plays a significant role in understanding protein HOS.[78−80] Early ideas of structural prediction were developed in the 1960s,[81,82] but it was not until 1974 that the first structural prediction algorithm was actually developed.[83] These demonstrations were restricted to secondary structure predictions owing to limited computing power and mechanistic understanding of protein folding. Later on, with higher resolution, protein 3D structures could be resolved, and template-based modeling (or homology modeling) became possible where available high-resolution structures serve as scaffolds for predicting HOS of unknown proteins.[79,84,85] Free modeling (ab initio), on the other hand, predicts protein HOS from scratch,[86,87] and feasibility was first demonstrated in 2005.[88] Thanks to the development of computational approaches and bioinformatics, protein HOS prediction has developed into an independent research subject and is contributing more and more to structural biology.[89,90]

Each of these biophysical approaches has both advantages and disadvantages. The throughput of X-ray crystallography and NMR is limited by the protein crystallization and data acquisition, respectively. Some NMR analyses also require selective isotopic enrichment and milligram levels of sample,

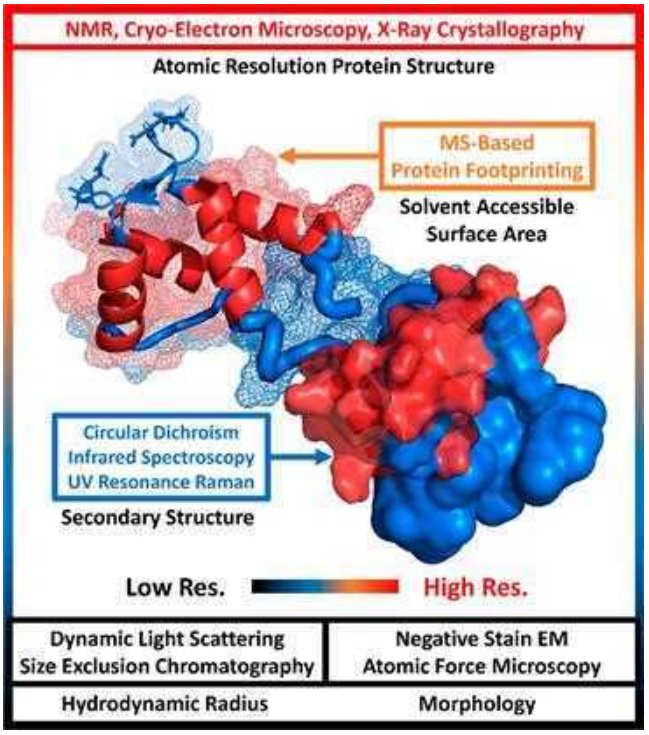


Figure 2. A summary of commonly used biophysical tools for characterizing protein HOS. Protein structure exemplified by calcium-free bovine calmodulin (PDB 1CFD[24]).

whose preparation is both labor and cost unfriendly. NMR is challenged by large proteins. Most protein samples require considerable signal averaging, and this can compromise the integrity of sensitive proteins. The structure from X-ray crystallography is solid-state and may be different than that in solution, where biochemistry occurs. Further, highly flexible regions diffract poorly and are not seen but can be tracked by footprinting. This limitation is particularly serious for intrinsically disordered proteins. Cryo-EM seems to overcome the throughput limitations; however, its broad adoption in analyzing small proteins remains to be established, and the resources are insufficient in number and expensive as of this writing.

Optical approaches, including CD, FT-IR, FRET, UV resonance Raman, etc., usually give high throughput. On the other hand, these approaches report on either global properties (e.g., alpha helical content in secondary structure from CD) or on the protein region near a label (e.g., fluorophore, donor/acceptor chromophore in FRET) and do not produce high-resolution spatial information. Although IR and Raman can give more specific information, they usually fail in probing a complete structure, particularly of large proteins and protein complexes. All of these limitations motivate the development of MS-based approaches for HOS analysis, as they fill in the gap by balancing throughput and spatial resolution. MS also contributes to protein HOS analysis from another perspective, as introduced in section 1.3.

1.3. Mass Spectrometry-Based Protein Structure Analysis

1.3.1. Historical Overview of Protein Mass Spectrometry. *1.3.1.1. Ionization.* To analyze a protein molecule in a mass spectrometer, ionization is the first step. The MS-based ionization methods were extensively developed starting in the 1960s and continuing to the 1990s; key developments include

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

chemical ionization,[91] electrospray ionization (ESI),[92] field desorption,[93] laser desorption,[94,95] californium-252 plasma desorption,[96] secondary ion mass spectrometry,[97] fast atom bombardment,[98,99] matrix-assisted laser desorption ionization (MALDI),[100,101] and many more. Several of these techniques involved sample desorption, that is, MALDI and its precursor laser desorption, plasma desorption (PD),[102,103] and fast atom bombardment (FAB).[104] The latter was extensively used for over nearly two decades (from 1980 to 2000) for studies of peptides and small proteins. Many of these approaches are vital steps in the long evolution of ionization methods, but none work as well as ESI or MALDI. PD preferentially ionizes the molecules on the surface of support, and its efficiency is low and affected by the homogeneity of the surface layer. FAB is a relatively hard ionization method with low ionization efficiency and a relatively high "chemical-noise" background and a propensity to fragment proteins upon desorption, limiting its further adoptions in studying peptides and small proteins as compared with ESI or MALDI.

Although FAB and PD convincingly demonstrated that proteins are amenable to MS analysis, the ionization scene improved significantly with two separate but complementary breakthroughs. In 1988, Tanaka and co-workers[105] improved MALDI by incorporating 30 nm cobalt particles to ionize successfully lysozyme and chymotrypsinogen and to obtain their masses, while Hillenkamp and Karas[106] were simultaneously developing the MALDI protocol that is used today, moving systematically from peptides to proteins. Meanwhile, Fenn[107,108] developed ESI and successfully ionized various proteins, showing that multiple charging brings the m/z of the ions into a mass range of most mass spectrometers where analysis was possible.

These two developments, particularly ESI, allow MS to become the enabling tool in the field of proteomics and subsequently in protein HOS analysis; these methods together became some of the most important biophysical tools in structural biology. Their efforts led to awards of the 2002 Nobel Prize in Chemistry to Koichi Tanaka and John Fenn, shared with Kurt Wüthrich (in the field of protein NMR), for their contribution in "the development of methods for identification and structure analyses of biological macromolecules". Among these two methods, ESI (and later nano-ESI[109,110]) has become more extensively used in proteomics and HOS analysis owing to its capability to install multiple charges in the analyte, its ease in setup, and most importantly its compatibility with liquid chromatography (LC), uniting in a concatenated way separation and analysis.

1.3.1.2. Instrumentation. Besides development of ionization methods, development of MS instrumentation also greatly elevated the detection limit and resolving power, empowering MS to work with larger proteins and deliver protein structural information with high mass resolving power. The advances in mass analysis show an evolutionary thread that included quadrupole mass filters[111] and double-focusing sector mass spectrometers;[112,113] both evolved into three[114] and four[115] (and even more[116]) analyzers for tandem-MS (MS/MS). Time-of-flight (ToF) instruments profited from developments of straight,[117] reflectron,[118] and orthogonal[119] ToF mass analyzers. Complementing and possibly exceeding them are quadrupole ion traps,[120] Fourier-transform ion cyclotron resonance (FTICR) instruments,[121] and more recently orbitrap mass analyzers.[122,123] Modern MS instruments are almost exclusively hybrid instruments that contain two or more different mass analyzers (represented by quadrupole-ToF (Q-ToF)[124] and quadrupole-orbitraps[125]), allowing rapid, accurate, and precise mass determination of both precursor and product ions, a long sought-after goal in MS.

1.3.1.3. Fragmentation. The use of MS in protein HOS analysis depends not only on the accurate measurement of the mass-to-charge ratio (m/z) but also on fragmentation methods that cleave the peptide backbone to provide sequence information that can locate sites of modification introduced as a protein footprint.[126] MS-based ion fragmentation required the introduction of collision induced dissociation (CID) and then tandem instruments, establishing the area of MS/MS.[126−128]

Owing to its rich history and high popularity at the time that peptides and proteins were first ionized, it is not surprising that CID was soon developed as a method to sequence peptides and to apply these ionization advances to unknowns.[129−131]

Over the years, different fragmentation methods, including surface induced dissociation (SID),[132,133] electron transfer dissociation (ETD),[134] electron capture dissociation (ECD),[135] electron detachment dissociation (EDD),[136] ultraviolet photodissociation (UVPD),[137,138] and infrared multiphoton dissociation (IRMPD),[139] were developed as means of activating peptides and proteins and causing them to fragment. Their unique characteristics, advantages, and limitations were reviewed in detail elsewhere.[126,140]

1.3.2. Choice of the Term "Protein Footprinting". Although other terms describe this area of analysis, including but not limited to "protein painting", "covalent labeling", "marking", "decorating", "oxidative labeling", "radical probe MS", and "chemical modification", we recommend the adoption of "protein footprinting" as a descriptive term that can potentially unify a diverse field and facilitate its organization and literature searching. Footprinting is a conceptually simple yet effective approach, during which the solvent accessible surfaces of a biomacromolecule are chemically labeled. The labels can be as simple as the replacement of an H for a D or as complex as adding a bulky photoaffinity tag to mark a protein. The labeling sites are then located at regional and even residue specificity. Those locations, at either the regional or amino acid levels, constitute a "footprint" to inform on protein HOS and its changes.

The advantages of "footprinting" over other terms lies in three aspects: the rich history, less ambiguity, and appropriate imagery. Protein footprinting was first proposed by Payne and co-workers[141] in 1988, and that term predates most of the terminologies listed above. Other terminologies such as "chemical modification" and "marking", although proposed prior to "footprinting", are less specific. Chemical modification of proteins is a broad term in the field of biochemistry. A good example of protein chemical modification is to attach a green fluorescence protein on to the target protein for visualization purposes, an application that is irrelevant to protein HOS elucidation. Lastly, footprints, by definition, are the impression left by a foot on the ground. Taking the foot as a symbol of the protein, we analogize that a protein footprint, as a reflection of HOS, is recapitulated by surface-sensitive labeling reagents in a similar way as the sand grains on the ground reflect the structure of the foot.

It is too simple to attribute all differences in a footprint to changes in surface area and solvent accessibility. There are examples of HDX where protection does not increase at a binding interface.[142] Other examples indicate that differences

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

are determined by "chemical accessibility" rather than surface accessibility.[143,144] There are also examples of radical precursors that interact by H-bonding and give a footprint that is disproportionate to accessible surface, challenging the simple notion that a reactivity only depends on surface area.[145] Although these scenarios indicate more complex factors behind the apparent modification extent, the labeling process can still be well represented by the term "footprinting". As walking on a sand or pebble beach produces different footprints, protein footprints will also vary from one approach or one reagent to another. All in all, a single term "protein footprinting" is preferable to many ad hoc terms that are currently used in papers and become difficult to locate and cite in a literature search.

One generalization that is applicable to this structural approach is that it is underpinned by chemistry, not by the physics of electromagnetic radiation absorption or diffraction; hence we favor the term "chemical footprinting". An experiment starts with the design and choice of "footprinting molecules" (e.g., D_2O in HDX, $^{\bullet}OH$ in radical footprinting), continues with the chemical reactions that label the protein, and with the proteolysis reactions to produce constituent peptides whose analysis produces the "footprint". To visualize the "footprint" (i.e., locate the modified sites and quantify their extents), ionization is required (proton transfer for ESI and MALDI, redox chemistry for EI or electron-based MS/MS, all are chemical processes) followed by MS/MS fragmentation. Interpretation of the MS data gives the "chemical footprint". It is a chemical approach: the chemistry of labeling, proteolysis, ionization, and fragmentation. These latter subjects have been under development for more than a half century, and when taken together at their current state-of-the-art, they combine to give a powerful approach in protein HOS analysis.

1.3.3. Labeling-Based Protein Footprinting. MS contributes to peptide and protein determination at all four levels of structure.[146] The primary structure of a protein can be determined through MS/MS-based de novo sequencing or partial sequencing and database searching.[147] Although protein sequencing was expedited mainly by sequencing its corresponding genome, there are still needs to sequence a protein when its corresponding genome is unknown or the sequence requires verification. MS/MS-based sequencing has become the nearly exclusive route to primary structure, replacing Edman degradation owing to the accuracy, certainty, and speed of MS.[148] In addition to sequencing, the MS/MS approach also enables profiling of post translational modifications (PTMs) by following shifts in m/z that pinpoint the modification type.[149−153] Indeed, this capability is the basis for protein footprinting, where finding the location of a chemical modification introduced purposely is the means of determining protein footprints that are indicative of HOS.

Footprinting probes secondary and tertiary structure by mapping solvent-accessible surface area (SASA) through various covalent-labeling approaches such as hydrogen/deuterium exchange (HDX). Some view HDX as a "noncovalent" approach and not consider it as "footprinting", but we take issue with this classification because N−H and N−D bonds, although reversible, are covalent, and the pattern of peptide N−H and N−D bonds is a "footprint." Nevertheless, there are examples of noncovalent footprinting,[154] but the adoption is far less than of those approaches involving covalent bonds.

Quaternary structures are often characterized by differential footprinting that locates the contacting region by responding to changes in SASA for two different states (e.g., bound and unbound). Native spray MS, a gentle version of ESI that uses aqueous media, can maintain the noncovalent interactions in the gas phase to yield information on topology and stoichiometry but not on interacting interface.[155−159] Ion mobility MS reports protein shape, a component of HOS, in the gas phase.[160−162] Chemical cross-linking, an in-solution approach, is now extensively used to provide some information on protein quaternary structures and to locate interfaces.[163−165] All of these approaches are the analytical basis for MS-based structural proteomics and provide higher-order structural information, although not as detailed as that from NMR, X-ray crystallography, and Cryo-EM.

In this review, we have adopted a focus on the emerging role of MS-based protein HOS determination by different covalent labeling approaches, which is becoming one of the most informative MS-based structural proteomics tools. Covalent labeling at solvent-accessible surfaces of the protein constitutes "footprinting", a term that was initially proposed and demonstrated for HOS analysis of nucleic acids in 1978[166] and later adopted to analyze proteins.[141]

Early protein footprinting approaches were mostly cleavage-based, where protein SASA is revealed by limited proteolysis occurring at sites of the highest SASA[167,168] or by chemical cleavage by reactions with reactive radical species.[169] Because these approaches do not afford high spatial resolution, cleavage-based footprinting was soon replaced by labeling-based approaches.

Over the years, two distinct categories of covalent labeling approaches have been developed and applied: reversible and irreversible. Reversible labeling is represented by HDX, which takes advantage of the hydrogen exchange between active hydrogens in the protein molecule and the hydrogen in the water solvent.[170] Upon dissolving a protein in deuterium oxide (D_2O), solvent-accessible and weakly H-bonded hydrogen atoms in the protein will exchange with deuterium atoms in the solution, where mass differences between hydrogen (1.0078 Da) and deuterium (2.0136 Da) atoms serve as a useful indicator for a mass spectrometer to report such exchanges. Although the exchanged H or D in backbone are covalently bonded in the protein, the exchange process is highly dynamic and reversible. Thus, for ex situ methods like MS, protein HDX analysis needs to be performed under constrained conditions to prevent back exchange, which would quickly "erase" the footprint and eliminate the structural information from the labeling. Details about protein HDX will be covered in section 2.

Another category is irreversible protein labeling. Labeling reagents generally react with solvent accessible amino acid side chains, leaving a chemical "mark" that can be identified in a forthcoming analysis. Nature already does elegant protein labeling, that is, by PTMs.[149,171,172] Mainly enzymatic processes introduce certain functional groups to proteins to modify them, often by adding or subtracting charges, so that new biological functions can be enabled by changing the conformation and the binding opportunities of a protein.[173]

Inspired by nature's use of PTMs, investigators have developed different chemical reagents that irreversibly label solvent accessible surfaces of proteins. One subclass of reagents that are readily available today are targeted reagents that react with only certain amino acid residues with high specificity but

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

low reaction rates.[174] Examples are acetylation of lysine or esterification of aspartic acid. The second category includes fast labeling reagents, mostly radical species, which are highly reactive but have lower specificity.[175] The advantage of irreversible labeling is that sample handling and postlabeling analysis can be slower and more efficient and effective than that of HDX, taking full advantage of all the separation advances made in proteomics. The different labeling time scales allow irreversible labeling to address several biological questions that are not amenable to HDX.

1.3.4. Bottom-Up and Top-Down. With MS/MS, locating labels and determining the protein footprint that reflects HOS can be performed with either bottom-up or top-down approaches.[176,177] Bottom-up protein analysis requires proteolytic digestion of the proteins prior to MS. Most approaches involving MS utilize a bottom-up approach as it has higher throughput, better compatibility with LC systems, greater sensitivity, and requires less advanced fragmentation technologies than top-down approaches. The key challenge for structural proteomics is to obtain high sequence coverage of the targeted protein, which is not a requirement for proteomics strategies that identify target proteins by using partial sequences determined for peptides that are "proxies" for the protein. Furthermore, PTMs are sometimes not accurately characterized owing to biased digestion efficiencies, as protease digestion efficiency is lowered when its cleavage site is post translationally modified.[178] A similar problem occurs in protein footprinting when a proteolytic cleavage site is modified by a footprinting reagent.

Top-down MS has been extensively developed in the past 20 years thanks to advances in MS fragmentation methods.[177,179] As the intact protein is submitted directly to the spectrometer, extensive fragmentation of the intact protein molecule must be induced to yield high spatial resolution. The major advantages of top-down is the capability of recovering different protein isoforms that are combinations of PTMs and sequence variants. High fragmentation coverage is needed for footprinting, but today this is achieved mainly for small proteins. The approach is especially attractive when enzymatic cleavage sites are limited.[180] All of these advantages make top-down a promising choice in proteomics.[181] Its drawbacks are primarily low throughput, limited sensitivity, and a lack of residue-level sequence coverage.[182] Reduced propensity for fragmentation in large proteins is due to the many degrees of freedom for delocalizing energy in a big molecule, making it challenging to investigate proteins with high molecular weights by collision-based fragmentation approaches.[183] Novel fragmentation methods such as ECD, ETD, and UVPD fragment the protein by different mechanisms, thus overcoming such drawbacks.

In the field of protein HOS elucidation by MS, there are fewer examples that couple protein footprinting with top-down MS,[184−186] than with bottom-up. The primary reason is that sequence coverage at residue level is critical for high spatial resolution, for which top-down currently fails to deliver. Combined developments of instrumentation, fragmentation methods, and sample handling strategies may afford a brighter future in protein footprinting.

There has been a growing need for MS-based protein footprinting in both academia and industry owing to its high throughput, prompt availability, and high spatial resolution. In the coming sections, we will cover the history, principles, designs, and mechanisms of covalent labeling-based protein footprinting approaches. We will largely confine our attention to reports over the last 10 years, emphasizing those that report advances in methodology, rather than applications, to highlight the biological questions that can be addressed. Our goal is to provide a review that will serve as both tutorial for MS-based protein HOS analysis and a reference for investigators seeking a MS-based tool to address their question in protein science.

2. "UNBIASED" PROTEIN FOOTPRINTING: HYDROGEN−DEUTERIUM EXCHANGE

2.1. A Historical Perspective

HDX is one of the most important and used covalent labeling protein footprinting approaches to date despite its reversible nature, although HDX is usually distinguished from other irreversible footprinting approaches by classifying the latter as covalent labeling or covalent footprinting, HDX labeling also involves covalent bonds. The distinction is not covalent labeling as the exchange but rather that the exchange of N−H for N−D in HDX is a reversible process. We argued earlier that the distinction between HDX and other footprinting is "reversible vs irreversible" not covalent vs an implied noncovalent bonding.

Protein HDX was first demonstrated in 1954 by Kaj Linderstrøm-Lang,[187] who was later recognized as "father of HDX".[188] In the very first HDX experiment, Hvidt and Linderstrøm-Lang[187] incubated dried pork insulin into D_2O to allow exchange, followed by snap-freezing and lyophilization. The protein powder was then suspended in H_2O. The deuterium content was measured by the density of the newly obtained protein−H_2O solution in a gradient tube. Later, they studied the pH and temperature dependence of H/D exchange rates,[189,190] made a connection between H/D exchange rate and the dynamics of protein molecules,[191,192] and formulated equations that underpin the basis for protein HDX.[190,192,193] Efforts by Linderstrøm-Lang and co-workers[188,194] are part of the early history of protein HDX, but in a display of scientific insight, these investigators demonstrated an approach of tremendous potential, opening a new era of covalent labeling-based protein HOS elucidation.

To enhance spatial resolution, HDX was later coupled with other characterization approaches (e.g., UV spectroscopy,[195] IR spectroscopy,[196,197] neutron diffraction[198]), among which NMR was the most successful in early applications. The NMR applications began with a series of studies in late 1950s, when Saunders and Wishnia[199−201] first demonstrated that the exchange of proteins and D_2O solvent can be measured by NMR with high spatial resolving power to give residue-level information at least for small proteins.[202] In 1976, these efforts were extended to the determination of individual amide proton exchange rates, as demonstrated by Wüthrich and co-workers.[30] Later, Wagner and Wüthrich[203] coupled two-dimensional correlation spectroscopy (2D-COSY) NMR with H/D exchange rate measurements whereby exchange rates for every amino protons in a protein can be obtained in a single experiment. Recent development of fast 2D $^{1}H-^{15}N$ correlation NMR demonstrates measurement of fast HDX for amide bonds within a few seconds of acquisition time.[204] Along with these developments, NMR measurements also contributed significantly to the mechanistic understanding of protein HDX phenomena.[205,206]

The use of HDX-NMR in studying protein−protein interactions was first demonstrated by Roder and co-workers[207] in 1990. Although powerful and promising, HDX-based

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

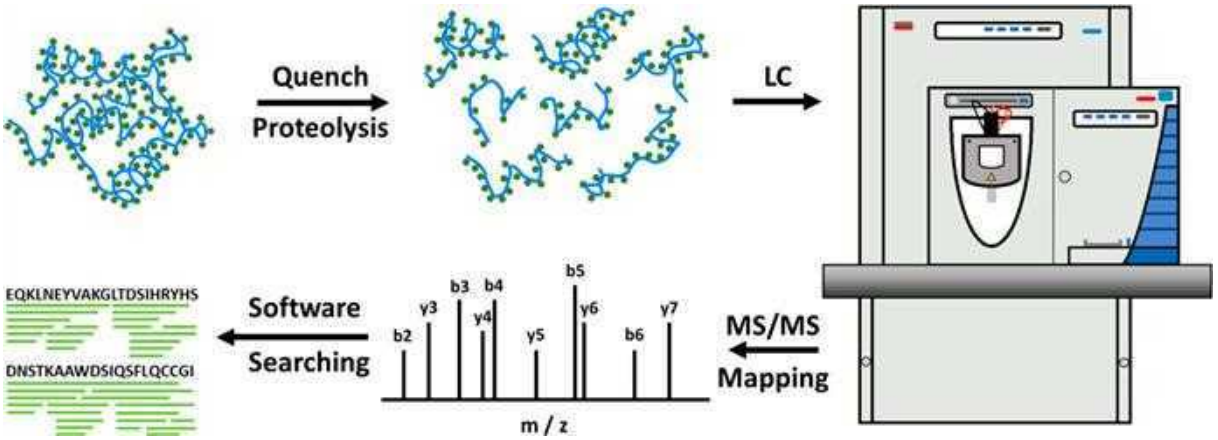


Figure 3. Schematic illustration of a bottom-up peptide mapping workflow, a necessary step prior to HDX. Green dots in the protein structure indicate hydrogen atoms in the peptide bonds.

protein−protein interaction studies are not often made in NMR. Early protein NMR approaches suffered from limited spatial resolution, making it even more challenging to characterize large proteins by NMR.

As a complement to NMR, MS began to "pick up the baton" to extend HDX further to larger proteins at lower and lower concentrations. The demonstration of HDX followed by ESI MS was first by Katta, Chait, and Carr[208] in 1991, who showed that the mass spectrum of a small protein (ubiquitin) changed significantly when it was sprayed in deuterated solvent. As the mass of the deuteron (2.014) differs measurably from that of hydrogen (1.008), MS monitors HDX by tracking the centroid mass of the isotopic envelop of a specific protein or peptide. As exchange-in continues, the centroid shifts to higher m/z and reports the average number of Hs that have exchanged with Ds.

The first example of off-line HDX and proteolytic digestion to reveal the regions of protein that exchange was by Zhang and Smith.[209] They termed their approach as "the protein fragmentation method", and they used fast atom bombardment (FAB) as their ionization method. This approach is the precedent for what is done today (except FAB has been replaced by ESI).

2.2. Mechanism of Exchange

To provide a basic understanding of the mechanism of HDX of the amide backbone, we will briefly discuss some fundamentals.[170,205,210] Among all hydrogens in a protein molecule, those that are part of O−H, S−H, and N−H bonds exchange most rapidly with solvent water (D_2O) molecules. In a commonly executed "exchange-in" scenario, the protein is first solubilized in H_2O. Upon diluting with D_2O, labile Hs on the protein exchange with the surrounding solvent Ds, leading to an increase in mass that can be measured by a mass spectrometer. There is also an "exchange-out" mode, where the protein is first incubated in D_2O, followed by addition of H_2O to cause a D-to-H exchange. Exchange-out is not as often used as "exchange-in," and it will not be discussed further. Although HDX happens for both protein backbones and amino acid side chains containing exchangeable Hs (e.g., NH_2, OH, SH), most experiments focus on backbone hydrogens. The hydrogens on the side chains are weakly involved in hydrogen bonding, making their exchange fast and not easily followed with most experimental setups. These exchanges do not confuse the experimental outcome because, as fast exchangers, they return to an −XH state during workup and proteolysis that uses H_2O, as expected in an ex situ measurement.

Because the HDX of protein side chains is fast and not readily measurable, the total exchange is nearly only that of the amide N−H in peptide bonds, making HDX an "unbiased" labeling method because every amino acid residue (except Pro) has an N−H hydrogen that can be exchanged. The rate of exchange is fastest for amides that are solvent-exposed and not involved in hydrogen bonding, for example, unstructured regions in a protein. HDX rates for these N−H hydrogens are represented by k_{ch}, which is a function of pD, the temperature of the solvent, and the nature of the side chains of neighboring amino acid residues. Previous studies reveal that HDX is both acid- and base-catalyzed and that k_{ch} minimizes at a pH around 2.5,[211,212] the pH that is chosen for quenching the reaction in a step that is essential for any ex situ measurement.

HDX for HOS determination is successful because the rate of HDX for backbone N−H hydrogen atoms is determined by local conformation. Protein local orders are stabilized by intramolecular N−H···O═C hydrogen bonding.[213] The breathing motion of these hydrogen bonds will alter the conformational states of the targeted hydrogens that comprise the N−H bonds from closed to open states with a rate constant of k_{op}. As the protein is highly dynamic, the open N−H bonds can also close with a rate constant of k_{cl}. A local view of HDX process for a backbone N−H, thus, can be represented by eq 1.

$$\mathrm{N{-}H_{closed}} \underset{k_{cl}}{\overset{k_{op}}{\rightleftharpoons}} \mathrm{N{-}H_{open}} \xrightarrow{\mathrm{D_2O},\ k_{ch}} \mathrm{N{-}D_{open}} \underset{k_{op}}{\overset{k_{cl}}{\rightleftharpoons}} \mathrm{N{-}D_{closed}} \quad (1)$$

In principle, every N−H possesses a unique combination of k_{op}, k_{ch}, and k_{cl}. The observed rate constant for HDX, k_{HDX}, is determined mainly by two factors: the solvent accessibility and the local hydrogen bonding in the region of the amide bond. For regions that are highly flexible, the effect of these two factors is minimized so that $k_{HDX} = k_{ch}$. For regions that are structured, HDX occurs under two limiting regimes, EX1 and EX2.[210,214] EX1 is characterized by $k_{ch} \gg k_{cl}$, for which $k_{HDX} = k_{op}$. In EX2, $k_{cl} \gg k_{ch}$, so that $k_{HDX} = K_{op}k_{ch}$, in which $K_{op} = k_{op}/k_{cl}$. Practically, EX2 is more prevalent than EX1.[192,206] There are also intermediate exchange regimes that are combinations of EX1 and EX2.[142,214]

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

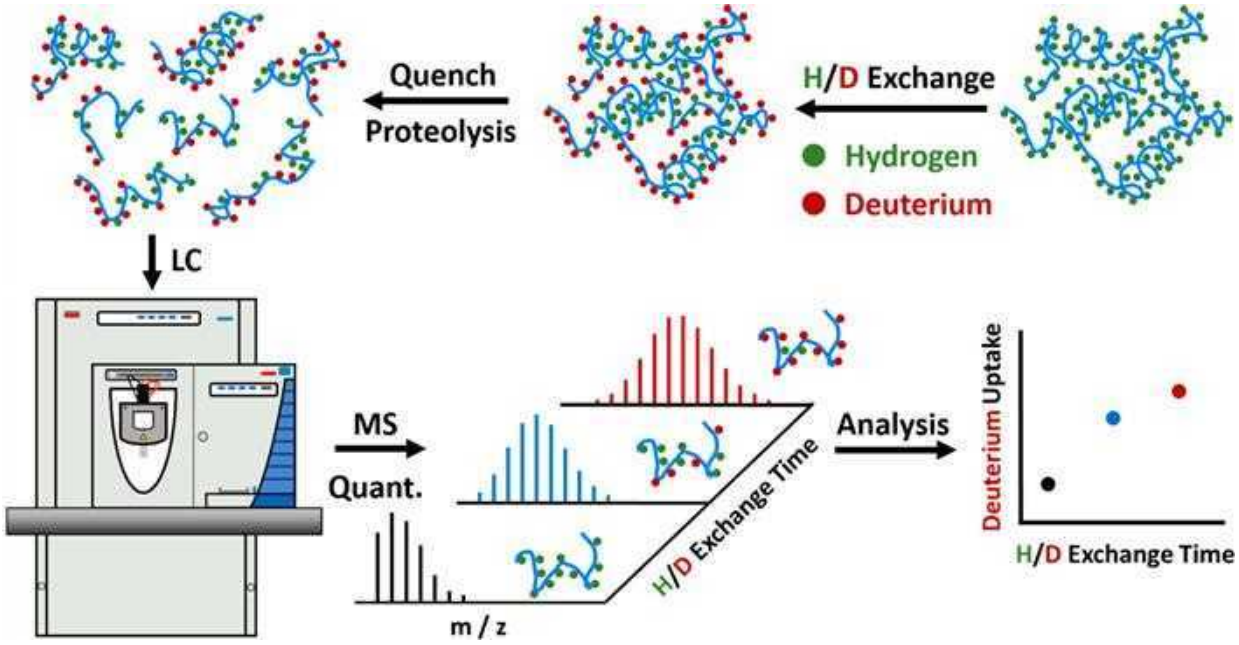


Figure 4. Schematic illustration of a bottom-up HDX workflow. Green and red dots in the protein structure indicate hydrogen and deuterium atoms in the peptide bonds, respectively.

2.3. Experimental Approach

Most HDX measurements compare two states of a protein (e.g., bound vs unbound, wild-type vs mutant); thus, this approach in footprinting is termed "differential". In a differential experiment, HDX for a ligand-unbound protein is compared to the ligand-bound state, and the effects of local hydrogen bonding and back exchange cancel, leaving the differences in HDX to represent changes in solvent accessibilities and H-bonding for the two different states as the only factors determining the relative differences in HDX.[215] Usually, ligand binding will induce a decrease in deuterium uptake at the binding region, but there are also cases that indicate a different scenario (i.e., increased deuterium uptakes at the binding sites upon interacting with ligands).[142]

Experimentally, HDX is most commonly executed in a "bottom-up" fashion,[212,216] which originated from early experiments of tritium exchange.[217,218] Bottom-up HDX was first introduced by Zhang and Smith[209] in 1993. In the peptide mapping mode, the protein or peptide (in H_2O) is denatured and, following quenching at pH ∼ 2.5 (in H_2O), is digested and made ready for LC-MS/MS analysis for peptide identification and m/z determination to establish a template for determining the average number of D taken up in the actual exchange (Figure 3). One goal is to optimize the quenching conditions to permit high sequence coverage and make available a peptide list comprised of peptide sequences, precursor masses, and their elution times under optimized quenching conditions. The quenching solution is usually composed of denaturant and an agent for disulfide bond reduction (if there are any) and is at pH = 2.5, where the rates of back-exchange (exchange of deuterium in the protein to hydrogen) are minimized. The quenching conditions, once optimized, are then used in further HDX experiments.

In the actual HDX experiment, the protein of interest is initially dissolved in H_2O, and the solution diluted with D_2O and incubated for several different times to allow measurement of the HDX kinetics (Figure 4). Upon exchanging, solvent-accessible amide hydrogens become replaced by D. The exchange is then quenched by decreasing the solution pH to 2.5. Quenching is followed by fast protease digestion and then MS analysis. The digestion is usually carried out online by flowing the protein solution through a pepsin column. The solvents used in HPLC are at acidic pH, and the LC-MS analysis is usually done at low temperature to minimize back exchange. Optimized experimental design shortens the time prior to the analysis of denatured peptides by MS to minimize back-exchange. The obtained data are analyzed by fitting the isotopic distribution for each peptide, from which a mass centroid is calculated, and the deuterium uptake is determined by the difference in peak centroids between the exchanged protein (or the peptides) and the unexchanged or control.

Measurement of HDX at the residue level with MS is challenging. Low-energy collisional activation found on orbitraps, ion traps, or Q-Tofs cannot be used because that activation induces deuterium scrambling,[219,220] erasing all the forward exchange that was scrupulously introduced. Novel fragmentation methods that minimize scrambling include ECD[135,221] and ETD,[222] making possible, in some cases, measurement of residue-level deuterium uptake or of HDX in a "top-down" fashion.[223,224]

2.4. Recent Advances and Applications

In this section, we will review several technical advances that increased the capabilities of HDX. One important goal is to increase its spatial resolution, ultimately to a single amino acid. Several approaches were developed to achieve a single-residue resolution, including vigorous proteolytic fragmentations to give overlapping peptides,[225,226] high-pressure online digestion to improve digestion efficiency,[227,228] incorporation of ETD and bottom-up HDX for enhanced spatial resolution,[229−231] and computation methods that facilitate processing the data from aforementioned methods.[225,232,233] Novel digestion protocols give shorter peptides and more overlapping peptides. Single-residue HDX information can be obtained by considering the absolute deuterium uptake levels of two

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

overlapping peptides that differ by a single amino acid residue.[225−228] This spatial resolution should also be achievable through fragmentation to give smaller peptides formed by ETD or ECD fragmentation that can occur with minimal deuterium scrambling. Quantification of the resulting fragments will give data that resemble the deuterium uptakes for those fragments in solution.[229] Novel computation methods either facilitate the automatic processing of the overlapping peptides[232] or utilize novel algorithms to deconvolute the residue-level HDX behavior through a peptide-level HDX curve (by either close examination of isotopic envelope shape information[220] or a Bayesian approach for deconvolution[226]).

To shorten further the exchange time, investigators developed theta capillary spray[234] and gas-phase HDX[235−237] that allow exchange times as short as 20 μs to be followed. With further optimization, it should be possible to track deuterium uptake on the amino acid residue side chains, which was not possible in conventional HDX (even with rapid mixing or stopped-flow) because the back-exchange of the active protons is rapid. It also allows HDX to capture, for the first time, fast dynamic processes.

To expand the capabilities of HDX to binding affinities, investigators developed titration-based HDX workflows including stability of unpurified proteins from rates of H/D exchange (SUPREX)[238] and protein−ligand interactions by mass spectrometry, titration, and H/D exchange (PLIMSTEX).[239] These approaches characterize ligand binding sites, binding orders (if there are multiple ligands that bind the protein of interest) and, most importantly, site-specific binding affinities. Both of these workflows were demonstrated at the peptide levels,[240] showing that it is possible to work at the single-residue level with the combination of the efforts mentioned above (ETD, ECD, or multiple proteolytic digestions).

To expand the ability of HDX to characterize membrane proteins, effective removal of detergent[241] and lipid,[242] optimization of HDX conditions for enhanced transmembrane domain sequence coverage,[243] and post-HDX deglycosylation[244] were developed. Post exchange detergent and lipid removal are based on either chromatographic separation and organic solvent extraction[241] or on the interaction between zirconium(IV) oxide and phospholipids.[242] These advances make possible the characterization of HDX for membrane proteins without incurring extensive back-exchange. Post-HDX deglycosylation removes highly heterogeneous glycans without compromising the deuterium labeling on the peptides, allowing an investigator to obtain a well-resolved, high-coverage HDX analysis.[244] All of these advances broaden the horizon of HDX.

Other than the experimental advances, we witnessed a burst of HDX data processing software in the past decade, some of the examples include HDExaminer,[245] DynamX (formerly HX-Express),[246] HDX Workbench,[247] AUTOHD,[248] Mass Analyzer,[249,250] HD Desktop,[251] HeXicon,[252] ExMS,[253] HDX-Analyzer,[254] HDX Finder,[255] and Mass Spec Studio.[256] There are also new HDX data visualization software include MSTools,[257] MEMHDX,[258] Deuteros,[259] and HDX-Viewer.[260] The development of these software tools contributes greatly to the broad application of HDX.

Undoubtedly, HDS-MS is now becoming relatively routine and is currently used in many applications in biochemistry and biophysics. The large number of these applications makes it challenging to cover them in this review. Instead, we will stick with our theme of protein footprinting and emphasize HDX as a footprinting method to provide HOS information. We will not review applications of HDX-MS because there are many specialized and comprehensive[170,261,262] reviews that cover most applications, including mapping epitopes and characterizing biotherapeutics,[263−266] monitoring protein folding dynamics,[267,268] locating protein binding sites,[215,269,270] examining conformations of individual proteins or of large complexes,[271−273] probing allosteric effects,[274] monitoring protein−membrane interactions,[275] and developing methods to minimize deuterium scrambling during MS fragmentations.[276] A collaborative effort from 2018 summarizes the key aspects of performing HDX experiments, providing valuable recommendations on how to carry out an effective HDX study properly.[216] Lastly, for more details on HDX theory, the reader should consult early reviews by Englander et al.[205] and Smith et al.[212]

2.5. Conclusion

The union of HDX and MS launched MS into the field of protein HOS determination. Because HDX causes a minimal perturbation on protein structure, now has a robust setup and analysis, aided by a number of data processing packages, and requires no design or synthesis of reagents, the adoption by the biochemistry community has been rapid. HDX covalently "labels" all protein solvent-accessible backbone amide hydrogens (except Pro). Unlike targeted- and the fast-labeling reagents, whose labeling efficiencies are hugely variable and determined by residue-specific reactivities, HDX occurs across the whole protein and its labeling efficiencies depend on local structure (solvent accessibility and hydrogen bonding). Moreover, deuterium labeling in HDX is less disruptive to the protein structure as compared with other labeling approaches. Any susceptibility is due to isotope effects, for which bond-making and bond-breaking could be susceptible. In general, HDX footprints proteins in a relatively "unbiased" way. The advantages of HDX, as envisioned by early pioneers, have been realized, and new applications are being implemented, making HDX an important technology for HOS determination nearly three decades after its introduction.

3. TARGETED-LABELING REAGENTS

The 20 amino acids have different kinds of functional groups on their side chains; most have functional groups (COOH, SH, NH_2, OH, $CONH_2$, aromatic ring) that can be labeled with chemical reagents that react specifically, usually with one or two amino acids. These modifying reactions can be used for footprinting provided the reactivity of these groups depends on SASA of the protein and the modifications in the early stages do not affect the protein structure, minimizing the biased report during the footprinting itself. Reagents react directly and usually slowly with specific solvent-accessible side chains in contrast to free radicals, which react rapidly. The product contains a characteristic mass tag that can be detected by MS analysis. Although numerous reagents have been developed to react with the amino acid side chains, a qualified protein footprinting reagent needs to label the protein under physiological conditions with reasonable efficiency and speed. The size and hydrophilicity of the reagent should be close to that of water to ensure its reactivity is an indicator of SASA. Most of these reagents developed to date are highly specific, targeting one or two side chains or functional groups although some can react with more than two.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

In this section, we will review commonly used targeted protein footprinting reagents, starting with their history in protein labeling and then discussing their chemistry and the products resulting from labeling. Many of the protein labeling reactions were discovered and implemented some time ago but not as footprinting reagents. That had to await the development of MS methods for sensitive, specific, and effective analysis to locate the modified sites. In early footprinting efforts, MALDI was chosen for analysis, but more recently the combination of ESI and LC-MS/MS has become the mainstay analytical tool. Our purpose in this section is to organize the reagents and credit their original development. We will not cover the numerous applications of these reagents.[174,277−281]

Because there are many reagents, this section is organized according to the target residues and the types of functional groups they contain. Cysteine is highly reactive owing to its highly nucleophilic thiol group. Tryptophan and tyrosine contain hydrophobic/aromatic side chains that are electron rich. Aspartic and glutamic acids are acidic residues for which many classic derivatization reactions are available. Arginine, histidine, and lysine contain basic residues, and they are more difficult to footprint. Although some of the residues are preferentially charged under physiological conditions, their charge state depends on pH. All schemes in this section depict the charge neutral forms for simplicity.

Chemical cross-linkers, although bifunctional derivatives of targeted labeling reagents, are designed to afford information on protein−protein interfaces. They can be viewed as bifunctional footprinters and, thus, we included them in a general discussion later in this section. Both footprinting and cross-linking have much in common, and viewing them together provides an opportunity to compare and contrast.

A unique function for targeted protein labeling is covalent inhibition, where a bulky labeling reagent may inhibit a biochemical reaction. The covalent attachment of the labeling reagent may occupy a binding pocket and block a functional site, inhibiting measurably protein function. Although operationally similar to footprinting, this form of protein labeling will not be covered here because this field was reviewed.[282−284] The covalent attachment of functional probes was also developed extensively during the past decade, as reviewed in a perspective by Tamura and Hamachi.[285]

3.1. Cysteine

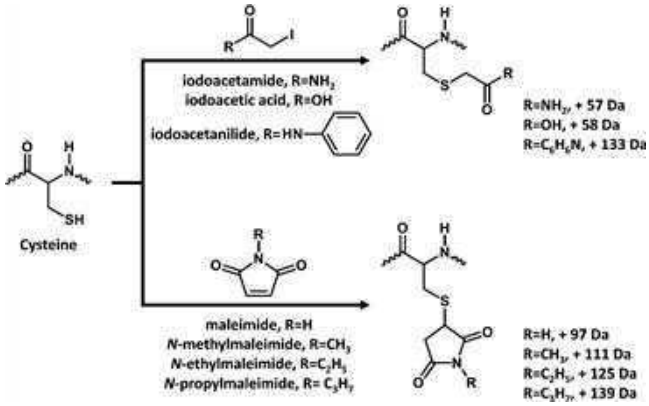
Cysteine (Cys) is one of the most reactive amino acids owing to the high intrinsic reactivity of the nucleophilic thiol side chain. Cys fulfills a variety of protein biological functions including forming disulfide bonds to stabilize structure, binding metal ions, transferring hydrogen and oxygen atoms, electrons and hydrides, catalyzing redox processes, and mediating hydrolysis.[286] Modifying Cys in a native protein, therefore, should be an informative footprinting goal as specific Cys groups throughout a protein are often important in biochemical processes.[287]

Cys labeling can be achieved through a large number of reagents, and many of them were developed as catalytic functional inhibitors (e.g., organomercurial compounds)[288] or affinity tags (e.g., functionalized biotin tag).[289] The use of these reagents is not the primary goal of this review and, therefore, the reagents will not be discussed further. Interested readers should refer to a recent book by Lundblad.[278]

The thiol group of the Cys side chain becomes an even stronger nucleophile when it deprotonates to a sulfide anion. Therefore, most of the Cys labeling reactions take advantage of the nucleophilicity of −SH, or better, $-S^-$.

The first class of Cys-labeling reagents, α-haloketo compounds, represented by iodoacetamide (IAM),[290,291] label Cys by S_N2 nucleophilic substitution (Scheme 1, top).

Scheme 1. Cys Labeling Pathways with Different Labeling Reagents



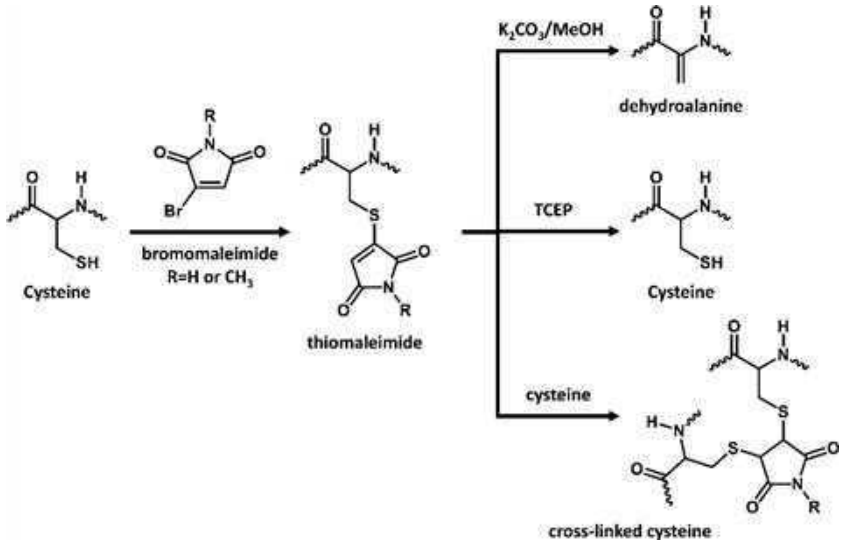
Other popular derivatizing reagents include iodoacetic acid,[292] iodoacetanilide,[293] and iodoacetate.[294] Other than labeling Cys, IAM and its derivatives also modify Met, His, Lys, Tyr, and Glu but with lower reactivity.[295−298] The reactivity of IAM not only depends on the solvent accessibilities of targeted residues but also on the local environments, a property that relates to most targeted labeling reagents.[299] The chloro- and bromo α-haloketo compounds are less reactive toward thiol groups (the leaving-group propensities for the halides are I > Br > Cl[300,301]) and, thus, chloro and bromo α-haloketo compounds are not commonly used in protein HOS analysis. These derivatizing agents, however, are more useful in affinity labeling where the reaction is facilitated because the reagent binds nearby, placing even a modestly reactive group in an appropriate position for reaction.[302]

The other class of well-established Cys labeling reagents is *N*-alkylmaleimide that labels Cys through a Michael addition (Scheme 1, bottom); a common example is *N*-ethylmaleimide (NEM).[303−305] Although NEM labels Cys residues with considerably higher reactivity and specificity than does IAM,[299] other nucleophiles including His and Lys can also be modified by NEM.[306,307] In the LC-MS analysis, however, the *N*-alkylmaleimide is prone to hydrolysis, forming an isomeric mixture of maleamic acid adducts.[298] The molar absorptivity of NEM is 620 M^{-1} cm^{-1} at 302 nm,[303] enabling UV−vis spectroscopy for analysis in NEM kinetic studies. The alkyl chain length in *N*-alkylmaleimide should be considered when the analysis is by MALDI, as that group can be responsible a nonuniform effect on signal intensity.[308]

Aside from NEM, another class of maleimide derivatives that may be effective as footprinter is bromomaleimides, which label Cys residues reversibly (Scheme 2).[309−311] In comparison to NEM, labeling by bromomaleimides involves nucleophilic substitution that preserves the double bond in the maleimide ring. Preservation of double bond allows a reaction with a second Cys residue (through Michael addition) that converts a bromomaleimide into a chemical cross-linker.[310,311] The reaction product, thiomaleimide, can be hydrolyzed to dehydroalanine or converted back to Cys by *tris*(2-

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

Scheme 2. Cys Labeling Pathways with Bromomaleimides



carboxyethyl)phosphine (TCEP) reduction. Such versatility enables novel workflows in labeling or footprinting free Cys residues.

IAM and NEM are widely used not only in MS-based footprinting that takes advantage of their reactivity with Cys but also for protecting free Cys residues, a standard practice in almost all MS-based proteomics experiments. In a pioneering study that used IAM as a Cys footprinting reagent, Wu and co-workers[312] investigated the active sites of human type II inosine 5′-monophosphate dehydrogenase (IMPDH). 6-Chloropurine riboside 5′-monophosphate (6-Cl-IMP) is a structural analogue of the IMPDH substrate and was used as an IMPDH inhibitor by modifying one of eight Cys residues in IMPDH. The results of IAM footprinting of IMPDH in its ligand-free and 6-Cl-IMP-bound states demonstrate that seven Cys residues show comparable IAM labeling, whereas Cys 331 was not labeled by IAM but by 6-Cl-IMP in the bound state, strongly suggesting that Cys 331 is critical in facilitating the enzymatic activity of IMPDH. IAM derivatives were also modified to serve as fluorescent probes for monitoring protein conformational changes in cells with relatively good spatial resolution.[313−315] Chumsae et al.[316] incorporated Cys fluorescence labeling with MS-based peptide footprinting to study the unpaired Cys residues in human immunoglobulin G subclass 1 antibody; details will be discussed in section 4.2.

NEM labeling as an MS-based protein footprinting was pioneered by Reich and co-workers,[317] who applied NEM to footprint the critical Cys residue in DNA methyltransferase. In the ligand-free state, Cys 25, 116, and 223 were modified by NEM, but upon incubation with DNA and sinefungin (a structural analogue of cofactor *S*-adenosylmethionine), Cys 223 remained essentially unmodified, demonstrating its critical role in preserving the enzymatic function of DNA methyltransferase. Notably, all the site-specific Cys labeling products were characterized by tandem MS with a magnet sector instrument, a demanding task at the time and now more readily accomplished with modern Q-ToF and orbitrap instruments.

Thiol groups in Cys residues can also be nitrosylated by active nitrogen compounds such as nitric oxide and tetranitromethane.[318,319] Reagents for Cys include cyanogen bromide,[320] selenium reagents,[321] *N*-(phenylseleno) phthalimide,[322] 5,5′-dithiobis(2-nitrobenzoic acid) (Ellman's reagent or DTNB) and its derivatives,[323] 2,2-dipyridyl disulfide and 4,4′-dipyridyl disulfide,[324] vinylpyridine,[325,326] and acrylamide. These reagents will not be discussed further as their utility in MS-based Cys footprinting remains to be established.[327]

Another group of reagents, reported by Thayumanavan and co-workers,[328] utilizes Michael acceptor α,β-unsaturated carbonyl scaffold to react effectively with nucleophilic residues. A proof-of-concept study demonstrates its promise for footprinting Cys, Lys, Tyr, and Thr residues.

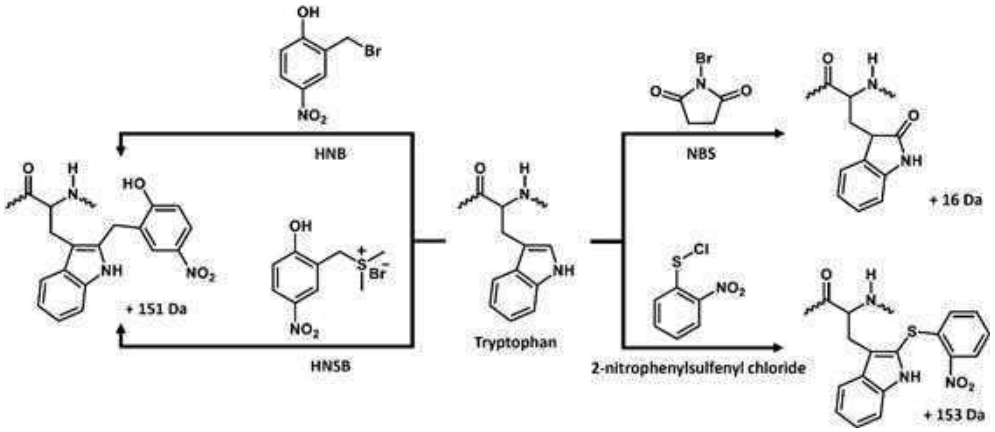
The oxidized form of cysteine, cystine, exists with a disulfide bond. Cystine footprinting, interestingly, originates from the analysis of wool and its oxidation.[329,330] To label disulfide bonds, cystine residues need to be reduced to cysteine. Many reducing reagents are available, including cyanide,[331] phosphorothioate,[332] (2*S*)-2-amino-1,4-dimercaptobutane,[333] bis-(2-mercaptoethyl)sulfone and *N*,*N*′-dimethyl-*N*,*N*′-bis-(mercaptoacetyl)hydrazine,[334] sodium borohydride,[335] and 2,3-dimercaptopropanol,[336] that can deliver effective disulfide reduction. Three reagents, however, are commonly used to reduce disulfide bonds in proteins. Dithiothreitol (DTT), originally synthesized by Evans et al.[337] in 1949, is able to reduce disulfide bonds as was first demonstrated by Cleland[338] in 1964 (DTT is also known as Cleland's reagent). DTT is prone to oxidation when used in air, with a half-life of 10 h in neutral solution (pH = 7.5, 20 °C), decreasing to 1.4 h when the pH = 8.5,[339] making it critical to prepare a fresh DTT reagent solution before use. The oxidized form of DTT has an absorbance maximum at 283 nm, which can be utilized to determine the extent of disulfide bond reduction.[340]

Unlike DTT that only functions under near-neutral pH, TCEP[341] reduces disulfide bonds rapidly under acidic conditions. TCEP is reasonably stable under acidic conditions, but its oxidation is significant at pHs greater than 7.[342] All of these features led to broad acceptance and application of TCEP in HDX workflows.

The last reagent is 2-mercaptoethanol (β-mercaptoethanol, BME), which is stable and functions under alkaline pH.[343] With the development of DTT and TCEP, BME is less used in disulfide bond reduction, but it is used in protein structural analysis.[278,344] This and the other reducing reagents are usually

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

Scheme 3. Trp Labeling Pathways with Various Reagents



employed under denaturing conditions for better disulfide reduction, which is not appropriate for footprinting.

In addition to conventional reducing reagents, a novel approach is to utilize electrochemistry to reduce disulfide bonds.[345] This development may be particularly useful in HDX experiments, as it enables fast and effective online disulfide reduction.[346] The reduction converts cystine residues to cysteine residues that are ready for labeling or "capping" to avoid disulfide bond reformation. MS-based disulfide mapping (cystine labeling) will be discussed in section 4.2.

3.2. Tryptophan

Tryptophan (Trp) is an electron-rich amino acid with a heterocyclic aromatic side chain; Trp is the lowest abundant amino acid in proteins.[347] As Trp is hydrophobic and rarely found on protein surfaces, its footprinting is not common. In the case of protein−protein and protein−membrane interactions, however, Trp contributes significantly,[348−351] motivating application of footprinting.

Trp and Tyr are responsible for most of the characteristic protein UV absorbance at 280 nm.[64] Trp is also responsible for most intrinsic protein fluorescence.[352,353] These unique features allow Trp residues to be photoactivated under UV irradiation.[354−359] In brief, Trp under irradiation by UV light results in a tryptophan radical and a hydrated electron.[354,355,357] The formation of tryptophan radical-mediated reactive oxidative species leads to hydrogen peroxide as a product,[356,360] which further oxidizes nearby residues.[359] A recent book by Lundblad[278] provides a detailed overview of Trp photoactivation.

The indole nitrogen atom in the side chain of Trp has a pK_a greater than 15,[278] making it difficult to activate for chemical modification. In limited demonstrations, Previero and Cavadore[361] developed formylation of Trp with hydrochloric acid and acetic acid, leading to *N*-formyl tryptophan as product. In another example, Linder and co-workers[362] employed 1,1,3,3-tetramethoxypropane to react with Trp residues under acidic conditions, producing a substituted acrolein with a free aldehyde. This chemistry was further developed by bringing together another immobilized hydrazide that reacts with the free aldehyde to afford chemical cross-linking.[363]

A better Trp labeling strategy is to target the C−H bond adjacent to the indole nitrogen, which is done by the four reagents as shown in the Scheme 3. 2-Hydroxy-5-nitrobenzyl-bromide (HNB) was first demonstrated to be an effective Trp labeling reagent by Koshland and co-workers;[364−366] HNB, therefore, is also known as Koshland's reagent (Scheme 3, top left). Over the intervening years, many HNB derivatives were developed to label Trp, including dimethyl(2-hydroxy-5-nitrobenzyl)sulfonium bromide (HNSB, Scheme 3, bottom left),[367] 2-methoxy-5-nitrobenzyl bromide,[368] and 2-acetoxy-5-nitrobenzyl chloride.[369] The applicability of HNB in protein HOS analysis was first demonstrated by Strohalm et al.,[370,371] where the investigators followed effective HNB labeling of Trp residues for test proteins lysozyme, cytochrome *c*, and myoglobin by using MALDI-MS. Only solvent-accessible Trp residues were labeled under physiological conditions, whereas all Trp residues did react in a denaturing environment, suggesting HNB to be a valuable Trp footprinter. Another widely used derivative of HNB is HNSB.[372−374] Unlike HNB, which must be prepared in an organic solvent (usually dry acetone), HNSB is a water-soluble sulfonium salt, which is an important factor that grants relevance in footprinting under physiological-like conditions.

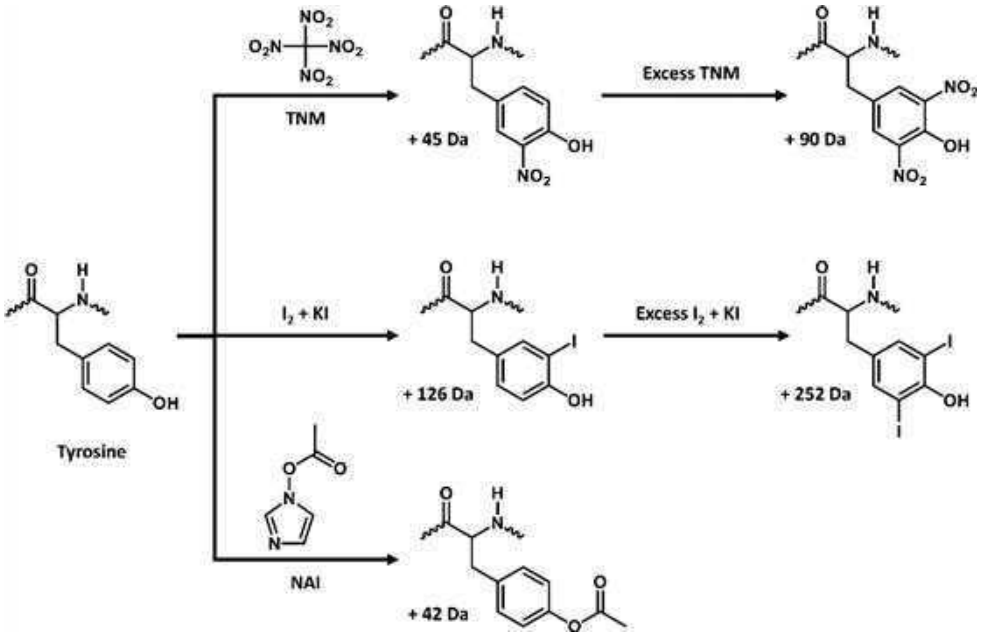
The third Trp labeling reagent is *N*-bromosuccinimide (NBS),[375,376] which converts the indole ring to an oxindole (Scheme 3, top right). Among many successful demonstrations of NBS in footprinting Trp residues, the pioneering study by Takahashi et al.[377] adopted NBS labeling to footprint Trp residues of an α-amylase inhibitor PHA-I of *Phaseolus vulgaris*. PHA-I adopts a tetrameric conformation of $(\alpha\beta)_2$, and two of eight Trp residues (Trp 188 on each β subunit) show significant oxidation by NBS. When complexed with porcine pancreatic α-amylase, however, none of the Trp residues in the PHA-I were modified, suggesting that Trp 188 in the β subunit is critical for its inhibitory activity. This is a fine example that illustrates not only the effectiveness of NBS in assessing protein HOS but also importance of Trp residues in assisting protein−ligand interactions.

The derivative of NBS, *N*-chlorosuccinimide (NCS), was also reported as an Trp labeling reagent.[378] In addition to Trp footprinting, all *N*-halosuccinimide reagents, including NBS,[375] NCS,[379] and *N*-iodosuccinimide,[379] are capable of cleaving peptide bonds that occur upon further oxidation of oxindole. NCS is the most adopted reagent for this purpose.[380] This peptide bond cleavage served as an early Trp footprinting approach but has not been developed further.

The fourth reagent that labels Trp residues is 2-nitrophenylsulfenyl chloride (Scheme 3, bottom right), which was first reported by Scoffone and Rocchi[381] in 1968. The reagent, 2,4-dinitrophenylsulfenyl chloride, was used together with 2-

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

Scheme 4. Tyr Labeling Pathways with Several Reagents



mercaptoethanol to attach covalently, on a Trp residue, a thiol group that can be further utilized for subsequent reactions.[382] The application of 2-nitrophenylsulfenyl chloride for MS-based Trp footprinting was first demonstrated by Simpson and co-workers[383] in 1993, where the investigators evaluated the effect of Trp derivatization on the biological activity and conformational stability of murine interleukin-6. Upon incubating with 2-nitrophenylsulfenyl chloride, Trp 36 and 160 showed significant labeling as identified by ESI-MS. The labeled interleukin-6 underwent a decrease in the biological activity in the murine-hybridoma growth-factor assay using 7TD1 cells and in binding affinity with its corresponding receptor. The overall conformation of labeled interleukin-6, however, remained unperturbed according to CD measurements. They concluded that Trp 36 and 160 are critical in maintaining biological activity but not the structural integrity of interleukin-6. The derivative of 2-nitrophenylsulfenyl chloride, 2,4-dinitrophenylsulfenyl chloride, was also used as a Trp footprinting reagent as demonstrated by Zhang et al.[384] in 1997.

In summary, Trp footprinting by targeted labeling reagents is less developed and used compared with that of other residues, which is, in part, because all the labeling reagents favor acidic environments.[174] Under physiological pH, these reagents are either less specific (HNB[365] and NBS[376]) or less reactive (NBS[385−387]). Other than the reagents introduced above, other Trp labeling reagents including acetyl chloride,[388] hypothiocyanous acid,[389] and chloroform (under photoactivation).[390] These reagents require harsher reaction conditions and, thus, are less suitable as reagents for protein HOS analysis.

3.3. Tyrosine

Tyrosine (Tyr) is a phenol-containing hydrophobic amino acid that is often found on the protein surfaces.[391,392] Tyrosine is a critical target of several PTMs, including nitration,[393−396] sulfation,[397,398] and phosphorylation.[399−402] These PTMs, especially phosphorylation, play an important role in cell signaling,[402] making Tyr an important target in protein footprinting.[400]

Among all the Tyr modification approaches, nitration has the richest history. An early demonstration of Tyr nitration dates back to the 1910s, when Johnson and Kohmann[403] determined the structure of nitrotyrosine and characterized Tyr nitration with nitric acid. Research over the years evolved to establish tetranitromethane (TNM) as the preferred Tyr nitration reagent, as was first reported by Wormall[404] in 1930 but not well established until the 1960s.[405,406] Tyr nitration by TNM is a representative of electrophilic aromatic substitution,[407] during which a nitro group is added to the electron-rich Tyr ring (Scheme 4, top). With excess amounts of TNM, a Tyr residue can be disubstituted as well. Although reaction between TNM and Tyr is reasonably specific, noticeable labeling of His,[406] Met,[406,408] Trp,[406,409] and Cys[410] at alkaline pH can also occur. Reactions between proteins and TNM will also induce intra- and intermolecular chemical cross-linking via Tyr residues under acidic pH,[407,411,412] providing a representative example of zero-length cross-linking[413] (see sections 3.9 and 4.4). A good illustration for TNM-facilitated zero-length chemical cross-linking is the dimerization of insulin, which was reported independently by several research groups.[412,414]

TNM labeling coupled with MS analysis as a footprinting platform has been used for protein HOS analysis. A pioneering study is by Ploug et al.,[415] where they characterized the interactions between urokinase-type plasminogen activator and its glycolipid-anchored receptor. Using MALDI-MS, they identified Tyr 57 in the receptor and Tyr 24 in the isolated growth factor-like module of the activator as the key interacting residues. In another study, Šantrůček et al.[416] evaluated the correlation between TNM labeling extents and the SASA of the Tyr residues in horse heart cytochrome *c*, hen egg-white lysozyme, and human serum albumin as model systems, using MALDI-MS as the characterization tool. Surprisingly, they found only weak correlations between the TNM labeling extents and the SASA calculated from the crystal structure. For example, Tyr 148 and 341 in human serum albumin are highly reactive toward TNM but are poorly

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

solvent-accessible. This study indicates that users should be aware of other effects on footprinting by targeted labeling reagents and understand the importance of differential experiments when the factors that govern apparent reactivities are not fully understood. Method development and evaluation studies are highly recommended.

Another Tyr nitration reagent is peroxynitrile, and it has become significant in the field of nitric oxide physiology and biological oxidations.[417,418] Its application in MS-based protein HOS analysis, however, is limited.[419] One possible reason is the low stability of peroxynitrile as compared with TNM. There are also reports demonstrating electrochemical[420] and copper-catalyzed[421] tyrosine nitration. Although promising, its application in protein HOS analysis remains to be established.

The second approach to label Tyr is iodination. As depicted in Scheme 4 (middle), iodination can lead to mono-iodotyrosine or di-iodotyrosine depending on reaction conditions. Although seldom used in protein footprinting, Tyr iodination is still valuable in protein radiolabeling[131] (has a radioactive decay half-life of 8 days[422]) and in pharmacokinetic studies.[423] To date, there are several methods for Tyr iodination, including sodium iodide (NaI) coupled with *N*-chloro-4-methyl-benzenesulfonamide (chloramine T),[424−426] I_2 with iodine monochloride (ICl),[427,428] 1,3,4,6-tetra-chloro-3a,6a-diphenyl-glycouril (iodogen),[429] lactoperoxidase,[430] and *N*-iodosuccinimide.[431] Iodination of Tyr by these reagents occurs as electrophilic aromatic substitutions, with iodonium ion or hypoiodous acid as the reactive species.[423] Other than Tyr iodination, oxidation of Met, Cys, and Trp by these iodinating reagents should not be ignored.[432]

To utilize Tyr iodination for protein footprinting, Hobba et al.[433] in 1996 first demonstrated its efficacy by probing the insulin-like growth factor binding sites of bovine insulin-like growth factor protein-2. When the latter protein was footprinted by chloramine T-mediated iodination, Tyr 60 showed protection upon interacting with its binding partner. Confirmatory evidence is that iodinated bovine insulin-like growth factor protein-2 has a significantly lower binding affinity with its binding partner, consistent with the importance of Tyr 60 whose iodination perturbs binding. Note that Tyr iodination can also be achieved by a radical-mediated pathway,[186] which will be discussed in section 5.6.

The last Tyr labeling reagent for discussion is *N*-acetylimidazole (NAI), which was first demonstrated to be a protein labeling reagent in 1963.[434] Subsequent study by the same group showed the applicability of NAI as a protein footprinting reagent by analyzing solvent accessible and buried Tyr residues in several proteins.[435] NAI acetylates phenolic hydroxyl groups through nucleophilic substitution, after which the Tyr residue becomes *O*-acetylated (Scheme 4, bottom). Acetylation of Ser and Lys also occurs occasionally.[436,437] Tyr labeling by NAI is optimized at a pH between 7.0 and 8.0 to minimize the hydrolysis of NAI and the product.[437−439] The application of NAI in mapping solvent-accessible Tyr residues was nicely demonstrated by Pucci and co-workers[440] in 1997, where the surface topology of a small de novo designed β-protein minibody was mapped by MS-based peptide footprinting. The NAI footprint revealed that Tyr 35, 39, and 59 are exposed for NAI labeling, whereas Tyr 15, 24, and 47 are silent. These results were further confirmed by TNM footprinting. Together with footprinting Lys (by acetic anhydride) and Arg (by 1,2-cyclohexanedione), the investigators obtained a comprehensive understanding of the minibody. Because NAI targets the hydroxyl group, whereas iodination and TNM activate the C−H bond on the aromatic ring, this combination of Tyr footprinting will presumably provide different conclusions when a specific Tyr residue is involved in hydrogen bonding.[440] When Tyr is involved in hydrogen bonding, the hydroxyl group is not available for NAI footprinting, whereas TNM is still capable of labeling the aromatic C−H bonds.

There have been two other acetylation reagents that are less specific as compared with NAI, 3-acetoxy-1-acetyl-5-methyl-pyrazole[441] and *N*-(2,2,5,5,-tetramethyl-3-carbonylpyrrolidine-1-oxyl)imidazole.[442] Both label aliphatic hydroxyl groups (Ser and Thr) as well.

Apart from the three major reagents introduced above, some other reagents for Tyr footprinting are acid anhydrides and acyl chlorides (target phenolic hydroxyl groups),[437,443] cyanuric fluoride (targets phenolic hydroxyl groups),[444,445] various diazonium salts (target phenolic C−H bonds),[446,447] polyhalogenated quinones (target phenolic hydroxyl groups),[448] *p*-nitrobenzenesulfonyl fluoride (NBSF, targets phenolic hydroxyl group),[449,450] diisopropylphosphorofluoridate (targets phenolic hydroxyl groups),[451,452] *p*-nitrophenyl acetate (targets phenolic hydroxyl groups),[453] and hemin-activated luminol derivatives (target phenolic C−H bonds).[454] Although this wide variety of reactive reagents demonstrates the opportunities of footprinting a given functional group, most await establishment in MS-based protein HOS elucidation.

3.4. Aspartic Acid and Glutamic Acid

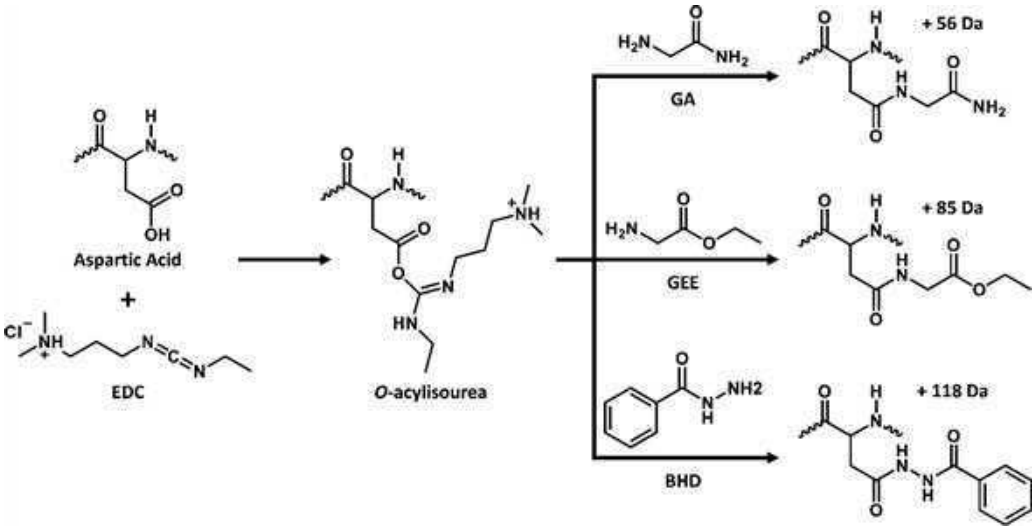
Carboxylic acid functional groups in aspartic (Asp) and glutamic acid (Glu) residues are the targeted functional groups in Asp and Glu labeling. By incubating with methanol under acidic conditions, carboxyl acids can be esterified, as first demonstrated for lysozyme in 1945.[455] Later, IAM,[456,457] *N*-ethylphenylisoxazolium tetrafluoroborate,[458] and various diazo derivatives[459−463] were also found to label Asp and Glu. These esterification reactions, however, are highly reversible and require methanol solvent and acidic conditions, making them inappropriate for footprinting under physiological conditions. Diazo reagents, although reactive under milder conditions, are not stable in aqueous solutions. For these reasons, the above reagents are less commonly used today in protein footprinting. Another class of reagent that was applied in Asp and Glu labeling is Woodward's reagent K (WRK).[464−466] Despite its success in covalent inhibition,[467,468] it has not been used for protein HOS interrogation because it may be too bulky and subject to steric hindrance to probe protein SASA in an interpretable way.

The emergence of carbodiimide-based reagents offers a welcome solution for footprinting acidic residues. The reagents are water-soluble and capable of labeling Asp and Glu residues under physiological conditions. Two early promising reagents are 1-cyclohexyl-3-(2-morpholinyl-(4)-ethyl) carbodiimide[469−472] and 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide.[471] On the other hand, their reaction products have been poorly characterized, owing in part to limited characterization techniques at the time of their study and to low yields.[472]

The breakthrough of carbodiimide Asp and Glu labeling was by Hoare and Koshland[473] in 1966, where carbodiimide serves as an activating reagent, and a second reagent is incorporated to accomplish the labeling, affording an improved labeling efficiency and higher confidence in identification. In the

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

Scheme 5. Asp and Glu Labeling Pathways with EDC-Activated Labeling Reagents As Demonstrated by Asp



original demonstration, *N*-benzyl-*N*′-3-(dimethylamino)-propylcarbodiimide (BDC) was the activating carbodiimide, and glycine methyl ester was the labeling reagent. Subsequent studies[474,475] revealed that the basis for quantification of Asp and Glu in proteins by carbodiimide-activated Asp and Glu labeling is activation of the carboxylic acid by carbodiimide to form an *O*-acylisourea intermediate (see Scheme 5). The intermediate then reacts rapidly with the nucleophilic labeling reagent to give a stable final product. In principle, any combination of water-soluble carbodiimide and a highly reactive nucleophile can facilitate such reactions. Currently, the activating carbodiimide used in Asp and Glu labeling is almost exclusively 1-ethyl-3-(3-(dimethylamino)propyl)-carbodiimide (EDC), thanks to its high solubility and stability. The labeling in the early work, measured by amino acid analysis of the protein constituents, was motivated by analytical needs to "count" the number of Asp and Glu in a protein by quantifying, through amino acid analysis, the increase of glycine labels that occurs as a consequence of labeling.

The evolution of Asp and Glu footprinting from basic chemistry to application is typical for development of footprinting other amino acid residues. The extension of the labeling to footprinting and the study of protein HOS was not imagined in the early work, probably because complete analysis of the labeling was difficult in the absence of modern MS. Now, applications are not only facilitated by improvements in analysis technology but also extended by reagent development of several nucleophiles that couple with EDC, including glycinamide (GA), glycine ethyl ester (GEE), and benzhydrazide (BHD), providing for the field a new birth (pathways for these reagents are in Scheme 5).

Asp and Glu labeling by GA (and possibly by isotopic encoded GA for more reliable identification and quantification) and glycine methyl ester[475] was not applied to study protein HOS until decades later. The applicability of GA as a MS-based Asp and Glu footprinting reagent was first demonstrated by Akashi et al.[476] in 1993. Using a differential bottom-up footprinting workflow, the investigators identified Asp 101 as a key interacting residue in hen egg-white lysozyme when binding with its inhibitor, tri-*N*-acetylglucosamine. Moreover, the investigators utilized multiple MS-based techniques, where they measured intact mass of the protein by ESI-MS, the digested peptides by LC separation, UV detection, and Frit-FAB MS, and conducted tandem-MS for selected peptides to identify the modified residues by a sector mass spectrometer. Taken together, these systematic characterization steps prefigured the modern protein footprinting workflow. Although multiple mass spectrometers were used at that time, the analysis can now be accomplished by a single LC-MS/MS system.

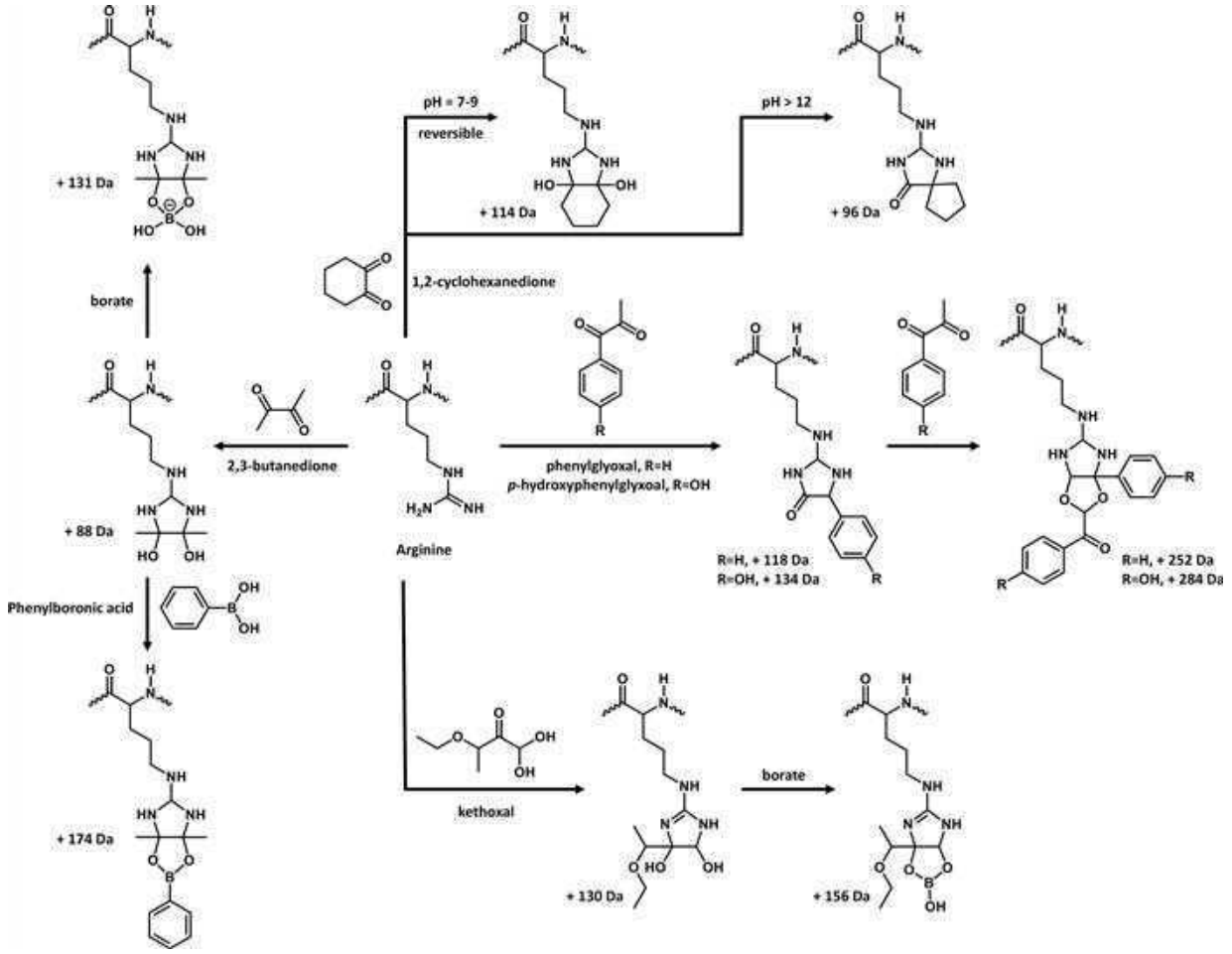
In perhaps the first EDC-facilitated GEE footprinting of Asp and Glu, Sanderson and Mosbaugh[477] investigated how PBS2 uracil-DNA glycosylase inhibitor (Ugi) protein inactivates uracil-DNA glycosylase (Ung) by forming a protein−protein complex in which Ugi mimics duplex DNA bacteriophage PBS2 uracil-DNA glycosylase inhibitor protein. Electrophoresis was used to separate forms of the protein that differ in the amounts GEE adduction. Subsequent biofunctional assays determined their activities whereas protein cleavage by cyanogen bromide and analysis by MALDI MS and amino-acid sequencing localized the GEE modification sites. Because species that are heavily modified are biologically inactive, they were able to identify in systematic studies the key Asp and Glu residues that retain the protein's activity to be Glu-28 and Glu-31 in α2-helix.

As mentioned above, modern approaches utilize online LC in combination with MS and MS/MS for better separation and quantification with a single analytical system. In an example from 2009, Wen et al.[478] utilized GEE to footprint the Fenna−Matthews−Olson (FMO) antenna protein from *Chlorobaculum tepidum*. By LC separation and MS and MS/MS identification, locations and extents of labeling of FMO protein located in a native membrane as well as on a chlorosome-depleted membrane were compared. The differential footprinting revealed for the first time the orientation of the protein associated with the cytoplastic membrane as the Bchl *a* #3 side. Later on, isotopic encoded GEE was implemented for a better and more convenient identification.[479]

Although effective, GEE has its disadvantages. The ester group in the GEE molecule will undergo hydrolysis under both acidic and basic conditions. Because the LC solvent usually contains acid to provide better ionization efficiency in MS, the hydrolysis problem can be significant.[478] Practically, both the

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

Scheme 6. Arg Labeling Pathways with Several Reagents. Mass of Boron is 11 in the Calculated Mass Shifts



GEE modified species and its hydrolyzed species were quantified, which not only complicates the data analysis but also increases errors in the quantification. To overcome these disadvantages, Guo et al.[480] developed BHD, which reacts with Asp and Glu under EDC activation (Scheme 5, bottom). Moreover, BHD labels the Asp and Glu residues with higher efficiencies than does GEE. The use of BHD also requires lower reagent amounts, resembling better the native-like environment of the target protein molecule.

3.5. Arginine

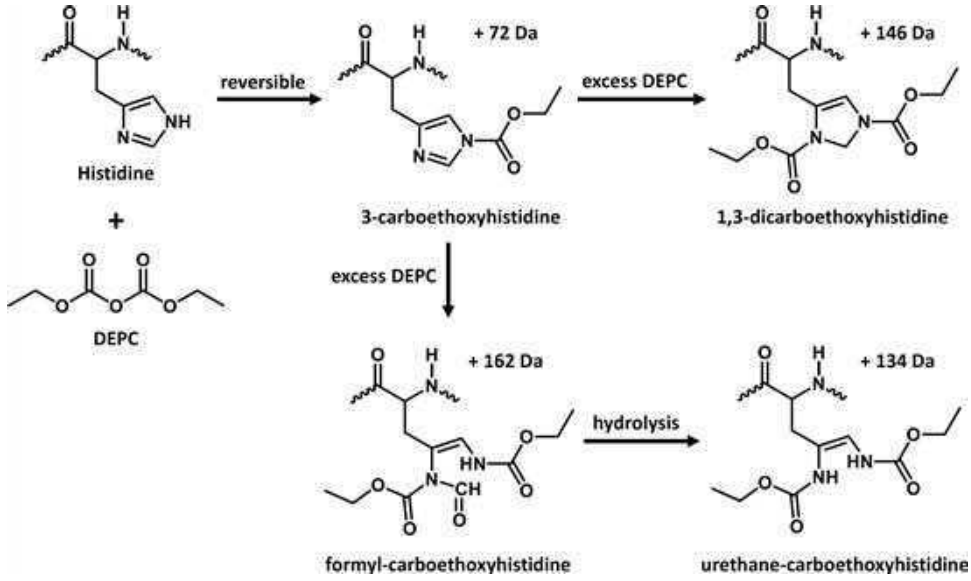
Arginine (Arg) labeling draws significant attention owing to its participation in salt bridges.[481,482] Moreover, Arg residues can form up to five hydrogen bonds with other partners, facilitate hydrophobic interactions through its three methylene carbons, and interact with other aromatic functional groups via its guanidinium group.[392] All these interactions make Arg critical in the protein folding that gives functional structures for protein and protein complexes. Post translational modification of arginine residues is also critical in regulating protein−nucleic acid[483] and protein−protein interactions.[484] Its chemical labeling, however, is challenging as the large pK_a (high basicity) of the guanidino group (greater than 12)[485] keeps Arg protonated at ∼ pH = 7 and decreases its nucleophilicity.

Arginine residues are mostly labeled by vicinal dicarbonyl reagents, resulting in a cyclic product (summarized in Scheme 6). Arg labeling by phenylglyoxal and its derivatives was first demonstrated by Takahashi[486] in 1968, where the reactivities of phenylglyoxal and the Arg residues in peptides and bovine pancreatic ribonuclease A were systematically studied. The reaction between phenylglyoxal and an Arg residue is a two-step process, as shown in Scheme 6, and both steps are reversible. The final product to which two phenylglyoxal molecules are added, however, is stable under acidic conditions but readily decomposes at neutral and basic pH.[486−488] Other phenylglyoxal derivatives, including *p*-hydroxyphenylglyoxal,[489] 4-hydroxy-3-nitrophenylglyoxal,[490] *p*-nitrophenylglyoxal,[491] and *p*-azidophenylglyoxal,[492] were also demonstrated to be Arg labeling reagents. Given a deep understanding of the reactivity and appropriate reaction conditions, phenylglyoxal and its derivatives are the most extensively used Arg labeling reagents.[278]

As modern HPLC systems when coupled with MS analysis use acidic mobile phases (if working in the positive-ion mode of the mass spectrometer), the stability of labeling products at acidic pH allow phenylglyoxal to be a MS-friendly Arg footprinter. The application of phenylglyoxal as an Arg footprinting reagent was pioneered in 1995 by Coggins and co-workers,[493] who used ESI-MS to identify Arg 19 and Arg 23 as critical functioning residues for type II dehydroquinases in *Streptomyces coelicolor* and *Aspergillus nidulans*, respectively. Because MS instrumentation with ESI was just emerging at the time of that study, key Arg residues had to be located by bottom-up MS analysis and confirmed by comparing the masses of resulting peptides with theoretical values instead of

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

Scheme 7. His Labeling Pathways with DEPC



by MS/MS. In a more sophisticated example later by McLafferty and co-workers,[494] phenylglyoxal Arg footprinting was coupled with tandem MS to allow confident identification of three key Arg residues that facilitate the function of rabbit muscle creatine kinase.

Another popular vicinal dicarbonyl reagent that labels Arg residues is 2,3-butanedione, which was first reported by Yankeelov and Crawford[495] in 1968. Upon reacting with Arg, a five-membered ring is the primary product (Scheme 6). The reaction between an Arg residue and 2,3-butanedione is highly reversible, and the labeling proceeds significantly slower than that of phenylglyoxal, limiting the application of 2,3-butanedione. In 1973, Riordan[496] discovered that borate has a dramatic effect on the reactivity of 2,3-butanedione with Arg residues, as borate stabilizes the initial diol product, thus pushing the reaction to a higher yield (Scheme 6). Similarly, Leitner and Linder[497−500] demonstrated that phenylboronic acid is also capable of stabilizing the initial adduct. This series of studies also represents an early applications of 2,3-butanedione as an Arg footprinting reagent in MS-based HOS analysis.[497−500] Currently, 2,3-butanedione labeling is mainly executed in borate buffers in the dark to minimize the photoactivation of 2,3-butanedione and to increase the labeling specificity and efficiency. 2,3-Butanedione is also capable of reacting with citrulline (the product of an Arg residue reacting with peptidylarginine deiminase), as demonstrated in a few MS-based footprinting studies.[501,502]

The last vicinal dicarbonyl reagent for Arg labeling is 1,2-cyclohexanedione, which was first shown to modify Arg residues under basic conditions (pH > 12 for effective deprotonation) in 1967.[503] This pH is not biologically relevant and greatly limits its application in protein labeling and footprinting. In 1975, Patthy and Smith[504,505] discovered that borate can stabilize the diol product between 1,2-cyclohexanedione and an Arg residue under physiological pH, a similar phenomenon as Riordan[496] observed for 2,3-butanedione.

The application of 1,2-cyclohexanedione in MS-based Arg footprinting was first reported by Przybylski and co-workers[506] in 1992, when they found that four of 11 Arg residues were modified in egg-white lysozyme. By comparing the labeled Arg residues with those in a crystal structure, the investigators found an inverse correlation between labeling efficiency and solvent accessibilities. Upon closely examining the local environments of the labeled Arg residues, they were able to show that the proton acceptor groups are in close proximity. These neighboring proton acceptors assist deprotonation of Arg residues (first step in 1,2-cyclohexanedione labeling) and facilitate Arg labeling in an intramolecular catalytic fashion. All in all, this work not only represents one of the pioneering efforts that demonstrate protein footprinting by targeted labeling reagents but also is an early example that addresses effects of local environment on the labeling efficiencies of targeted labeling reagents.

An Arg labeling reagent that is not a vicinal dicarbonyl compound is represented by 3-ethoxy-1,1-dihydroxy-2-butanone (kethoxal), which is a popular RNA footprinting reagent that reacts with the guanine group.[507,508] Kethoxal was later demonstrated to react with Lys residues in proteins.[509−511] MS-based Arg footprinting by kethoxal was first demonstrated by Fabris and co-workers[511] in 2005, where they used Arg-containing derivatized amino acids, di- and tripeptides, and model proteins to investigate the reaction product between kethoxal (structures depicted in Scheme 6) and an Arg residue and to demonstrate its capability as a Arg footprinting reagent. As the initial addition of kethoxal forms a diol product, the reaction can also include borate for higher yield.

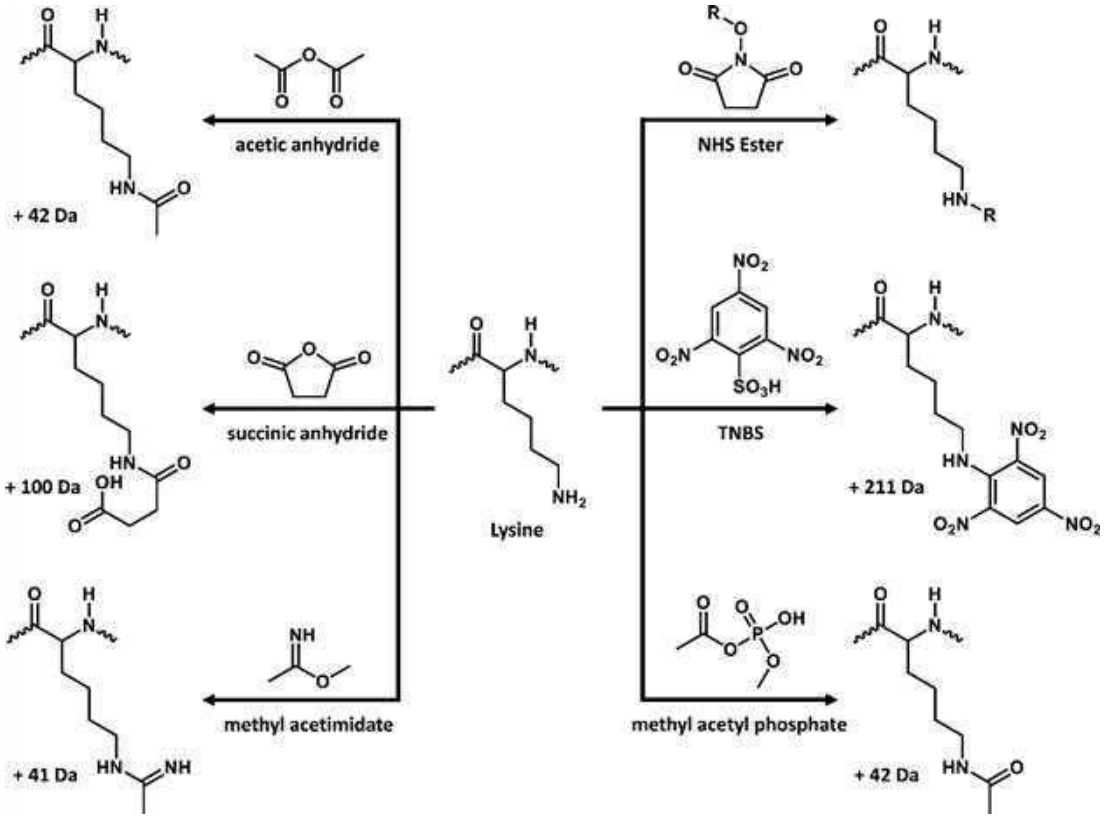
In addition to the reagents introduced above, there are other Arg footprinters, including diphenylethanedione,[512] malondialdehyde,[513−515] methylglyoxal,[516,517] and ninhydrin.[518,519] These reagents, although they show potential for reacting with Arg, have found little application thus far in Arg footprinting.

3.6. Histidine

Targeted labeling of histidine (His) has drawn considerable attention owing to its high abundance in many enzyme active sites.[520] The two nitrogen atoms on the imidazole ring impart a unique character as an acid−base catalyst,[521] which further facilitates various enzymatic transformations. In general, His residues are not critical for structural stability of proteins,[522] and surface-accessible His residues are not involved in catalytic

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

Scheme 8. Lys Labeling Pathways with Several Reagents



domains.[523] These general observations make His footprinting important for assessing structures of enzymes because buried His residues often play critical roles in biocatalytic processes.

Early His modifications were achieved by using bromoacetic acid and bromoacetate.[295,300,524] Two nitrogen atoms on the imidazole are potential nucleophilic sites for modification. As halogen atoms are good leaving groups, many α-halo carboxylic acids and amides were developed and utilized in His labeling, including but not limited to bromoacetate,[295,525−530] chloroacetate,[295,300,526] iodoacetate,[526,528,531] bromoacetamide,[532] and IAM.[533] Although effective, these reagents soon fell into disuse, primarily due to their lack in specificity. These reagents were later found to be reactive with other amino acid residues, most notably cysteine, which can be modified even more rapidly than histidine.[296] IAM today is one of the most widely used Cys alkylating reagent as mentioned in section 3.1. His labeling by bromoacetate, a reagent that can be used for Cys-free proteins, however, requires days to complete, greatly limiting broad application[525,530] and causing concerns about structural change during the footprinting and misleading overlabeling. A similar situation applies to methyl-*p*-nitrobenzenesulfonate, which can label His residues but now principally serves as a alkylation reagent for Cys.[534,535]

Another reagent that is related to the α-halo carboxylic acid is 2-bromo-1-(4-bromophenyl)ethenone (*p*-bromophenacyl bromide), which irreversibly labels His residues; this reagent is similar to other α-halo carboxylic acids. *p*-Bromophenacyl bromide was first reported to react with carboxyl groups in Asp and Glu at the active sites of pepsin.[456] Subsequent studies demonstrated its capability for labeling His residues.[536,537] Although still being used today, *p*-bromophenacyl bromide has been almost exclusively used in labeling phospholipase A_2.[278,538,539]

The most extensively used His labeling reagent to date is diethylpyrocarbonate (DEPC), whose reaction pathway with His is described in Scheme 7. Development of His labeling by DEPC reactions with intact proteins was pioneered by Fedorcsák and co-workers[540−542] in the 1960s. DEPC has advantages over the other reagents because it has relatively high specificity[278] and reactivity.[543] In the pH range between 5.5 and 7.5, DEPC primarily labels His residues, but it is also reactive with Arg, Cys, Lys, Ser, Thr, and Tyr residues.[544−547] As illustrated in Scheme 7, monomodification of His by DEPC is reversible in the presence of other nucleophiles[547] as well as under both acidic and basic conditions (half-life for ethyl 1*H*-imidazole-1-carboxylate, side chain of 3-carboethoxyhistidine, is 55 h at pH of 7, 2 h at pH of 2, and 18 min at pH of 10[548]).

With excess DEPC, the His residue can be dimodified irreversibly. The reactivity between His and DEPC increases in acidic buffers lower than pH 6−7.[549] The most significant disadvantage of DEPC in His labeling is that DEPC is insoluble in water, making it essential to add water-miscible organic solvents (e.g., methanol, acetonitrile). The amount of organic solvent needs to be carefully tuned to prevent protein denaturation, which can be measured in preliminary experiments using, for example, CD. A recent study also suggests that the presence of hydrophobic residues increases the local concentration of DEPC at the protein surface, increasing the reactivity.[143]

The potential of DEPC in MS-based protein footprinting was explored through a demonstration study that used FAB-MS to characterize the reaction product between DEPC and five model peptides.[550] By taking advantage of fragmentation

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

accompanying FAB ionization, the investigators identified major reaction products between DEPC and His residues as well as their fragmentation pathways. In one of the pioneering MS-based His footprinting studies that uses DEPC as labeling reagent, Glocker et al.[551] in 1996 mapped solvent-accessible His residues in recombinant human macrophage colony-stimulating factor β by using MALDI-MS for analysis. Among four DECP-modified His residues, His 9 and His 15 are directly correlated with biological activity whereas the activity is not affected by labeling of His 176 and His 210. In another study, Halsall and co-workers[552] used LC ESI-MS to determine the location and modification extent of solvent-accessible His residues in α_1-acid glycoprotein. In addition to its biological significance, this work nicely illustrates a modern MS-based protein footprinting workflow, where the digested peptides are separated by HPLC, followed by MS/MS analysis to determine the modification sites, and quantification by integrating extracted ion chromatograms. Recent advances of DEPC labeling were mostly by Vachet and co-workers,[143,174,544,553−556] and some of their contributions will be reviewed in section 4.

Other than the reagents introduced above, epoxides can also label His, but these reagents are not specific toward His. Very recently, Joshi and Rai[557] scanned a series of nucleophiles and found 2-cyclohexenone reacts with the His residue in different proteins with high specificity. Like DEPC, 2-cyclohexenone is poorly soluble in water, necessitating the undesirable addition of organic solvents during labeling. Although not without drawbacks, this study nicely demonstrates the potential of 2-cyclohexenone and its derivatives in His footprinting.

3.7. Lysine

Lysine (Lys) residue side chains and the N-terminus are the sites of primary amine groups. As these primary amines are most commonly found on protein surfaces,[174] labeling Lys residues becomes valuable for probing protein surface structure and its changes with respect to ligand binding or environmental change.[506]

Several different reagents were developed to label the primary amine group in Lys residues (summarized in Scheme 8). The most widely adopted approach is acylation by acid anhydrides, as exemplified by acetic anhydride in Scheme 8. Reaction between acetic anhydride and a Lys residue (Scheme 8, top-left) is highly pH-dependent, as the acetic anhydride readily hydrolyzes under acidic pH. Acetic anhydride can also react with hydroxyl groups found on Ser and Thr sites and aromatic Tyr sites and with the imidazole ring (His). Formation of *O*-acetyl tyrosine can be minimized under alkaline pH or in the presence of acetate.[558,559] The other side reactions (with Ser, Thr, and His) are also reversible and can be minimized by optimizing the conditions.[560] Although acetic anhydride is one of the first Lys labeling reagents,[561] it still is used today in MS-based protein HOS analysis.[506,562−570] Lys footprinting by acetic anhydride was first demonstrated by Fenselau, Vestling, and co-workers in 1991,[571] where they studied the melittin binding sites in calcium-bound calmodulin. By comparing Lys modifications of calcium-bound calmodulin for both melittin-bound and unbound states, they identified Lys 21, 75, and 148 as interacting Lys residues by FAB-MS analysis. In another pioneering demonstration, Przybylski and co-workers[506] analyzed the surface topology of egg-white lysozyme. By footprinting Lys residues in egg-white lysozyme and quantifying by ^{252}Cf plasma desorption MS, they obtained reactivities for six Lys residues and showed the reactivities correlate well with the SASA values derived from a crystal structure. The results highlight early examples of the potential of targeted amino acid labeling in combination of MS analysis, primitive by today's opportunities, for analyzing protein HOS.

Another popular anhydride for Lys modification is succinic anhydride (Scheme 8, middle-left).[572] Under physiological pH, Lys is usually positively charged as the primary amino group is protonated. When Lys reacts with succinic anhydride, the product is a carboxylic acid that deprotonates under neutral pH (Scheme 8, middle-left). In other words, Lys modification by succinic anhydride results in charge reversal, which usually leads to dissociation of multimeric protein aggregates and has been widely utilized to solubilize insoluble proteins.[573−576] Another unique feature of succinic anhydride is that it can efficiently modify Lys residues at acidic pH,[174] which is significantly different than reactions of acetic anhydride. Other than amino groups, succinic anhydride is also found to be reactive toward Tyr residues.[577] In the earliest example of Lys footprinting with acetic anhydride for MS-based protein HOS analysis (1996), Przybylski et al.[578] studied the *Rhodobacter capsulatus* general diffusion porins. MS together with X-ray crystallography were used to identify the N-terminus and three other modified solvent-accessible Lys residues (Lys 46, 298, and 300). Charge reversal upon succinylation induced an increase in cation selectivity and single-channel conductance of the porin.

Other than acetic and succinic anhydrides, the acylation of Lys can also be done with other anhydrides, including citraconic anhydride,[579,580] maleic anhydride,[581,582] trimellitic anhydrides,[583] phthalic anhydride,[584] 3-hydroxyl phthalic anhydride,[583] hexahydrophthalic anhydride,[585,586] methyltetrahydrophthalic anhydride,[585,587] succinic anhydride,[588,589] *cis*-aconitic anhydride,[588,590] and fatty acid anhydrides.[591]

Unlike succinic anhydride that reverses the charge of Lys upon labeling, methyl acetimidate labeling retains the positive charge on the Lys residue at physiologically relevant pH (Scheme 8, bottom-left).[592] Methyl acetimidate is a representative reagent in the imidoester family that targets α-amino groups. Other popular imidoester reagents include ethyl acetimidate,[593] methyl isonicotinimidate,[594] *S*-methylthioacetimidate,[595] isethionyl acetimidate,[596] and *S*-sulfethylthioacetimidate.[597] The latter two are sulfonic salts, are membrane-impermeable, and can be utilized to study soluble domains of cell membrane and membrane-bound proteins, providing spatial specificity (inner and outer membrane sides).[598] *S*-Methylthioacetimidate also labels Cys, resulting in a methyl disulfide with the sulfhydryl group.[599] The most significant drawback for imidoesters is their limited half-life in water owing to the rapid hydrolysis of the esters.[592,600] Therefore, imidoesters have not been often used in protein HOS analysis. The bis-imidoesters, however, are a popular cross-linking reagent, and details on this application are in section 3.9.1.

N-Hydroxysuccinimide ester (NHS-ester, Scheme 8, top-right) derivatives are another class of reagents for Lys labeling. NHS-ester was adopted in protein modification by Blumberg and Vallee[601] in 1975. Its sulfonic salt, sulfo-NHS ester, was developed by Staros[602,603] in 1982 and was widely adopted owing to its improved solubility.

Other than Lys, NHS-esters also react with Ser, Thr, and Tyr residues.[604] The NHS-ester is also frequently used for biotinylation of proteins, a process that is of great importance in protein purification because the modified protein binds

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

tightly with streptavidin (and also avidin).[605,606] The biotin derivative of the NHS-ester has been widely used in MS-based Lys footprinting and in locating binding sites.[607−612] In pioneering work, Knock et al.[608] in 1991 footprinted *Aplysia* egg-laying hormone with a biotin-functionalized NHS ester. Time-dependent labeling of the hormone revealed that among three amino groups (Lys 36, Lys 8, and N-terminal NH_2) that were labeled by NHS-ester; Lys 36 is the most solvent-accessible, whereas the N-terminal amino group is the most protected. In combination with data from a bioactivity assay, they concluded that Lys 8 and the N-terminal amino group are critical for preserving bioactivity.

Another biologically significant reagent that is based on NHS-ester chemistry is the Bolton−Hunter reagent, succinimidyl-3-(3[^{125}I],4-hydroxyphenyl)propionate, which plays an important role in protein radiolabeling (^{125}I is radioactive).[613] The fatty acid of sulfo-NHS is also used for footprinting because it can modify membrane proteins.[614] NHS-esters are also widely used in chemical cross-linking of proteins, as discussed in section 3.9.1.

2,4,6-Trinitrobenzenesulfonic acid (TNBS) reacts with primary amines with reasonable specificity (Scheme 8, middle-right), and the incorporation can be easily measured by utilizing its optical absorbance at 420 nm, a characterization strategy that is readily adoptable.[615−617] These unique features to footprint free amino groups (N-terminal α-amine and ε-amine in Lys residues) in proteins at the time (1960s) were utilized when routine spray or desorption ionization coupled with high resolution MS analysis were not available. Moreover, protein HOS analysis and conformational changes can also be followed with TNBS coupled with absorbance spectroscopy in a protocol that is similar to protein HOS analysis with fluorescence spectroscopy (absorption versus emission).[618−620] TNBS footprinting of the N-terminal α-amino group induces a strong hydrophobic shift in the LC separation, and this feature was utilized in peptide separations of complex proteomic samples.[621] The combination between TNBS labeling and MS detection in protein HOS interrogation seems promising but is not yet well established.

Similar to TNBS, another potentially favorable yet poorly demonstrated footprinting reagent is methyl acetyl phosphate (Scheme 8, bottom-right).[622] As the phosphate group carries negative charges under physiological conditions, methyl acetyl phosphate becomes a useful probe for characterizing anion binding sites.[623−625] A similar phosphate-containing Lys labeling reagent is pyridoxal-5′-phosphate (PLP), which also preferentially binds to anion binding sites prior to Lys labeling.[626,627] Its fluorescence[628] and UV absorbance[629] properties were utilized to study protein conformational changes. The application of PLP in MS-based protein footprinting, however, is not well tested. One problem is that the phosphate group will cleave to lose phosphoric acid during normal MS/MS analysis, preempting useful sequence-specific fragmentation. This problem is resolvable by using ETD in the MS/MS mode.

Other than the reagents discussed above, dinitrofluorobenzene,[6] 4-chloro-3,5-dinitrobenzoic acid,[630] cyanate,[631] and aldehydes (formaldehyde, glycolaldehyde) coupled with reducing reagents (sodium borohydride, sodium cyanoborohydride or amino boranes)[632,633] can also label Lys residues. Reducing sugars are also capable of reacting with the ε-amino groups in Lys.[634] These reagents, however, were not designed or used for protein HOS analysis and, thus, will not be covered further.

3.8. Other Residues

Methionine (Met) is an attractive footprinting target, yet its labeling by targeted reagents is challenging. Despite being easily oxidized under various conditions, Met is only reactive with targeted labeling reagents under acidic conditions, where Met functions as a nucleophile.[635] Alkylating reagents including iodoacetate,[297] methyl iodide,[636] iodoacetamide,[637] and iodoacetic acid,[638] all of which react with Met in a pH range of 2−4. Furthermore, there are only a few demonstrations where these reactions occur under gentle conditions (iodoacetate with Met residues in pig kidney general acyl-CoA dehydrogenase pH = 6.6).[639] Thus, the likely disruption of protein native structure by placing it in an acidic environment make this reagent too risky as a footprinter.

Site-specific backbone cleavages of Met-containing peptides by cyanogen bromide (CNBr) was extensively used in the 1980s.[278,640,641] Upon reacting with CNBr, Met forms a five-membered ring with methyl thiocyanate as the leaving group. The five-membered ring subsequently cleaves to yield a peptide homoserine lactone and subsequently a peptide homoserine after hydrolysis. This reaction was usually coupled with top-down MS for analysis,[177,642] but top-down analysis becomes less necessary owing to the increasing effectiveness of current methods for inducing fragmentations via MS/MS.

Another study by Reid et al.[643] utilized ω-bromoacetophenone as a Met alkylating reagent, and subsequent CID activation in MS/MS analysis caused a neutral loss that leaves a single characteristic product ion for Met. All in all, Met footprinting by targeted labeling reagents seems to be ineffective in contrast to Met oxidation[644−646] and the fast-labeling by hydroxyl radicals in the synchrotron and fast photochemical oxidation of protein platforms (see sections 5 and 6).

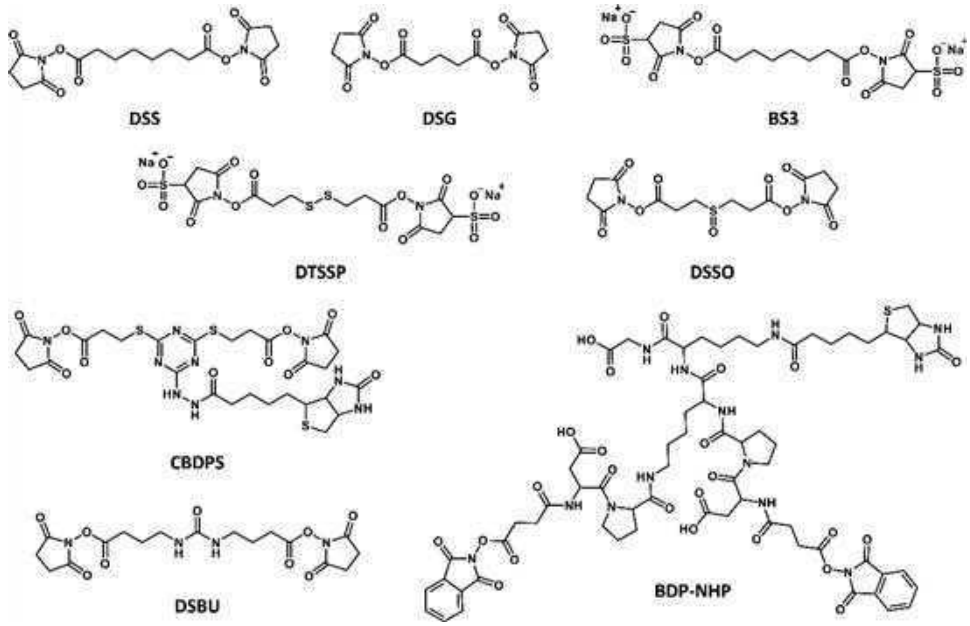
Serine (Ser) and threonine (Thr) both contain hydroxyl groups in their side chains, making these two residues as likely targets for different enzymatically introduced PTMs including *O*-linked glycosylation[647−649] and phosphorylation.[650] Introducing chemical modifications, however, is quite challenging. Both Ser and Thr can be acylated at lower reaction rates by *p*-nitrophenyl acetate,[651] a reagent that primarily reacts with the stronger nucleophilic Tyr[652] and Lys.[653] When located at the N-terminus, Ser and Thr can be oxidized by periodic acid to form an aldehyde.[654] Formic acid will esterify Ser and Thr residues at high concentrations in a reaction known as *O*-formylation.[655] NHS-esters[604] and DEPC[544] also react with Ser and Thr as side reactions. These reactions, however, are not of sufficient yield to footprint Ser and Thr residues in proteins under physiological conditions.

Asparagine (Asn) and glutamine (Glu) are the amide derivatives of Asp and Glu, respectively. The primary modification they undergo is deamidation, which occurs naturally as a PTM[656] and is also observed during cyanogen bromide/formic acid cleavage.[657]

Glycine (Gly), alanine (Ala), valine (Val), leucine (Leu), isoleucine (Ile), and proline (Pro) are generally inert to most chemical modifications that target functional groups as their residue side chains consist of unreactive H, alkyl, and cycloalkyl groups. Although the phenyl group of phenylalanine (Phe) can undergo electrophilic substitution, its chemical activation in a protein is also challenging under psychological

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

Scheme 9. Common NHS-Ester Cross-linkers



conditions because the phenyl group is not reactive. Although Leu, Ile, Pro, and Phe are not activated by normal chemical processes, remarkably they can be modified by free-radical reactions, as will be covered in section 5.

3.9. Chemical Cross-linkers

Chemical cross-linking occurs via the formation of covalent bonds between interacting proteins and within a protein on adjoining regions.[163−165] This structural proteomics tool often uses a bifunctional chemical reagent and can be viewed as bifunctional footprinting. Many reagents have been developed, promoting more and more applications of chemical cross-linking. The character of a cross-linking reagent is determined by the nature of chemistry, spacer length, and built-in functional groups. Cross-linking can be an effective complement to monofunctional reagents that serve as footprinters, and therefore the subject is briefly covered in this review particularly to make that point. The emphasis is on the commonly used reagents, which fits the theme of footprinting. Indeed, the location of monolinks can be viewed as a footprint. Our discussion is organized around the cross-linking reagents as follows.

3.9.1. Amine-Reactive Cross-linkers: NHS-Ester and Imidoester. *3.9.1.1. N-Hydroxysuccinimide Ester.* NHS-esters are the most widely used cross-linkers in the field. The ester group undergoes attack by nearby nucleophilic sites (e.g., primary amines on Lys side chains or at the N-terminus of a protein/peptides, hydroxyl groups on Ser or Thr, and even sulfhydryl groups on Cys) to form C−X bonds (X = N, O, S). The differing reactivities with various nucleophilic groups, however, depends on reaction conditions and can introduce bias. Primary amines possess the highest reactivity at physiological pH, and their reactivity can be further enhanced when the pH is >7.[603,658] Hydrolysis of NHS-esters has a half-life of 4−5 h at pH 7 and 0 °C;[659] the hydrolysis, however, is accelerated under alkaline conditions and at higher temperature.[660,661] Other possible products (i.e., esters and thioesters formed with Ser/Thr or Cys) are less stable and undergo more rapid hydrolysis.[661] Although the reaction is favored under acidic conditions (pH = 6.0), cross-linking at pH = 7 still occurs. The current consensus is to consider all possible residues as sites for reactions in simple protein systems, whereas only Lys residues and the N-termini are recommended for cross-linking a large complex or in a whole proteome study to minimize dispersion of cross-links and maximize the probability for detecting the cross-links.[662]

The first NHS-ester originated in late 1970s as a homobifunctional cross-linker.[663] Since then, many NHS-esters possessing useful physical and chemical properties have become suitable for answering a range of biological questions. Designs include cross-linkers with high hydrophobicity and zero charge, being lipophilic and membrane-permeable and ideal for intramembrane cross-linking (e.g., disuccinimidyl suberate (DSS) and disuccinimidyl glutarate (DSG), Scheme 9). Other cross-linkers incorporate a sulfonate group, which imparts water-solubility and avoids steps of predissolution in organic solvents that could perturb aqueous conditions and cause some protein denaturation.

Currently, sulfo-NHS esters have become the dominant cross-linkers in characterizing soluble proteins and their interactions. bis(Sulfosuccinimidyl)suberate (BS3, Scheme 9), the most extensively used sulfo-NHS cross-linker, can be encoded with deuterium to facilitate better cross-link assignment. The spacer length is 11.4 Å, allowing cross-linking for residues separated by 27 Å measured between two α carbons in an amino acid residues that are linked (termed Cα).[664]

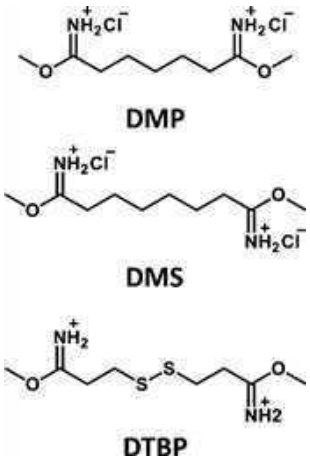
New developments in NHS-esters resonate well with the rapid growth in proteomics research enabled by advanced MS. More complicated systems with larger-scale protein candidates require better sequencing in MS/MS for a confident cross-linking identification. The disadvantage of using conventional cross-linkers is that the user must sequence species made up of two cross-linked peptides and the cross-linker, introducing complexity into fragmentation and challenging the acquisition of high-quality, readily interpretable MS/MS data. Cleavable cross-linkers are, therefore, designed to address that problem.[665] NHS-esters have incorporated chemical-cleavable motifs (e.g., S−S bond in 3,3′-dithiobis-

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(sulfosuccinimidylpropionate) (DTSSP)[658]) and CID-cleavable functional groups that include C−S bonds (e.g., disuccinimidyl sulfoxide (DSSO)[666] and cyanurbiotindimercaptopropionyl succinimide (CBDPS)[667]) and C−N bonds (e.g., disuccinimidyl dibutyric urea (DSBU)[668] and *N*-hydroxyphthalamide ester of biotin aspartate proline (BDP-NHP)[669]). Chemical structures of these cross-linkers are shown in Scheme 9. In addition, some NHS-cross-linkers contain biotin tags (e.g., CBDPS and BDP-NHP) to provide a means of enriching the cross-linked species by affinity purification.

3.9.1.2. Imidoester. Imidoesters, introduced in 1966 as one of the oldest reagents for protein cross-linking, are water-soluble reagents that react specifically with primary amines (see Lys footprinting, section 3.7).[670] The functional imidate group reacts to form an amide bond via intermediates at an optimized pH range of 8−10.[660] One major advantages of using imidoester cross-linkers is that the reaction product, an amidine, carries one positive charge. Charge removal that occurs with most lysine-targeting cross-linkers may disrupt intramolecular and intermolecular interactions and distort protein conformation, giving a biased result. The lifetime of imidoesters, however, is limited by rapid hydrolysis, being less than 30 min.[671,672] The most commonly used cross-linkers in this class are dimethyl pimelimidate (DMP), dimethyl suberimidate (DMS), and a cleavable analogue, dimethyl 3,3′-dithiobispropionimidate (DTBP), as depicted in Scheme 10.

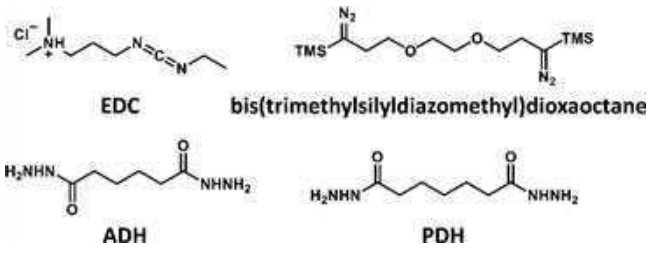
Scheme 10. Common Imidoester Cross-linkers



3.9.2. Carboxylic Acid-Reactive Cross-linkers: Carbodiimide and Dihydrazides. The most widely used carbomiimide is EDC (Scheme 11), also known as a "zero-length" cross-linker.[413,673] When this reagent facilitates cross-linking of carboxylate groups and primary amines, there is no spacer chain inserted between the targeted proteins but rather an amide bond (~3 Å) between COOH and NH_2-containing side chains.[413,673] One critical reaction intermediate, *O*-acylisourea, is not stable in aqueous condition (see section 3.4 and Scheme 5), and it continues to degrade back to the carboxyl group. Therefore, sulfo-NHS is often incorporated in the cross-linking protocol to transform the *O*-acylisourea into a stabilized NHS-ester for more efficient conjugation. EDC cross-linking shows the highest reactivity at pH 4.5 and still a moderate reaction efficiency at neutral pH.[163] 2-(*N*-Morpholino) ethanesulfonic acid (MES) and phosphate buffer are compatible with carbodiimide reagents; however, more concentrated reagents should be used in the latter buffer owing to the reduced reactivity.

Other commonly used cross-linker for Asp and Glu are dihydrazides. In 2008, Kruppa and Novak[674] first reported its use in combination with EDC activation. The reaction was carried out in acidic conditions (pH = 5.5), which is not physiologically friendly, to give a low cross-linking yield. Later in 2014, Aebersold and co-workers[675] reported a new coupling reagent (other than EDC), 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methyl-morpholinium chloride (DMTMM), and achieved significantly increased reaction yields and better biocompatibility at neutral pH. Recently, Lei and co-workers[676] developed a coupling reagent-free cross-linker, bis-(trimethylsilyldiazomethyl)dioxaoctane (Scheme 11), that offers selectivity and efficiency under physiological conditions. Isotopic encoded dihydrazides are now commercially available with different spacer lengths (e.g., adipic acid dihydrazide (ADH) and pimelic acid dihydrazide (PDH), Scheme 11).

Scheme 11. Common Carbodiimide and Dihydrazide Cross-linkers



3.9.3. Sulfhydryl-Reactive Cross-linkers: Maleimide. Maleimide cross-linkers react specifically with sulfhydryl groups at pH 6.5−7.5 to form stable thioethers.[305,677] Under alkaline conditions (pH > 8.5), primary amines are also possible targets but not Tyr and His.[678] One major concern in maleimide cross-linking is to conjugate free sulfhydryl groups, requiring prior reduction of existing disulfide bonds; breaking the −S−S− bond has high potential to distort protein native structure. A more common way of employing maleimide chemistry is in combination with NHS-esters, where both functional groups are coupled onto a heterobifunctional cross-linker. These cross-linkers can be incubated with a protein sample in a stepwise manner, reducing disulfide bonds after reaction with primary amines.[660] In addition, to ensure the best cross-linking performance, many thiol-containing compounds (e.g., DTT and BME) should be eliminated in the reaction buffer.

3.9.4. Carbonyl-Reactive Cross-linkers: Hydrazide. Hydrazides are carbonyl-reactive reagents that exhibit highest reactivity at pH 5−7. Aldehydes and ketones are major targets that can be found in glycoproteins introduced by oxidation of the polysaccharide.[660] The hydrazone bonds thus formed are moderately stable in aqueous solution and can be further secured by reducing the double bond to a secondary amine.

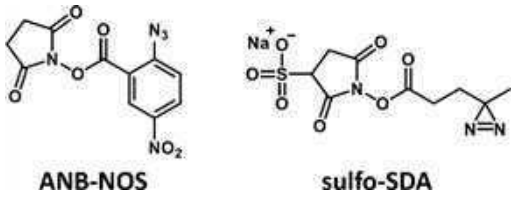
3.9.5. Photoreactive Cross-linkers: Aryl Azide and Diazirine. Photoreactive cross-linkers are generally non-specific owing to the high reactivity of the intermediate species, a nitrene, carbene, or free radical. The most adopted chemistry utilizes nitrenes and carbenes, whose precursors are azides or diazirines, respectively. Among the three major categories, phenyl azides play a dominant role in current

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

applications. Different substituents on the aromatic rings shift their UV absorption dramatically; therefore, they are selected with a biological question in mind.[660] Nitrophenyl azides, which can be activated at 300−460 nm, are compatible with most studies because this long wavelength causes minimal damage to protein molecules. Upon UV activation, the nitrene diradical can insert into most chemical bonds but with a preference for active C−H and N−H sites.[679] Diazirines are a relatively new class, first reported in the 1990s, and they have better photostability than phenyl azides.[680]

Carbene diradicals are usually generated photochemically at ~355 nm, and they show reactivity with both single and double bonds. Like the nitrene diradical, heteroatom-H bonds undergo easier insertion. Some diazirine reagents are designed as analogues of amino acids (e.g., photo-Leu and photo-Met[681]) that can be incorporated into the protein sequence during translation for in situ radical generation. Such idea sets the foundation of photoaffinity labeling (PAL). High reactivity of carbenes and nitrenes can also cause problems in cross-link identification. The use of homobifunctional photoreactive cross-linkers tends to make product analysis complex; therefore, one radical precursor is usually combined with an amine- or sulfhydryl-reactive motif. The commonly used cross-linkers are *N*-5-azido-2-nitrobenzoyloxysuccinimide (ANB-NOS)[682] and sulfosuccinimidyl 4,4′-azipentanoate (sulfo-NHS-diazirine, sulfo-SDA), as seen in Scheme 12.

Scheme 12. Chemical Structure of Common Photoreactive Cross-linkers, ANB-NOS, and Sulfo-SDA



3.9.6. Summary. We have seen extensive development of chemical cross-linking during the past two decades, enabled by evolving MS instrumentations, new data processing software, and novel cross-linkers. Chemical cross-linkers are bifunctional protein footprinters, which are governed by similar principles as protein footprinting by targeted reagents. The bifunctional nature of the reagent allows generating distance restraints between two cross-linked residues, and such restraints can be further utilized to locate protein−protein binding interfaces, to assess protein topologies, to characterize the protein interactome in a large complex, and to facilitate protein docking and modeling. Some applications of chemical cross-linking are discussed in section 4.4.

3.10. Conclusion and Perspective

Protein footprinting by targeted reagents takes advantage of early research that provided many effective chemical reactions for protein labeling. These reagents can now serve as probes to characterize protein SASA and as a basis to reason about protein HOS. The sites of modification can now be efficiently measured by MS, owing to the rapid development of LC separations and hybrid mass spectrometers with MS/MS capabilities. The chemical cross-linker, which is viewed as a bifunctional protein footprinter, when connected to two proteins offers valuable distance restraints between the proteins. These targeted footprinting reagents coupled with MS detection now are utilized to answer several biological questions and are becoming increasingly significant in structural biology, as will be discussed in section 4.

A limitation of targeted footprinting is that reagents primarily react with one or a few residues that contain functional groups, including bases, acids, nucleophiles and aromatic rings, leaving a sizable fraction of the amino acid residues to be "silent". Bulky reagents may be unable to penetrate an interface, suggesting that the lack of reaction may not signal the lack of an interface. Moreover, the footprinting reactions are relatively slow, making most targeted footprinters unable to characterize fast protein dynamics (e.g., protein

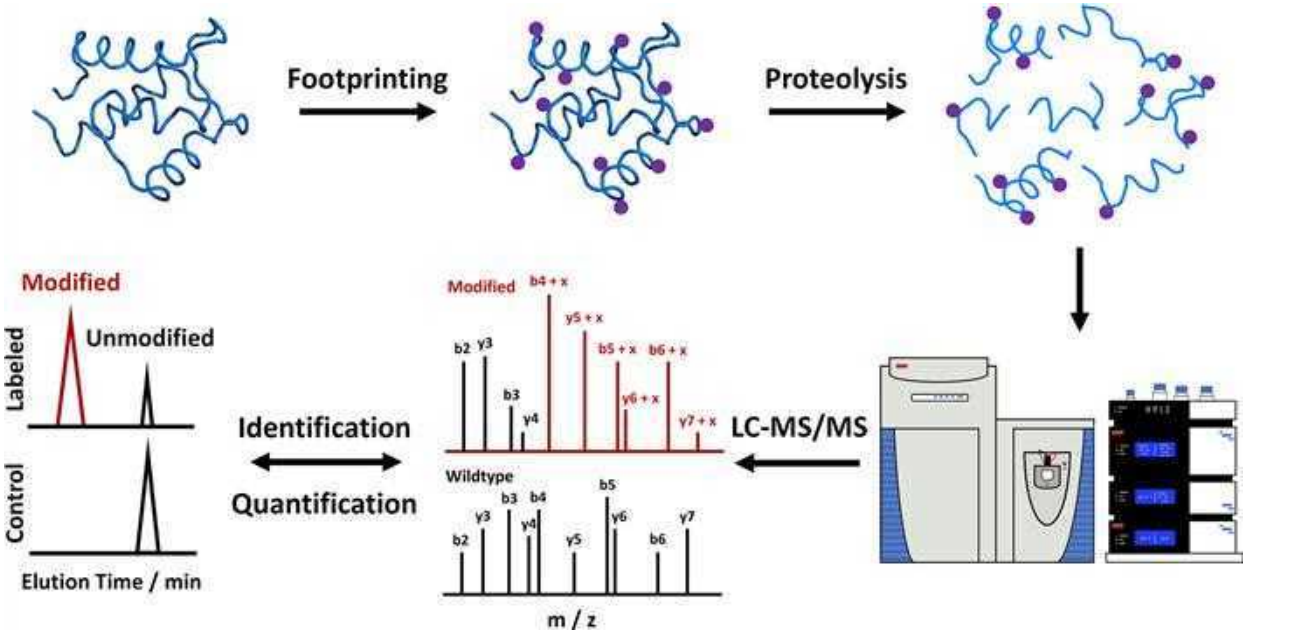


Figure 5. Schematic illustration of bottom-up protein HOS analysis through irreversible labeling approaches.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

folding and protein aggregation). To overcome these drawbacks requires a new approach that has broader residue coverage and higher reaction rate. Fast footprinting reagents represented by reactive radical species meet these requirements as discussed in section 5.

4. APPLICATIONS OF TARGETED-LABELING REAGENTS

Targeted labeling has become a mature field in structural proteomics. Research in targeted-labeling has evolved from developing reagents and methods to addressing biological questions or, in many cases, developing a method and using it to solve a biological problem. The maturation of the field runs parallel with MS instrumentation and method development. The key enabling technologies are (1) chemical reactions that are suitable for protein labeling and (2) high sensitivity, specificity, and speed analysis available with modern mass spectrometer systems. These steps are mature and permit careful application where perturbation of structure by the inserted reagent can be minimized or even eliminated.

A general workflow for bottom-up protein HOS analysis through irreversible protein labeling (through either targeted labeling reagents or reactive radical species) is shown in Figure 5.

In the general workflow, a protein of interest is first labeled in a native or native-like environment, during which the labeling reagent reacts with the solvent-accessible amino acid side chains. Proteolytic digestion cleaves the protein molecules into peptide fragments that are then submitted to LC separation and MS analysis. By comparing the mass-to-charge (m/z) ratios of the precursor and its corresponding fragments, the investigator can assign chromatographic peaks to either the unmodified or the modified regions of the protein via its "proxy" peptides. Quantifying the modification extent can be achieved by integrating the corresponding chromatographic peaks (via extracted ion chromatograms) and comparing areas assigned to "modified" and "unmodified" peptides. When LC is insufficient to separate modified peptide isomers, fragmentation by ETD to allow comparison of the resulting product ion intensities may be an alternative approach for accurate quantification.[683] Assignment of the precise location of the modification can often be ascertained by MS/MS, online with the chromatography. The underlying assumption is that a residue side chain that is more solvent-accessible will be modified more heavily as compared with a residue that is buried inside the 3D structure of a protein. Coarse aspects of protein HOS can be elucidated by analyzing the modification fractions of the intact protein, "proxy" peptides, and even single-residue levels, taking advantage of the specificity of the labeling chemistry.

Conventionally, protein footprinting is almost exclusively executed in a differential manner for which the labeling fractions for a specific region or residue are compared for different protein states. The difference in states can be achieved by incorporating a binding partner, changing the temperature, varying the denaturant concentrations or pH, or modifying the protein through mutagenesis. The changes in protein conformations that occur with these perturbations can be elucidated with spatial resolution limited by the number and spacing of the targeted residues. Recent developments also forecast that elucidation and prediction of protein HOS with protein footprinting data for a single state can be done, as will be covered in later sections.

Generally, the labeling is irreversible, allowing a wide range of postlabeling sample handing procedures (i.e., sample enrichment, deglycosylation, and longer LC gradients) as compared to that used for HDX, where labeling is reversible and precautions are necessary to minimize back exchange. The irreversibility feature also allows addressing problems that HDX cannot.

Two disadvantages can limit the approach. First, the coverage is limited to the targeted residue(s), and the number of them in the protein of interest can be low. Second, because the labeling chemistry is generally slow (minutes), there is always a question whether the targeted protein undergoes a conformational change in the early stages of labeling and then continues to be modified to yield a composite and misleading footprint.

Given the maturity of the field and the large number of applications in the literature, we will select several applications of protein footprinting by targeted labeling reagents. The applications will be grouped by the biological questions that are addressed. We will limit our review to publications since 2009 and primarily highlight demonstrations that either embody novel developments or implement improved workflows. For work prior to 2009, we recommend a comprehensive review by Mendoza and Vachet[174] for a detailed historical perspective and early applications. For a discussion of reagent and method development, the reader is directed to section 3, where pioneering work is discussed.

4.1. Analyzing Metal-Ion Binding

Metal ion−protein interactions facilitate many cellular functions, especially signal transduction, as 25−50% of all proteins found in organisms contain metal ions.[684−687] Metal ions interact with proteins by coordination whereby electron pairs from the amino acid side chains bond with the vacant orbitals of the metal ions.[687] We will classify the protein-binding metal ions into two distinct categories based on their coordination properties. Soft metal ions represented by Zn(II), Cu(II), Cd(II), and Fe(II) bind with His and Cys, whereas hard metal ions (e.g., Ca(II) and Mg(II)) usually coordinate with oxygen-containing functional groups in Asp and Glu.[688,689]

Many biophysical approaches have been applied to characterize protein−metal ion interactions, including MS,[690] magnetic resonance,[691] X-ray crystallography,[686] cryo-EM,[692] and fluorescence.[693] One approach is native MS whereby complexes are directly observed, but this begs the question of whether the gas-phase structures are secure and trustworthy reporters of those in the liquid phase. Footprinting avoids this question by using the mass spectrometer only as a measurement approach, not a reaction vessel. Footprinting can be conducted by HDX,[239,694] fast irreversible labeling,[695,696] or slow irreversible labeling. All can reveal metal-ion binding sites and the corresponding protein conformational changes induced by the metal binding. HDX probes the binding-induced backbone conformational changes, whereas Asp and Glu, commonly present in a binding site of a hard metal ion, do not react well with most of the fast-labeling free radicals in that version of footprinting. Although capable, HDX and fast free-radical footprinting (next section) report the consequences of metal-ion binding by modifying regions of the protein (i.e., nearby peptide bonds in HDX) that may not be directly involved in binding. A direct footprinting experiment seems better suited for probing metal-ion binding because reagents

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

that react with ligand binding sites probe directly the metal ion binding unlike indirect methods that report on conformational changes occurring nearby. All of this assumes that the structural perturbation caused by the footprinting is minimal.

As examples of a direct, targeted footprinting reagents, DEPC affords modification of His, carbodiimide-activated reagents (e.g., GEE) label Asp and Glu, and succinimide derivatives, iodoacetic acid, and *N*-alkylmaleimides, for example, label Cys.[174]

4.1.1. Protein Binding with Soft Metal Ions. Protein binding of soft metal ions, Cu(II) and Zn(II), is the most extensively studied. In a pioneering assessment of Cu(II) binding sites in a prion protein (PrP, from human), Qin and Westaway[697] in 2002 applied DEPC to footprint histidine-dependent Cu(II) binding sites and used MALDI for the analysis. No histidines are protected upon Ca(II), Mn(II), or Mg(II) binding, one or two residues are protected by binding of Zn(II) or Ni(II) ions, and five histidines are protected upon Cu(II) binding. Through a classical bottom-up approach, the investigators pinpointed the five histidine residues in mouse PrP23−231 that show reduced DEPC footprinting upon chelating with Cu(II), revealing the Cu(II) binding sites, which were further confirmed by a separate electron paramagnetic resonance (EPR) study.[698] Subsequently, the same group successfully extended the method to identify the Cu(II) binding sites in doppel (a glycosylphosphatidylinositol-anchored protein).[699]

In 2007, Sarkar and co-workers[700] applied DEPC footprinting in combination with EPR, UV, and fluorescence spectroscopies to pinpoint the critical Cu(II) binding residues of human copper metabolism gene MURR1 (mouse U2af1-rs1 region1) domain (61−154) as His101, Met110, and His134. By adopting similar strategies, Zhao and Waite[701] investigated the metal-binding properties of mcfp-4 matrix protein in load-bearing junctions, Binolf et al.[702] characterized the Cu(II) anchoring sites in α and β-synuclein, Vachet and co-workers[553,554] located the Cu(II) binding sites in β-2-microglobulin and its preamyloid oligomers, and Karmakar and Das[703] identified the copper binding sites in αA Crystallin.

Another relevant yet distinctive demonstration in 2009 was by Ginotra and Kulkarni,[704] who characterized the solution structure of $Cu^{2+}(His)_2$ under physiological conditions. The coordination between Cu(II) and His in proteins occurs through a $Cu^{2+}(His)_2$ complex, whose coordination regime is either N4O4 or N3O2 donor atoms. The former was supported by various spectroscopic techniques in solution, whereas the latter was favored from X-ray crystallography at a pH of 3.7. By DEPC labeling of $Cu^{2+}(His)_2$ complex under physiological conditions and MS identification, the investigators found the most abundant species was a three DEPC adduct, indicating that the $Cu^{2+}(His)_2$ complex exists in solution as a neutral five-coordinate structure with N3O2 donor atoms.

Protein binding with Zn(II) usually involves the canonical Zn(II) binding motif, termed a zinc finger, whereby a Cys thiolate is the chelating functional group.[705] Binding with Zn(II) greatly reduces the solvent accessibility of the cysteine side chains and slows the footprinting. Forest and co-workers[706] in 1999 investigated Zn(II) binding of the ferric uptake regulation protein from *Escherichia coli* (*E. coli*) by using IAM footprinting. Among three reactive Cys residues, Cys92 and 95 are involved in the Zn(II) binding. The remaining Cys132 alkylates with the fastest reaction rate among the three, implying that it is surface-exposed and unbound.

In 2001, Apuy el al.[707] used cleverly designed NEM-based Cys labeling in a pulsed fashion. Their experiment established the relative cysteine thiolate reactivities of six zinc fingers in metal-response element-binding transcription factor-1 as F5 > F6 ≫ F1 > F2 ≈ F3 ≈ F4, indicating lower binding affinities for zinc fingers F5 and F6. This study is also an early example of the use of isotopic encoding in protein footprinting whereby the investigators used h_5-NEM and d_5-NEM to allow a better quantification.

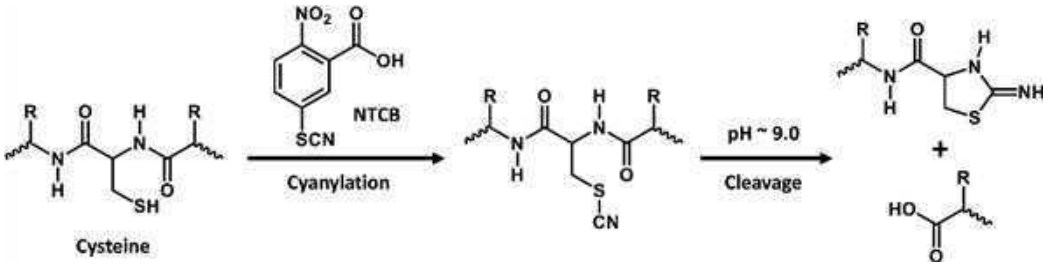
Atsriku et al.[708] in 2005 investigated the relative reactivities of Cys residues in the zinc finger of the estrogen receptor by iodoacetic acid labeling and found that Cys 240 is more reactive than Cys 237, although both are located in the second zinc finger of the estrogen receptor. The differences in reactivities are explained by the unequal positioning toward basic amino acids that affects the thiol pK_a. Hanas and co-workers[709,710] in 2005 used a Cys_2His_2 zinc finger peptide, Sp1−3, as a model system to investigate the interaction between Zn(II), Au(I), and the zinc finger protein. In addition to identifying the binding site through Cys footprinting by IAM, the investigators determined the relative binding affinities between the zinc finger and Zn(II) and Au(I). At a lower IAM concentration (100 μM), Zn(II) binding becomes distorted, whereas the Au(I) binding was not affected at 100 μM IAM and was diminished at 500 μM IAM, a concentration that eliminates Zn(II) binding. These results indicate that the zinc finger in Sp1−3 binds tighter to Au(I) than to Zn(II).

Millhauser and co-workers[711] in 2007 characterized the Zn(II) binding properties of PrP (from Syrian hamster) by DEPC footprinting. Syrian hamster PrP combines Zn(II) and Cu(II) binding in one protein: the Zn(II) binding, however, can change the overall Cu(II) redox properties by altering the distribution of copper ions among various binding modes. More importantly, DEPC footprinting coupled with a Zn(II) titration revealed the zinc binding affinity of prion protein 60−91, demonstrating new opportunities for targeted labeling reagents in the quantitative understanding of protein metal-ion interactions. Upon binding with Zn(II), the number of DEPC-reactive sites decreases, leading to a decreased DEPC footprinting as measured by MS. The system composition as a function of Zn(II) concentration (changed via a titration) can thus be monitored by quantifying the extent of DEPC footprinting and can be fit by modeling to extract binding affinity.

Karmakar and Das[712] in 2012 footprinted α-crystallin by DEPC followed by MALDI-ToF analysis and identified His79, 107, and 115 as key Zn(II) binding residues in αA-cystallin, whereas His104, 111, and 119 are critical for Zn(II) binding of αB-crystallin. Their method is limited proteolysis; regions bound by Zn(II) show missed cleavages in trypsin proteolysis because the Zn(II) binding stabilizes the region in which it binds and causes missed cleavages. In 2013, Russell and co-workers[713] used NEM labeling to map nonbinding sites and to detect directly peptides containing the metal to locate the Zn(II) and Cd(II) binding to human metallothionein-2A. By combining bottom-up, top-down, and ion mobility approaches, the investigators not only identified the metal binding sites and binding orders but also provided insights on the structural transitions during the metal binding and demetalation in the gas phase. The idea of footprinting the metal chelating groups with targeted labeling reagents was also adopted to label

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

Scheme 13. Free Cys Footprinting by NTCB-Mediated Cleavage



specific cysteine residues by using reversible metal protection, which is important in protein engineering.[714]

In another study, Ramakrishnan et al.[715] characterized the Zn(II) binding sites of hepatitis B virus X protein (HBx) through HDX and NEM footprinting. HDX locates four peptides that exhibit decrease in deuterium uptake upon binding with Zn(II), whereas NEM footprinting pinpoints the critical Cys residues that are responsible for the Zn(II) binding. Moreover, the NEM footprinting in this study was executed as a kinetic study, where the use of multiple NEM labeling times increased the confidence for identification of sites as is found with HDX kinetic curves and can provide some assurance that the protein structure is not perturbed (perturbation would give a change in kinetics). Time-dependent footprinting by targeted labeling reagents is discussed in section 4.3.

4.1.2. Protein Binding with Hard Metal Ions. For hard metal ions, Ca(II) is one of the most characterized owing to its biological significance in signaling.[716,717] The canonical Ca(II) motif, an EF-hand, utilizes the negatively charged carboxyl groups in Asp and Glu inter alia to chelate Ca(II) ions.[684,718] In a study by Zhang et al.[719] in 2012, the Asp and Glu residues in calmodulin (CaM) were footprinted by GEE in a method evaluation experiment. The Asp and Glu located on the EF-hands of CaM show reduced GEE footprinting upon binding with Ca(II), which is consistent with a SASA calculation of these residues. This study successfully demonstrates the efficacy of targeted labeling reagents in identifying hard-metal ion binding sites with high spatial resolution. By similar approach in the same year, Arata and co-workers[720] identified the Ca(II) binding sites in sarcoplasmic reticulum Ca^{2+}-pump ATPase by DEPC footprinting.

Very recently, Guo et al.[480] examined the Ca(II) and Mg(II) binding properties of CaM by BHD footprinting and MS analysis. Besides demonstrating the efficacy of BHD, a new footprinter, in identifying metal binding sites, the investigators compared the modification extents between Ca(II)-bound CaM and Mg(II)-bound CaM to determine that the conformational changes of CaM induced by Mg(II) binding are less significant than those by Ca(II). Because the high resolution structure of Mg(II)-bound, full-length CaM is not currently available,[721] this study shows the opportunities for footprinting to assist in HOS studies where X-ray crystallography is difficult. Successful probing of the low affinity interactions between Mg(II) and CaM also shows the potential of using targeted labeling reagents in protein−metal ion analysis and illustrates again opportunities afforded by isotope encoding.

Although powerful, the uses of targeted labeling reagents in probing protein−metal interactions are not numerous. An important reason is that the metal-binding pocket is usually buried, making the size of the labeling reagent critical. Another reason is that reaction rates between targeted labeling reagents and the residues are usually slow, making it challenging to characterize metal-binding systems with low binding affinities (whose metal ion off-rates are usually fast). Footprinting by using fast-labeling reagents, although an indirect method, can also overcome the above disadvantages to characterize protein−metal ion interactions, a subject to be covered in section 6.4.

4.2. Mapping Cys and Disulfide Bonds

Disulfide bonds are critical in protein folding and protein structural stabilization, making localization of disulfide bonds an important task. MS-based approaches contribute significantly to the determination of disulfide bonds.[722] Although many methods including plasma-induced oxidative cleavage,[723] ETD-based disulfide fragmentation,[724] electron-transfer higher energy dissociation (EThcD)-based disulfide fragmentation,[725] UVPD-based disulfide fragmentation,[726,727] online electrolytic cleavage,[728] and ion mobility[729] were applied to characterize disulfide bonds, they either require advanced instrumentation or suffer from disulfide scrambling. Researchers also developed novel algorithms that identify disulfide bonds by examining MS/MS fragmentation patterns of disulfide-linked dipeptides.[730,731] Its application in unknown systems, however, is limited because the fragmentations are poor and the enrichment is challenging, problems that also confront cross-linking. Although classical, disulfide bond identification by targeted Cys labeling may be an effective and reliable approach, with potential for mapping disulfide bonds. Therefore, to illustrate opportunities in this area, we will review a few examples that are either pioneering studies or show novel workflows

The potential of MS in disulfide mapping was first demonstrated by Morris and Pucci[732] in 1985 remarkably with FAB-MS, where they identified disulfide-linked dipeptides by aligning the peptide masses with the protein sequence. Some of the early FAB-MS-based efforts were reviewed by Smith and Zhou[733] in 1990. These methods evolved into quantitative assessments of free Cys and disulfide bonded by MALDI-MS, where a "count" of the free Cys residues in a protein can be made by analyzing the mass shift upon modifying the free Cys residues.[734,735]

To locate disulfide bonds is challenging and cannot be done universally with MS/MS because fragmentation is not always cooperative. In a pioneering study in 1996, Watson and co-workers[736] utilized 2-nitro-5-thiocyanobenzoic acid (NTCB) to footprint free Cys and cystine. NTCB is a cyanylation reagent developed by Vanaman and co-workers[737] in 1973, and NTCB labels free Cys by replacing the hydrogen of the thiol with a cyanide. The cyanylated Cys undergoes backbone cleavage under alkaline pH (Scheme 13). The method is

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

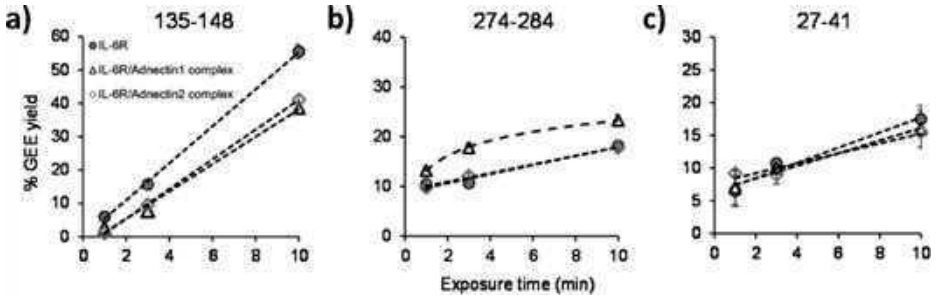


Figure 6. GEE labeling kinetics for selected IL-6R peptides in the ligand-free (gridded circle), adnectin 1-bound (triangle) and adnectin 2-bound state (diamond) state. (a) Region 135−148 shows decreased GEE incorporation upon adnectin 1/adnectin 2 binding, whereas (b) region 274−284 shows increased GEE modification upon adnectin1 binding. (c) Representative peptide region without differentiable GEE modification extent between bound and unbound as a control. Dashed trend curves in (a−c) are generated by linear or 2nd degree polynomial fitting. Reproduced with permission from ref 743. Copyright 2017 American Chemistry Society.

utilized by choosing specific peptide cleavages, guided by location of free Cys residues and determining the location of free Cys residues by analyzing the resulting peptides, here by MALDI-MS. Its application in disulfide mapping was by a differential approach, where the protein is footprinted by NTCB under both native and reduced conditions (TCEP was used for disulfide bond reduction). The Cys residues that are not cleaved under native condition but cleaved under reducing condition can be assigned to disulfide bonds.

Modern disulfide-bond footprinting makes use of the silence of Cys residues toward labeling reagents when they are covalently bound to form a disulfide bond. To distinguish disulfide-bonded and solvent-inaccessible Cys requires careful design and suitable control experiments. In one of the early applications in 2000, Glocker and co-workers[738] developed a general workflow to map disulfide bonds with targeted labeling reagents. They applied their approach to heat shock protein Hsp33, a redox-regulated molecular chaperone with six Cys residues. In its reduced state, all six Cys are in the thiol form, and the protein is not biologically active. Exposing the protein to oxidizing conditions promotes conversion of Hsp33 into its active form, during which the protein folds, assisted by formation of disulfide bonds. Upon IAM footprinting followed by MS analysis, all six Cys in reduced Hsp33 can be alkylated. In the oxidized state, however, Cys 232, 234, 265, and 268 are silent in IAM footprinting. Peptide-level MS analysis suggests that two disulfide bonds are Cys232−S−S−Cys234 and Cys265−S−S−Cys268. Cys141 was found to be highly reactive in both protein states, indicating that it is not involved in disulfide formation. In contrast, the modification extent of Cys 239 decreased significantly but did not disappear for the oxidized state, suggesting that Cys 239 is not involved disulfide bonding but rather part of a significant conformational change, burying it upon Hsp33 activation.

In a 2009 study, Chumsae et al.[316] incorporated Cys fluorescence labeling with MS-based peptide mapping to study the unpaired Cys residues in human immunoglobulin G subclass 1 antibody (IgG1). Although all 32 Cys in IgG1 can be involved in disulfide bonds, the presence of small amounts of free Cys is a common occurrence in recombinant and wild-type IgG1 antibodies. To differentiate the free and disulfide-bonded Cys residues, the IgG1 was first labeled with 5-idoacetamidofluorescein (5-IAF), a fluorescent reagent that reacts with free Cys residues. Subsequently, the 5-IAF labeled IgG1 was submitted to disulfide bond reduction and iodoacetic acid labeling. As a result, the free Cys residues are labeled with a fluorescence probe, whereas the cystine residues are labeled with iodoacetic acid. Upon digestion, the peptides were submitted to LC separation, fluorescence detection, and MS analysis. As the free Cys residues are less abundant in IgG1, the fluorescence labeling improves the detection of free Cys-containing peptides. This study also is a good example of an integrated approach that addresses an important biological question (especially in pharma and biotechnology).

One year later in 2010, Weerapana and co-workers[739] developed an isotopically encoded IAM derivative with an attached phenyl ring. The carbons of the phenyl ring can be either ^{12}C, the light form, or ^{13}C, the heavy form, to allow isotope encoding. A subject protein was first labeled to give the light form, tagging the free Cys resides. TCEP was then added to reduce the disulfide bonds, followed by a second stage of labeling where any Cys formerly in a disulfide bond is now labeled by the heavy form. Although originally demonstrated for quantitative profiling of Cys reactivities in primary structure proteomics,[739,740] the idea of isotopic-encoded reagents is readily applicable in footprinting disulfide bonds. It avoids the problem of using two different reagents (one to label free Cys and another to label the disulfide-bonded Cys after reduction). The reactivities of two different reagents are seldom the same, leading to ambiguities in quantification. Isotopic-encoded alkylating reagents nicely overcome this drawback and add precision and efficiency.

More examples of targeted Cys footprinting in combination with MS peptide mapping with discussions of experimental protocols and issues of disulfide scrambling can be found elsewhere.[174,741,742]

4.3. Characterizing Protein−Ligand and Protein−Protein Binding Interfaces

Among all biological questions that MS-based protein HOS elucidation can address, locating protein−protein binding interfaces is the most widely asked. Upon interacting with a binding partner, the solvent accessibility of the binding interface usually decreases, leading to a decreased footprinting. Here we present three examples of targeted protein footprinting where improved methodology was developed. Each example shows unique features from an analytical perspective.

During the past decade, carboxyl group footprinting by GEE has been adopted in characterizing the structure of antibodies as well as their interaction with antigens and drugs.[743−748] In an informative example, Li et al.[743] applied orthogonal MS-based footprinting approaches to characterize the interaction between the extracellular region of human interleukin-6

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

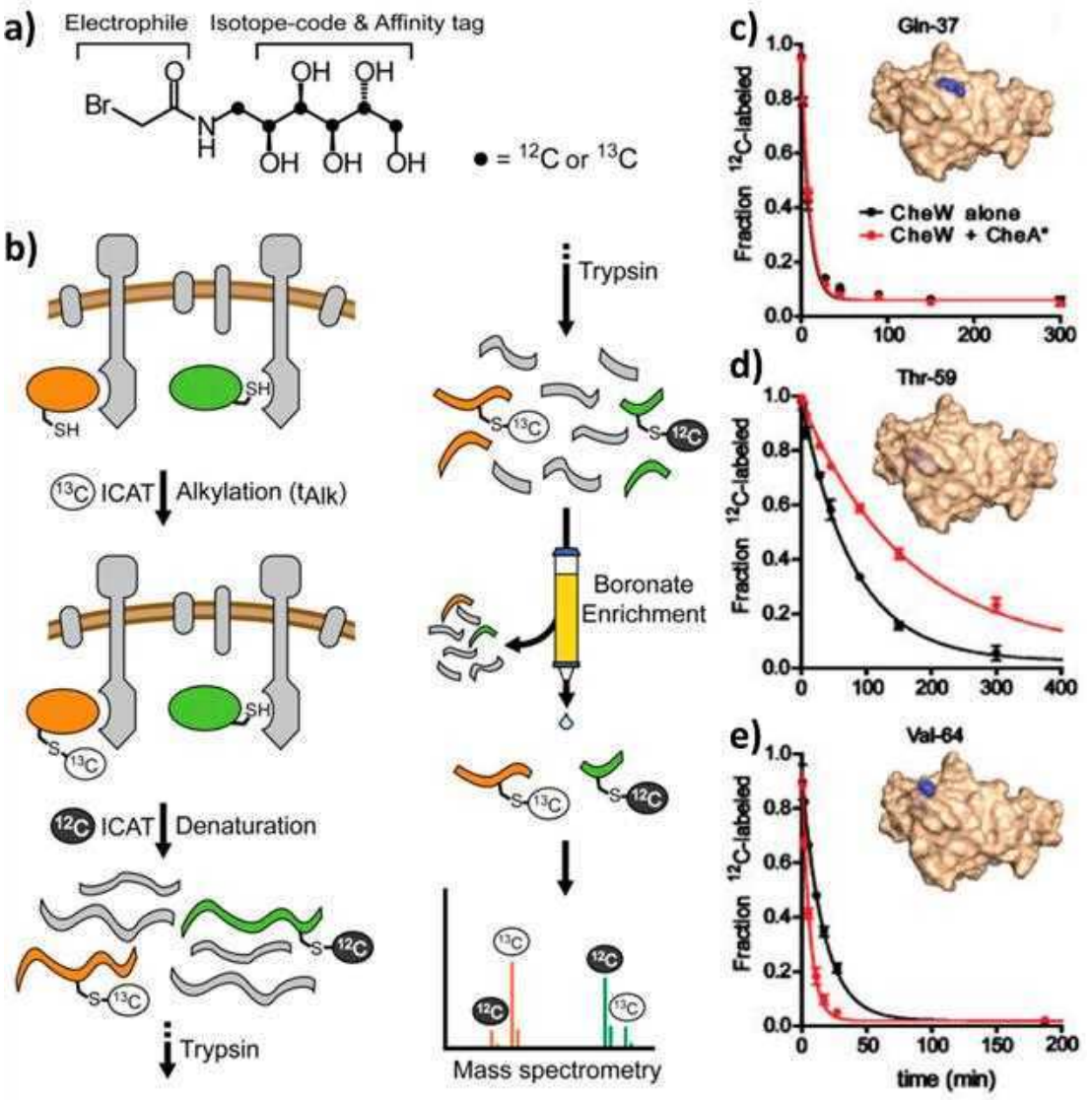


Figure 7. Overview of Cys footprinting with an ICAT reagent. (a) Structure of bromoacetamide ICAT reagent. Dots represent ^{12}C for light reagent and ^{13}C for heavy reagent. (b) Workflow for alkylation rate determination by ICAT isotope pairs. (c–e) Representative alkylation time courses of CheW Cys variants footprinted in the presence (red) or absence (black) of CheA*. Adapted with permission from ref 749. Copyright 2011 Elsevier.

receptor α-chain (IL-6R) and two types of adnectin (adnectin 1 and 2) that bind with picomolar and nanomolar affinity, respectively. Different from labeling in many other studies, GEE footprinting was executed in a time-dependent manner (Figure 6a–c) as commonly seen in HDX studies. Region 135–148 becomes protected upon ligand binding (Figure 6a), and the differences are made more apparent by obtaining kinetic curves rather than relying on a single time point. Subsequent MS/MS analysis indicates the GEE labeling is exclusively on residue Glu 140.

In comparison to region 135–148, region 274–284 becomes deprotected upon binding with adnectin 1 but does not change significantly when bound to adnectin 2 (Figure 6b). Region 27–41 was not involved in binding as the GEE labeling extents remain comparable with and without IL-6R binding. The investigators concluded that the adnectin 1 and 2 binding region is a flexible loop between two β-strands in the cytokine-binding domain of IL-6R, which was verified by other MS-based footprinting in this study. Besides demonstrating the advantages of applying orthogonal methods in characterizing a single protein system, the protocol executed the targeted labeling in a kinetic way, significantly enhancing the confidence in the data interpretation as compared with the commonly used single time point labeling.

In examples in 2008 and 2011, Kiessling and co-workers[749,750] developed isotope-coded affinity tag (ICAT) coupled targeted labeling reagent for protein footprinting. The bromoacetamide ICAT reagent, often used to alkylate free Cys residues, was repurposed for footprinting (Figure 7a). The idea of ICAT originates from proteomics,[751,752] where isotopic encoding enables absolute quantification and the affinity tag facilitates postlabeling enrichment. Here the goal is to footprint the binding interface of adaptor protein CheW and the multidomain histidine kinase CheA (and its Cys-depleted form CheA*) in the *E. coli* chemotaxis signaling system. As protein–protein interactions in signaling systems are usually transient and low in affinities, it is necessary to add excess binding partners to the system to shift the binding equilibrium to the

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355–4454

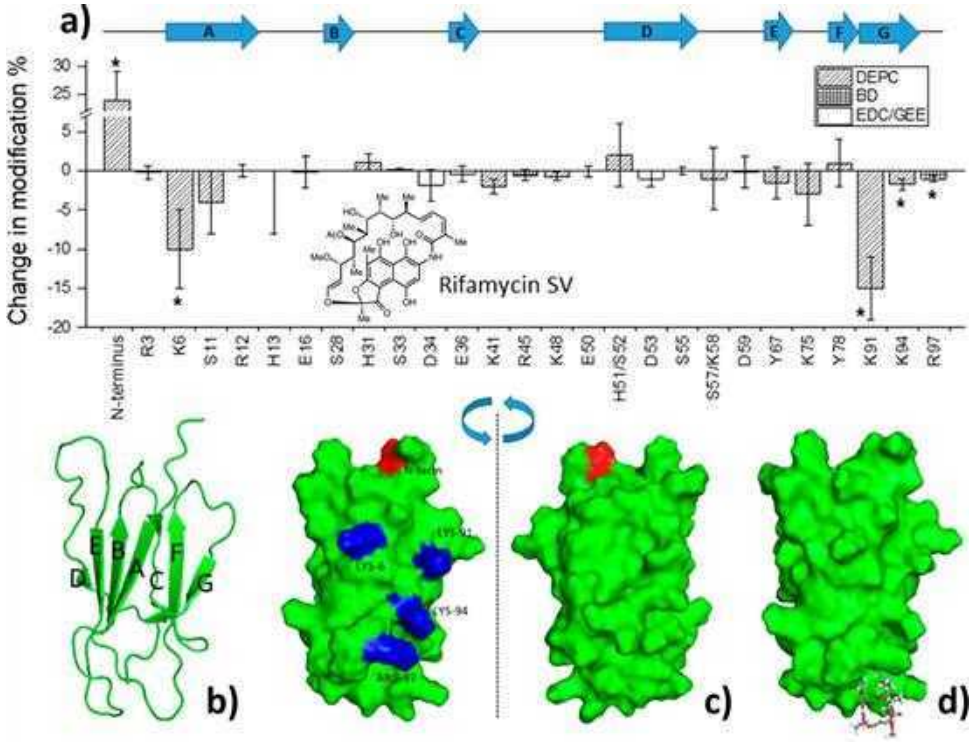


Figure 8. Covalent labeling results of β2m binding with rifamycin SV. (a) Changes in labeling modification percentages with rifamycin SV bound to the Cu(II)−protein complex. BD is 2,3-butanedione. Error bars represent standard deviations from triplicate measurements. Asterisks above the bars represent the residues that undergo a significant change in modification level at 95% confidence as determined by a two-sample unpaired Student's *t* test. The arrows at the top of the graph indicate the locations and directions of the seven β strands in β2m. (b) Ribbon structure of β2m, showing the seven β strands labeled A through G. (c) β2m surface structure illustrating the sites undergoing significant changes in covalent labeling induced by rifamycin SV. Sites that increase in labeling are colored red, whereas those that decrease in labeling are colored blue. (d) Protein−ligand docking results for verification. Reproduced with permission from ref 754. Copyright 2017 American Chemistry Society.

bound state. When the binding partner is a large protein, as exemplified by CheA, excess protein will create an intense protein/peptide background and potentially overwhelm the MS detection. The ICAT footprinting reagent enables postlabeling enrichment, making it possible to label the CheW under high CheA concentration without concern for the high CheA background during MS analysis.

The isotopic encoding enables accurate quantification. As depicted in Figure 7b, the complex is first footprinted by the heavy reagent, where all the solvent accessible Cys residues can be labeled. After a certain labeling time, the protein complex is denatured and the light-reagent is added, and those Cys residues that were originally protected can now be labeled by the light reagent. The sample is then digested and submitted to affinity enrichment. In this study, the Cys footprinting by heavy reagent was executed under multiple time points (t_{Alk}), and the relative fraction of a specific Cys residue under solvent-accessible and protected states was quantified by comparing the signal intensities from a single injection of the heavy- and light-labeled species. Alkylation rates of specific Cys residues can be obtained by plotting the fraction of labeling by the light reagent as a function of t_{Alk}, where a more abundant light-labeled fraction indicates a decrease in solvent accessibility upon binding with CheA.

In the experiment, CheA* is used instead of native CheA to minimize the effect of free Cys from CheA on the alkylation kinetics of CheW Cys residues. As there is no Cys in the CheW sequence, several mutants were expressed and purified, each with a single-site Cys mutation to facilitate the alkylation footprinting. Cys residues in CheW that show decreased alkylation rates upon mixing with CheA* are the binding sites, as exemplified in Figure 7d. Among nine mutants that were analyzed, four showed significant decreases in alkylation rates upon binding with CheA* (46, 48, 59, 60), indicating that these four positions of CheW are likely involved in CheA binding. The enrichment by ICAT not only offers a better quantification but also enables the footprinting of transient and weak-binding partners. Taken together with the two earlier examples by Ramakrishnan et al.[715] and Li et al.,[743] footprinting of targeted residues that were executed at multiple labeling times will significantly enhance the confidence of the data interpretation.

In a third study, Vachet and co-workers[554,753] studied the preamyloid dimer and tetramer of β-2-microglobulin (β2m) by footprinting with different reagents. This study nicely demonstrates a canonical workflow and data presentation of protein HOS analysis by multiple targeted labeling reagents, where the focus is the differences between ligand-bound and ligand-free states, and the results are mapped onto existing structural models or used to construct a new model based on the footprinting results.

Like many other amyloid-forming proteins, β2m self-associates into fibrillar amyloid in the presence of Cu(II) and can be deposited in the musculoskeletal system. Because most targeted labeling reagents are reactive with specific amino acid residues, a strategy of using several labeling reagents is effective to increase the amino acid coverage. By combining sulfo-*N*-hydroxysuccinimide acetate, DEPC, and 2,3-butane-

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

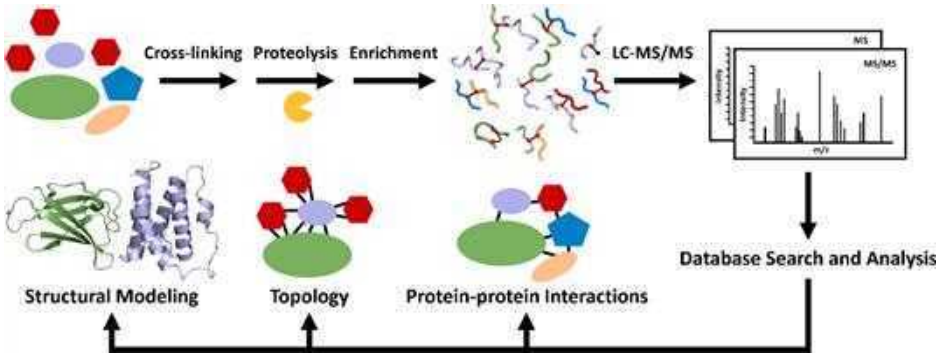


Figure 9. Workflow of bottom-up cross-linking mass spectrometry.

dione labeling, the investigators interrogated the N-terminus and Asn, Arg, His, Lys, Ser, Thr, and Tyr residues in β2m (as indicated by the filling of the bars in Figure 8a). Comparison between the footprinting of β2m in monomeric and dimeric states suggests that the dimer interface is formed by antiparallel stacking of ABED β-sheets by two β2m monomers. Further, in-solution footprinting data rule out the interaction of D−D strands in the dimer, which is not consistent with a previously resolved X-ray crystal structure, emphasizing the need to use complementary HOS structural characterizations. In a subsequent study with similar experimental procedures, the interface of β2m tetramer was identified as D strands from one dimer unit and G strands from another dimer unit.[753]

Because β2m tends to aggregate, the same group also studied the binding sites of its amyloid inhibitors (e.g., rifamycin SV, structure depicted in Figure 8a), and the results are in Figure 8.[754] The changes in modification fractions upon binding with rifamycin SV (Figure 8a) show that four residues become protected (K6, K91, K94, and R97), whereas the N-terminus becomes more exposed. These results were mapped onto the β2m surface structure (Figure 8c) and compared with protein−ligand docking. The comparison strongly suggests a binding pocket between Lys75 and Arg 97 (Figure 8d). Considering all the results, the authors concluded that rifamycin SV likely binds near Lys 94 and Arg 97 and perhaps near Lys 91, which is along the G-strand of β2m (Figure 8b). The mismatch between docking and footprinting is likely due to the structural perturbation of β2m induced by rifamycin SV; however, the similarity is encouraging. Two other ligands were also evaluated in a similar fashion.

Lastly, targeted labeling reagents were also adopted to footprint protein−nanomaterial interfaces to provide critical insights in the toxicity and potential biomedical applications of nanomaterials.[755−757] Determining protein−protein and related interfaces with fast labeling reagents is also attracting considerable attention, and this subject will be covered in section 5 and 6.

4.4. Assessing Topology and Stoichiometry through Chemical Cross-linking

Chemical cross-linking has developed as a complementary area of research owing to its capability to probe HOS and to locate and define protein/protein interfaces. Our intention, however, is not to review this topic comprehensively but to describe the workflow, show how cross-linking can be viewed as a means of footprinting (i.e., "double footprinting"), discuss its role in integrated approaches, and highlight some recent developments to unify the methodology described in this work.

Cross-linking mass spectrometry (XL-MS) is effective at assessing protein complex topologies and elucidating protein structures besides capturing protein−protein interactions. Conceptually, XL utilizes bifunctional labeling reagents to cross-link (footprint) the constituent proteins forming an interface, thus providing information on the interacting species and their interfaces. When dealing with large proteins, XL can also report on the overall protein conformation.

To date, most of the MS-based XL studies have been by a bottom-up approach, whose workflow is presented in Figure 9. Briefly, proteins are first incubated with chosen cross-linkers under the optimized conditions that are established usually by monitoring with techniques simpler than MS (e.g., gel electrophoresis). The tethered proteins are then submitted to enzymatic digestion, enrichment of informative peptides, and LC-MS/MS for separation and detection. The MS data are further analyzed by using search engines to identify cross-linked peptides with an uncertainty specified by mass tolerance and false discovery rate (FDR). Significant advances since 2008 in data analysis have facilitated this approach; new software includes but is not limited to pLink,[758] xQuest,[759] XlinkX,[760] and StavroX.[761] Generated cross-link maps not only identify the connectivity of the adjacent protein subunits but also provide distance restraints as given roughly by the molecular separation between the two reactive functional groups of the cross-linker.

XL-MS alone is a middle-to-low spatial resolution approach because usually there are limited number of reactive residues at or near an interface of two proteins, and few cross-links form. Young et al.[762] in 2000 showed that combining XL distance restraints and computational modeling is a compelling way to improve the resolution for elucidating protein structures. The workflows can be adapted for integrative modeling[763,764] and de novo structure prediction.[765,766] The major challenges of computational studies are the sampling and scoring of the generated models.[767] Larger data sets derived from XL-MS effectively decrease the size of the sampling space and benefit the scoring function, thus promoting efforts to find high-confidence models. Given there can be various reactive groups and spacer lengths on cross-linkers, a strategy involving multiple reagents should be taken to ensure comprehensive information.

In 2014, Chait and co-workers[764] demonstrated the use of DSS and a zero-length EDC in characterization of the nuclear pore subcomplex of Nup84. Two data sets generated from the use of two reagents delivered different but complementary structural information that significantly benefited the subsequent model construction. Many other studies adopted

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

similar ideas by combining amine-targeting reagents and carboxyl-targeting reagents (e.g., dihydrazides[675,768]) or non-selective cross-linkers.[769,770] In addition, the use of reagent combinations with different spacer lengths is recommended because short cross-linkers yield fewer cross-links but with narrower distance restraints, whereas long cross-linkers afford more cross-links but less structural definition because the distance assignments are over a broader range.[767]

New sample preparation methods also have developed rapidly (e.g., on-bead cross-linking[766]), improving the cross-linking chemistry and providing better analysis sensitivity. Many examples are discussed in recent reviews.[665,771,772]

By way of contrast, footprinting by targeted and free-radical reagents and/or by HDX maps solvent accessible regions and reflects differences in protein dynamics or binding events. This information provides deeper understanding of the entire protein structure, not just interfaces, and can adjudicate constructed 3D models from XL for in-solution protein.[773−775]

XL is not limited to structural characterization of a single protein complex, but rather it can promote protein−protein interactions (PPIs) studies in proteome-wide investigations.[758,776−778] To meet the need for understanding the roles, functionalities, and mechanistic behavior of many protein complexes through their PPIs maps, cleavable XL reagents have been developed,[779] simplifying and increasing the accuracy for identification of cross-linked peptides. Incorporation of affinity groups and sophisticated enrichment procedure have led to success; examples are studies performed on *E. coli*[777] and *Caenorhabditis elegans* (*C. elegans*).[778] For example, the Heck group[760] reported a XL-MS study of whole human cell lysates in which they identified 2179 unique cross-links. These protein complexes (e.g., 80S ribosomal core complex) reveal novel interactions and provide new structural insights.

To summarize, the increasing number of publications based on XL-MS continue to demonstrate its utility in MS-based biophysics, particularly in combination of other techniques. Integrating XL with other methods enables characterization of dynamic biological systems even of heterogeneous systems. The continuing development of cross-linkers will increase applications. Reagents that react more rapidly and target more and more amino acids, instrumentation advances that provide new fragmentation methods in MS/MS (e.g., ETD, EThCD, and UVPD) and software are expected for the future. Furthermore, improved separation and enrichment procedures will accommodate the increasing mixture complexity following XL and digestion and allow detection of low abundant yet informative cross-linked species.

4.5. Footprinting Fast Kinetic Processes

Given the high reactivity of the Cys thiol group, specific labeling reagents were also developed to follow protein folding in addition to footprinting a static state.[780] An early application addressed slow events in the folding of recombinant human macrophage-colony stimulating-factor β, which takes tens of hours to complete.[781] Melarsen oxide (*p*-(4,6-diamino-1,3,5-triazin-2-yl)aminophenylarsonous acid) used in this study footprints Cys in 10 min, which is sufficiently fast for this slow folding system but far too slow to footprint most protein folding.

To shorten the footprinting time scale and make Cys footprinting capable of monitoring protein folding pathways in general, other reagents, methylmethanethiosulfonate (MMTS)[782] and DTNB,[323] were applied because they react with Cys on the time scale of milliseconds. Furthermore, the steric sizes of these reagents can provide insight on the structural changes during conformational opening/closing.[783] These reagents were successfully applied in a pulse-labeling approach to follow unfolding of apo-myoglobin (aMb).[784,785]

In an elegant demonstration of specific amino acid footprinting, Jha and Udgaonkar[786] in 2007 applied pulsed Cys labeling by MMTS to study the fast folding reaction of barstar. Tracking the labeling of 10 Cys residues as a function of refolding times showed that four remain solvent-accessible throughout the folding. The other six Cys residues become protected upon folding, and their rates of reaction vary by 3-fold and are dependent on their locations in the protein. The investigators showed the general workflow of MS-based protein folding analysis and demonstrated footprinting with a single residue can provide sound information. Unfortunately, Cys is the only residue that can be labeled on the millisecond time scale, probably because both the −SH group and the reagent are reactive. A workaround may be to express the protein with more Cys residues located strategically to inform on the roles of various regions. The downside is that the proteins may respond to these substitutions and adopt a non-native state. Thus, a general approach to study folding by specific amino acid labeling will be difficult to achieve. Moreover, labeling times are still too long for probing short-lived intermediates and fast folding. Fast free radical labeling reagents can overcome these drawbacks, as will be discussed in section 6.2.

In an unusual approach to fast reactions, McLuckey and co-workers[787−790] starting in 2009 demonstrated the covalent labeling of peptides in the gas phase. They illustrated the reactions between peptides and 4-formyl-1,3-benzenedisulfonic acid (reacts with primary amines),[787] NHS-ester (reacts with primary amine and carboxylates),[788,789] and Woodward's Reagent K (reacts with carboxylic acids).[790] Although not yet applied in gas-phase protein footprinting or folding, this approach is potentially advantageous because gas-phase ion reactions are rapid, sometimes even collision controlled. A disadvantage is that reactions in the gas phase may not be relevant to in vitro and in vivo biochemistry.

4.6. Footprinting in Cells

It can be questioned whether protein biophysical properties and their HOS obtained in test tubes are truly reflective of those in vivo as the native environment in which a protein functions is challenging to reproduce outside a living cell. Thus, protein HOS analysis in a cell and even in vivo is a long-sought goal of protein biophysics. Note that throughout this review we reserve the word "in vivo" to research done in a complete organism or in an animal. Macromolecular crowding within the cell limits the diffusion of the protein molecules and that further affects the biophysical properties of proteins including their interactions with nucleic acids and their reaction kinetics and association rates with their binding partners.[791,792] As HOS analysis of a protein in a complete organism or animal is challenging, protein footprinting in living cells may be the solution. When one recognizes that the expanding capabilities of MS-based proteomics analysis permit deeper and deeper analysis of an entire proteome, the goal of in vivo footprinting comes closer to realization.

Early developments of protein in-cell labeling primarily focus on the localization and quantification of proteins in the cellular

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

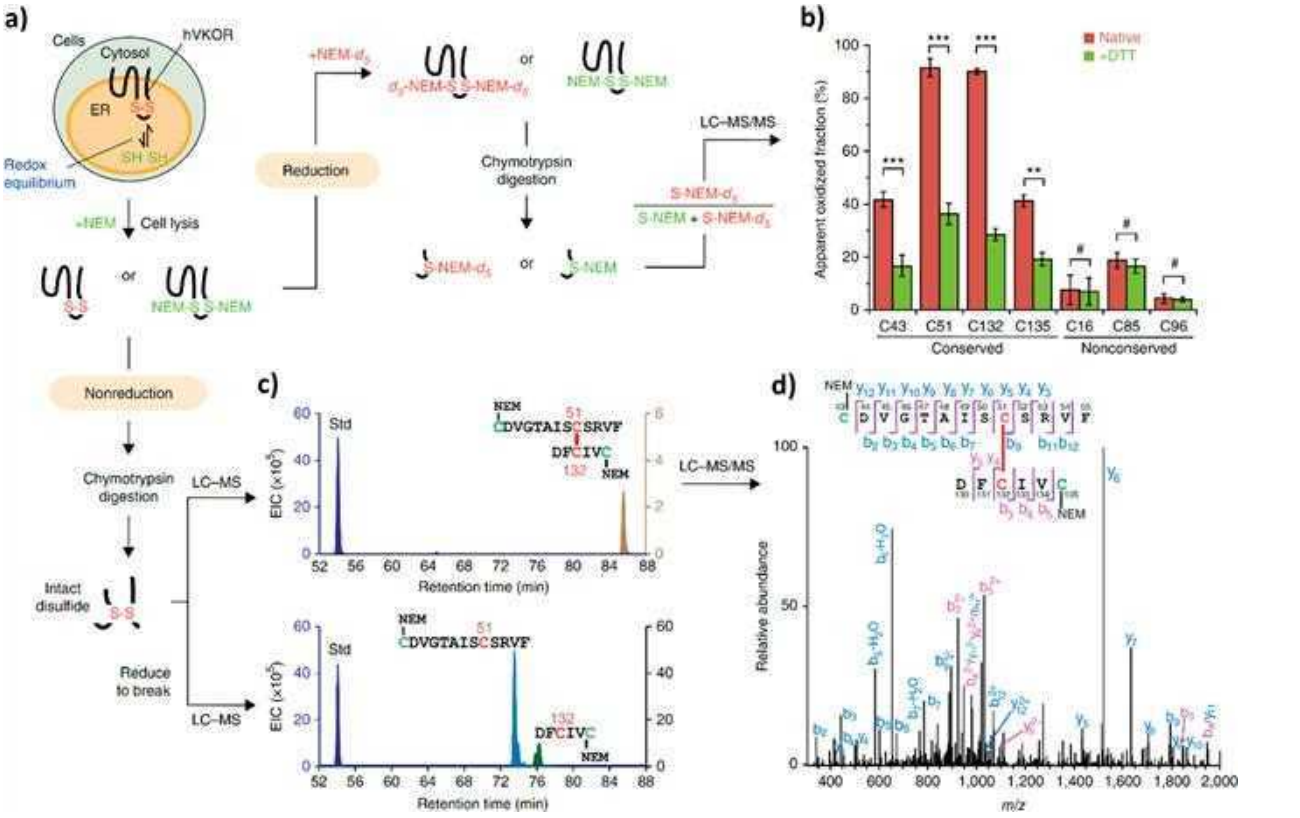


Figure 10. Schematic illustration of in-cell Cys footprinting workflow. (a) Scheme of MS analysis. The redox status of cysteines was quantified by differential isotope footprinting with NEM before and after a reduction step (right). Intact disulfide bonds were detected under nonreducing conditions (bottom). (b) Intracellular cysteine status was determined by quantitative MS before (red) and after DTT reduction (green). Bar graphs show means and errors of multiple peptides from three independent experiments $^{**}P < 0.001$; $^{***}P < 0.0001$; $\#P > 0.05$ by two-tailed Student's t test. (c) Extracted-ion chromatograms showing a Cys51−Cys132 disulfide-linked peptide (top), which separates into two individual peptides (bottom) after reduction of the disulfide. A reference peptide is used as the standard (Std) for peak comparison. (d) MS/MS spectrum of the Cys51−Cys132 linked peptide (from c). Reproduced with permission from ref 819. Copyright 2017 Springer Nature.

environment. For localization purposes, a common practice is to attach fluorophores covalently to the target proteins.[793,794] The fluorophores can be either small organic molecules[795] or proteins,[796] which then illuminate under a fluorescence probe. The labeling can be noncovalent and mediated by an engineered ligand that has a fluorophore attached.[797,798] Here the labeling efficiencies are generally low but relatively specific. Recent demonstrations couple FRET with protein in-cell labeling (semisynthetic protein switches) to allow characterization of PPIs and protein conformational changes.[799] Despite success in engineering proteins and controlling their functions, these approaches have low structural resolution and sometimes low sensitivity that limit applications in protein HOS analysis.[793]

Protein labeling in-cell and in vivo coupled with MS are used for quantification in proteomics and metabolomics. These approaches almost always involve isotope encoding for increasing accuracy and speed. Although these methods are used more and more, their goal is not structural information. For this reason, and because these approaches were reviewed elsewhere,[800−802] they will not be discussed further.

High resolution protein HOS analysis in-cell is challenging. Among all three high-resolution approaches, NMR has the richest history of characterizing proteins in cells.[803] Because a protein's X-ray structure cannot be acquired under native conditions, that approach holds little promise. Macromolecular crowding makes it almost impossible to obtain cryo-EM images for a specific protein of interest under in-cell conditions.[804] Protein footprinting in cells was first done by cross-linking, which can elucidate topological features of protein and protein complexes and capture PPIs to reveal their biological functions. Implementation of cross-linking chemistry in the native environments is challenged by the heterogeneity and complexity of the cross-linked products.[805,806] The Bruce lab,[807] in pioneering efforts starting in 2009, developed a series of MS-cleavable cross-linkers (i.e., protein interaction reporter (PIR)) that undergo characteristic fragmentation during CID, thus providing more precise cross-links identification. To optimize the acquisition of fragments for MS3 sequencing, the investigators developed a real-time analysis for cross-linked peptides technology (ReACT).[669,808] In addition to biotin-based enrichment, ReACT has been successfully applied to XL in many organisms (e.g., *Shewanella oneidensis*,[807] *E. coli*,[809] *Pseudomonas aeruginosa*,[810] *Acinetobacter baumannii*,[811] and *human cells*[812]). Other cleavable cross-linkers are also available (e.g., *bis*(succinimidyl)-3-azidomethylglutarate (BAMG)[813] and azide-tagged, acid-cleavable disuccinimidyl bis-sulfoxide (azide-A-DSBSO)[814]).

Recently, Liu et al.[815] characterized the interactome of native mouse-heart mitochondria with DSSO and identified 3322 unique physical contacts, the largest survey to date of mitochondrial proteins interactions. In-cell cross-linking experiments reveal localization of many subunits and provide evidence for coexisting respiratory supercomplex assembly.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

In other demonstrations, Blankenship and co-workers[816,817] combined in-cell XL-MS with time-resolved spectroscopy and characterized a fully functional megacomplex composed of a phycobilisome antenna complex and photosystems I and II from the cyanobacterium *Synechocystis* PCC 6803. More applications and progress were reviewed lately.[806,818] Although XL-MS has shown its effectiveness, there are still many challenges raised by the complexity of biological systems. More and better methodologies in the sample preparation, enrichment, and MS detection will allow more comprehensive characterization of interactomes, providing another dimension of information to understand the molecular machinery that impacts living systems.

Another approach to in cell footprinting is illustrated by the investigation of the membrane protein human vitamin K epoxide reductase (hVKOR) by Li, Gross, and co-workers.[819] Membrane proteins are appropriate targets for in-cell footprinting because retaining their native structures in vitro can be challenging, especially for complexes with multiple subunits.[820] Although native MS allows probing noncovalent interactions associated with membrane proteins, structural information is hard to obtain with this approach.[157,821,822] In the study of hVKOR, Li and co-workers[819] targeted an intramolecular disulfide bond (Cys51−Cys132) that stabilizes a key intermediate redox state, a state that can be perturbed by binding with warfarin, a highly prescribed drug to prevent blood clotting. In brief, h_5-NEM was added to penetrate live cells and footprint free Cys in the native environment. After lysing the cells, the remaining disulfide bonds were reduced and labeled with d_5-NEM to footprint the newly reduced Cys (Figure 10a), a methodology discussed in section 4.2. The sample was submitted to LC-MS and MS/MS for identification (Figure 10c,d). The partition of a specific Cys residue in its free and disulfide-bounded states can be obtained by quantifying the ratio of its h_5-NEM and d_5-NEM labeled species (by peak intensities in the mass spectra), defined as "apparent oxidized fraction" (Figure 10b).

To facilitate a differential footprinting to illustrate whether a specific Cys residue prefers the free or disulfide-bounded state in the native cellular environment, the experiments were executed in two different conditions, illustrated in Figure 10b as "native" and "DTT". The "native" state denotes the footprinting under native cellular environment as mentioned above. The "DTT" state, however, was executed under reducing conditions. Upon admitting DTT to the cell, the disulfide-bound Cys residues in the native state were reduced and footprinted by h_5-NEM. As a result, the apparent oxidized fractions for the disulfide-bonded Cys residues decrease in the "DTT" state as compared to the "native" state, as seen for Cys43, 51, 132, and 135. The remaining three Cys residues (Cys16, 85, and 96) do not show decreased apparent oxidized fractions, indicating that these three Cys residues are in a reduced state (free Cys) in the native cellular environment. By such an approach, the investigators were able to elucidate the native architecture of hVKOR, reveal the electron-transfer pathway, and propose an hVKOR catalysis inhibition mechanism by the ligand warfarin. A subsequent paper summarizes the precautions, technical difficulties, and advantages of in-cell membrane protein footprinting.[823]

In another demonstration, Li, Gross, and co-workers[824] applied in-cell carboxyl group footprinting to probe formation and breakage of the salt bridges that are hypothesized to transport carbohydrates of human glucose transporters (GLUTs). By footprinting with isotope-encoded reagent GEE, they identified key salt bridges in GLUT1 on both its intracellular and extracellular sides. Subsequent mutation of the residues involved in salt bridges lead to a significant substrate uptake deficiency of the cell, providing confirmation that salt bridges serve as molecular switches. The investigators conducted footprinting under different substrate conditions to reveal the gating mechanism of the transporter as well.

MS-based in-cell footprinting, as a probe of protein solvent accessibility, offers structural evidence that can be directly related to functional states in vivo. On the other hand, challenges remain. The first is to ensure that the excess amount of labeling reagent will not compromise the native environment of the target protein. This can be tested by checking the function in a biological assay of the target protein upon adding the labeling reagent under native cellular environment. Such an approach works well when the function of target protein can be examined through indirect measurements; for example, a membrane transporter whose function can be characterized by the intra/extracellular substrate concentration, or a product that can be monitored by the substrate conversion. It is, however, challenging but still essential for those cases when viability of the target protein cannot be easily determined in the native cellular environment. The second challenge is the complexity introduced by the high crowding of intracellular components, requiring use of large quantities of reagent that must be admitted to the cell. The dynamic range of cell proteins is high, and that challenges the analysis of low-abundance proteins. Nevertheless, the success in answering focused questions in a few examples demonstrate that protein footprinting in cells by targeted labeling reagents has significant promise. Like in-cell cross-linking, better strategies in postlabeling sample enrichment will enable broader adoption of this method.

4.7. Integrating Footprinting with Other Approaches

Each of the methods reviewed here has limitations that may be overcome by combining with other methods. Integrating footprinting with other approaches is most seen with XL, owing to the limited structural resolution of XL but also its distance restraints that can be readily used as inputs in modeling.

To construct better a 3D model to near-atomic resolution, the XL platform can be coupled with electron microscopy[825,826] and X-ray crystallography.[764,765,827] Electron density maps often show areas that diffract poorly when a protein region is flexible or heterogeneous. To localize each subunit accurately, restrained distances can be contributed by XL-MS. One compelling example is reported by Greber et al.,[828] who determined the architecture of the porcine 39S large proteome in the mammalian mitochondrial ribosome. A previous Cryo-EM structure was resolved at 4.9 Å. Many of the protein extensions (a sequence change extending the amino acid sequence at the N- or C-terminal end with one or more amino acids) that do not have homology models cannot be refined at this resolution. In a combined approach, distance restraints from XL can be used in refining structure considerations, leading to a higher spatially resolved model (at 3.4 Å). This better-described protein assembly now reveals many interaction "hotspots" (e.g., the active site of peptidyl transferase and the path of its idiosyncratic exit tunnel). Integration of native MS and XL-MS can lead to remarkable structure elucidations[829−832] of macromolecular assemblies

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

that are usually difficult to purify and crystallize by traditional methods. Native MS can determine topology and stoichiometry of protein complexes. The approach is even more effective when combined with ion mobility MS, although direct determination of the critical binding residues is nearly impossible. On the other hand, XL-MS reveals the topology and connectivity of interacting regions by using solution chemistry and gas-phase analysis (via MS). The two strategies complement each other by providing architectural information on multiprotein assemblies.

Because cross-linking occurs via strong, covalent bonding, integrating it with other approaches may allow detection of transient protein−protein interactions that are hard to investigate otherwise. For example, the yeast initiation factor 3 (eIF3) protein, consisting of six subunits, regulates translational processes as a scaffold for many initiation factors. The molecular architecture, however, is elusive owing to its dynamic nature. Politis et al.[833] in 2015 determined the composition of the complex by sequentially dissociating subcomplexes, monitoring them by native MS, and coupling that information with physical contacts and approximate distances taken from cross-linking. The integrated information was encoded as connectivity restraints for scoring purposes in the model generation. The outcome is a high-resolution, 3D topological model of eIF3 assembly in complex with eIF5. To preserve the transient and dynamic interactions, a two-step cross-linking protocol was utilized, in which formaldehyde fixation is implemented at substoichiometry levels prior to initiating the cross-linking chemistry.[834] Structural stabilization of this type greatly facilitates cross-link formation and can be reversed at high temperature before MS identification.

Other MS-based footprinting methods have also been used in conjunction with XL-MS (e.g., footprinting by targeted labeling reagents[835] and HDX[773,836−838]). In one example, Zhang et al.[773] evaluated the combination of HDX, XL-MS, and molecular docking for characterization of the binding interface of interleukin 7 and its α-receptor. HDX reports widespread protection in the receptor. There is, however, no differential evidence of binding-induced protection or remote conformational change. By employing different reagents in XL, the investigators significantly increased the spatial resolution of binding site assignment. To generalize the integrated approach, protein−protein docking was executed with different number of cross-links as distance restraints. With model clustering and HDX adjudication, a high confidence model could be obtained with only two cross-links (root-mean-square deviation (RMSD) below 2.0 Å with respect to the known crystal structure).

Integrated methods can now play a role in protein structural modeling. By utilizing the structural restraints from DEPC labeling, cross-linking, and native MS for protein structural modeling, Politis and co-workers[835] in 2017 successfully constructed structural models of several protein complexes, tryptophan synthase, carbamoyl phosphate synthetase, and the double heterohexameric ring RvB1/2 with high-resolution subunit structures. A comparison of a representative model for tetrameric tryptophan synthase derived from all available structural restraints and its corresponding crystal structure (Figure 11) show agreement with a RMSD of 9.6 Å. Systematic evaluation of the modeling show that restraints from DEPC labeling and chemical cross-linking markedly increase the predictive power.

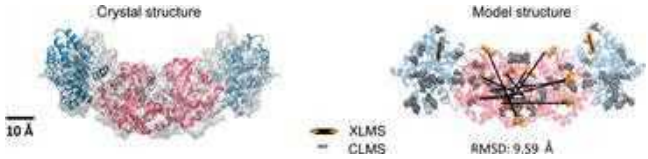


Figure 11. Representative model of the tetrameric tryptophan synthase and its corresponding crystal structure. Inter-residue proximities (XL-MS) and residue solvent accessibilities (covalent labeling MS, CL-MS) are highlighted. Reproduced with permission from ref 835. Copyright 2017 American Chemistry Society.

The workflow was tested by employing three model systems and then applied to a study of the F-type ATP synthase from spinach chloroplasts, a protein complex without a high-resolution structure. By integrating these experimental measurements with molecular dynamics simulations, the conformational states of the peripheral stalk as well as the flexible regions within the complex were revealed. This work demonstrates the potential of MS-based footprinting data to aid computer modeling.

In other work in 2019, Aprahamian and Lindert[839] describe in detail a Rosetta script that utilizes DEPC footprinting data to provide better restraints in protein structure predictions. Restraints from cross-linking can direct the modeling of protein complexes with multiple subunits; however, the short supply of modifiable residues for DEPC labeling constrained the SASA information. Fast labeling reagents (radicals) generally modify multiple residues in a single footprinting experiment that overcomes this drawback. More details about computer-aided protein structural modeling and radical footprinting-aided protein structural predictions can be found in section 6.6.

4.8. Conclusion

To conclude, protein footprinting by targeted labeling reagents has been utilized to address many biological questions and is considered as a powerful tool for MS-based protein HOS analysis. As mentioned at the beginning of the section, two major disadvantages limit its broad application. The limited number of targeted residues can be overcome by developing and adopting novel labeling reagents. Fast radical species can also complement such an approach, as discussed in sections 5 and 6.

Another question is whether the targeted protein undergoes a conformational change in the early stages of labeling and is now labeled in that new state (termed overlabeling), leading to potentially composite and misleading structural information. In a study of interferon-β1a, Kaltashov and co-workers[840] demonstrated that alkylation of Cys17 induces steric clashes within the native structure, as seen by the unfolding of helix D containing a 88−102 segment. On the contrary, there are also numerous studies to support the validity of targeted protein footprinting, either by a combination of multiple footprinting approaches[743,841,842] or by comparing the targeted footprinting results with crystal structures.[174,549,843] In a particular example, Li et al.[743] used HDX, GEE footprinting and hydroxyl radical footprinting to study the binding region of IL-6R to adnectin 1 and 2. All three footprinting approaches suggest the same binding region of 135−141. In other words, the validity of the slow but targeted GEE labeling is supported by two other protein footprinting approaches that function on different time scales (20 s to 10 min for HDX, submillisecond for hydroxyl radical footprinting).

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

The contradicting judgements on the effectiveness and validity of protein footprinting by targeted reagents are system dependent. In selected systems, modification of critical residues can induce protein conformational changes. It is thus important to characterize a protein system with multiple footprinting approaches or investigate the effect of modification by other methods (e.g., CD, NMR) before embarking on an extensive study. Different labeling time scales of these approaches can also be utilized to provide complementary and synergistic understanding of a protein's HOS.[842]

5. FAST LABELING REAGENTS: REACTIVE RADICAL SPECIES

Thus, far, we reviewed two means of protein footprinting: HDX and specific amino acid labeling. Both involve covalent labeling, but HDX footprinting is reversible, imposing constraints on protein and peptide analysis to minimize back exchange. HDX is less disruptive of protein structure than specific amino acid labeling because it occurs in medium in which H_2O has been replaced by D_2O. The concentration of the "reagent" D_2O is over 50 M, a concentration that changes negligibly over the time of footprinting. Exchange occurs by pseudo-first-order kinetics on the microsecond-to-many-hours time scale. The coverage is nearly complete, including all the common amino acids except Pro, and the modifications that are "counted" occur on the peptide backbone.

In contrast, specific amino acid labeling has low coverage, modifying one or a few amino acid side chains. Some reagents that modify $-COO^-$ or $-NH_3^+$ groups change the local charging of the protein; other reagents add bulky steric hindrance to the protein (even as large as another protein), both modifications posing a danger of structure disruption. Although a few footprinting reactions occur on the millisecond time scale, most require minutes or more during which the modified protein can change conformation and continue to react and produce a footprint that is composite. The concentrations of reagents are 10^2–10^4 times greater than that of the protein (there are some demonstrations in which the reagent is only 2–4 times in excess[754]), unlike HDX, where the concentration of D_2O is ~10^7 times greater than that of the protein. Considering that a large protein could have 10^3 potentially reactive amino acid residues, the reagent concentration may change, causing the reactions to have mixed kinetics.

The third class of footprinting reagents, fast labeling regents, enjoys features of HDX and specific amino acid labeling but has some important advantage over both. The modifying reagents are usually radical species that react with solvent-accessible amino acid residues to report on protein SASA. This form of footprinting shares characteristics with both HDX and specific amino acid labeling. Given the high reactivity of radical species, the labeling time can be as short as microseconds,[844] which makes possible the fastest footprinting available to follow protein dynamics, providing a "snapshot" of the protein SASA with residue-level resolution for those residues that react. Fast footprinting can also probe protein folding/unfolding. The theoretical coverage is broader than that of specific amino acid labeling, but it is nearly impossible to realize the coverage of HDX because the reactivity between radicals and the side chains can vary by ~10^3, and the footprinting will emphasize the more reactive side chains. The reagent concentrations are 10^2–10^3 times to that of a protein; for large proteins and reactive reagents, the free radicals will become limiting reagents. Both specific amino acid labeling and free radical labeling usually produce irreversible modifications, thereby facilitating effective protein isolation and digestion prior to the MS analysis.

In this section, we will review all available fast labeling reagents and describe their generation, development, biological relevance, and application in protein footprinting. We will also discuss their advantages and limitations.

5.1. Hydroxyl Radical

Hydroxyl radical ($^\bullet$OH) was initially used for nucleic acid footprinting, where it reacts with solvent accessible hydrogens on deoxyribose (DNA) and ribose (RNA), cleaving their nucleic acid chains.[169] The motivation is to footprint nucleic acids, to reveal their SASA and determine their HOS, and to map protein DNA/RNA interactions. Later, the idea was extended to protein footprinting. For DNA/RNA footprinting, the approach is based on selective cleavages, whereas with protein footprinting, the approach is based on chemical modifications. Today, $^\bullet$OH has become the most used fast labeling reagent.

5.1.1. Biological Relevance. The hydroxyl radical is a naturally occurring reactive oxygen species (ROS) that participates in several cycles that regulate several physiological functions of living organisms.[845−847] Early studies identified the important signaling role of $^\bullet$OH in regulating several physiological functions of living organisms, especially its contribution to retrograde redox signaling from organelles to the cytosol and nucleus.[845,847] The size and hydrophilicity of $^\bullet$OH are similar to those of the water molecule,[848] allowing it to probe the SASA of biomacromolecules (to travel where water can travel). The oxidative damage in many pathologies caused by $^\bullet$OH naturally and in radiation damage attracted early attention, motivating determination of its rate constants[175,849] and mechanisms[850] in amino acid reactions. The reaction rate constants between $^\bullet$OH and free amino acids range from 1.7×10^7 to 3.5×10^{10} M^{-1} s^{-1}, with preference toward aromatic, heterocyclic, and sulfur-containing side chains (Table 1). The broad residue coverage in $^\bullet$OH reactions contrast significantly with the targeted or specific labeling reagents covered above. The early fundamental studies, likely driven by radiation concerns, form a solid

Table 1. Rate Constants for Reactions between Free Amino Acids and $^\bullet$OH[175,849]

substrate	rate constant (M^{-1} s^{-1})	pH[a]	substrate	rate constant (M^{-1} s^{-1})	pH[a]
Cys	3.5×10^{10}	7.0	Pro	6.5×10^8	6.8
Trp	1.3×10^{10}	6.5–8.5	Gln	5.4×10^8	6.0
Tyr	1.3×10^{10}	7.0	Thr	5.1×10^8	6.6
Met	8.5×10^9	6–7	Lys	3.5×10^8	6.6
Phe	6.9×10^9	7–8	Ser	3.2×10^8	~6
His	4.8×10^9	7.5	Glu	2.3×10^8	6.5
Arg	3.5×10^9	6.5–7.5	Ala	7.7×10^7	5.8
cystine	2.1×10^9	6.5	Asp	7.5×10^7	6.9
Ile	1.8×10^9	6.6	Asn	4.9×10^7	6.6
Leu	1.7×10^9	~6	Gly	1.7×10^7	5.9
Val	8.5×10^8	6.9			

[a]"pH" indicates the conditions where the rate constants were measured.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

foundation for the use of •OH in protein footprinting, contributing to its growing applications.

5.1.2. Generating •OH in Solution. The generation of •OH has a long history, during which many different methods were developed. In this section, only methods that are useful for footprinting biomacromolecules will be reviewed. Early approaches made use of •OH-induced protein backbone cleavages,[169] similar to those in nucleic acid footprinting. Upon cleavage, gel electrophoresis was used to determine the cleavage sites. Low cleavage efficiency, lack of accurate mass determination, and low precision greatly limit the further development of such an approach; thus, it will not be covered in detail.

5.1.2.1. Fenton and Fenton-like Chemistry. The catalytic property of Fe(II) in promoting the oxidation of tartaric acid by hydrogen peroxide (H_2O_2) was first reported by Fenton in 1894.[851] Later on, Cu(I), Ti(III), Co(II), and Cr(III) were found to behave similarly.[852−855] Such processes were later called Fenton reactions. Although it was supposed that •OH is the key reactive species in Fenton chemistry, it was not well established until Haber and Weiss[856,857] first proposed a chain reaction mechanism in 1932. The understanding of the mechanism was later expanded by Barb and co-workers,[858,859] and it is now referred to as the "classical Fenton pathway", as summarized in eqs 2−8.

$$Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + {}^{\bullet}OH + OH^- \quad (2)$$

$$Fe^{3+} + H_2O_2 \rightarrow Fe^{2+} + HO_2^{\bullet} + H^+ \quad (3)$$

$$H_2O_2 + {}^{\bullet}OH \rightarrow HO_2^{\bullet} + H_2O \quad (4)$$

$$H_2O_2 + {}^{\bullet}OH \rightarrow HO_2^{-\bullet} + H_2O \quad (5)$$

$$Fe^{3+} + HO_2^{\bullet} \rightarrow Fe^{2+} + O_2 + H^+ \quad (6)$$

$$Fe^{3+} + O_2^{-\bullet} \rightarrow Fe^{2+} + O_2 \quad (7)$$

$$Fe^{2+} + {}^{\bullet}OH + H^+ \rightarrow Fe^{3+} + H_2O \quad (8)$$

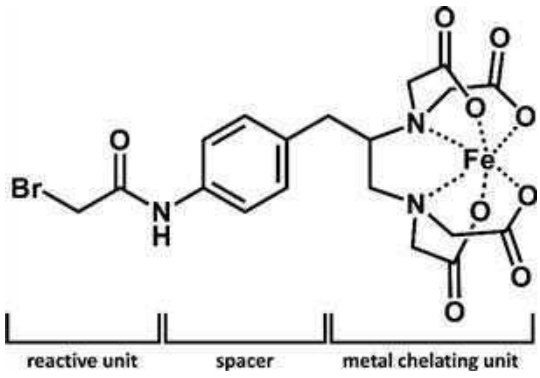
In Fenton chemistry, •OH is generated by oxidation of Fe(II) and Fe(III) with H_2O_2 (eqs 2 and 3). This chemistry is not optimum for footprinting biomacromolecules. For example, the reaction has an optimal rate at pH of 3−4,[858,859] which is clearly denaturing for many biomolecules. Under physiological conditions, however, Fe(III) readily precipitates.

In 1985, Tullius and Dombroski[860−862] proposed an elegant system that utilizes Fenton chemistry to map DNA−protein binding sites. By incubating Fe(II)-ethylenediaminetetraacetic acid (Fe(II)-EDTA), H_2O_2, and ascorbate with DNA, they achieved an effective footprint after a few minutes. The approach exploits Fenton chemistry in two ways. First, with EDTA as chelators for Fe(II) and Fe(III), the solubility of these ions under physiological pH increases significantly. EDTA also minimizes the binding of these ions to biomacromolecules and increases their catalytic efficiency.[863] Second, ascorbate can reduce Fe(III)-EDTA back to Fe(II)-EDTA, fulfilling a cycle. Thus, the Fe(II-EDTA)/H_2O_2/ascorbate Fenton system had become the usual approach for footprinting, and it was applied in many subsequent DNA footprinting studies.[864]

For protein footprinting, the system was first applied to footprint cAMP receptor protein under backbone-cleavage conditions by Heyduk and co-workers[865] in 1994 and then by Sharp and Hettich[866] in 2003, who extended oxidative labeling to aMb by using •OH to map amino acid side chains as a measure of their SASA. A more recent variation on this theme by Monroe and Heien[867] combines electrochemistry with this classical reaction in which an electrical flow cell was used to reduce Fe(III) to Fe(II), thus completing a cycle. Footprinting of ubiquitin demonstrated the efficacy of such method in probing protein SASA.

Another powerful approach using Fenton chemistry was first demonstrated by Rana and Meares[868−870] in 1990. As a replacement of Fe(II)-EDTA in a classical Fenton system, they synthesized a novel class of reagents represented by Fe-(*S*)-1-(*p*-bromoacetimidobenzyl)-EDTA (Fe-BABE). Fe-BABE is composed by three structural units (Scheme 14), a metal-

Scheme 14. Structure of Fe-BABE



chelating unit that binds with reactive metal centers to facilitate •OH generation, a sulfhydryl reactive unit that reacts with cysteine residues on proteins, and a linker separating them. The design localizes the Fe-BABE in a specific position of a protein. Upon incubating with H_2O_2, •OH is only generated in close proximation with the localized metal center. Fe-BABE has the advantage of high conjugation rate with proteins, high anchoring yield, and compatibility with neutral pH, making such site-directed •OH footprinting effective in probing spatial relationships in protein−nucleic acid[871−873] and protein−protein[874] complexes. Very recently, this approach was adopted for characterizing protein−carbohydrate interactions,[875] whose details will be discussed in section 6.8.

Fenton chemistry is probably the most easily accessible method for •OH protein footprinting, as it can be achieved without the help of sophisticated instruments. The biological relevance of iron makes it "friendly" with most biological systems, as demonstrated by in-cell footprinting of the membrane protein porin OmpF.[876] Moreover, given that multiple reactive metal centers can initiate Fenton-like reactions, successful •OH protein footprinting can be achieved by using Co(II),[877] Cr(III),[877] Ni(II),[878] Cu(II),[879,880] and Mn(II)/Mn(III) for generating •OH upon reaction with H_2O_2.[854] This general property makes Fenton-like chemistry worth considering to footprint proteins that bind to these active metal centers.

On the other hand, •OH generation by Fenton chemistry usually takes minutes, during which the protein may get over labeled and alter its conformation during the footprinting owing to the changes in hydrophilicity or perhaps to some modification of amino acid residues with low SASA.[881] Once

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

the protected region becomes exposed, additional $^{\bullet}$OH can further label the newly exposed sites, generating a misleading readout. Newer $^{\bullet}$OH generation approaches usually label proteins on the time scale of milliseconds or less to minimize overlabeling. Although a subsequent study employing Fenton chemistry achieves a steady $^{\bullet}$OH concentration at 2 ms, the amount of H_2O_2 (30%) required is certainly stressful to most biomacromolecules.[882]

5.1.2.2. Synchrotron Water Radiolysis. A synchrotron X-ray source is capable of delivering a continuous spectrum of 10^{14}–10^{15} photons per second with energies ranging from 5 to 30 kV.[883] Implemented by Chance and co-workers,[884−886] synchrotron-based $^{\bullet}$OH footprinting was first demonstrated for nucleic acids and later for proteins.[887] The latter initiates the modern era for protein footprinting by fast-labeling reagents because, for the first time, the labeling time scale was reduced to milliseconds.[848] Mechanistically, synchrotron X-rays utilize water as the $^{\bullet}$OH precursor. High energy photons ionize water to produce hydrated electrons (e_{aq}^-) and activated water (H_2O^*), and these initiate the subsequent reactions as summarized in eq 9−11.[175]

$$2H_2O + e^-(\text{ionizing irradiation}) \rightarrow H_2O^{+\bullet} + H_2O^* + e_{aq}^- \quad (9)$$

$$H_2O^{+\bullet} + H_2O \rightarrow {}^{\bullet}OH + H_3O^+ \quad (10)$$

$$H_2O^* \rightarrow {}^{\bullet}OH + H^{\bullet} \quad (11)$$

Besides labeling solvent-accessible amino acid side chains in the protein, $^{\bullet}$OH undergoes quenching reactions owing to its high reactivity. Under anaerobic conditions, primary quenching is a self-recombination at a diffusion rate limit (eq 12−13).[888]

$$^{\bullet}OH + e_{aq}^- \rightarrow OH^- \quad (12)$$

$$2{}^{\bullet}OH \rightarrow H_2O_2 \quad (13)$$

Under aerobic conditions, oxygen is involved in the quenching process, leading to other ROS (eqs 14−15).

$$e_{aq}^- + O_2 \rightarrow O_2^{-\bullet} \quad (14)$$

$$H^{\bullet} + O_2 \rightarrow HO_2^{\bullet} \quad (15)$$

In subsequent studies, $^{\bullet}$OH dosimetry, the effects of different buffers, different beam currents, and other additives in the system were thoroughly investigated.[889] These efforts were comprehensively reviewed elsewhere.[175]

Synchrotron-based water radiolysis has significant advantages in protein footprinting, as it utilizes water as the $^{\bullet}$OH precursor. Water concentration is ∼55 M, and it is not limiting. Although water interacts with proteins, there are many more molecules that constitute bulk, and this may not be the case for other radical sources that are at lower concentrations. More importantly, water is the medium for biological systems, and its use simplifies the experimental procedure. The fast labeling speed, easy dose control (by controlling exposure time with a shutter[889]), and high reproducibility all contribute to its utility. An obvious downside is there are limited number of synchrotron sources. A recent development of a new beamline should make the method more powerful and more accessible than in the past to the general research community.[890]

5.1.2.3. Laser Photolysis of Hydrogen Peroxide. Upon exposure to UV around 250 nm, H_2O_2 homolytically cleaves into two $^{\bullet}$OH with a primary quantum yield of 0.4−0.5.[891,892] In the absence of reactive substances, the resulting $^{\bullet}$OH reacts with the H_2O_2 that did not undergo photolysis as shown in eqs 16−18.

$$H_2O_2 \rightarrow 2{}^{\bullet}OH \quad (16)$$

$$^{\bullet}OH + H_2O_2 \rightarrow H_2O + HO_2^{\bullet} \quad (17)$$

$$HO_2^{\bullet} + H_2O_2 \rightarrow H_2O + O_2 + {}^{\bullet}OH \quad (18)$$

The rate constants for the latter two reactions are 2.7×10^7 and 7×10^9 M^{-1} s^{-1}, respectively.[892] Large reaction rate constants guarantee fast labeling.

In 2004, Sharp et al.[893] first footprinted lysozyme and β-lactoglobulin by $^{\bullet}$OH generated from H_2O_2 photodissociation. For this protocol, the protein sample contains 15% H_2O_2, and the irradiation (by a UV lamp) takes up to 5 min. This high H_2O_2 concentration is stressful to many proteins. Moreover, the long labeling time is a risk that the system will be overlabeled, as discussed above.

To reduce the labeling time and lower the H_2O_2 concentration, Aye and Sze[894] replaced the UV lamp with a pulsed Nd:YAG laser that operates at 266 nm (frequency quadrupled). They used a light pulse energy of 2 mJ/pulse and a pulse width of 3−5 ns in a static system. H_2O_2 was added to 0.3% just prior to laser irradiation to minimize exposure of the protein to this mild oxidizing agent. The residual H_2O_2 after irradiation was removed by snap-freezing and lyophilizing the aliquot (possibly a source of error as protein oxidation by H_2O_2 can occur in the solid state at low T as reported later[895]). These improvements over Fenton chemistry combine to achieve a moderate oxidation of ubiquitin with a single laser shot. As compared with Fenton chemistry, the H_2O_2 concentration and labeling time were both reduced significantly, providing higher confidence of the result.

At the same time, Hambly and Gross[844] reported a method that couples H_2O_2 laser photolysis and a flow system to footprint proteins oxidatively, which was later named fast photochemical oxidation of proteins (FPOP). As compared with using a Nd:YAG laser, the KrF excimer laser they chose has an output of 248 nm, which is closer the λ_{max} of H_2O_2. The laser pulse width was 17 ns, and the power was 50 mJ/pulse, both readily achieved with such a laser. These efforts combine to afford measurable oxidation levels of proteins with 0.04% H_2O_2. The flow system, as illustrated in Figure 12, ensures that most proteins in the aliquot are only irradiated once. The flow rate was calculated in correspondence with laser pulse width and frequency to allow a 20−25% exclusion volume, which minimizes any "double shots" of the protein solution.[896,897] Leftover H_2O_2 was decomposed by catalase in the collection tube. Moreover, the reaction duration was controlled and became tunable by adding a radical scavenger (Glu[844] and later His[186]), which not only limits the labeling time to 0.5 μs for minimizing overlabeling but also allows better control of the labeling process[881] and makes possible probes of fast kinetic processes including protein folding/unfolding.[898,899] These advantages make FPOP a widely adopted platform for protein footprinting by fast labeling reagents.

As compared to synchrotron water radiolysis, H_2O_2 photolysis retains most of its advantages in generating $^{\bullet}$OH. Although H_2O_2 is a necessary precursor, low concentrations

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

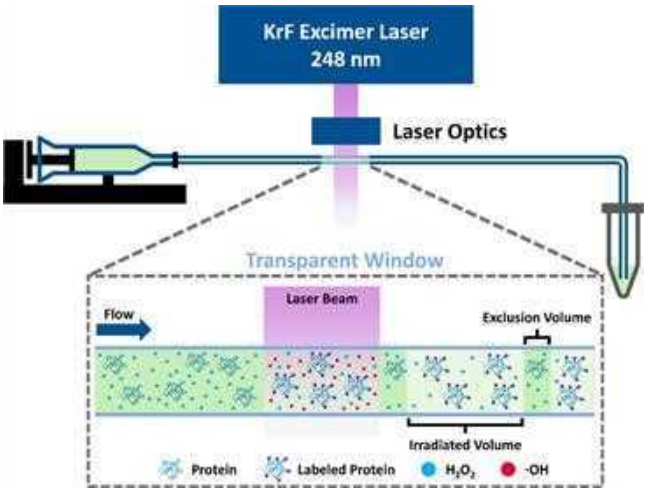


Figure 12. Schematic illustration of FPOP setup.

minimize any oxidation interference to protein native states. Most importantly, H_2O_2 photolysis makes the ultrafast •OH protein footprinting more accessible to the general research community, although efforts are underway, principally by a commercial developer, to incorporate a discharge lamp to replace the laser, which imposes safety requirements and some expertise for handling. Nevertheless, subsequent developments take advantage of the strengths of FPOP, and a description of several applications will be given in section 6.

On the other hand, H_2O_2 is a relatively strong oxidant, which may make this approach troublesome for proteins that are prone to oxidation. A modified FPOP apparatus with a mixer located just before the laser window largely minimizes the H_2O_2-induced protein oxidation.[900] A recent report suggests that H_2O_2 can interact with the protein, giving local high concentrations of both the H_2O_2 and the radicals on regions of the protein surface.[145]

In response to the original claim that the radicals have been scavenged in ~1 μs, Konermann and co-workers[901] employed a reporter dye to follow the FPOP kinetics and showed that the free radical oxidation of the dye extends over tens of ms, indicating a longer radical lifetimes than the original prediction. This phenomenon likely applies to the secondary radicals formed as •OH reacts. It remains to be seen if these less reactive radicals react with the protein. The assertion about μs lifetime is probably correct for the primary radicals (i.e., •OH).

Another consideration for this platform is the potential migration of initially formed protein radicals to other reactive residues that are buried below the surface. The investigators addressing that question found that two solvent-inaccessible residues were modified by FPOP, possibly in accord with a potential long-distance radical transfer (radical jumping) to buried residues.[902] This and other considerations need further attention to make FPOP more robust and applicable to more biological systems.

5.1.2.4. High Voltage Electrical Discharge. High voltage electrical discharge takes advantage of the requirement to use mass spectrometry in the analysis by a simple adaption of a typical ESI source to increase the spray voltage to 6−8 kV.[903] To make the radicals, Downard and co-workers[903−906] increased the potential difference between the emitter tip and the grounded collection plate (either the MS inlet or a collection vessel) to induce a corona discharge, during which oxygen in the nebulizer gas will be activated to give a plasma at the tip of the emitter (eq 19−20).

$$O_2 \rightarrow O_2^{+\bullet} + e^- \quad (19)$$

$$O_2^{+\bullet} + 2H_2O \rightarrow H_3O^+ + {}^\bullet OH + O_2 \quad (20)$$

Meanwhile, water from the ESI spray can be activated as well (eq 21).

$$2H_2O \rightarrow H_2O^{+\bullet} + e_{aq}^- \rightarrow H_3O^+ + {}^\bullet OH + e^- \quad (21)$$

By switching the nebulizer gas from air to pure O_2, the oxidation efficiency can be doubled.[906] The approach is also compatible with both positive and negative polarities.[907] The oxidatively labeled protein is then submitted directly to the mass spectrometer for top-down analysis (proteins are dissolved in ammonium acetate to facilitate sufficient ionization), or the protein can be accumulated in a collection vessel for further treatment and proteolysis. In subsequent studies, an O_2-assisted electrical discharge was successfully utilized to study protein complexes between ribonuclease S-protein−S-peptide[908] and calmodulin−melittin.[909] In a more recent twist, Maleknia and Downard[910] joined MALDI and electrical discharge protein footprinting to afford an even higher throughput.

Given the popularity of ESI in protein analysis, high voltage electrical discharge makes •OH protein footprinting readily available to almost every MS lab. The fast-labeling time frame and high labeling efficiency seem to be promising. Given that charge-induced protein unfolding occurs in an ESI source,[911] a question arises whether the protein of interest can maintain its high-order structural integrity during the oxidative footprinting. It is always preferred to perform protein footprinting under native or near-native states to minimize these ambiguities, where a high-voltage ESI spray is not optimal. Moreover, it is challenging to footprint binding systems with weak interactions through this method, as weak interactions are less likely to be preserved in an ESI droplet than in neutral solution. The labeling time frames and the radical lifetimes also remain to be determined, and the online approach will be difficult to apply to large proteins.

The concern of charge-induced unfolding by electrical discharge-based •OH labeling was recently addressed in 2017 by Minkoff and Sussman, who developed a novel experimental approach named plasma induced modification of biomolecules (PLIMB, Figure 13).[912] Instead of delivering the protein through an ESI source, the protein aliquot was placed in a grounded tube with a charged needle on top of the liquid. A plasma discharge is induced when supplying 1−31 kV of potential to the needle with frequencies of 0−15 kHz. Over a period of 60 s, denatured and native bovine serum albumin were oxidatively labeled to achieve modification fractions up to 15% at the peptide level. As part of the initial demonstration, PLIMB was adopted to characterize the epidermal growth factor-induced structural changes in the extracellular domain of its receptor. These examples demonstrate efficacy in facilitating •OH-based protein footprinting on this platform.

PLIMB offers a benchtop solution for radical protein footprinting by overcoming the major drawback for conventional electrical discharge methods, as the protein of interest is now footprinted in solution under near-native conditions. On the other hand, the plasma discharge in PLIMB only happens

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

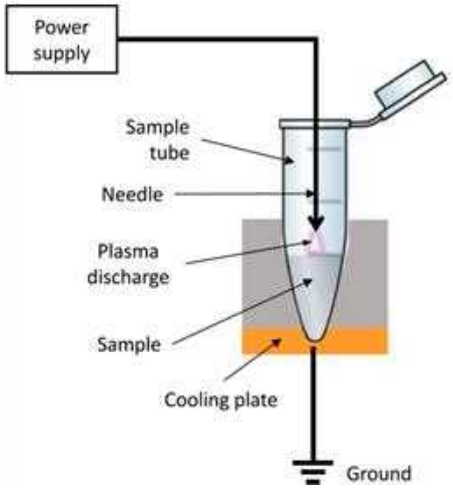


Figure 13. Schematic illustration of PLIMB setup. Adapted with permission from ref 912. Copyright 2017 Springer Nature.

at the surface of the aliquot. Given the short lifetime of •OH,[175] most of the •OH labeling occurs in limited regions of the solution. Means to uniformly label the sample aliquot and to minimize the over labeling are likely to be future developments of PLIMB.

5.1.2.5. Gamma-Ray Water Radiolysis. Gamma ray (γ-ray) is high energy electromagnetic radiation. Similar to synchrotron X-ray water radiolysis, γ-rays excite water molecules and trigger reactions that produce •OH (eqs 2−8).[913] γ-ray-based •OH footprinting, first applied to study nucleic acids,[914] was implemented by Nukuna et al.[915] to footprint cytochrome C in 2004. This approach was adopted by a few research groups in subsequent years,[916−921] but more recently, the approach is seldom used due to safety concerns of the highly penetrable nature of γ-rays and the emergence of other, more readily used approaches.

5.1.2.6. Other Hydroxyl Radical Generation Methods. Besides the methods introduced above, there are other methods for •OH footprinting, including ultrasound sonolysis,[922,923] fast neutrons,[924−926] peroxynitrous acid decomposition,[927,928] pulsed electron beam,[929] photolysis of *N*-hydroxypyridine-2(1*H*)-thione,[930,931] boron-doped diamond electrochemical surface mapping,[932] and ozonolysis.[906] Because these methods were only used in nucleic acid footprinting or not developed for protein footprinting, they will not be covered further in this review.

5.1.3. Residue Specificity and Proposed Reaction Pathways. One of the unique advantages of •OH in protein footprinting is its broad reactivity. Xu and Chance[175,917,933−935] conducted systematic studies and found that the reactivity between •OH and amino acid residues rank as Cys > Met > Trp > Tyr > Phe > cystine > His > Leu, Ile > Arg, Lys, Val > Ser, Thr, Pro > Gln, Glu > Asp, Asn > Ala > Gly. Among all these residues, Gly, Ala, Asp, and Asn are unlikely to serve as useful probes owing to their low reactivity (also see Table 1). Although the reactivities for Ser and Thr are higher than that of Pro, which has been found to be a reactive and useful substrate, their oxidation products are not easily detectable and may be dispersed among several pathways (e.g., oxidation of −OH to =O). From rate constant determinations and reports of experience with •OH and protein modification, the general conclusion is that 14 out of 20 amino acid residue side chains are active in •OH-based protein footprinting experiments. In our experience, the presence of highly reactive residues (Cys, Met, Trp, Tyr, Phe, His) may siphon most of the reagent radicals to their modification, reducing the coverage. The products resulting from oxidative modifications are residue-dependent, and they have been covered in detail in the Xu and Chance review;[175] most are +16 Da oxidations as summarized in Table 2.

Table 2. Primary Oxidation Produces (upon •OH Footprinting) and the Corresponding Mass Changes (in Reactivity Order) For Various Amino Acid Residue Side Chains[917]

side chain	modification and mass changes
Cys	sulfonic acid (+48), sulfinic acid (+32), hydroxy (to serine, −16)
Met	sulfoxide (+16), sulfone (+32), aldehyde (−32)
Trp	hydroxy- (+16, + 32, + 48, etc.), pyrrol ring-open (+32)
Tyr	hydroxy- (+16, + 32, etc.)
Phe	hydroxy- (+16, + 32, etc.)
cystine	sulfonic acid (+48+H), sulfinic acid (+32+H)
His	oxo-(+16), ring-open (−22, −10, +5)
Leu[a]	hydroxy- (+16), carbonyl (+14)
Ile[a]	hydroxy- (+16), carbonyl (+14)
Arg	deguanidination (−43), hydroxy- (+16), carbonyl (+14)
Lys	hydroxy- (+16), carbonyl (+14)
Val[a]	hydroxy- (+16), carbonyl (+14)
Ser[b]	hydroxy- (+16), carbonyl (−2, or +16-H2O)
Thr[b]	hydroxy- (+16), carbonyl (−2, or +16-H2O)
Pro	hydroxy- (+16), carbonyl (+14)
Gln	hydroxy- (+16), carbonyl (+14)
Glu	decarboxylation (−30), hydroxy- (+16), carbonyl (+14)
Asp	decarboxylation (−30), hydroxy- (+16)
Asn	hydroxy- (+16)
Ala	hydroxy- (+16)
Gly	N/A

[a]For aliphatic side chains, +14 Da products are normally much less than +16 Da products. [b]For Ser and Thr, only trivial amount of +16 and −2 Da products were found.

Mechanistically, •OH activates the solvent accessible side chains by either removing H• from an activated site or adding a •OH to a double bond, both of which result in a protein radical that is subsequently quenched differently depending on the origin of the radical (H_2O or H_2O_2).

For synchrotron radiolysis, Xu and Chance,[175,933−935] in a comprehensive review, summarized the modification pathways of •OH-active amino acid residue side chains.[175] Leu, Ile, Val, Pro, Arg, Lys, Glu, and Gln side chains are activated through H• abstraction, whereas His, Phe, Tyr, Trp, and Met are primarily by •OH addition. The resulting protein radicals are subsequently quenched by dissolved O_2 in aerobic conditions. Reaction pathways between •OH and the extremely reactive Cys and cystine residues are complicated and not well understood. Alternatively, a likely reaction pathway that covers major oxidative products was proposed for Cys.[935] Similar pathways should apply to the electrical discharge-based •OH labeling, as the precursor for •OH is also water.

In contrast to the two systems mentioned above, where uniformly distributed water is the •OH precursor, any H_2O_2 photolysis approach, as represented by FPOP, follows slightly different reaction pathways. A recent quantum calculation

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

study revealed the hydrogen bonding between H_2O_2 and amino acids including His, Arg, Tyr, Cys, Thr, Gln, Asp, Lys, Met, and Trp.[936] Such hydrogen bonding will induce a preformed H_2O_2−amino acid residue complex, which will result in a local fluctuation in •OH concentrations upon laser photolysis. Liu, Gross, and co-workers[145] studied the reaction pathways between •OH and 13 different amino acid residues on a FPOP platform. By using ^{18}O-enrichment of all three available oxygen sources (H_2O, H_2O_2 and dissolved O_2) on an FPOP platform, one at a time, they differentiated three classes of residues based on their oxygen uptake preferences. His, Arg, Tyr, and Phe preferentially take oxygen from H_2O_2. Met competitively take oxygen from H_2O_2 and dissolved O_2, whereas Leu, Ile, Val, Pro, Lys, Asp, Glu, and Gln take oxygen exclusively from O_2. Given that the •OH activation pathway is very similar to that in water radiolysis, such differences in oxygen uptake preferences reveal a different pathway for some activated residues than described in the Xu and Chance[175] review. Besides quenching by dissolved O_2, selected activated residues can also be favorably modified by reaction with another •OH in the FPOP setup, owing to hydrogen bonding with H_2O_2 and preferential localization of the reagent and, of course, the radicals.

As mentioned earlier, •OH is also capable of cleaving protein backbones, whose pathways are reviewed elsewhere.[175,850]

Summarizing the role of •OH in footprinting, we recognize unprecedented advantages as a fast protein footprinter. Biological relevance, similar size and hydrophilicity as water, fast labeling time, high labeling efficiency, multiple generation methods, broad residue coverage, well established platforms, and mature understanding of the labeling pathways make it the most utilized fast-labeling reagent. Its demonstrated utility forecasts many more promising applications (see section 6). On the other hand, the +16 Da from oxidative labeling can be problematic in rare cases, as oxidation is common in native biological systems, and it is challenging to distinguish the oxidations that exist prior to labeling and those that are from •OH reactions. Other oxidations besides those producing +16 can add background of many low abundance products to complicate the analysis. Furthermore, even though 14 amino acid residues are reactive with •OH, those that are most reactive (S-containing, aromatics) can be too competitive, reducing the overall coverage and spatial resolution achieved for the protein. These drawbacks motivate the development of other fast labeling reagents that provide complementary coverage to •OH labeling and employ an MS tag other than +16 Da.

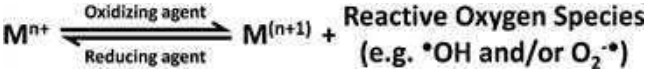
5.2. Sulfate Radical Anion and other Sulfur Containing Radicals

5.2.1. Biological Relevance. Sulfate radical anion, $SO_4^{\bullet-}$, is a potent oxidant with a standard reduction potential of 2.43 V at neutral pH.[937] Its strong oxidation capability can cause considerable damage to different cellular components, including lipids, carbohydrates, proteins, and DNA/RNA.[938] It is an exogenous factor that promotes other oxidant formation in biological systems,[939] and it has been applied to inactive pathogenic microorganisms in disinfection.[938] Many investigators hold that the oxidation pathways of biomolecules by $SO_4^{\bullet-}$ are like those of the •OH; however, detailed mechanistic studies are still needed.

5.2.2. Radical Generation in Solution and Applications in Biology. The production of the sulfate radical anion in vitro primarily utilizes two precursors, persulfate ($S_2O_8^{2-}$) and peroxymonosulfate (HSO_5^-). Activation methods are needed to cleave the O−O peroxide bond and give $SO_4^{\bullet-}$ generation. UV irradiation,[6] thermolysis, radiation,[2] and transition-metal catalysis[940,941] are possibilities. As a promising oxidant, $SO_4^{\bullet-}$ may be useful for answering biological questions. In 1983, the Werbin group[942] cross-linked DNA with lysosome to study reactive mechanism of a carcinogen, *N*-acetoxy-*N*-acetyl-2-aminofluorene (*N*-AcO-AAF). The Kodadek[940] and Jovin groups[941] both demonstrated efficient $SO_4^{\bullet-}$-generation from a ruthenium complex, Ru(bipyridine)$_3$, for cross-linking initiation. The latter study further shows that $SO_4^{\bullet-}$ induces tyrosyl radical formation, allowing inter cross-links between alpha-synuclein (αSyn) aggregates. The covalently bound large oligomers exhibit distinct biophysical properties and increased toxicity.

5.2.3. $SO_4^{\bullet-}$-Based Protein Footprinting: Residue Specificity and Proposed Reaction Pathways. Bridgewater and Vachet[943−945] in 2005−2006 first used sulfate radical anion to determine binding and map the SASA of proteins. Their generation of $SO_4^{\bullet-}$ utilizes a metal-catalyzed oxidation (MCO), a process that enables radical formation at the site of a redox-active metal when bonded to a peptide/protein (Scheme 15). Cu/Zn superoxide dismutase was chosen

Scheme 15. General Scheme That Produces Reactive Oxygen Species during Metal-Catalyzed Oxidation Reactions



as a model system in the presence of ascorbate and persulfate. High concentrations of ascorbate (i.e., 100 mM) rapidly initiate the redox cycle through Cu(II) reduction to provide scavenger-generated $SO_4^{\bullet-}$, which reacts by pseudo-first-order kinetics. The study showed that at an optimized $SO_4^{\bullet-}$ concentration (i.e., one generated from 1 mM persulfate) allows the oxidation to occur only on the nearby Cu-bound amino acids. Increasing the persulfate concentration allows modification of non-Cu-bound residues within 10 Å radius of the Cu center. UV−vis absorption spectroscopy and oxidation kinetics further confirmed the survival of the intact protein structure throughout the analysis, showing promise for future applications under their conditions. Although the steric effects of a specific residue may alter its oxidation extent, this tunable $SO_4^{\bullet-}$-based approach provides another means for 3D mapping of the protein environment around a reactive site. This approach may be particularly important to map metal-binding residues that have low solvent accessibility, and it may be incorporated into an FPOP platform.

In 2010, the Gross group[946] generated sulfate radical anion on the FPOP platform by photolysis of $^-OSO_2-O-O-O_2SO^-$ at a quantum yield of 0.55. As a stronger oxidant than hydroxyl radical, the sulfate radical anion gives comparable extents of modification on the protein-level with 3−5 times lower amounts of precursor (Figure 14a). Incubation with persulfate anion without laser irradiation does not introduce observable oxidation on the protein, showing negligible background from reactions of the precursor reagent. In addition, the investigators showed that the reactivity and specificity of $SO_4^{\bullet-}$ is similar to those of •OH for residues Met, Trp, Glu and Ser; the reactant radical favors

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

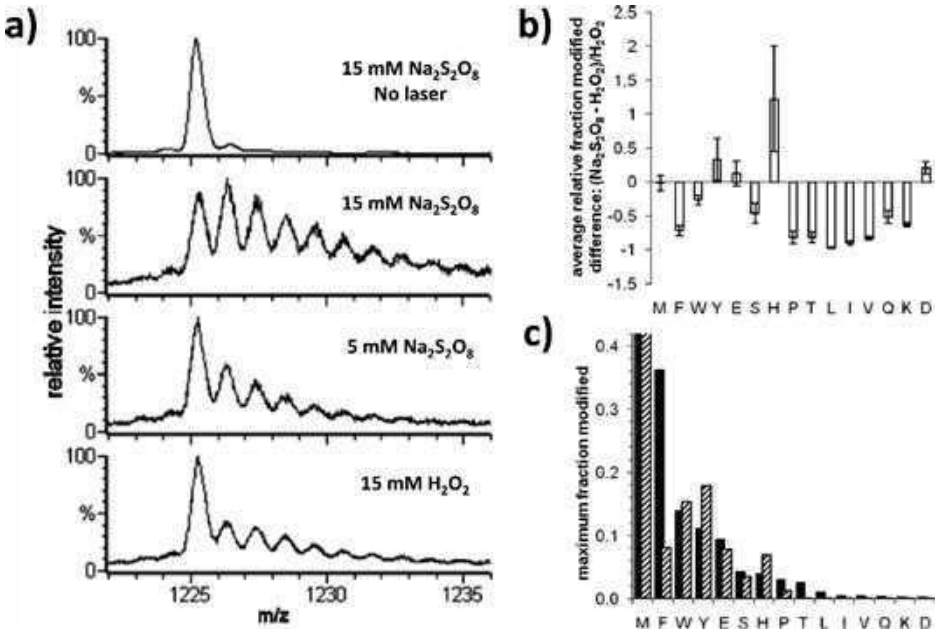


Figure 14. (a) The ESI-quadrupole time-of-flight mass spectra of the 15th charge state of β-lactoglobulin submitted to different labeling conditions as shown in the figure. (b) The relative difference in fraction modified between persulfate and the H_2O_2 approaches for CaM, aMb, bradykinin, and angiotensin II residues are averaged per amino acid type. The error bars denote the average pairwise-comparison standard error. (c) The maximum fraction modified among all the same amino acid residues is plotted. Black bars denote H_2O_2 FPOP. Reproduced with permission from ref 946. Copyright 2010 American Chemical Society.

His and Tyr (Figure 14b,c). The overall reactivity ranking of $SO_4^{-\bullet}$ is

Met > Tyr = Trp > Phe = Glu = His > Ser > Pro > Asp = Thr > Lys = Gln > Leu = Val = Ile

Sulfate-radical-induced oxidation can happen through multiple pathways. Aromatic and Met side chains are the main targets for the strong electrophile $SO_4^{-\bullet}$ and form a radical cation intermediate, which subsequently undergoes either fragmentation or hydration.[947,948] $SO_4^{-\bullet}$ can also oxidize a carboxyl anion ($R\text{-}COO^-$) to give acyloxyl radicals ($R\text{-}COO^\bullet$) followed by decarboxylation to generate carbon-centered radicals.[948] Moreover, hydrogen abstraction by $SO_4^{-\bullet}$ is another pathway to oxidative modification, likely occurring on Leu, Ser, and Phe.[949]

Surprisingly, although sharing some similarities, persulfate FPOP and hydroxyl FPOP also have different chemical outcomes. For systems where there is need to maintain more "physiologic-like" conditions, the lower amount of reagent for persulfate anion makes it a more desirable choice. Different physical properties of the two radicals also determine their preferred circumstances for surface mapping. Negatively charged $SO_4^{-\bullet}$ is prone to form electrostatic interactions with other positively charged groups, introducing a possible bias. One example is shown by Gau et al.,[946] who compared per-residue fractional modification for six histidines in aMb to their SASA values calculated from an X-ray and NMR study. A good correlation (R^2 = 0.83 in Figure 15b) shows the potential of $SO_4^{-\bullet}$ footprinting to be an effective experimental measure of SASA. The interaction of $SO_4^{-\bullet}$ and H64, an axial ligand of the heme iron, however, contributes to the enhanced reactivity of H64 (Figure 15a). On this matter, the hydroxyl radical, which bears no charge and has a comparable size to water, may be a better choice as footprinter that responds to SASA. In addition,

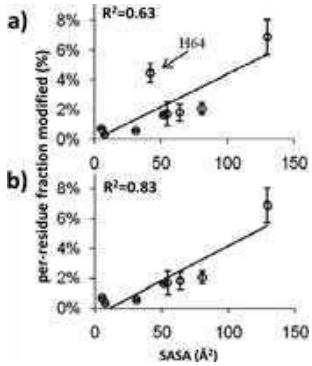


Figure 15. Modification yields of apomyoglobin His residues, where (a) all histidines were included and (b) His64 was omitted, are plotted against their calculated SASA, shown with least-squares straight-line fits. The SASA values were calculated by using a 1.4 Å probe. Reproduced with permission from ref 946. Copyright 2010 American Chemical Society.

the hydroxyl radical is more membrane-permeable, suitable for studies on micelles, liposomes, and nanodiscs.[950] The sulfate radical anion, on the other hand, reacts more efficiently on extracellular-accessible residues. The combination of the two in a tandem format may be effective in addressing complicated protein systems consisting of phospholipid bilayers imbedded with protein analyte.

5.3. Carbonate Radical Anion

5.3.1. Biological Relevance. The carbonate radical anion, $CO_3^{-\bullet}$, is formed primarily from the bicarbonate−carbon dioxide pair that exists in cells to maintain physiological pH.[951] Superoxide dimustase (SOD),[952] xanthine oxidase (XO),[953] and hemoprotein in an iron(III) state can adopt conditions whereby the buffer components are sources of $CO_3^{-\bullet}$, causing oxidative damage on biomolecules.[954] Owing to its longer half-

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

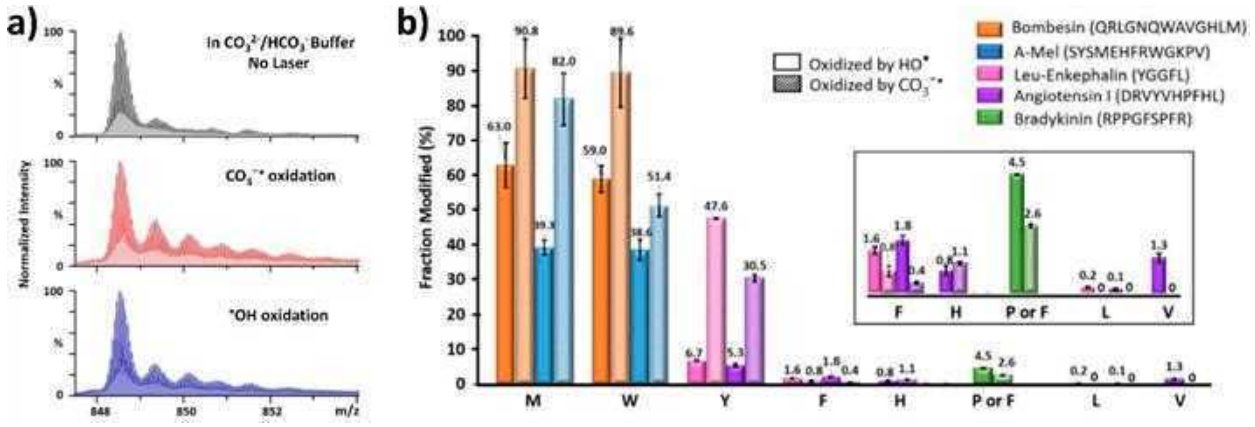


Figure 16. (a) Quadrupole time-of-flight mass spectra of the +20 charge state of apo-myoglobin under different labeling conditions as given in the figure. (b) Comparison of the residue-level fraction modified (in percentage) of a "peptide cocktail" by $CO_3^{-\bullet}$ and •OH. Different colors represent different peptides, and the corresponding sequences are shown. Solid bars denote the •OH oxidation extent, and patterned bars denote the $CO_3^{-\bullet}$ oxidation. The data were corrected with respect to a negative control. Error bars are the standard deviations of three independent runs. The inset is an enlarged portion of the figure. For bradykinin, only F undergoes oxidation with $CO_3^{-\bullet}$, whereas F and P are both oxidized by •OH but not distinguishable because the chromatograms of the two modified peptides overlap. Reproduced with permission from ref 962. Copyright 2019 Elsevier.

life compared to those of other radicals (e.g., •OH), $CO_3^{-\bullet}$ can diffuse farther and cause oxidation distant from its site of formation. One of the well-recognized roles of $CO_3^{-\bullet}$ is to modulate peroxynitrite activity by abstracting hydrogen from amino acids, primarily tyrosine. The sequential NO_2 addition will lead to nitro-substituted residues, which may be involved in many chronic diseases.[955]

5.3.2. Radical Generation in Solution and Applications in Biology. The carbonate radical anion can be produced in vitro in many ways including UV photolysis[956] of a cobalt complex, $[Co(NH_3)_4CO_3]ClO_4$, in phosphate buffer, enzymatic production from SOD1 in a bicarbonate buffer,[956] pulse radiolysis[957] of an N_2O saturated $NaHCO_3$ solution, and laser flash photolysis of persulfate anion in bicarbonate buffer.[958] To monitor the production of $CO_3^{-\bullet}$, EPR spin trapping is often engaged.[951] In 2001, Geacintov and co-workers[958] generated $CO_3^{-\bullet}$ as a secondary radical from the sulfate radical anion and, for the first time, characterized its site-selective oxidation of guanine in double-stranded oligonucleotides. Subsequently, many groups focused on the consequences of protein oxidation in biology. One of the foundation studies was done in 1973,[959] where the rate constants of the carbonate-radical reactions toward some biochemical compounds, including many free amino acids, were measured. Tyrosine, tryptophan, and methionine are the most reactive, with rate constants on the order of 10^8 M^{-1} s^{-1}. Gebicki and co-workers[957] showed that horseradish peroxidase (HRP) and other proteins that are rich in Tyr, Trp, and Met residues experience a 20−30% loss in their activity after being exposed to 100 μM $CO_3^{-\bullet}$. Paviani et al.[956] characterized a ditryptophan cross-linked product of lysozyme under carbonate radical anion mediated reactions; the study delineated the origin of a previously detected ditryptophan in the nonamyloid aggregation of human SOD,[960] a process that may be pathogenic for amyotrophic lateral sclerosis.[961]

5.3.3. $CO_3^{-\bullet}$-Based Protein Footprinting: Residue Specificity and Proposed Reaction Pathways. The carbonate radical anion as a protein footprinter was first evaluated by Zhang et al.[962] and demonstrated on a FPOP platform. The hydroxyl radical generated from hydrogen peroxide photodissociation reacts with carbonate/bicarbonate (CO_3^{2-}/HCO_3^-) buffer to form $CO_3^{-\bullet}$. To ensure its dominant presence in solution, highly concentrated CO_3^{2-}/HCO_3^- (e.g., 700 mM) is necessary, supported by a numerical simulation based on second-order kinetics. Experimentally, Zhang et al.[962] showed no modification in a no-laser control and a similar oxidation pattern as that produced by •OH FPOP upon laser irradiation (Figure 16a). The last thorough study of $CO_3^{-\bullet}$ prior to this work took free amino acids as substrates,[959] and that may not be relevant for proteins in solution particularly because the N-terminal amines of the free amino acids may complicate the measurements of side-chain reactivity. Therefore, a systematic study of model peptides and proteins was conducted, taking into consideration the residue context, solvent accessibility, and local environment. From the results, the reactivity of the carbonate radical anion can be assigned as follows:

$$\text{Met} \approx \text{Trp} > \text{Tyr} > \text{His} \approx \text{Phe}$$

The outcome is similar to that reported previously.[963]

Given that $CO_3^{-\bullet}$ (E° = 1.58 V at pH = 7)[964] is a weaker oxidant than •OH, the former showed more specificity toward residues with electron-rich side chains (i.e., Met, Trp, and Tyr), which underwent even more oxidation than with •OH (Figure 16b). Phe and His are also reactive but only modestly. Other aliphatic amino acids are basically inert. The pH also has a considerable effect on the reactivity of a specific residue (e.g., His where the imidazole ring carries more charges at pH ∼ 6 than under neutral or basic conditions). The electrostatic interaction between the positive side chain and the negative charged $CO_3^{-\bullet}$ may contribute to a higher local radical concentration that promotes the oxidative chemistry. The oxidation pathway is proposed in the *Summary and Perspective* (section 5.7, Figure 22), where the $CO_3^{-\bullet}$ forms a protein-centered radical that undergoes addition of hydroperoxyl radical. Water hydrolysis finally leads to an oxidized product.

The carbonate radical anion can be a good candidate to characterize structural dynamics and binding interfaces of protein complexes, especially when the involved residues are reactive with it. The more specific modifications not only

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

enable higher oxidation levels to reveal subtle differences but also allow faster data analysis and higher throughput. A limitation of its use is also obvious; compared to $^{\bullet}OH$, $CO_3^{-\bullet}$ has a narrower range of reactivity, and the larger size of $CO_3^{-\bullet}$ compared to $^{\bullet}OH$ may afford lower spatial resolution footprints. In addition, the pH of the relatively high concentration of the CO_3^{2-}/HCO_3^- buffer used in the FPOP experiments is basic, possibly not friendly for maintaining a protein's native state. Furthermore, its generation is not straightforward, diminishing its convenience as a footprinter. It may be that the other production methods can relieve the concerns. Nevertheless, the outcomes show the value of the FPOP platform in fundamental studies of radicals and radical ions that are relevant in biology.

5.4. Carbenes

5.4.1. Biological Relevance and Carbene Generation in Solution. Diazirines, common precursor of carbene diradicals, were synthesized in 1960[965] and emerged as a versatile PAL agents in the late 1970s.[966] PAL reagents usually are comprised of a binding motif and a reactive motif. Practically, the binding motif of PAL reagent reversibly binds to the active site of the target protein. Upon photoactivation, the reactive motif is activated to react with an adjacent site, after which the PAL reagent is covalently attached to the target protein, serving as a footprint.[967−969] Investigators typically activate diazirines with a UV laser at approximately 350 nm to cause release of N_2 molecules and give an equimolar amount of carbene diradicals. Highly reactive carbenes form irreversible covalent bonds with proteins, allowing stringent downstream affinity purification and target identification.

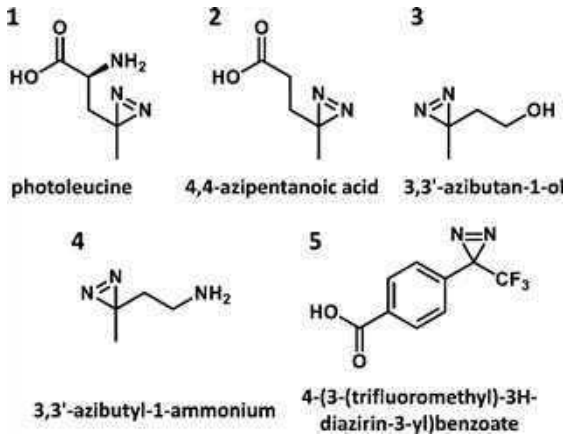
The use of diazirine-based PAL is an effective strategy to understand protein−drug or other small molecule interactions and identify new drug binding sites. To address specific questions, investigators have designed and synthesized several diazirine analogues to adapt the affinity agent to the protein sample environment. For example, adamantylidene, a lipophilic reagent and an analogue of adamantane, was used to obtain topological information of Na/K ATPase[970] and Ca-ATPase[971] in membranes as early as 1983. In addition, H-diaziflurane, an analogue of halothane, allowed the examination of binding sites of inhaled anesthetics and their action mechanism.[972]

More recently, reagents containing diazirines were adopted as a new class in photoactivatable cross-linkers, and several were described.[163,973] The ability to react with many bond types or amino acid residues makes carbenes powerful reagents to capture protein dynamics and intermolecular interactions. Obtaining multiple cross-links magnifies the information by providing more distance restraints on an interacting protein system and thereby furnishing vital data for molecular simulation or docking to give a more complete description of the system.

5.4.2. Residue Specificity and Proposed Reaction Pathways. Richards et al.[974] first described the carbene diradical as a footprinting reagent in 2000. Methylene was generated from diazirine gas (CH_2N_2) upon UV irradiation and allowed to footprint α-lactalbumin. The labeling yield, however, was low, owing to the limited solubility of gaseous CH_2N_2 in aqueous media. Furthermore, the explosive gas requires conscientious preparation, storage, and safe handling, limiting wide application.

Later in 2011, Schriemer and co-workers[975] reported a new diazirine-based reagent, photoleucine (reagent **1** in Scheme 16), that has higher water solubility and stability than CH_2N_2. The reaction platform incorporates a Nd:YAG pulsed laser

Scheme 16. Diazirine-Based Footprinting Reagents



(355 nm, 1000 Hz) for radical initiation and a 96-well plate cover with lids containing slits, allowing the laser beam to enter the protein solution. With this experimental setup, the investigators obtained a maximum conversion of the diazirine to the carbene diradical by using an irradiation time of 2 min in the absence of other competitive chromophores. Furthermore, they found that photoleucine does not react with targeted proteins without laser activation. Upon photolysis, irreversibly labeled products exhibit a characteristic +115.03 Da mass shift (Figure 17a,b).

$$\Delta\text{mass}_{\text{av}} = \left[\frac{\sum m_i I_i}{\sum I_i}\right]_{\text{labeled}} - M_{\text{unlabeled}} \tag{22}$$

To examine the labeling sensitivity of carbene diradicals for different accessible surfaces of proteins, two calmodulin systems (i.e., with/without Ca^+ and bound/unbound to the peptide M13) were investigated. The average number of labels on each protein molecule can be estimated by determining the difference in the centroid masses between the labeled and unlabeled proteins as seen in a deconvolved mass spectrum of the intact protein (eq 22 and Figure 17c,d), where m_i and I_i represent mass and signal intensity, respectively. A reduction in protein surface area resulting from Ca^+ binding or M13 binding is reflected clearly as the measured labeling extent over 2−10 min of irradiation, namely holo-CaM is 45 ± 7% less labeled and M13-CaM is 39 ± 5% less than the unbound forms.

Residue-level quantification of carbene-induced modification needed attention, and that was discussed in a sequel study.[976] Theoretically, carbene diradicals can insert into X−H bonds and C═C bonds.[978] Insertion into carboxylic acid functional groups will form labile esters that can be lost in CID fragmentation, thus complicating data interpretation and even losing information. The Shriemer group[976] chose ETD fragmentation as an alternative to CID and compared the two fragmentation methods by reporting the fraction modified for each y/z ion. Although most fragments shared the same trend, the y_{10} ion gave poor precision because it undergoes a neutral loss of the modifying group (Figure 18a). Notably, reducing the collision energy reduces the loss, but the abundance of the peptide fragments (product ion in MS/MS) is also reduced. ETD is a superior fragmentation method

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

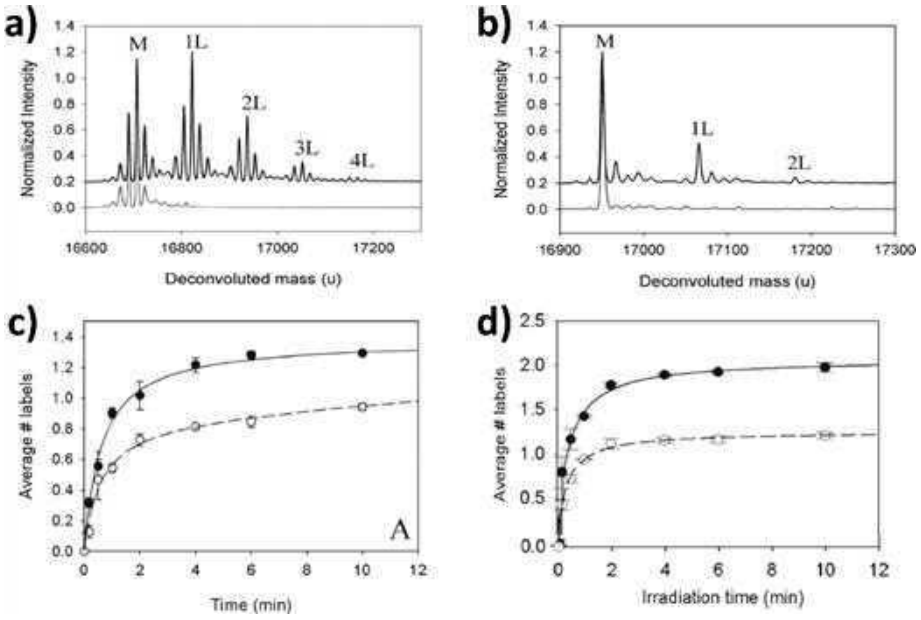


Figure 17. (a) Deconvoluted mass spectra of myoglobin after footprinting with carbenes. (b) Deconvoluted mass spectra of labeled and unlabeled holo-CaM. (c) Ca^{2+}-binding induces conformational change on calmodulin where apo-calmodulin (closed circles) and holo-calmodulin (open circles) were labeled with 100 mM photoleucine in phosphate buffer and monitored as a function of time. (d) Free holo-calmodulin (filled circles) is referenced to M13-bound holo-calmodulin (open circles). Reproduced with permission from ref 975. Copyright 2011 American Chemical Society.

for retaining labile modifications to afford more comprehensive information than CID. ETD, however, performs poorly when peptides are low in charge and small in size. The best approach might be to use a combination of the two modes of MS/MS.

Another issue is the electrostatic interaction between the carbene precursor and various amino acid side chains; the interaction may concentrate the precursor molecule around the site, promoting more modification and a biased residue preference. Switching to another carbene precursor reagent, 4,4-azipentanoic acid (**2** in Scheme 16), that contains no positively charged amine group, shows the effect. Although both reagents give similar results for Tyr, Lys, Glu, and Asp, positive-charged photoleucine shows higher reactivity with the negatively charged glutamic acid (Figure 18b). In addition, the ionic strength of the buffer and solution temperature also affect the electrostatic interaction, where increasing ionic strength decreases electrostatic interactions[979] and higher temperatures will lower the dielectric constant.[980]

A related approach was also reported by the Gross group[981] in 2015, when they adapted carbene generation in solution on the FPOP flow system. Careful control over the exclusion volume guarantees that photoleucine and the protein CaM are mainly irradiated once, thus diminishing concerns of perturbing the solution equilibrium by the generated nitrogen gas and causing conformational change with excess labeling. The outcome is less modification for holo-CaM and significant labeling on Try, Asp, and Glu, possibly owing to interactions of the protein with photoleucine, concentrating the reagent on the protein surface.

In 2017, Schriemer and co-workers[977] refined the carbene platform to employ a single-shot laser with higher energy (i.e., 150 mJ) to avoid nitrogen perturbation and protein conformational change at the induced air−water interface. Prior to irradiation, the sample solution was snap-frozen in liquid nitrogen before laser irradiation to restrict radical diffusion, to maintain protein HOS, and to minimize quenching of carbene radical and increase modification. To establish residue specificity with carbene chemistry, the investigators, in a Herculean study, footprinted 777 peptides by using three different precursors, a negatively charged precursor **2** (Scheme 16), a neutral reagent **3** (3,3′-azibutan-1-ol, Scheme 16), and a positively charged reagent **4** (3,3′-azibutyl-1-ammonium, Scheme 16). The labeling trends are similar for the three reagents, and the bond insertion propensities generally are a function of side chain polarity and size. Reagents **2** and **3** both favor Arg, Glu, and Asp, whereas the neutral reagent **4** shows higher reactivity with His (aromatic and neutral) in addition to the three residues, Arg, Glu, and Asp. Remarkably, many hydrophobic amino acids show noticeable modification (Figure 18c), supporting the high reactivity of carbenes, even with aliphatic groups. The nature of the carbene precursor also impacts the residue selectivity by participating in complex molecular interactions to increase the local concentration of the precursor ion at the protein surface.

Manzi et al.[982] tested a carbene from the precursor, 4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)benzoate (reagent **5**, Scheme 16) that is more reactive than those from reagents **1**−**4**. The investigators' design involved installing an adjoining trifluoromethyl group (reagent **5**, Scheme 16), leading to the use of less reagent and less irradiation time (i.e., 10 mM for 4 s irradiation). Further, the labeling efficiency improved in comparison to that with photoleucine at 100 mM and 16 s irradiation (Figure 19a). The improved reactivity is due to increased stabilization of the carbene radical by the added trifluoromethyl group and the increased hydrophobicity but little change in zwitterionic character. They tested the new reagent on an unknown protein complex (i.e., the deubiquitinating enzyme ubiquitin specific protease 5 (USP5) upon binding with diubiquitin (di-Ub)). USP5 is a multidomain cysteine protease including two ubiquitin-associated domains (UBA) and a Zn-finger ubiquitin-binding domain (ZnF-UBP). Previous studies showed the binding stoichiometry between

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

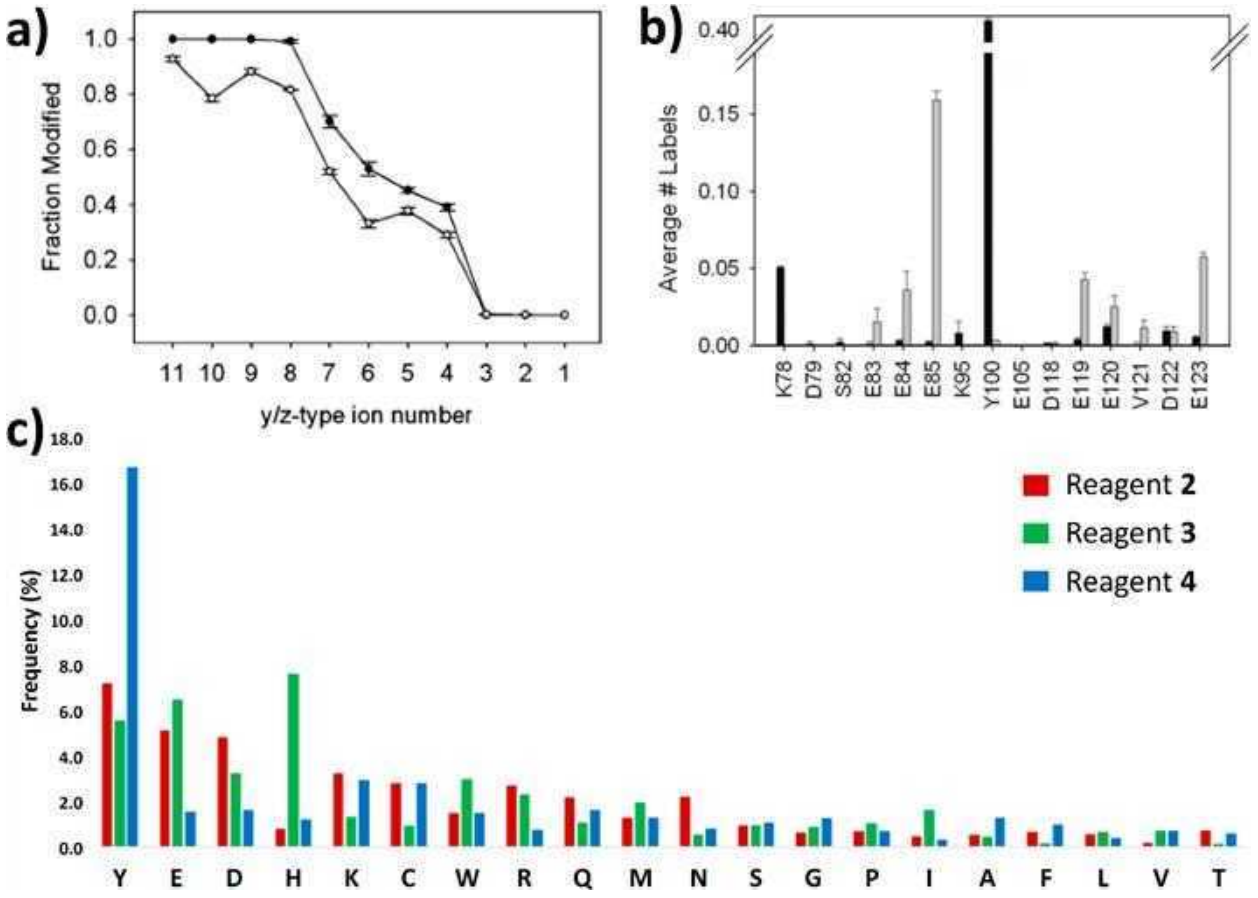


Figure 18. (a) Fractional distributions of carbene label derived from reagent **1** as determined by ETD (*z*-ions, filled circles) and HCD (*y*-ions, open circles) for peptide LTDEEVDEMIR (117−127) in CaM. (b) Residue-level reagent incorporation for select residues, based on ETD fragmentation of MKDTDSEEEIR (77−87), VFDKDGNGYISAAELR (92−107), and LTDEEVDEMIR (117−127) with reagent **1** (gray bars) and reagent **2** (black bars). Error bars are ±1 standard deviation. Reproduced with permission from ref 976. Copyright 2012 American Chemical Society. (c) Average frequency of carbene insertion at each residue generated from the photolysis of reagents **2**, **3**, and **4** in the presence of protein digests (777 peptides). Site of the label insertion was located with MS/MS with a Fusion Lumos with EThcD fragmentation and analyzed with Mass Spec Studio software. Reproduced with permission from ref 977. Copyright 2017 Springer Nature.

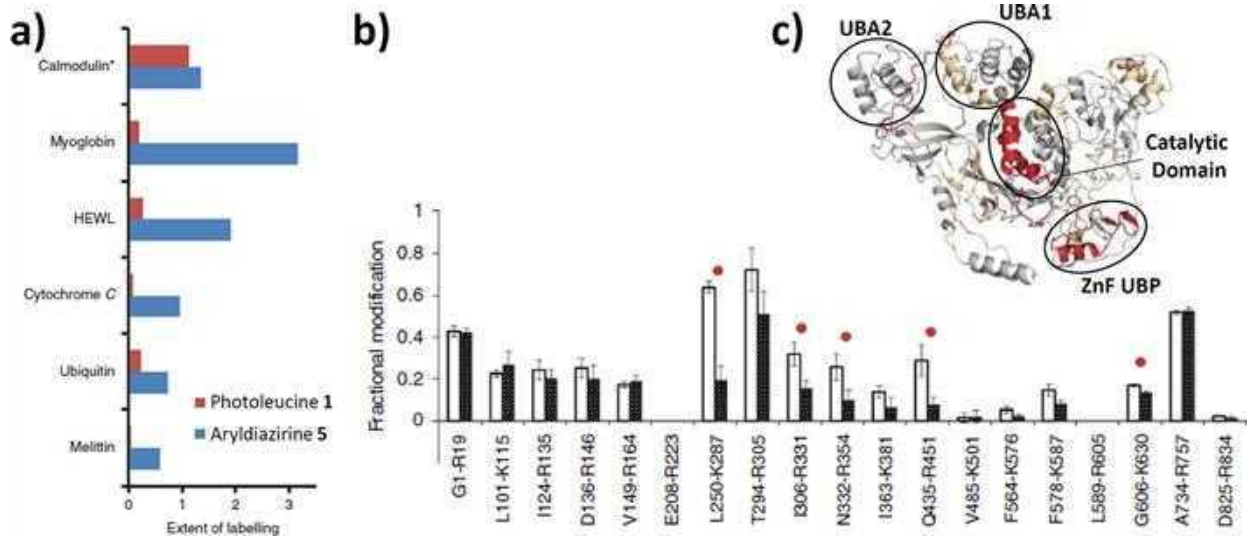


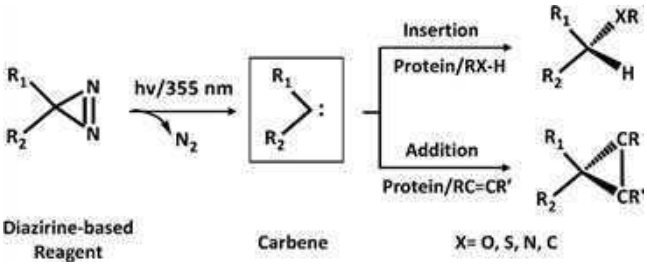
Figure 19. (a) Extent of labeling of a range of proteins with reagent **1** (100 mM, 16 s irradiation), and reagent **5** (10 mM, 4 s irradiation (*1 s in the case of CaM)). (b) Fractional modification by reagent **5** of USP5 peptides in the presence (black bars) and absence (white bars) of diubiquitin. Error bars are ± standard deviations (n = 3) and significant differences (Student's t test, $p < 0.05$) are highlighted with a red dot. (c) Model of USP5 (based on PDB 3IHP) showing the locations of the five peptides (red) that are masked from labeling by diubiquitin binding and their relative locations to the ZnFUBP and catalytic domains. Reproduced with permission from ref 982. Copyright 2016 Springer Nature.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

USP5 and di-Ub to be 1:1, suggesting additional binding sites than those in Znf-UBP. These sites were not identified. The investigators footprinted the C335A mutant of USP5 with carbenes in the presence and absence of one equivalent di-Ub and observed distinct binding regions (Figure 19b), which were mapped onto the X-ray structure of USP5 (Figure 19c). The catalytic domain was shown to be the other binding site and, additionally, a remote conformational change for the region represented by peptide G606-K630 was found. The design of new, successful radical precursors indicates that there are more opportunities for improvement and application, emphasizing the potential of carbene footprinting as an effective and accurate structural probe for HOS of proteins.

The pathways for carbene chemistry may involve several radical intermediates whose structures and reactivities can be tailored by using different precursors. For example, it is possible to generate by photolysis not only singlet[983] and triplet carbenes[984] but also diazo isomers that further decompose into carbocations.[985,986] Although a singlet carbene preferentially inserts into O−H, N−H, and S−H bonds,[985] it is challenging to pinpoint the dominant pathway just from the nature of the modified products. Insertion into Thr can be done by singlet carbene through O−H bond insertion or by triplet insertion into the methylene group located on the amino acid side chain. In addition, the small energy difference (i.e., ∼2 kcal/mol[987]) between the two states add complications because mixtures of products can form. Because heteroatom-containing residues are usually the favored sites of reaction (e.g., Glu, Asp, Tyr, and Arg), the singlet state may be favored; however, blended pathways are more likely. A general route is shown in Scheme 17.

Scheme 17. Proposed Carbene Reaction Pathway



Carbene footprinting has high potential in structural biology. The labeling time is shorter than 10 ns,[988] faster than most protein folding. Irradiation on a flow-system or excitation after snap freezing eliminates deceptive modifications originating from carbene insertions in a protein that has undergone a protein conformational change. Compared to the hydroxyl radical, the short survival time of carbenes owing to reaction with solvent water obviates the need for a scavenger. Carbene generation by diazirines is at a less damaging wavelength to proteins (i.e., ∼350 nm). In addition, most carbene precursors do not react with proteins prior to laser irradiation (unlike H_2O_2 for $^{\bullet}OH$), and this lack of reactivity minimizes background interference. Once a carbene inserts, there are no reactive biproducts as there are when other radical reacts. Furthermore, the resultant mass shift for carbene modification can be adjusted to be bio-orthogonal by tailoring the precursor design. The physical properties of reagents, however, can favor preconcentration on the surface of the protein, possibly delivering biased residue preference or even information loss. Those properties can also be chosen advantageously to promote binding in lipid membranes, permitting footprinting of transmembrane proteins. As this field develops and more diverse carbene reagents are implemented, a better understanding of interactions will emerge to permit rational design of new carbenes. New footprinting reagents that can target specific residues or provide comprehensive coverage are expected in the future.

5.5. Trifluoromethyl Radical

5.5.1. Biological Relevance and Applications. Fluorine-containing compounds are extremely rare in biology; only five entities containing F have ever been identified.[989] The most common molecule is fluoroacetate, which is found in many tropical plants as a toxin. Footprinting reactions that insert either fluorine or fluorine-containing substrates may be advantageous because fluorine is the most electronegative[990] (Pauling electronegativity = 4.0) common substance and has a small radius (1.33 Å), not so dissimilar to that of H,[991] allowing F to be a surrogate for H.

Introduction of fluorine alters existing compounds to be more metabolically and thermally stable, more lipophilic, and possibly more interactive with a targeted protein[992] through contact with hydrophobic patches. Therefore, anthropomorphic fluorinated compounds already play a role in pharmaceuticals, up to 20−25%.[989] One indispensable integrant is the trifluoromethyl group (CF_3) found in several drugs (e.g., Celebrex (Pfizer) and Sarafem (Eli Lilly)).

The CF_3 group was utilized early in several applications in structural proteomics. In 1962, Singer[967] developed affinity labeling, in which a targeted protein noncovalently interacts with a probe molecule (e.g., a drug). The probe is comprised of the drug plus a reactive motif installed by chemical synthesis. The probe is activated so that it covalently binds with the protein in a similar way as the drug to "mark" the binding site.

The Brunner group[993] advanced the approach by introducing a CF_3 group in a diazirine-containing reagent, 3-trifluoromethyl-3-phenyldiazirines (TPDs). TPDs are chemically stable and exhibit high quantum yields for carbene formation upon UV irradiation. Even though the linear diazo rearrangement is inevitably formed upon UV irradiation, CF_3 and phenyl stabilize the linear diazo and prevent it from undergoing intramolecular rearrangements (e.g., to an arylalkyl diazo derivative that competitively diffuses away from the active site before decomposing to the carbene, leading to a complex labeling pattern) and diminished side-chain reactions.[978]

5.5.2. Radical Generation in Solution. Trifluoromethyl radicals were generated in late 1940s from the reaction of homogenized CF_3I[994] in the presence of ethene. Because gaseous CF_3I is not easy to control and store, other radical precursors including $Te(CF_3)_2$[995] and trifluoroacetic anhydride[996] were later developed. High temperature and low radical yield limit their application, however. Alternatively, Togni's[997] and Umemoto's[998] reagents are better choices, and they require activation by various transition metals (e.g., $Ru(bipyridine)_3Cl_2$[999,1000] and CuI[1001]). The increased efficiency of generation and the reactivity of the radical depend largely on the metal-containing compounds that are used.

Whereas, organic solvents are usually necessary to retain the activity of the catalysts. A more biological compatible $^{\bullet}CF_3$

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

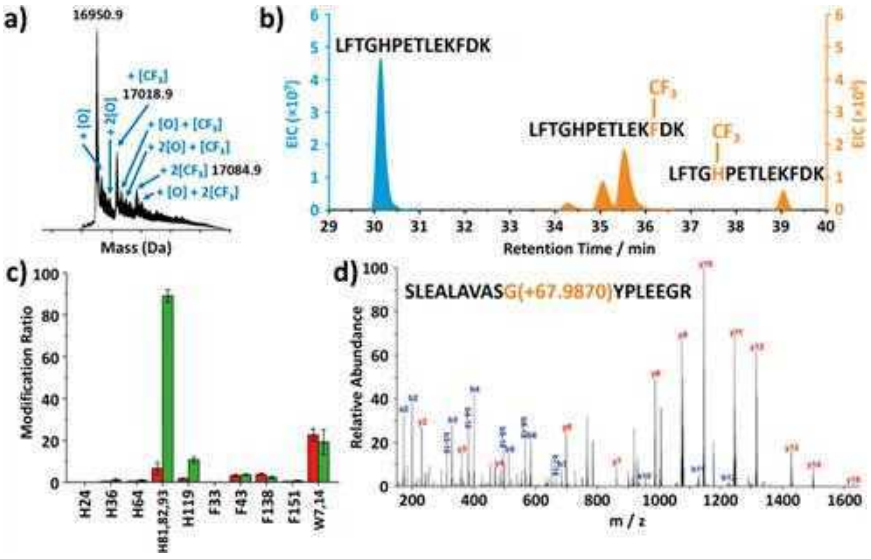


Figure 20. (a) Deconvoluted mass spectrum of CF_3-modified aMb. (b) EIC of unmodified and CF_3-modified peptide 32−45 of aMb. (c) Comparison of the extent of modification at the residue level for hMb and aMb. (d) A key mass shift of +68 Da is consistent with an unusual Gly labeling, here CF_3 modification on Gly-276 of peptide 267−83, a glycine with large SASA. Reproduced with permission from ref 1006. Copyright 2017 John Wiley and Sons.

precursor is the Langlois reagent, first reported in 1991[1002] and promoted by Baran and co-workers.[1003,1004] The system consists of sodium trifluoromethanesulfinate (CF_3SO_2Na) and *t*-butylhydroperoxide (TBHP) as an oxidant. In the absence of transition metals, $^{\bullet}CF_3$ can be initiated from the oxidant and generated in an aqueous buffer. Trifluoromethylation on various heteroarenes including pyridines, uracils and xanthines occurs with moderate to good yields (33%−96%). In addition, a protein system, *β*-lactamase, could be successfully labeled in Tris buffer while maintaining the protein's functionality. More recently, Fennewald et al.[1005] incorporated an addition catalyst, TPGS-750-M, to achieve a better yield under milder reaction conditions.

5.5.3. $^{\bullet}CF_3$-Based Protein Footprinting: Residue Specificity and Proposed Reaction Pathways. Trifluoromethyl radical as a footprinting reagent was implemented by Cheng et al.[1006] in 2017 to be used on the FPOP platform. The radical precursor is the water-soluble salt, $NaSO_2CF_3$ (Langlois reagent). The $^{\bullet}CF_3$ formation is initiated by the $^{\bullet}OH$, from photolysis of hydrogen peroxide. A likely mechanism is the $^{\bullet}OH$ displaces $^{\bullet}CF_3$ by attack on the S═O bond to form $HOSO_2^-$, the conjugate base of sulfurous acid, a mechanism different than that suggested in the original paper. The subsequent trifluoromethylation of proteins likely occurs by capping a protein radical (generated through $H^{\bullet}$ abstraction by either $^{\bullet}OH$ or $^{\bullet}CF_3$ (Figure 22).

The investigators tested the ability of the $^{\bullet}CF_3$ to footprint neuropeptide Y (18−36), apo-/holo-myoglobin (aMb/hMb), and VKOR, a transmembrane protein, on the FPOP platform. After laser irradiation, multiple CF_3 adducts were formed for the intact protein as evidenced by mass shifts of +67.987 Da for each CF_3 addition (Figure 20a). Notably, the trifluoromethylation of VKOR occurs selectively on the solvent-exposed region rather than in the transmembrane regions. Although $^{\bullet}CF_3$ is hydrophobic, it is generated from a water-soluble precursor that is not membrane permeable. Given the hydrophobicity of $^{\bullet}CF_3$, its generation in the "unfriendly" aqueous environment explains, in part, its high reactivity. The addition of a CF_3 group to the protein increases the hydrophobicity of the peptides formed by digestion, shifting their reversed-phase retention times to longer values, a phenomenon that is different for $^{\bullet}OH$ footprinting, where the peptides have higher hydrophilicity and usually elute earlier than their corresponding unmodified peptides (Figure 20b).

The $^{\bullet}CF_3$ can react with 18 out of 20 different amino acids residues, except Met and Cys, showing its complementary nature with $^{\bullet}OH$, which reacts rapidly with Met and Cys. There is some evidence that residues containing aliphatic side chains, which are inert with $^{\bullet}OH$ (e.g., Gly and Ala), are reactive with $^{\bullet}CF_3$ (Figure 20d). The reactivities with various residues, however, are not identical. $^{\bullet}CF_3$ is highly electron deficient and, thus, preferentially reacts with aromatic side chains (e.g., Trp, Tyr, His, and Phe). Furthermore, $^{\bullet}CF_3$ successfully reports the structural difference in aMb/hMb, as expected for a good footprinter. As found in previous studies, region 80−96 undergoes significant conformational changes in the conversion between the two states, and that change is confirmed by $^{\bullet}CF_3$ footprinting on His 81/82/93 (Figure 20c). Similarly, His 119 also undergoes more modification in the apo-state, whereas other residues react to comparable extents in both the apo and holo states.

Trifluoromethyl radical footprinting appears to have interesting advantages. The broader and complementary residue coverage compared to $^{\bullet}OH$ makes it a useful candidate in the structural biology "tool-box". Combination of $^{\bullet}CF_3$ and $^{\bullet}OH$ in tandem allows more comprehensive characterization of the residues on a targeted protein than does either radical alone, therefore providing a better opportunity to capture subtle structural changes. In addition, $^{\bullet}CF_3$ can survive in the presence of typical radical quenching reagents[1003,1006] like dimethyl sulfoxide and retain its reactivity. For many membrane protein systems where a detergent is necessary for solubilization, $^{\bullet}CF_3$ has the potential to be a membrane footprinter provided a precursor can be redesigned that will partition from water to the membrane (detergent). One limitation of using the Langlois reagent as precursor is the need for initiation by $^{\bullet}OH$. Reagents that quench $^{\bullet}OH$ will also

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

lower the trifluoromethylation yield. We expect to see alternative fluorination reagents in the future.

5.6. Iodine Radical

5.6.1. Biological Relevance. In physiology, iodine plays essential roles in metabolic regulation of thyroid function, especially hormone production.[1007,1008] Nearly all ingested iodine is carried and circulated as iodide,[1009] which in the presence of ROS or thyroid peroxidase (TPO)[1010] can be oxidized into molecular iodine, which oxidizes tyrosine residues on thyroglobulin to form thyroxine and triiodothyronine hormones.[1011] The reactive species causing iodination, however, are not settled, although a likely one is the iodine radical.[1012]

Iodination strategies have been widely adopted in radioimmunology.[1013] Iodide radical appears to react also with histidine residues[1014] under more alkaline conditions but giving a lower yield than with Tyr.

5.6.2. Radical Generation and Applications in Biology. Generation of iodine radical in a biorelevant condition is usually achieved by TPO[1015,1016] or UV photolysis[1017] of iodine-containing compounds. When iodine is directly incorporated into DNA and proteins, the photolabile C−I bond can produce useful radicals to address mechanistic and structural questions concerning the region surrounding the C−I bond. The generated iodine radical, on the other hand, seems to serve as only as a leaving group.

In 2005, Hiroshi and co-workers[1018] utilized 5-iodoracil as a substitute reagent for DNA synthesis. UV photolysis of the modified DNA causes a strand break to give products that serve to probe local structure.[1019] A subsequent study further elaborated the method to include interactions of DNA and proteins.[1020] A photochemical reaction forms cross-links that report on interactions between proteins and the targeted DNA both in vivo and in vitro.

Ly and Julian[1021,1022] implemented starting in 2008 a similar strategy for footprinting proteins. Starting with electrophilic iodination of a protein, the investigators isolated and photolyzed the intact protein at 266 nm in a linear ion trap mass spectrometer to generate odd-electron species that modify nearby Tyr and His resides and direct subsequent CID fragmentation. Radical-directed dissociation dominates the cleavages via characteristic pathways, primarily backbone fragmentations at modified Tyr residues. Secondary backbone cleavages were also observed in proximity of modified Tyr, especially if Pro is present. Similar chemistry occurs via His iodination in the absence of Tyr. The informative fragments can report on the presence of D-amino acids,[1023] elucidate protein tertiary structure, and assist with identification of proteins in proteomics experiments. Later on, Ranka et al.[1024] characterized the free iodo-tyrosine rearrangement pathways under UVPD to show that loss of iodide radical leads to a high-energy radical in the aromatic ring that engages in hydrogen/proton rearrangements. This study provides fundamental understanding of UVPD-based top-down analysis of iodinated proteins. More applications and detailed analysis are expected for these novel approaches.

5.6.3. I•-Based Protein Footprinting: Residue Specificity and Proposed Reaction Pathways. The iodine radical has potential to be an effective fooprinting reagent given its specific reactivity toward tyrosine and histidine. In footprinting, one precursor for the free radicals is 4-iodobenzoic acid, an organic iodide with some water solubility. The I• is formed presumably in concert with a $^{\bullet}C_6H_4COOH$ (represented by R• in Figure 22) by photolysis of I-C_6H_4-COOH at 248 nm. The carboxylphenyl radical likely abstracts an H• from the OH of Tyr or from the NH of the imidazole ring to give a stabilized radical that is subsequently "capped" by reaction with I• to give an iodinated protein (Figure 22), although there may be other mechanisms. This chemistry can be initiated on an FPOP platform, as shown by Chen et al.[186] in the presence of a histidine scavenger to control the I• lifetime and thereby reduce labeling-induced conformational changes.

After laser irradiation, the hMb and aMb were analyzed as intact proteins (Figure 21), where a single addition of I• leads

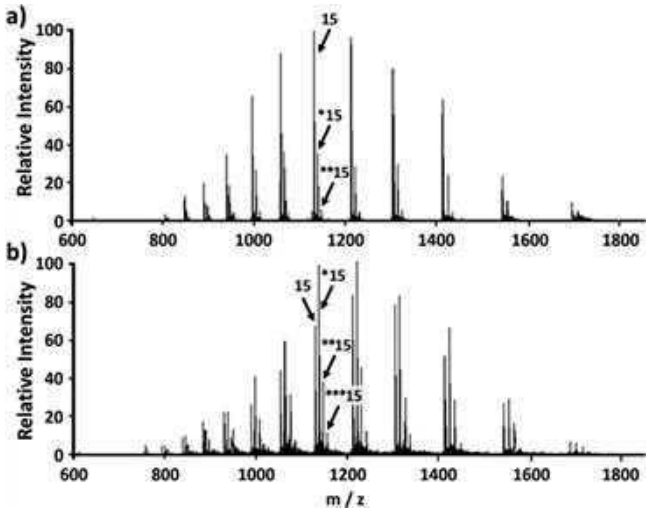


Figure 21. (a) Full ESI mass spectra of iodinated hMb. (b) Full ESI mass spectrum of iodinated aMb. The unmodified, mono-, di-, and tri-iodinated species of the 15th charge state are indicated by the number of stars. Reproduced with permission from ref 186. Copyright 2012 Springer Nature.

to 125.90 Da mass shift. Modifications occur to give more mono-, bi-, and tri-iodinated aMb than for hMb, consistent with a more exposed conformation for aMb resulting from heme removal. In contrast to •OH oxidation, the nearly exclusive iodination on Tyr and His simplifies the interpretation and data analysis, making possible a coupling with top-down MS for residue-level information. Top down ECD fragmentation on a 12 T FT-ICR gave a $*z_8^+$ ion that showed that iodination of Tyr146 was 26% for aMb but not detectable for hMb. On the contrary, iodination of Tyr103 (difference between $*z_{55}^+$ and $*z_{44}^+$) was similar for both proteins. The results can show that Tyr146 must be involved in heme-binding.

Owing to the limited sequencing efficiency of ECD technology at the time of the research, modified residues located in the middle region of a protein would be better analyzed with traditional bottom-up methods. His residues, although 30−100 times less reactive than Tyr,[1025] can also be addressed via a bottom-up strategy. Iodine radical footprinting may become a sensitive structural probe as is seen in an application to two other protein systems (i.e., apo-/holo-carbonic anhydrase II and EDTA-/zinc-insulin).[186] The capability of this approach to yield structural information is seen by the abundant labeling of the region that carries the

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

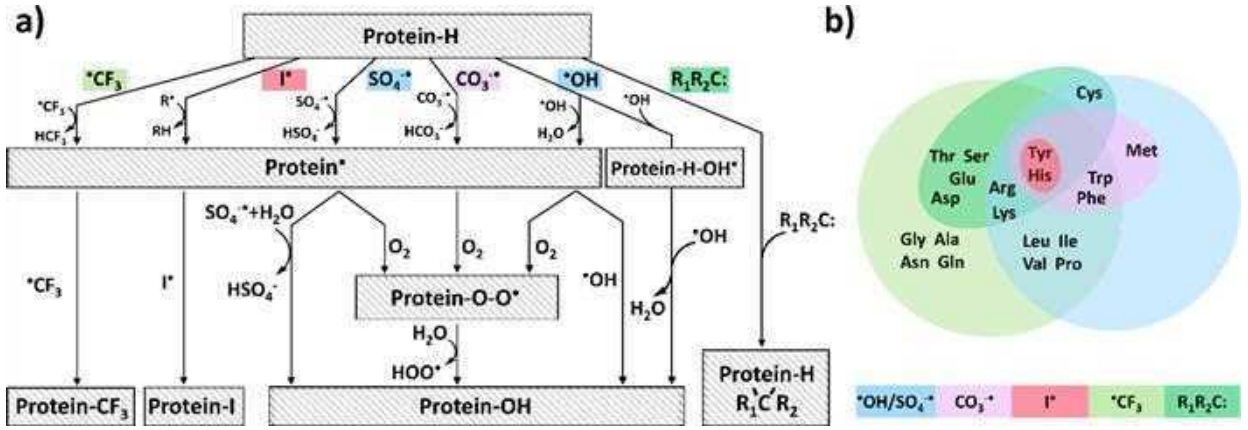


Figure 22. Summary of radical-based footprinting reagents of (a) proposed pathways and (b) residue specificity.

incipient f helix in the aMb and the remarkably low to nondetected footprinting of that region of the hMb.

Although the coverage afforded by the iodide radical is limited, footprinting two targeted residues can answer specific questions with easier data analysis. Compared to •OH footprinting, the unique and larger mass shift increases the confidence of assigning the modification sites. Moreover, the precursor of the iodine radical does not react with proteins, which minimizes the background modification and allows simplified postlabel sample handling. Unlike specific amino acid labeling, the modifications on the FPOP platform are fast, alleviating concerns about labeling-induced conformational changes.

5.7. Summary and Perspective

Radical footprinting is an effective means to acquire HOS information for proteins. Radical reactions are much faster (nano to milliseconds) than reactions of a conventional chemical reagent that modifies one or a few amino acid residues. Therefore, free radical footprinting will deliver fast, broad, and less biased information by largely avoiding questions of labeling-induced perturbation of protein structure. In addition, the short time scale can allow a sensitive report of subtle structural and even dynamical changes. An advantage of the labeling speed is the successful temperature jump, two-laser experiment on the FPOP platform, where •OH footprinting can successfully track the folding of barstar at times as short as a few tenths of millisecond.[898] An example of versatility is the tracking of multiple intermediate oligomeric states during amyloid beta aggregation.[1026]

Although different radicals react via distinct pathways, they share similar features (Figure 22a). Many radicals (e.g., $^{\bullet}CF_3$, $I^{\bullet}$, $SO_4^{\bullet-}$, $CO_3^{\bullet-}$, •OH) permit modification by forming a protein-centered radical that is subsequently quenched, whereas others (e.g, •OH and carbene diradicals) can directly add to a protein molecule. Whatever the process, the informative modifications are usually by an irreversible covalent bond to an amino acid side chain, facilitating downstream sample handling and analysis compared to that of reversible labeling. A protein footprinting-based "toolbox", that is being built for the future, will contain radicals that show different residue specificity to address specific biological questions (Figure 22b). For example, when protein systems are rich in Glu and Asp (e.g., as for calcium-binding proteins), GEE or carbene footprinting may be a good choice with quick and simple data analysis and high throughput. More importantly, investigators can customize reagent combinations for targeting specific residues or for obtaining high coverage. Other radical species (e.g., nitric oxide radical) will be tested, and new radical precursors will emerge soon to footprint both soluble and transmembrane proteins. A "toolbox" with diverse reagents and associated modeling software will aid investigations of complex biological questions with effectiveness, sensitivity, and accuracy.

6. APPLICATIONS THAT UTILIZE FAST LABELING APPROACHES

Since the initial introduction of protein footprinting by free radicals in 1999,[887] combinations of different reactive species and several experimental designs have successfully demonstrated the reliability of fast-labeling approaches and their ability to address various biological questions. In this section, we will discuss recent advances since the Chance review in 2007[175] and highlight key developments.

6.1. Mapping Epitopes

Monoclonal antibodies (mAb)[1027] are important biomacromolecules owing to their broad applications in analytical assays[1028−1030] and their profound potential in therapeutics.[1031,1032] Antibodies bind to antigens with high specificity and strong affinity, and this property has been utilized for some time in immuno-histochemistry[1028] and in the enzyme-linked immunosorbent assay (ELISA),[1029] assays that are now standard practice around the world. The highly specific recognition also makes it possible to develop targeted therapeutics.[1031,1032] These advances rely on a deep understanding of antibodies from both a structure and functionality perspective.

Interacting surfaces in antibodies and antigens are termed as paratopes and epitopes, respectively. Paratopes mainly consist of loop regions of 10−15 amino acid residues and are termed "complementarity determining regions" (CDRs). Paratopes of a typical mammalian antibody are found in six CDRs, three from the heavy chain and the remainder from the light chain.[1033] Although the combinations of amino acid residues that comprise the six CRDs are nearly unlimited,[1034] the structure and position of the paratopes are usually well-defined in accord with the rigid scaffold of the antibodies.

Antigens, however, differ dramatically from antibodies and from each other. Epitopes in antigens do not exist in a specific structure. The binding site for the antigen (epitope) depends on the location of the actual binding interface that interacts

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

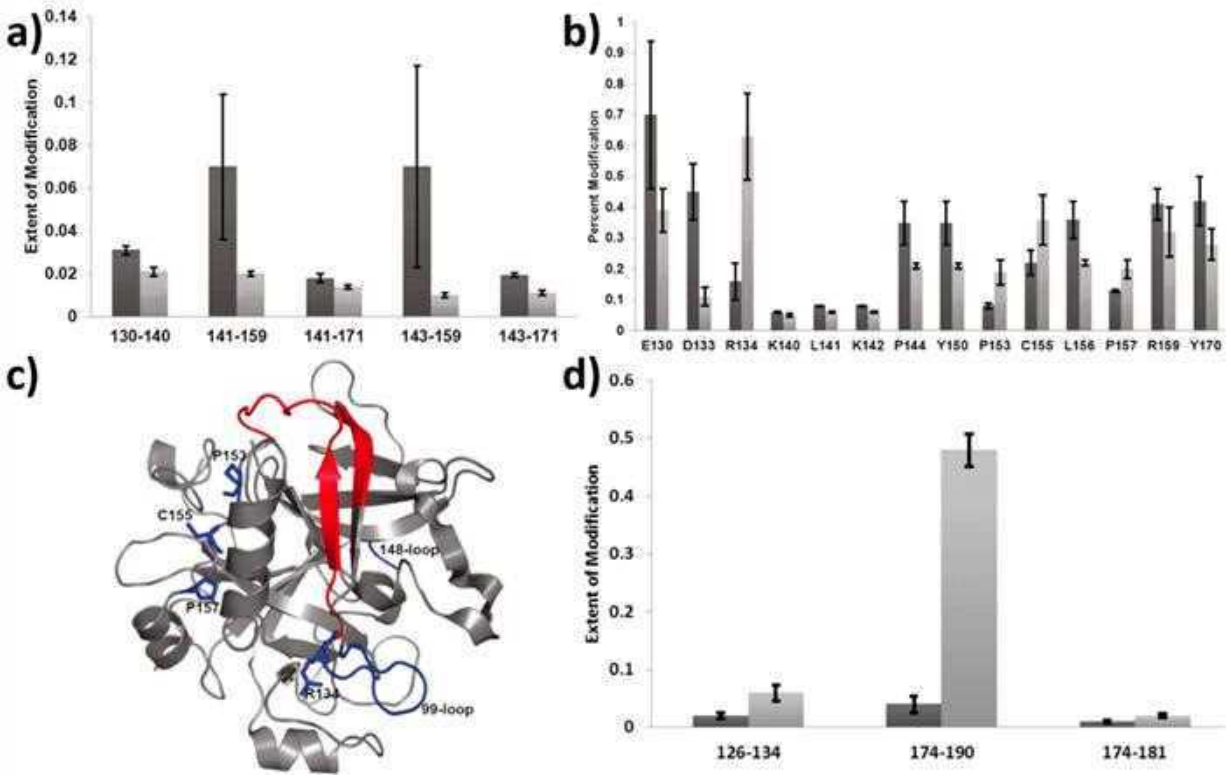


Figure 23. (a) Extent of modification of thrombin alone (darker bars) and antibody-bound thrombin (lighter bars) for the five peptides that span the 130−171 region. (b) Extent of modification of thrombin alone (darker bars) compared to antibody-bound thrombin (lighter bars) at the residue level. (c) Structural model of thrombin (PDB 2AFQ[1049]) with the proposed epitope colored in red and the loop regions colored in blue. The individual residues that show increased modification for antibody-bound thrombin are specified. (d) Extent of modification of peptides from thrombin alone (darker bars) and antibody-bound thrombin (lighter bars) show increased solvent accessibility in the antibody-bound form. Reproduced with permission from ref 1050. Copyright 2011 American Chemical Society.

with antibodies. In other words, epitopes can be located anywhere on the antigen surfaces.[1035] Moreover, an epitope can be either linear (consecutive) or conformational (assembled).[1036] A consecutive epitope is composed of a single continuous structural motif, whereas an assembled epitope consists of multiple and nonadjoining binding motifs. As the majority of reported epitopes are assembled,[1035,1036] a complex binding scheme usually pertains and requires responsive methods for mapping epitopes.

X-ray crystallography,[1037] NMR,[1038] and alanine-scanning mutagenesis[1039] are the most commonly used high-resolution biophysical approaches in mapping epitopes. X-ray crystallography and NMR reveal the binding regions, allowing assignment of epitopes as surfaces on the basis of proximity of the antibody and antigen. Such contact, however, may not always lead to structural recognition. Mutagenesis maps the epitope from a functional aspect, but that approach may not be informative for higher-order structure, and the amino acid substitutions may affect the protein HOS. Alanine scanning and shaving are the most adopted mutagenesis approaches, whereby amino acid residues in the protein are replaced one at a time (scanning) or in a group (shaving). The newly constructed proteins are then submitted for binding assays to judge the effect of the mutation, through which functional epitopes can be confidently located. Other functional assays including but not limited to studies of synthetic peptides representing the antigen[1040] or represented by antigen truncation[1041] were also developed as high throughput but low spatial resolution approaches as compared to the alanine scanning.

To complement these approaches, MS-based epitope mapping methods were developed substantially and now are regarded as reliable approaches.[1042−1044] Two different schemes were originally used; that is, epitope excision[1042] and epitope extraction.[1045] Both approaches utilize the specific binding between antigen and immobilized antibodies. In an epitope excision workflow, an immunocomplex is formed in solution, after which a protease is added to digest the antigens. The peptides from the antigen that contain the epitope remain bound with an immobilized antibody, and they are then extracted and characterized by MS. Epitope extraction, however, starts by digesting antigens and utilizes the resulting peptides in a peptide-immobilized antibody complex. In this case, those peptides that bind the antibody represent epitopes.

In both schemes, the investigator identifies epitope peptides by using MS and then sequences them by MS/MS. The amount of antigen required can be as low as subpmol.[1046] As usual, MS-based approaches significantly lower the sample amount and reduce the need for sophisticated sample preparation and data analysis as compared with X-ray crystallography and NMR, while preserving mid-to-high spatial resolution from a structural view. MS-based approaches locate potential epitopes with significant less effort, providing valuable guidance for designing mutagenesis and other biological functional assays to locate functional epitopes confidently.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

More recently, MS-based epitope mapping approaches utilize HDX to label the protein chemically and monitor changes in SASA that result when antigens bind with antibodies.[1047] Thanks to the high binding affinities for many immunocomplexes, off-rates do not heavily distort the HDX kinetics. Both the antigen and the immunocomplex can be footprinted under native-like conditions, minimizing potential perturbations of structure. Although the HDX rates are significantly slowed (quenched) at a pH of 2.5 and 0 °C,[189,190] back-exchange cannot be avoided. Proteases that function under acidic environments are generally limited to pepsin and fungal protease XIII,[1048] both of which are nonspecific, and the resulting peptide mixture is complicated and challenges the data analysis. Moreover, H/D scrambling during MS/MS analysis limits the spatial resolution to the peptide level; that resolution can be increased possibly to the residue level by using ECD[135] and ETD,[222] which presumably do not scramble the D labels,[221,222] or by digestion to smaller and overlapping peptides.[225]

Epitope mapping by fast labeling methods was first demonstrated by Jones and Gross[1050] in 2011, where the epitope of serine protease thrombin was footprinted with hydroxyl radicals on the FPOP platform. Two regions show binding-induced protection as evidenced by the decrease in extent of modifications (representative results are in Figure 23a), where five distinct peptides exhibit increased protection upon binding (darker bar represents unbound and lighter bar represents bound). The protection spans 42 amino acid residues, but such protection seems too extensive considering the size of thrombin. Taking advantage of the irreversible labeling of FPOP chemistry, the investigators could further assign protection at the residue level (results in Figure 23b). Among 14 resolvable residues, protection in the antibody-bound form occurs between residue D133 and Y150; these results are mapped in Figure 23c. Proposed epitopes agree well with those from an earlier HDX study.[1047]

Moreover, four residues (Figure 23b) and three peptides (Figure 23d) show deprotection upon forming the immunocomplex. This deprotection, not observed with HDX, is assigned as a remote conformational change. HDX monitors both backbone amide hydrogen bonding and solvent accessibility whereas FPOP emphasizes the latter. If allosteric or remote conformational changes do not involve changes in hydrogen bonding but occur as reorientation of amino acid residue side chains, they will not be seen by HDX.

Subsequent studies of similar design reveal the epitope of human epidermal growth factor receptor (EGFR)-Adnectin 1,[1051,1052] IL-6R-Adnectin 1 and 2,[743] and human immunodeficiency virus 1 gp120 envelope glycoprotein,[1053] demonstrating even more the efficacy of fast labeling approaches and making a case for generality.

As mentioned earlier, changes in SASA/hydrogen bonding may not be directly associated with biological recognition. Binding-induced protection can occur either from binding at the epitope region or from remote conformational change sites, which can be differentiated by site-specific mutagenesis and functional assays or possibly by cross-linking.

A 2017 study by Li et al.[1054] applied orthogonal methods, including FPOP, HDX, and alanine-shave mutagenesis and determined energetic epitopes of an antibody/interleukin-23 interaction. By examining the system with FPOP footprinting, they found that five regions of IL-23 p19 domain show protection (results mapped onto the crystal structure in Figure 24a). Similarly, HDX reveals four slightly different protected regions, three of which overlap with those reported by FPOP

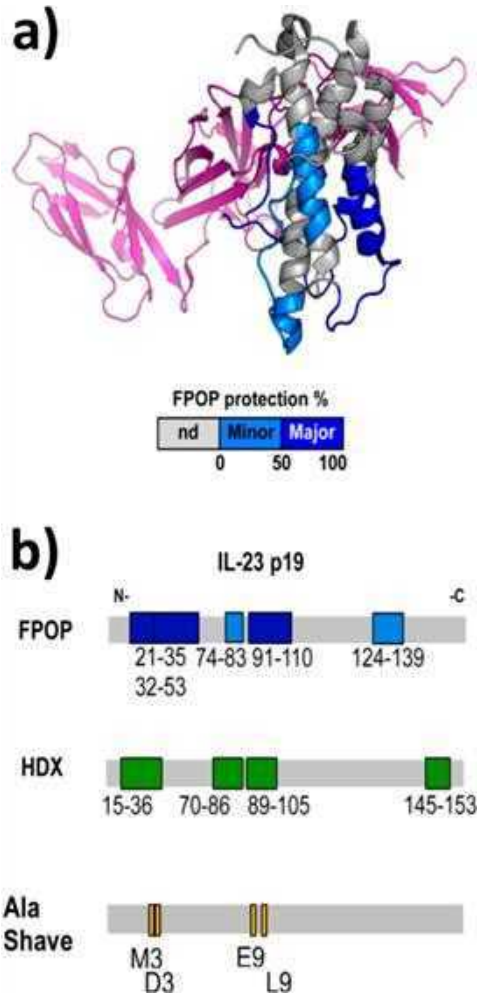


Figure 24. (a) Epitope regions determined by FPOP mapped on the crystal structure of IL-23 p19 domain. Color code: no significant difference (gray), minor epitope region (cyan), and major epitope region (blue). The p40 subunit is colored in purple. (b) Epitope regions determined by FPOP, HDX, and Ala shave energetics are mapped on the linear sequence of the IL-23 p19 domain. M3, D3, E9, and L9 stand for M35A, D36A, E93A, and L97A, respectively. Adapted with permission from ref 1054. Copyright 2011 American Chemical Society.

(Figure 24b). Imperfect overlap motivates subsequent alanine shave analysis, where a series of mutants can be designed based on MS-analysis, enabling confirmation of the residues/peptides identified by the two MS methods. Four distinct mutants show significant lower binding affinities as compared with the wild-type protein. Note that all four mutants retain near wild-type-like conformations as confirmed by optical and NMR approaches. Thus, this comprehensive approach points to the first two regions of IL-23 p19 domain as epitopes, whereas protection for regions 124−139 (by FPOP) and 145−153 (by HDX) are from remote binding-induced conformational change. This study is a good example of the workflow of fast radical labeling and HDX MS to guide mutagenesis and functional assay for epitope mapping.

In addition to the advantages of low sample amount, mid-to-high resolution spatial information, and high-throughput, fast-

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

labeling approaches can bring amino acid residue level spatial resolution to epitope mapping without the need of specialized MS instrumentation of ECD or ETD fragmentations in HDX or the effort needed for alanine-scanning mutagenesis. Irreversible labeling is not subject to the back-exchange of the label as applies to HDX, and the sample can be examined with any fragmentation method, usually without concern for scrambling of the label. Moreover, the labeled protein can be digested with a variety of enzymes at their preferred pHs, whose cleavages can be highly specific (e.g., trypsin), simplifying the data analysis. Fast labeling approaches are sensitive to subtle changes in the SASA of amino acid side chains, which may not be accessible to HDX.

FPOP, which generates $^{\bullet}$OH through laser photolysis of H_2O_2, controls the lifetime of the radicals to avoid perturbation of native protein conformation. The concentration of $^{\bullet}$OH needed for comprehensive footprinting is not large and is readily generated by the photolysis (~1 mM of radicals can be formed). In a FPOP-based epitope mapping approach, protection is evaluated through changes in the extent of modification, requiring that addition of the antibody to the antigen solution does not increase reactive sites so extensively as to quench the radicals and lower the modification extent of the antigen. This is particularly a problem when the antigen is small compared to the large antibody. This potential bias can be overcome by either introducing small molecule radical dosimeter[1055]/reporter peptide[1056] to calibrate the effective $^{\bullet}$OH concentration or by keeping the mass of total protein nearly the same in the control and the test experiments. An excellent way of doing this would be to include a nonbinding antibody when footprinting the antigen as control. To generate $^{\bullet}$OH through other approaches like synchrotron water radiolysis whose $^{\bullet}$OH precursor amount is not limited may also be a good alternative.[887]

6.2. Tracking Protein Folding/Unfolding

Dating back to 1961, the ribonuclease refolding experiment by Christian B. Anfinsen[2] showed that small globular proteins can fold, without assistance from other biomolecules, to their free energy minimum state. Soon after, Cyrus Levinthal[1057] argued that such protein folding would be too slow because the protein has too many possible conformations to "search" before finding its energy minimum, known today as the Levinthal Paradox. Following the ribonuclease refolding experiment, Anfinsen later proposed a thermodynamic hypothesis wherein the protein native structure is only determined by the protein's amino acid sequence.[1058] Decades of structure determination resulted in the accumulation of over 160 000 high resolution structures in the PDB. It is now accepted that protein folding is driven primarily by interactions including hydrogen bonding, van der Waals interactions, backbone dihedral angle preferences, hydrophobic interactions, and electrostatic interactions.[1059] Understanding protein folding pathways and their corresponding intermediates, however, remains a huge problem that now has become a field of research onto itself.

Modern views of protein folding take an energy landscape as a foundation for the folding problem. Rather than a defined folding pathway, multiple routes facilitate protein folding from an unstructured chain of amino acids to a native conformation; that is, a protein negotiates a folding funnel.[1060,1061] Modern structural prediction algorithms pioneered by Critical Assessment of Protein Structure Prediction (CASP) also contribute greatly to addressing this problem.[1062] All these developments were reviewed previously.[1063−1065]

Experimentally, folding or unfolding is followed, after introducing a perturbation, as a function of time, through characterization by several methods;[1066] perturbations that can be used are a sudden change typically in temperature,[898,899,1067,1068] denaturant concentration,[1069,1070] pH,[1071] or pressure.[1072] As secondary structural motifs including α-helices and β-sheets usually form in 0.1−10 μs,[1073] and some small proteins can fold as fast as tens of microseconds,[1074] it is important that the characterization/labeling method is fast enough to follow the process and capture any intermediates. For decades, several approaches, including circular dichroism,[1075] fluorescence,[66] FT-IR,[1076] and NMR,[1077,1078] have been most extensively used. Although lower in spatial resolution, the detection time limit for optical approaches can be as low as 10 fs,[1079] a time that is ideal for tracking rapid structural transitions. On the other hand, using rapid mixing and multidimensional NMR coupled with HDX has been successful in characterizing folding processes on the time scale of milliseconds with high spatial resolution.[1078] Recent developments of NMR instrumentation allow observation of folding intermediates on the time scale of 100 μs,[1072] greatly elevating the spatial resolution over that achieved by optical approaches. Modern single molecule measurements[1080] and mutational approaches[1081] also shed light on folding.

Following protein folding/unfolding by MS was first demonstrated by Chait and co-workers[1082] in 1990, who electrosprayed cytochrome *c* at different pHs. They observed different charge-state distributions when the protein was introduced in different conformations; an unfolded form favors higher charge states. This observation enables MS to follow, in a simple and global way, the protein conformational changes upon perturbation, a representative one being myoglobin reconstitution coupled with online continuous-flow labeling.[1083] In the same decade, HDX was married to MS,[208,1084] and the combination was used soon after to address lysozyme folding.[1085] Further development of a "rapid mixing" apparatus made possible the measurement of folding kinetics by HDX-MS[1077,1078,1086] and, in combination with protease digestion, allowed peptide-level spatial resolution.[70] There are also demonstrations that utilize methionine oxidation as a probe to follow protein folding/unfolding thermodynamics upon chemical denaturation (stability of proteins from rates of oxidation, SPROX).[1087,1088] In brief, a protein of interest is titrated with denaturant, during which H_2O_2 is also added into the system to facilitate methionine oxidation, which is a reporter of protein conformational change. As a result, methionine oxidation extent can be plotted as a function of denaturant concentration, and the transition midpoint from the titration curve is used to extract thermodynamic information during denaturant-induced protein unfolding. SPROX can work with complex biological mixtures and can be executed at the proteome level;[1089] it was successfully applied to identify protein−drug[646] and protein−ligand[1088] interactions, characterize the ATP interactome,[1090] and follow changes in protein expression in breast-cancer models[1091] and conformational changes in breast cancer-related proteins.[1092] All of these efforts empower MS in protein-folding analysis and were reviewed elsewhere.[267,780,1065]

The application of free-radical footprinting in folding studies was first demonstrated by Chance, Woodson, and co-

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

workers[883,884] for RNA, where a "stopped-flow" apparatus combined with synchrotron-generated •OH radicals probed ribozyme folding dynamics. Proteolysis to form peptides and MS to locate the footprinting provided there is some regional specificity for the folding.

This approach was later applied to protein unfolding. By varying urea concentration, Maleknia and Downard[1093] followed aMb denaturation through a series of •OH footprinting experiments. Upon protease digestion, they showed that helices A and B/C unfold in a cooperative fashion whereas helix G unfolds locally at lower denaturant concentration. Thermodynamic parameters were obtained, and they agree with those from fluorescence-based measurements. Although measured under equilibrium conditions, this study for the first time demonstrated the possibility for radical labeling in addressing protein folding/unfolding. Later studies by Poor et al.[1094] utilized this idea and demonstrated successfully the refolding of the paramyxovirus fusion protein by changing equilibrium temperatures and determining the free energy landscape of bacterial immunity protein (Im7) folding by combining mutational analysis with free radical labeling.[1095]

Although the approaches in equilibrium measurements are useful, it is challenging and likely more informative, to examine protein folding kinetically. Taking advantage of the Hambly and Gross[844] demonstration of •OH protein footprinting through laser photolysis of H_2O_2 in a flow system, Konermann and co-workers[1096] first integrated continuous-flow rapid mixing with •OH labeling to study the protein folding/unfolding in a kinetics experiment (Figure 25a). The rapid mixing apparatus has two mixing "tees". Using three syringes,

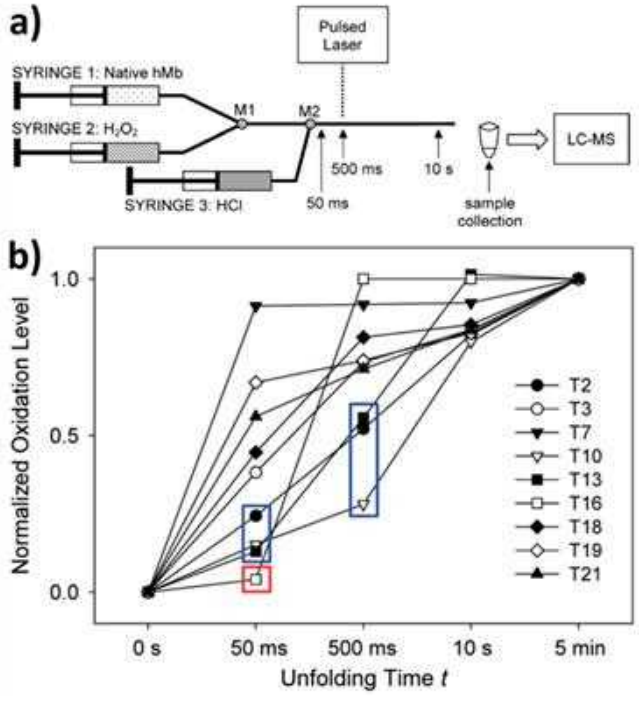


Figure 25. (a) Schematic illustration of the three-syringe, continuous-flow, rapid mixing setup for oxidative •OH labeling. (b) Normalized oxidation levels of tryptic peptides plotted as a function of unfolding time *t*. Blue boxes highlight three peptides that retain considerable protection at 50 and 500 ms, whereas the peptide highlighted by a red box is protected at 50 ms but not at 500 ms. Adapted with permission from ref 1096. Copyright 2009 American Chemical Society.

the investigators mixed native protein (hMb), H_2O_2, and denaturant HCl in a continuous fashion. By tuning the flow rate, pulsed laser frequency, the distance between the second mixing tee M2, and the window for laser irradiation, the delay time between protein denaturation and •OH labeling can be manipulated from 50 ms to 5 min. Using tryptic digestion, the investigators followed oxidative modifications as a function of denaturing/unfolding time of hMb with regional specificity. Peptide T7 represents a region that unfolds initially upon adding denaturant (Figure 25b). Peptides T2, T10, and T13, on the other hand, stay relatively protected up to 500 ms after mixing with HCl. Peptide T16 represents the region that unfolds between 50 to 500 ms. The hMb becomes completely unfolded after mixing with HCl for 5 min. On the other hand, by starting with the denatured protein and replacing HCl with a renaturing buffer, the investigators could study the folding of cytochrome *c*,[1097] the folding and dimerization of S100A11,[1098] the folding of an integral membrane protein bacteriorhodopsin,[1099] and even submillisecond folding of aMb.[1100]

Recently, the delay time was further shortened to the microsecond time scale by a newly designed microfluidic mixer, and the unfolding was followed by FPOP. The data reveal two kinetic phases of egg lysozyme that occur before 1 ms.[1101] The requirement for low sample amounts, the achievement of regional specificity, and the capability for kinetics open new possibilities for protein folding studies even on the submillisecond time frame.

In 2010, Chen, Rempel, and Gross[898] reported a design that combined FPOP with a laser-induced temperature-jump to probe submillisecond folding. As demonstrated in Figure 26a,[898] two lasers (KrF excimer laser at 248 nm for FPOP and Nd:YAG laser with a Raman shifter at 1900 nm for ∼20 °C temperature jump) were controlled by a signal generator and a delay circuit. The laser beams were aligned to intersect sequentially at the same transparent window on the flow capillary. By controlling the delay time between firing the two lasers, it is possible to track the protein folding/unfolding kinetics with high spatial and time resolution. The delay time of this two-laser approach is no longer limited by fluid dynamics, as in rapid mixing, but by the time separation between two lasers. By tuning the reaction conditions, •OH should be able to footprint the protein irreversibly in times as short as 1 μs[881,898] or less, making it possible to follow folding dynamics of fast folding proteins.

The investigators successfully footprinted the first intermediate state of barstar folding, which happens within 2 ms of the temperature jump.[899] Changes of the modification fractions at the amino acid residue level revealed the key residues that serve as a nucleus for barstar early folding as seen in Figure 26b,[899] where residues H17, L20, and L24 can be assigned as closely associated with the hydrophobic core around which barstar folds; residues I5 and F74 are weakly associated with the hydrophobic core; residues W53, L88 and many others, however, are not involved in the fast folding.

Radical species have the advantage of short reaction times, which are ideal for following fast processes. The •OH FPOP-based two-laser approach brings high resolution to submillisecond protein folding studies,[898,899] which was not easily accessible prior to this work. Secondary structural motifs for some fast-folding proteins, however, can form as fast as 0.1 μs,[1073] thus requiring an even faster probe. Carbene diradicals label proteins on the nanosecond time scale,[981] suggesting that

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

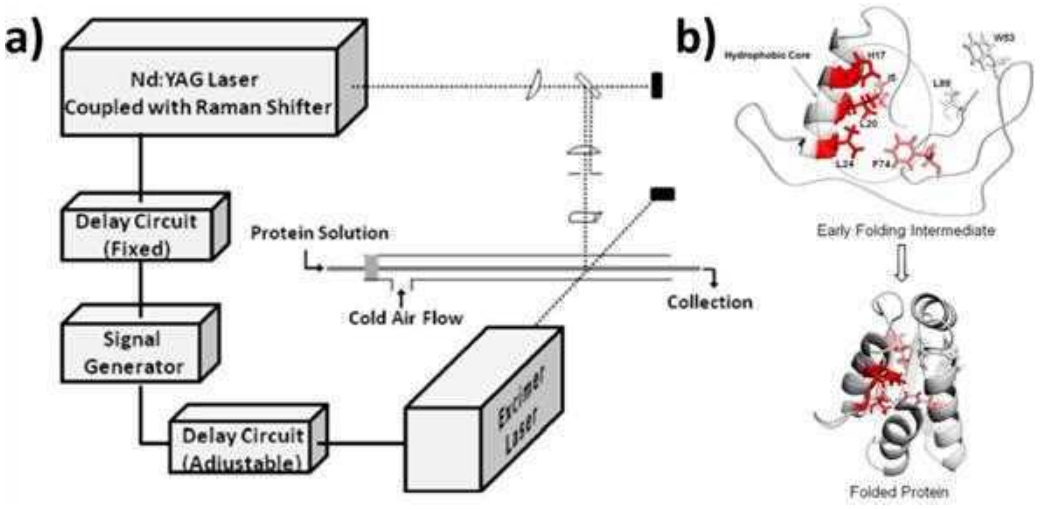


Figure 26. (a) Schematic illustration of the flow system intersected by two laser beams at a transparent window, as part of a temperature-jump apparatus. Adapted with permission from ref 898. Copyright 2009 American Chemical Society. (b) Proposed intermediate for barstar early folding. Residues colored in red, pink, and gray represent residues that are closely associated with hydrophobic core, weakly associated with hydrophobic core, and not involved in early folding intermediate, respectively. Adapted with permission from ref 899. Copyright 2009 American Chemical Society.

those footprinters may be appropriate to address such problems.

Despite these advantages, the two-laser apparatus is only applicable if the temperature jump is sufficient to cause the protein to fold or unfold. Folding is more restrictive, as few proteins are denatured at low temperature and become renatured upon heating. The change in temperature that the laser can deliver is also limited. From this perspective, rapid mixing, although relatively slow[1101] with certain exceptions, is more universal as it is compatible with using several perturbations of cause proteins to change their fold. Recent development of theta-capillary emitter for the ESI source greatly shortens the mixing time and pushes the limit down to 270 ns−27 μs.[1102] It was demonstrated that this novel development can be used in both protein folding studies[1103] and HDX[234] of proteins on the μs time scale. These exciting developments open new possibilities for protein folding studies. By spraying a denatured protein and renaturing buffer through a theta-capillary emitter and analyzing the resulting protein in an ETD-based top-down fashion, it may be possible to analyze protein folding with high spatial resolution. A caution is that the protein conformational change occurring upon ESI is not physiologically relevant.

6.3. Assaying Protein Aggregation

Studies of protein aggregation and its relationship to neurodegenerative diseases date back to 1910, when Fritz Heinrich Lewy first observed unusual protein aggregates in the brains of patients with Parkinson's disease.[1104,1105] These aggregates were later termed Lewy bodies[1106] and became the diagnostic for Parkinson's disease. Now, it is well understood that a Lewy body is composed of misfolded α-Syn. Alzheimer's disease, as first described by Alois Alzheimer,[1107] is another major neurodegenerative disease induced by protein misfolding and aggregation of amyloid beta (Aβ).[1108] A 2016 report indicates there are approximately 6.2 million people globally who were affected by Parkinson's disease.[1109] The number is 29.8 million for Alzheimer's disease.[1109] Unfortunately, neither of these diseases have cures, and their pathologies are not well understood.[1105,1108,1110,1111] Thus, it is important to develop novel approaches to locate and quantify protein aggregates, to study the aggregation mechanism, and to develop therapies.

Characterization of protein aggregates is challenging, especially at high spatial resolution. The final state of these aggregates, solid Aβ fibrils, for example, can be characterized with high resolution approaches in solution (labeled by D_2O and solubilized in dimethyl sulfoxide)[1112] and by solid-state NMR,[1113] X-ray crystallography,[1114] and modern cryo-EM.[1115,1116] The low molecular weight soluble oligomeric intermediates that "come and go" during aggregation, although physiologically important, are not amenable to those techniques owing to their low solubility, high heterogeneity, intrinsic disorder, instability, and high aggregation propensity. Some of these aggregates can be investigated by atomic force microscopy (AFM) to give a high-resolution morphological picture of early aggregation intermediates.[1117] No site specificity or spatial resolution, however, are available when using such approaches.

Among all characterization approaches, fluorescence is the most widely adopted to follow protein aggregation. With the discovery of thioflavin T (ThT),[1118] a dye that when it binds β sheet-rich structures exhibits an enhancement in fluorescence, has become the classic way of following aggregation.[1119] The Aβ aggregation process can then be tracked and quantified both in vivo and in vitro. Given its broad applicability for characterizing different aggregation stages, ThT fluorescence is regarded as a characteristic of amyloid fibril formation.[1120] Nevertheless, limited spatial resolution and possible distortion of the aggregation process by adding the dye have always been major concerns.[1121]

MS-based approaches can contribute to this field from two distinct points of view. MALDI-imaging enables label-free quantification of Aβ aggregates.[1122] Ion mobility MS is capable of separating different Aβ oligomers based on their collision cross sections in the gas phase.[1123] HDX coupled with MS also reveals structural information on Aβ fibrils,[1124] protofibrils,[1125] and even Aβ aggregation kinetics.[1126] With top-down sequencing and ECD fragmentation, H/D scrambling is minimized, allowing residue-level information on Aβ

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

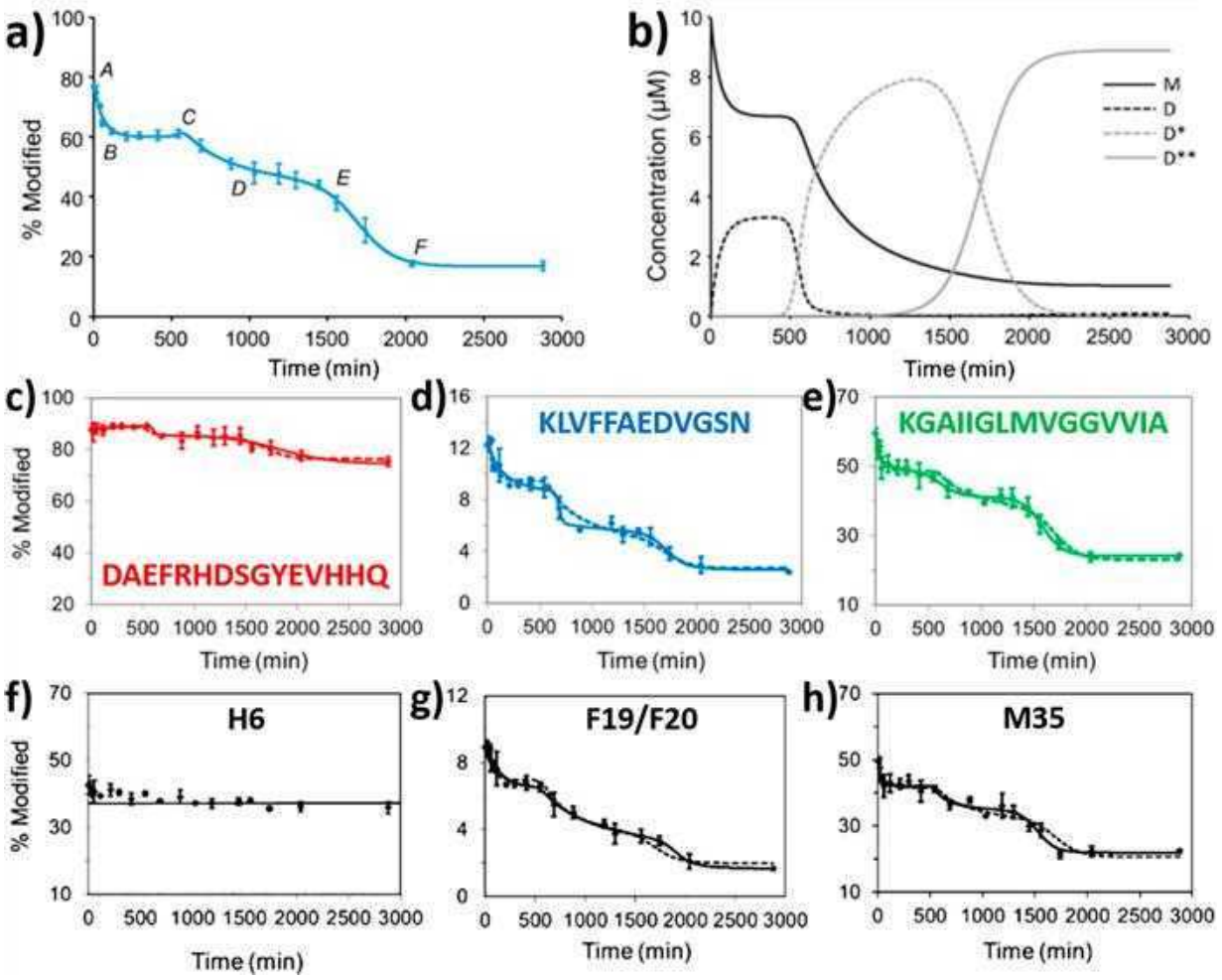


Figure 27. FPOP and kinetic modeling results for time-dependent aggregation of 42-residue Aβ (Aβ_{1-42}) at (a,b) global, peptide, and residue levels. (a) Global (whole protein) level results for Aβ_{1-42} aggregation. Solid curve is from fitting by a model based on two autocatalytic reactions. (b) Concentrations for representative species (M-monomer, D-paranuclei, D*-protofibrils, D**-fibrils) as a function of incubation time based on kinetic simulation. (c−e) Peptide-level results for Lys-N digested Aβ_{1-42}, N-terminal region 1−15, middle region 16−27, and C-terminal region 28−42, respectively. (f−h) Residue level results for three representative residues, H6, F19/F20, and M35, respectively. Points in each plot represent experimental data (10 μM, pH 7.4, no agitation) and error bars are standard deviations from three independent trials. Solid and dashed lines in (c−h) are model fits independent of or constrained by the global rates, respectively. Adapted with permission from ref 1026. Copyright 2009 American Chemical Society.

oligomers.[1127,1128] These efforts were recently reviewed.[1129−1131]

The ability to study Aβ aggregates by radical footprinting was first demonstrated in 2014, when Klinger, Chance, Axelsen, and co-workers[1132] utilized synchrotron-based hydroxyl radical footprinting (HRF) to examine the fibrils and prefibrillar forms of the 40-residue Aβ (Aβ_{1-40}). By comparing the protection factors for selected residues in prefibrillar and fibril Aβ_{1-40}, they mapped the footprinting data onto high-resolution solid-state NMR models and showed that the solution information obtained from HRF-MS is consistent with several core filament structural models elucidated in solid state, but not with other models, which they rejected. Moreover, protection factor analysis supported the linear heterogeneity of the Aβ_{1-40} fibrils; that is, two- and three-filament assemblies alternating along the length of the fibril.[1132] Measured protection factors of the residues on the flexible loop regions of Aβ_{1-40} are also consistent with the structural restraints from solid-state NMR studies.

Later on, Li et al.[1026] applied FPOP to study the aggregation process of 42-residue Aβ (Aβ_{1-42}). By following the Aβ_{1-42} as a function of incubation time with FPOP, the investigators found a clear decrease in modification fraction (Figure 27a) consistent with a decrease in SASA owing to aggregation. By modeling the results with a two-nucleation/two-autocatalytic mechanism, they successfully fit the data (solid line in Figure 27a) and constructed a system composition plot as a function of incubation time, shown in Figure 27b. Plotting relative fractions for four components, the investigators proposed an Aβ_{1-42} early aggregation mechanism, where Aβ_{1-42} monomer rapidly assembles into paranuclei and further accumulates until a certain threshold. Upon passing that limit, the paranuclei self-catalyze a structural reorganization to deplete the monomers and form readily the mature fibrillar aggregates.[1026]

Upon digesting the Aβ_{1-42}, spatial resolution is elevated to regional and even to some residue levels. The N-terminal region remains solvent accessible throughout aggregation (Figure 27c) as seen by the minimal change in modification in that region during aggregation. The central (Figure 27d)

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

and C-terminal regions (Figure 27e), however, actively participate in aggregation, as evidenced by significant decreases in modification fraction. Similar logic applies to residue-level interpretation, where SASA of H6 (Figure 27f) changes little during aggregation, whereas residues F19/F20 (Figure 27g) and M35 (Figure 27h) are actively involved in aggregation. The investigators proposed that F19/F20 contributes a driving force for $A\beta_{1-42}$ aggregation by serving as a hydrophobic nucleation interface.[1026] Although some residue-level resolution was obtained, not all residues could be studied, reaffirming the need for alternative free-radical footprinters that prefer to react with other residues than those reactive with $^{\bullet}OH$ and give complementary information.

MS-based approaches make possible the characterization of early aggregation states, especially transient existed intermediates, which was not possible by the earlier low-resolution optical approaches. MS-based radical footprinting deepens the understanding of $A\beta$ folding mechanisms with some residue-level resolution, and this approach can be further extended to other aggregating systems. Although the footprint cannot be assigned to a specific oligomer, the approach is sensitive to the broad range of states and their changes during aggregation. Nevertheless, the approach can further evolve for aggregation studies of proteins alone or with small organic molecules or metal ions that either inhibit or promote aggregation, thus affording a platform to assist the development of inhibitors.

From a broader perspective, the two MS-based footprinting approaches, pulsed-HDX and radical labeling, complement each other. Both methods report changes in SASA; HDX is sensitive to backbone hydrogen bonding and solvent accessibility, whereas radical footprinting monitors changes in side chain orientations. Moreover, radical footprinting can shorten the detection time scale to milliseconds or less. By combining pulsed HDX and fast radical labeling, one can obtain an understanding of protein aggregation intermediates and mechanisms, and these insights should prove invaluable in developing novel therapeutics.

6.4. Probing Ligand Binding, Affinity, and Dynamics

Numerous proteins interact with ligands to facilitate biological processes for which quantitative understanding is vital.[1133] Protein−ligand binding affinity measurements quantitatively characterize these processes; by definition, affinity is the equilibrium constant for the reaction of ligand binding to a protein. It can be expressed as a K_a or K_d (see 23−25). To date, three approaches have been used in the majority of applications to characterize such interactions. Measurements include circular dichroism,[1133] fluorescence and fluorescence polarization,[1134,1135] FT-IR,[1136] and NMR,[1038,1137] that provide read-out of the system composition at given ligand concentrations. These readouts are further utilized, via eqs 24 and 25, to derive macroscopic binding affinities. Surface plasmon resonance (SPR) measures the ligand-on and -off rates, providing binding affinity through a kinetic approach.[76,1138] Isothermal titration calorimetry (ITC) measures heat flow during a ligand titration, and binding affinities are obtained employing the van't Hoff equation.[74] All these methods are considered to be standard in binding affinity determination, and SPR and ITC are commonly used for their speed and ease of operation.[1139] On the other hand, these methods give limited spatial resolution, sometimes providing little structural information (e.g., SPR and ITC), or by the requirements of large sample amounts (e.g., mg for NMR) or special sample preparation (e.g., chip-loading for SPR).

$$\text{Protein}(P) + \text{Ligand}(L) \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} \text{Protein−Ligand complex}(PL) \tag{23}$$

$$K_a = \frac{k_{on}}{k_{off}} = \frac{[PL]}{[P][L]} \tag{24}$$

$$K_d = \frac{k_{off}}{k_{on}} = \frac{[P][L]}{[PL]} \tag{25}$$

MS-based approaches can advance this field by providing another direct approach that significantly lowers the sample amount. A caveat is whether direct MS can preserve the biological-relevant environment.[1139] Under optimal conditions, a MS-based binding assay only requires high-picomole amounts of samples.[1140,1141] Similar to concentration-measurement approaches, MS examines the binding system by measuring its composition (i.e., the concentrations at equilibrium) either directly or indirectly. The direct approach takes advantage of native spray, a gentle form of ESI whereby the protein can be sprayed in a native or near-native state, retaining noncovalent interactions[157,1142,1143] and allowing direct measurements of the various equilibrium concentrations;[1139,1144] this was first demonstrated by Loo et al.,[1145] who determined the affinities for binding between ribonuclease S-protein and S-peptide. Early applications were in several disciplines including protein−protein,[1146,1147] protein−peptide,[1148] protein−oligonucleotide,[1149] protein−small molecule,[1150] peptide−antibiotic,[1151] and small molecule−RNA[1152] complexes.

This direct method has been carefully developed for quantifying protein−glycan interactions, as pioneered by Klassen and co-workers.[1139,1153−1155] In 2019, Nguyen and Donald[1156] measured small molecule−protein interactions through nanoscale ion emitters coupled with native spray, which was not possible previously owing to the high salt concentration needed to stabilize such weak interactions. Although the method is convenient, there is always a question of whether the measured gas-phase concentrations truly represent those in solution; further, the ionization mechanism is complex and not yet well understood.[1139] The glycan/protein interactions agree well with those determined by other solution approaches, possibly because the binding systems, including the ligands, are similar.

Indirect approaches, on the other hand, require predetection labeling. Using this form of labeling in MS was first done by using HDX, where two separate titration methods were demonstrated in the early 2000s, namely SUPREX[238,1157] and PLIMSTEX.[239] Both utilize the differences in deuterium uptake between ligand-free and ligand-bound states at a selected exchange time. SUPREX uses denaturant as the titrant, and the titration midpoint is extracted by modeling the data; those data are further extrapolated to give folding free energies and binding affinities.[238,1158] PLIMSTEX, on the other hand, uses the ligand as titrant, and binding constants are derived by fitting the titration curve.[239] Both approaches have been extended to the peptide level, as seen in a study that characterizes interactions between a small-molecule drug candidate and apolipoprotein E3, providing regional specificity and improving spatial resolution.[240]

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

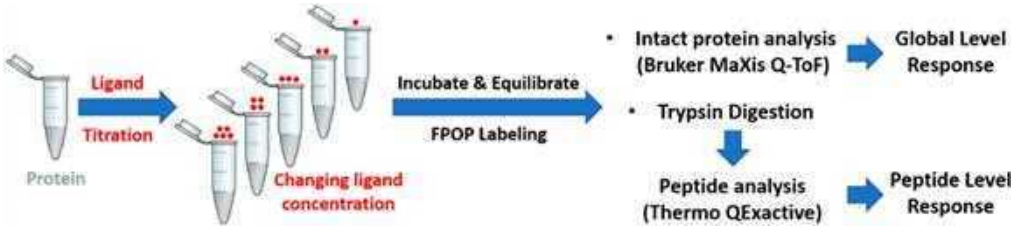


Figure 28. Schematic illustration LITPOMS workflow. Reproduced with permission from ref 1141. Copyright 2019 Springer.

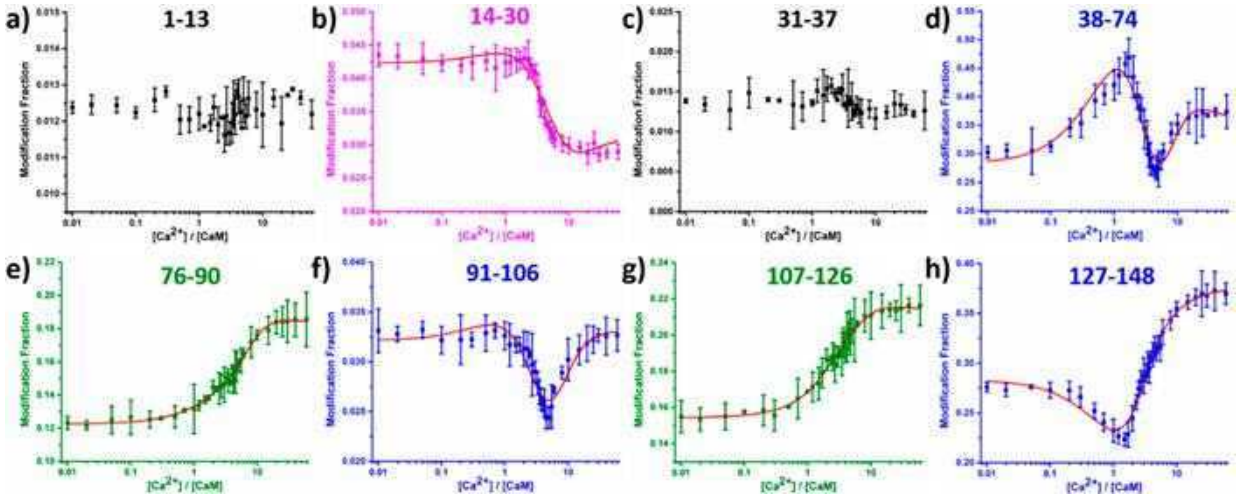


Figure 29. LITPOMS response of Ca^{2+}-CaM titration at peptide level, where modification fractions were plotted as a function of calcium:calmodulin concentration ratio. Four different classes of behaviors are shown in black (a,c), magenta (b), blue (d,f,h), and olive (e,g). Red solid lines in (b,d,e,f,g,h) are from fitting by a model used in PLIMSTEX. Reproduced with permission from ref 695. Copyright 2019 American Chemical Society.

Owing to the relatively long times used for HDX, neither of these approaches are compatible with systems with high ligand-off rates. They also suffer from postlabeling back exchange. Footprinting by radical species overcomes these drawbacks owing to the irreversible and fast labeling that pertain. In 2019, Liu et al.[1141] proposed a method named ligand titration, fast photochemical oxidation of proteins, and mass spectrometry (LITPOMS) that for the first time marries fast labeling species to binding affinity determination. As shown in Figure 28, a series of aliquots with different ligand concentrations comprise the titration process. Each aliquot is FPOP footprinted with $^{\bullet}OH$, and the modified samples further analyzed by MS. LITPOMS can be executed under two different concentrations, where the high-concentration experiment reveals binding stoichiometry and the low-concentration experiment can be modeled to extract binding affinity, just as with PLIMSTEX.

LITPOMS was quickly adopted, also in 2019, to study a more complex binding system, Ca^{2+}-CaM, in which CaM binds to four Ca^{2+} ions in a cooperative fashion.[695] Other than the simple binding behavior shown in Figure 29b, composite binding curves in Figure 29d,f,h show a combination of binding and conformational changes induced remotely. For example, Ca^{2+} binding at the region represented by peptide 127−148 (Figure 29h) triggers a structural opening at another region represented by peptide 38−74 (Figure 29d), preparing this site to take its own Ca^{2+} and revealing allosteric behavior that was not well understood before. Such composite LITPOMS curves can be further dissected by using residue-level analysis together with several different protease digestions to increase spatial resolution.[1159] Moreover, binding order can be assigned through ranking the onset point of LITPOMS decrease as EF4 > EF-3 > EF-2 > EF-1,[695] and was confirmed by a subsequent simulation study.[1160] Site-specific binding affinities can also be obtained by modeling of eight LITPOMS curves together. This measurement strategy may be particularly effective for signaling proteins that undergo numerous complex conformational changes in executing their functions.

The capability to make binding affinity measurements through radical labeling and MS is promising, as it provides most of the key information about the protein−ligand interactions via a single approach that reveals not only the binding sites and binding affinities but also the binding orders, allosteric behavior, and protein binding dynamics at high spatial resolution.[695,1159] To obtain all this information previously required a combination of several sophisticated methods. LITPOMS should have utility in studying complex binding systems, signaling proteins as mentioned above, as it fills the gap that exists between methods offering no spatial resolution to high-resolution X-ray or NMR structures with an approach that offers mid spatial resolution for ligand-free and ligand-bound states.

Challenges remain, however, as there is not yet a radical species that can react with all 20 amino acid residues in a single experiment. For example, $^{\bullet}OH$ reacts with all 20 amino acids, but the range of reactivity is too large to allow the least reactive residue to compete with the most reactive in a single experiment, as was discussed earlier in this review. Thus, development of novel and complementary labeling reagents is vital, whereby one footprinter can react with the residues with

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

competitive reactivity and several footprinters can cover all the amino acids in several experiments, or even in a single experiment utilizing multiple reagents. In that way, all residues can be involved in binding to report binding-induced SASA changes. Moreover, modeling of titration data can be challenging, especially for proteins that bind multiple ligands and yield titration curves whose fitting requires construction of complex mathematical models.

6.5. Labeling in Cells and in Vivo

As discussed in section 4.5, in-cell labeling has unique advantages, forecasting a bright future. In-cell labeling by radical species was first demonstrated with nucleotides, where Ottinger and Tullius[1161] used hydroxyl radicals to footprint the lambda repressor-DNA complex in live *E. coli* cells. Hydroxyl radicals in this study were generated by ionizing water molecules with a ^{137}Cs gamma ray source, which required up to 15 min to produce a sufficient concentration of radicals. To minimize heat damage to the cells during the long irradiation, Woodson and co-workers[1162,1163] utilized synchrotron X-rays to reduce the exposure time to as short as 100 ms and froze their samples (−34 to −38 °C) to minimize damage.[1164] Recently, similar experiments were done with a flow system to minimize double labeling and further reduce the exposure time to 10−20 ms.[1165] Better X-ray dose control also improves the accuracy and reproducibility.

In-cell protein footprinting by radical species, however, was not demonstrated until 2009, when Zhu and Sze[876] used hydroxyl radicals generated by Fenton chemistry. The protein of interest was OmpF porin from *E. coli*, an outer membrane channeling protein that facilitates aqueous passive transport in Gram-negative bacteria. Their approach began with incubating live *E. coli* with Fe(II)-EDTA and H_2O_2 and then allowing reactions of $^{\bullet}$OH from Fenton chemistry to label in-cell the solvent-accessible residues. As demonstrated in Figure 30c, solvent accessibilities of two loop regions, L1 and L3, do not change significantly upon opening and closing the channel, suggesting that the loops are not responsible for manipulating the channel.[876] In other words, entrance of the channel is not blocked by the extracellular loops. Peptides from regions inside the channel (Figure 30d), however, are heavily oxidized when the channel is opened and barely oxidized when the channel is closed. Data support the hypothesis that the three resolved β-sheets are in the channel pore, as demonstrated by Figure 30a,b. Although voltage gating of OmpF was shown previously in vitro, this study for the first time reports the observation of gating phenomena in a native cellular environment.[876] Moreover, the novel footprinting also shows the capability of in-cell protein footprinting by radical species in addressing real biological questions.

In subsequent work, Shcherbakova et al.[882] shortened the time scale of $^{\bullet}$OH generation. By increasing the concentration of Fe(II)-EDTA to 2 mM, a 40% loss of fluorescence intensity was observed after 2 ms, suggesting that $^{\bullet}$OH can be rapidly generated and that this approach can label other short-lived intermediates.

Another widely used $^{\bullet}$OH generation method, FPOP, was also adopted for in-cell protein labeling.

In the first demonstration by Jones and co-workers,[1166] they mixed African green monkey kidney cells with H_2O_2 and submitted the cells to KrF laser irradiation for $^{\bullet}$OH generation and protein labeling. H_2O_2 crosses the membrane through both passive diffusion and via channeling proteins such as

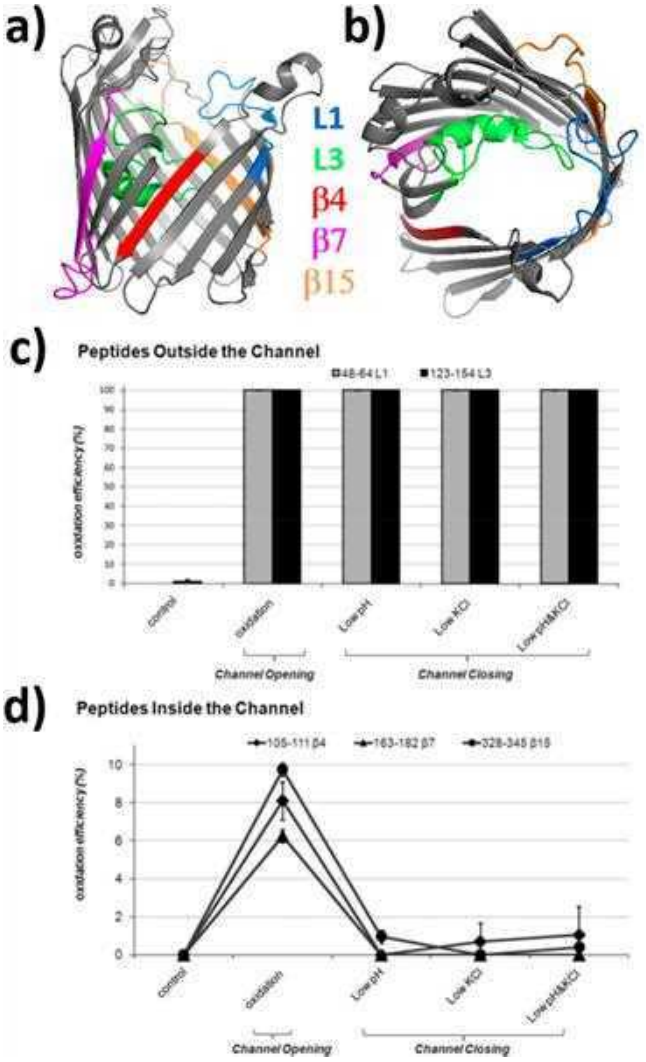


Figure 30. In cell labeling of OmpF porin by hydroxyl radicals. 3D structure of OmpF (PDB 1OPF) from (a) side view and (b) top view. Structural components of loop 1, 3, and β-sheet 4, 7, 15 are highlighted in blue, green, red, magenta, and orange, respectively. Oxidation efficiency for (c) loops and (d) β-sheets under different experimental conditions are plotted. Error bars are standard deviations from duplicates. (c,d) Reproduced with permission from ref 876. Copyright 2009 American Society for Biochemistry and Molecular Biology.

aquaporin, producing sufficient H_2O_2 inside the cell to give a good yield of $^{\bullet}$OH upon laser irradiation. Although H_2O_2 is toxic to live cells, the time required to execute the experiment seems to be short enough to avoid serious toxicity. Indeed, viability tests prior to labeling suggest that over 70% of cells remain alive under the experimental conditions used for the footprinting. The investigators observed 105 proteins that are oxidized by $^{\bullet}$OH. These proteins were from different subcellular compartments, as shown in Figure 31a, indicating an outstanding dynamic range for protein detection.[1166]

When zooming into a specific protein (e.g., actin), FPOP modification fractions can be correlated with SASA of the two different states of actin, namely, open and tight. Results in Figure 31b show a better correlation with the open state (R^2 = 0.89), suggesting that majority of actin molecules in the native Vero cells are in their open state.[1166] Later, the same group improved the experimental apparatus by introducing a sheath

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

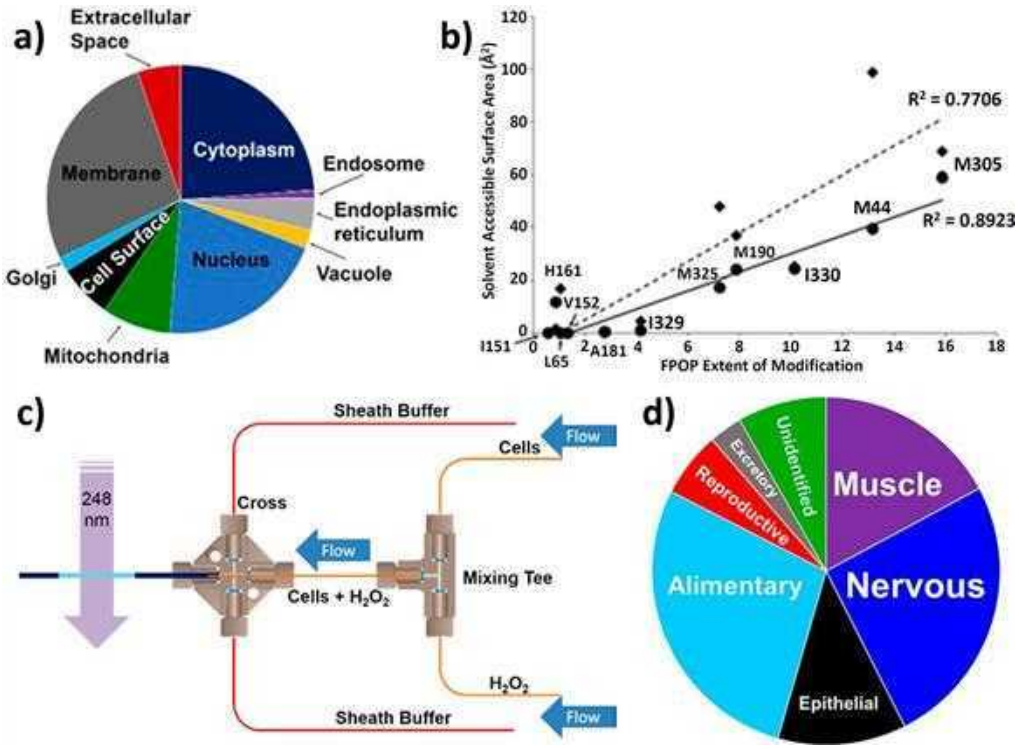


Figure 31. (a) Pie chart of the subcellular compartment location (African green monkey kidney cell) of oxidized proteins that were identified by LC-MS/MS. Proteins that are present in multiple compartments are represented multiple times. Reproduced with permission from ref 1166. Copyright 2015 American Chemical Society. (b) Correlation of residue-level FPOP modifications with SASA in the open (circles and solid line) and tight (diamond and dashed line) states of actin. Reproduced with permission from ref 1166. Copyright 2015 American Chemical Society. (c) Schematic illustration of an improved flow system for in vivo FPOP labeling. Blue arrows indicate flow and colored lines represent tubing. Reproduced with permission from ref 1167. Copyright 2016 American Chemical Society. (d) Pie chart of the oxidatively modified proteins within different body systems of *C. elegans*. Reproduced with permission from ref 1168. Copyright 2019 American Chemical Society.

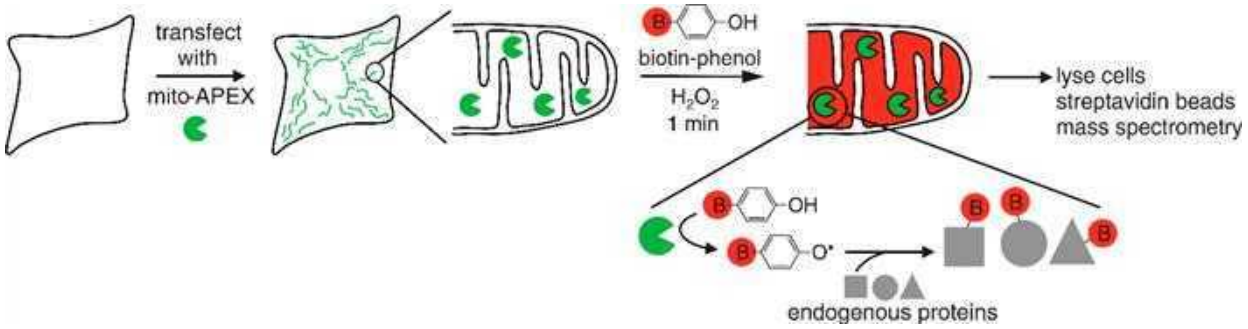


Figure 32. Schematic illustration of APEX labeling workflow. Reproduced with permission from ref 1170. Copyright 2013 American Association for the Advancement of Science.

buffer (Figure 31c) to reduce cell aggregation greatly and tube clogging.[1167] Their design also ensures that radiation exposure of each cell in the flow system remains comparable. As a result, they were able to increase the number of identified oxidized proteins by 13-fold.[1167]

In 2019, Jones and co-workers[1168] introduced small, live worms, *C. elegans* into the flow system, and they were able to identify oxidatively labeled proteins from different body systems as shown in Figure 31d. A closer look at the myosin chaperone protein UNC-45 suggests that the modification fractions obtained in this in vivo study correlate well with the SASA calculated from the crystal structure.

Owing to incomplete understanding of the absolute extent of modification, most MS-based protein footprinting approaches have been executed in a differential way, where the extent of labeling for a peptide/residue in different states (e.g., ligand-bound versus ligand-unbound) can be compared in a valid way. Although promising, work by Jones et al. fails to reveal such differences up until now.

In one of the pioneering demonstrations of in-cell radical footprinting that used a differential approach, Zhu and Sze[1169] footprinted the structural changes of epidermal growth factor receptor upon binding with epidermal growth factor (EGF) in live *E. coli* cells. Hydroxyl radicals were generated by an FPOP approach, and the oxidation extents of EGF-free and EGF-bound states were compared. The results are consistent with crystal structures of these two states. Their study marries conventional bottom-up structural proteomics with in-cell free radical labeling, probing differential structural changes in a real cellular environment. Moreover, their workflow, including

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

enrichment by immunoprecipitation and in-gel digestion, is applicable to other studies.

Another important class of in vivo protein footprinting by radical species is proximity labeling, pioneered by Ting and coworkers starting in 2013.[1170,1171] Unlike all other labeling approaches introduced prior to this study, proximity labeling is usually facilitated by enzymes. The first enzyme used to facilitate the labeling is ascorbate peroxidase (APEX), and the method was appropriately termed APEX labeling. In brief, APEX is genetically tagged onto the protein of interest, whose location in the cell is generally known. Upon adding H_2O_2 and biotin−phenol into the cell, APEX catalyzes the transformation of biotin−phenol into biotin−phenoxyl radicals, which then footprint nearby proteins (Figure 32). The biotin tag facilitates postlabeling protein enrichment. As a result, the investigators were able to identify 495 proteins within the human mitochondrial matrix, 31 of which were not previously linked to mitochondria. APEX labeling is highly specific, as the radicals are short-lived (<1 ms)[1172] and the labeling radius is within 20 nm.[1173] APEX labeling was optimized in 2015 by introducing different enzymes[1174] and labeling reagents,[1175,1176] making APEX labeling widely applicable to various cellular systems. On the other hand, APEX labeling is mainly used for proteomics purposes (i.e., determining primary structure). Its potential in protein footprinting for HOS elucidation remains to be explored. The idea of proximity labeling inspired another investigation of protein−carbohydrate interactions,[875] which is covered in section 6.8.

All in all, in-cell and in vivo labeling appears to have unprecedented advantages in structural proteomics to complement in vitro approaches that usually do not accurately reproduce the cellular environment. Another advantage for these approaches is their applicability in studying membrane proteins and their complexes, which have poor solubility or instability in vitro. The short time scale of $^{\bullet}$OH footprinting enables the detection of many transient structural intermediates, especially with the FPOP approach. Modern MS instruments also allow the detection of thousands of unique proteins in a single run. All of these advantages forecast a bright future for in-cell and in vivo footprinting.

On the other hand, current demonstrations of this footprinting fail to reveal as much structural information as conventional in vitro bottom-up approaches. As there are thousands of proteins in a single cell, most current studies must still emphasize protein identification. It is challenging to track structural changes of a single protein during a biological event. This requires better protein separation and enrichment, further improvements in MS instrumentation, and better data processing software. Previous efforts demonstrated the correlation between absolute modification fractions and SASA. By using computer-based protein structure modeling, footprinting can produce restraints that can be used to predict protein structure. In this way, it may be possible to bypass the differential approach without sacrificing structural or spatial resolution. Several studies show promising results, and they will be covered in the next section.

6.6. Footprinting in Support of Computer-Based Structural Predictions

So far, most applications described in this section have been based on differential experiments, where the modification fractions for specific peptides and residues were tracked as a function of protein states. Different protein states are achieved, for example, as ligand-bound vs unbound in epitope mapping, as native vs mutant, after different folding times, at different aggregation states, at various ligand concentrations in affinity determination, and after conformational changes (e.g., opening and closing a channel). Although straightforward, the differential approach has some limitations. Although various systems can be characterized, comparisons require relatively pure systems where different protein states can be clearly differentiated.

Ideally, HOS would be determined directly by footprinting a single protein state. Radical labeling and even other MS-based approaches, however, are not capable of providing enough restraints to construct a high-resolution protein structure as NMR, Cryo-RM, and X-ray crystallography. Computer-based protein HOS prediction is powerful and has great potential, yet the incomplete understanding of protein folding does not allow accurate prediction of protein HOS.[79,1064] Data from NMR,[1177] small-angle X-ray scattering,[1178] Cryo-EM,[1179,1180] and site-directed spin labeling EPR[1181] overcome such limitations because they are based on structural restrictions or restraints.

An MS-based approach that is particularly amenable to modeling is chemical cross-linking.[766,1182] Most of HOS elucidated from this workflow, however, is of protein complexes, because only with large and complex systems is it possible so far to obtain sufficient cross-links for modeling. Further, the principal HOS outcome is a protein−protein interface and not the full structure, although more detailed structure determinations are becoming possible. It may be that footprinting combined with cross-linking will lead to more reliable models than from either method itself.[773]

Structural prediction based on radical footprinting data was not possible until it can be properly quantized and reasonably correlated to the absolute SASA. Gerega and Downard[1183] in 2006 developed a docking algorithm named PROXIMO that adopts directly modification fractions for each resolvable residue as restraints. They were able to dock properly the calmodulin−melittin and ribonuclease S-protein−peptide complexes. Although straightforward, their approach failed to consider the different reactivities between $^{\bullet}$OH and different amino acid residues, and that failure may produce bias in the footprinting data. This pioneering work expands the possibilities of radical footprinting and motivates more precise data quantification.

To date, three groups have demonstrated the efficacy of modeling footprinting results. In 2015, Huang and Chance[1184] first proposed a protection factor (PF) that successfully correlates single-state footprinting data with absolute SASA. The observed reaction rate from a dose−response curve from synchrotron-based $^{\bullet}$OH footprinting (modification fraction versus X-ray irradiation time) is a function of both intrinsic reactivity of $^{\bullet}$OH and the amino acid residue and its solvent accessibility.[166] Teasing out a structural contribution can be done by normalizing the observed reaction rate constant with respect to the residue-specific intrinsic reactivity. As a result, solvent accessibilities (calculated from X-ray structures) for three model systems and local structural contacts exhibit a quantitative agreement with calculated PFs, motivating future developments.

Recently, the investigators adopted this idea for characterizing the human estrogen receptor alpha (hERα), which contains a DNA-binding domain and a ligand-binding domain.[1185] Structures of these two domains are known

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

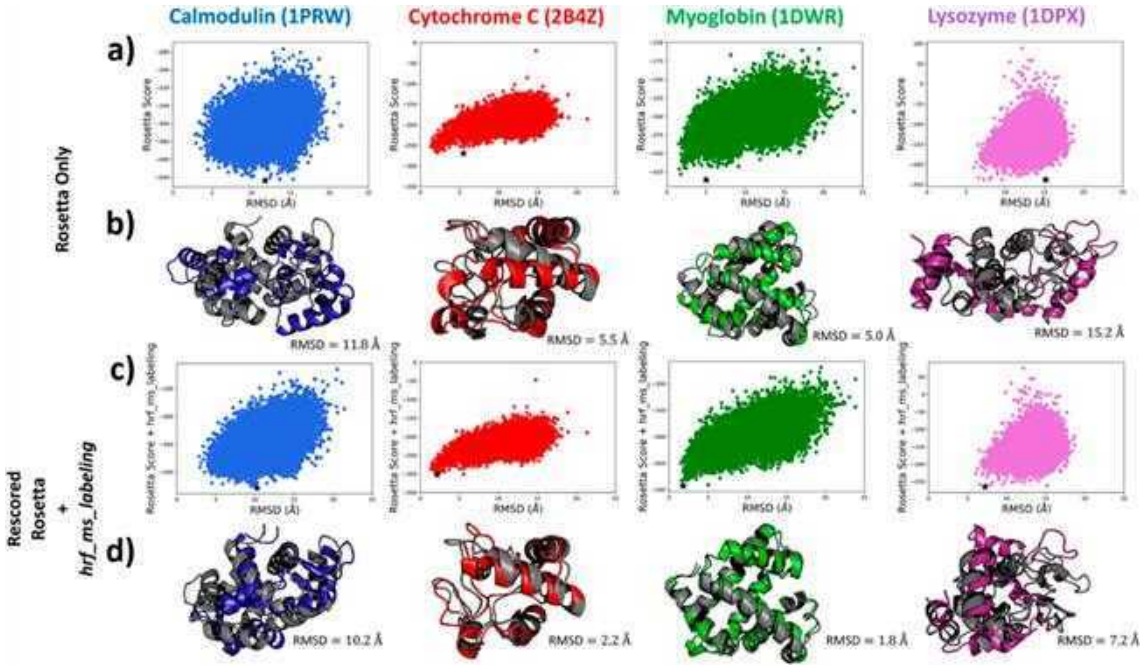


Figure 33. (a) Rosetta score versus RMSD (with respect to the native structure) plots for 20 000 models generated from Rosetta ab initio for each of four proteins. Top scoring model is highlighted by a star in each plot. (b) Top scoring models from the Rosetta score versus RMSD distributions in (a) (color) superimposed on the respective native model (gray). PDB ID for these native models are depicted at the top of the figure. (c) Rosetta score + *hrf_ms_labeling* versus RMSD (with respect to the native structure) plots for each of the four proteins after rescoring with the new score term. The top-scoring model is highlighted by a star in each plot. (d) Top scoring models from the Rosetta score + *hrf_ms_labeling* rescoring distributions in (c) (color) superimposed on the respective native model (gray). Reproduced with permission from ref 1187. Copyright 2018 American Chemical Society.

individually but not in the complex. Using results from •OH radical footprinting, PF calculations of resolved residues, and data from SAXS, the investigators generated sufficient structural restraints to guide successfully molecular docking. The outcome is a successful structure determination of an asymmetric L-shaped structure of the multidomain hERα, revealing key mechanisms that facilitate allosteric function.

In another elegant experiment, Xie and Sharp[1186] introduced a method for assessment of side chain absolute SASA values by using an FPOP-based •OH footprinting platform. By incorporating adenine as •OH dosimeter, they obtained reaction rates for each residue through multipoint FPOP experiments (different •OH doses). These data were further normalized with respect to intrinsic reaction rates of •OH and free amino acids. Normalized protection factors obtained through these efforts were compared with fractional SASA, and a good linear regression was obtained.[1186] Further, two SASA prediction models were constructed for the soluble proteins myoglobin and lysozyme, and the predicted SASAs were compared with a calculated SASA from crystal structures by using a protein-unfolding MD simulation. Although the incorporation of radical protein footprinting data in some structural modeling is successful, use of the footprinting data directly in protein structure prediction is yet to be established.

In 2018, Lindert and co-workers[1187] developed a new Rosetta score term named *hrf_ms_labeling* that utilizes residue-level PFs from •OH footprinting data as restraints to predict protein structures. PFs were calculated based on the formula proposed by Huang and Chance.[1184] With Rosetta, they generated 20 000 structural models with an ab initio method for four different proteins (i.e., calmodulin, cytochrome *c*, myoglobin, and lysozyme). The Rosetta score for each model was plotted against their RMSD with respect to their native crystal structures as shown in Figure 33a. The 20 000 models for each of the proteins covers a broad range of RMSDs. Aligning the top-scoring model with the native structure allows better visualization of the differences, as there are even topological mismatches for calmodulin and lysozyme (Figure 33b). Although there are models that have near-atomic-level matching (RMSD $\leq$ 2 Å), they were not recognized as the top scoring models by Rosetta. When rescoring the same 20 000 models with restraints from experimentally determined PFs, similar distributions were identified (Figure 33c). With the help of new scoring function *hrf_ms_labeling*, the RMSDs for the top-scoring models improve significantly for all four proteins. A visualization of the outcome (Figure 33d) shows that all models identify the correct protein topology.[1187] Structure models for cytochrome *c* and myoglobin come with near-atomic resolution, indicating the efficacy and applicability of the newly developed scoring function. This work demonstrates for the first time that incorporation of •OH protein footprinting data can greatly enhance model quality in protein structural prediction, elevating the use of MS-based footprinting from a qualitative description to a quantitative evaluation. If these efforts continue to be successful, protein footprinting may no longer limited to differential experiments, and the results for a single protein state may be transformed into a high-resolution structure. The script of this newly developed scoring function and the corresponding instructions are freely available with Rosetta.

In another example, the same group[1188] developed novel Rosetta scoring function that utilizes SID MS data and demonstrated that SID data significantly enhance the confidence of structural predictions of protein complexes.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

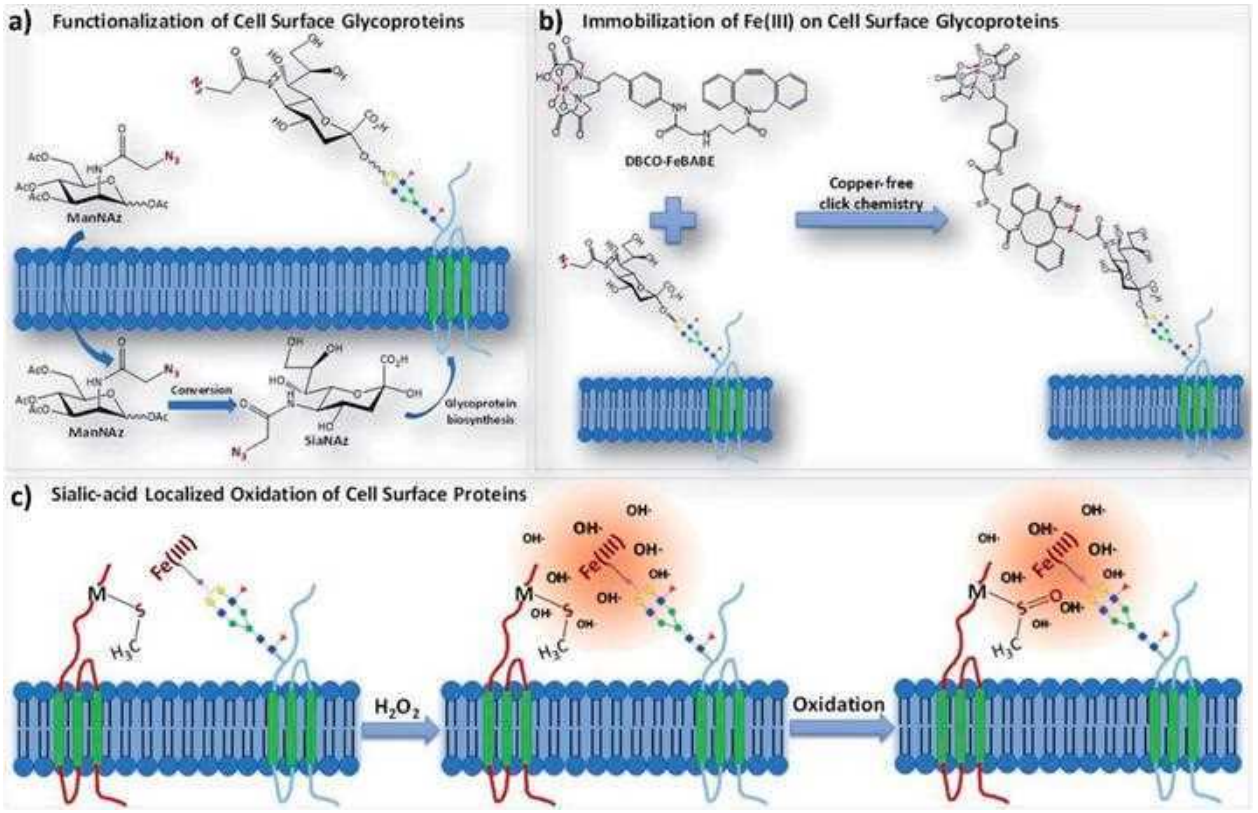


Figure 34. Schematic illustration of POSE workflow. (a) The incorporation of ManNAz into cell surface glycoproteins as SiaNAz. (b) The bioorthogonal reaction of the DBCO-FeBABE probe with the SiaNAc group conducted via click-chemistry. (c) The oxidation of proteins in the sialic acid environment by $^{\bullet}$OH generated by Fenton chemistry. Reproduced with permission from ref 875. Copyright 2019 Royal Society of Chemistry.

Although still in its early stages and requiring further work, structural prediction utilizing footprinting data shows significant promise. It will be necessary to generalize these workflows and make them compatible with other footprinting approaches. Further, it will be even more significant to develop a new structural prediction algorithm that is guided by footprinting data rather than by rescoring the existing models, as hinted in a subsequent study by Lindert and co-workers[839] in 2019.

6.7. Revealing Pathways in Biological Systems

ROS are those that form upon incomplete reduction of oxygen; they include superoxide anion ($O_2^{-\bullet}$), H_2O_2, singlet oxygen ($^1O_2$) and $^{\bullet}$OH.[847,1189] These species play essential roles in regulating various functions in biological systems.[847] On the other hand, ROS are highly reactive and are likely to induce oxidative damage at unfavorable locations. An interesting 2017 study takes a lead from free-radical damage and goes on to reveal successfully the water channels in photosystem II (PSII).[1190]

PSII is a membrane-bound oxidoreductase that catalyzes the conversion of water oxidation to molecular oxygen and simultaneously the reduction of plastoquinone. Sunlight is converted into chemical energy, and electron transfer in photosynthesis is thus initiated. A high resolution structure of PSII identified the catalytic center as Mn_4CaO_5, where substrate water is oxidized to O_2.[1191] The pathway for the substrate water molecule into the catalytic center, however, is poorly understood. As $^{\bullet}$OH is the side-product of such conversion, and the size and hydrophilicity of $^{\bullet}$OH are similar to those of H_2O,[848] as discussed earlier, it is possible to map the H_2O channel by tracking the movement of $^{\bullet}$OH and its oxidative damage. In other words, the $^{\bullet}$OH modified residues leave a trail or "footprint" of natural oxidative damage that line the wall of H_2O channel, revealing the structural aspect of the PSII. Using an approach that includes analysis used for protein footprinting, the investigators identified three distinct water channels, two of which aligned well with those identified in previous studies. This work resembles that of Ting,[1170,1171] described in section 6.5, except the oxidation occurs naturally.

This simple experimental design utilizes the naturally occurring $^{\bullet}$OH and related species as labeling reagents to footprint PSII in a cell. The structural information obtained in this work increases the mechanistic understanding of the PSII system. Moreover, the outcome motivates further development of radical protein footprinting and ROS localization to understand the chemistry of $^{\bullet}$OH and to augment, even supplant, other low-resolution approaches for detection of $^{\bullet}$OH.[1189] MS-based in situ radical footprinting offers an elegant way to localize the $^{\bullet}$OH by tracking its damage trails on the surrounding amino acid side chains. The utilization of naturally occurring radical species in protein footprinting may also provide insights into "natural" oxidative modifications and whether they introduce perturbations to native biological systems.

Inspired by this idea, a recent study focused on $CO_3^{-\bullet}$, a naturally occurring radical, and used the FPOP platform to characterize the reactions of $CO_3^{-\bullet}$ that may occur in vivo.[962] Although a reactive species, the $CO_3^{-\bullet}$ radical anion does not qualify as a good footprinter owing to ambiguity in its generation from $^{\bullet}$OH or other strong oxidizing radicals, underscoring that novel methodologies and other radical species are necessary in the footprinting field.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

6.8. Other Applications

In addition to the applications covered above, there are other promising studies showing the capability of free radical protein footprinting to address other structural concerns.

The first is the identification of bound waters in the proteins. Water plays an essential role in protein folding, structure, and stability.[1192,1193] X-ray diffraction[1194] and NMR[1195] can detect highly conserved water molecules that are located bound with polar or charged residues of proteins. Few experimental approaches, however, can differentiate waters that are bound to the surface and the interior of proteins. Synchrotron-based •OH protein footprinting is one option because it ionizes water to give •OH.[175] Using a careful experimental design, Gupta and Chance[1196] footprinted cytochrome *c* under two different temperatures of 25 and −35 °C.

Freezing the sample during synchrotron irradiation limits the diffusion of •OH. A comparison of modification fractions of a certain peptide or residue under two different temperatures enables the differentiation of types of water that locate in these regions. A significant decrease (13−200-fold) suggests that the modification is from bulk water that gives •OH and labels the protein through diffusion. Sites that suffer a moderate decrease (3−10-fold) are located on the protein surface, and these sites have bound water molecules. Residues that have minimum changes (decrease less than 2-fold) upon footprinting at two different temperatures must have internal, closely located water to give •OH radicals that react with adjoining residues without significant diffusional motion. Results from •OH footprinting are consistent with those seen in an X-ray structure.[1196] Further experiments utilized a time-resolved $^{18}O/^{16}O$ exchange apparatus and probed the water dynamics on the submillisecond time scale,[1196] offering a unique view of protein−water interacting dynamics that was only possible with NMR before.[1195] Synchrotron-based •OH footprinting has unique advantages in studying protein−water interactions, as water serves as the precursor of •OH, and wherever water is located, •OH can be produced and react with protein sites, even within the lipid bilayer of membrane proteins.[1197,1198]

The second application focuses on protein−carbohydrate interactions. Cell membranes are covered with heterogeneous glycans that create an interactive environment and govern many cellular functions.[1199] Many cellular functions are based on protein−glycan interactions and glycan-mediated protein−protein binding, yet characterization of the protein−glycan interactions is mainly limited to glycan arrays.[1200] Very recently, Li and Lebrilla[875] reported a novel workflow termed protein oxidation of sialic acid environment (POSE) that footprints those proteins in close proximity to sialic acids, an idea that shares features with APEX labeling introduced earlier. As sialic acid usually terminates the glycan chains in the glycoproteins, the first step is to functionalize the sialic acid with an azido group (Figure 34a) and to introduce the active iron species through click chemistry, the glycan chain now serving as a probe (Figure 34b). Upon adding H_2O_2, •OH is generated by Fenton chemistry, thus oxidatively labeling the surrounding proteins (Figure 34c). POSE allows in situ labeling of proteins that are located closely to glycans, providing an effective and efficient way to screen possible glycan-binding targets.

Finally, there are other promising applications of fast labeling approaches including footprinting membrane proteins,[950,1006,1201,1202] following the early onset of ROS oxidative damage,[1203] probing protein hidden conformations,[1204] mapping protein−DNA interactions,[1205] and many more. These efforts all provide valuable information in their respective fields, but they will not be covered in detail in this review.

7. CONCLUSIONS AND PERSPECTIVES

MS-based protein HOS analysis has increased the throughput and decreased the sample amount requirements of protein HOS determination while becoming an effective complement to traditional biophysical characterization methods. More importantly, different protein footprinting approaches allow MS-based approaches to view protein HOS from different perspectives. For example, HDX usually measures the kinetics of exchange and infers HOS from the "labeling" of protein backbone amide bonds. Targeted or specific labeling reagents exploit various organic reactions to report on side chain solvent accessibilities. Fast labeling reagents react with the residue side chains on the time scale of nanoseconds to milliseconds, affording a "snapshot" of protein structure and dynamics. Taken together, MS-based protein labeling approaches "paint" the protein solvent accessible surfaces over a time frame from nanoseconds to days to afford a comprehensive understanding of the protein of interest. Moreover, modern proteomics digestion (bottom-up) workflows, top-down fragmentation techniques, together with ultrasensitive MS instruments, yield not only HOS information at mid-to-single-residue level but also dynamics of the protein under various conditions that are hardly accessible by a single technique. This is important because a footprint should be done with high coverage, unlike primary structure identification in traditional proteomics where 100% coverage is not needed. Over the past decades, MS-based structural proteomics approaches have grown extensively and now can answer important and challenging biological questions. None of these developments were possible without the impressive technical innovations in the field of MS that have occurred over 40 years.

7.1. Hydrogen−Deuterium Exchange

During the past decade, HDX has grown to be a mature tool for MS-based protein HOS characterization in both academia and in the biopharmaceutical/biotechnology industry. Interestingly, 79 out of 80 electronically submitted biological license applications to the United States Food and Drug Administration between 2000 and 2015 include MS data, and three of them utilize MS for protein HOS analysis, all done by HDX.[1206] Despite its growth and success, HDX methods still face multiple challenges. One is the tedious, complex measurement, which is being met with robotic automation of the HDX measurement, making bottom-up HDX more readily applicable to most simple protein systems.[1207] Titration-based HDX workflows including SUPREX and PLIMSTEX provide site-specific binding affinities in addition to locating the binding sites.[238,239] Incorporation of ion mobility offers an orthogonal dimension in the separation of constituent peptides prior to MS measurements, facilitating the deconvolution of overlapping isotopic patterns for coeluting peptides.[1208] Novel fragmentation methods including ETD and ECD together with multienzyme digestions allow investigators to reach out for residue-level HDX information.[226] Exchange in theta capillaries[234] and in the gas phase[235−237] may allow HDX to be measured on the microsecond time scale. Innovations in MS instruments permit HDX to be conducted in a top-down

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

fashion. Although promising, these novel advances still require additional demonstrations and testing to spur even broader adoption.

Most current applications of HDX focus on soluble protein systems. As HDX is reversible, its measurement by MS is ex situ, setting experimental constraints that minimize back exchange. These constraints limit application of HDX in complex protein systems; for example, glycoproteins that require postlabeling deglycosylation,[244] membrane proteins need postlabeling lipid or detergent removal,[243] and structurally rigid proteins that are challenging to denature and digest. Although pioneering studies demonstrate feasibility, routine and robust workflows still need to be developed and demonstrated before generalization and widespread acceptance will occur.

Another important aspect of HDX-based protein HOS analysis is the proper utilization of absolute deuterium uptake for structural modeling. Although most current HDX measurements are conducted in a differential manner to locate binding sites or determine the effects of mutation, absolute uptake reports on backbone dynamics and structural features of the protein. Fast exchange is typical of protein loops or intrinsically disordered regions, whereas very slow exchange is characteristic of either structurally rigid or deeply buried regions. Although the concept of absolute protection estimation via HDX measurement is well appreciated, it is poorly demonstrated when combined with computer modeling to predict protein HOS.[1209−1211] The quantitative assignment of protection not only requires high-quality data acquisition but also reliable data processing. As it is challenging, the outcoming is also rewarding. Incorporating protection extent from HDX measurements into protein structure prediction will greatly elevate the contributions of HDX to structure determination, as is seen in recent radical labeling demonstrations.[1187]

7.2. Irreversible Labeling

As compared with HDX, protein HOS analysis by irreversible labeling and MS has a shorter history even though other applications of specific amino acid labeling originated 60 years ago. Many examples of specific amino acid labeling and free radical footprinting, however, have been appearing during the past decade. One advantage is the irreversibility of the labeling, making it compatible with sophisticated postlabeling sample treatment such as long gradients in HPLC or extensive signal averaging as in top-down studies. These features enable approaches that answer broad biological questions about many types of proteins.

Despite its broad compatibility, a missing piece of the picture is membrane protein footprinting. Early studies incorporating nanodiscs[1212] and other membrane mimetic systems[1213] show the feasibility of footprinting membrane proteins. Although these demonstration studies used both targeted labeling reagents[819,823,824] and fast radical species,[876,950,1201] the primary focus was the soluble domain. An effective footprinting approach for the transmembrane domains of membrane protein remains to be established, although there are hints that productive footprinters exist. Unlike HDX, whose labeling reagent, D_2O, cannot be extensively partitioned into biomembranes, the targeted labeling reagents and the radical precursors can be chemically tailored to favor the membrane where they can footprint the transmembrane regions and yield information on those regions that are buried and those that adjoin the membrane.[1214]

Another fundamental aspect of targeted and radical protein labeling that remains to be addressed is the effect of microenvironment; that is, the effect of adjacent residues on the reactivities between targeted residue and the labeling reagents. Such effects are mainly explained as fluctuations of local reagent concentrations, local hydrogen-bonding schemes, or the electronic or steric environments determined by the nearby functional groups. The effect of adjacent residues on the exchange rates in HDX was thoroughly studied and is now well established.[211] For footprinting by targeted labeling reagents, a few studies hint that adjacent residues play a role.[143,506,1215−1219] In the case of radical footprinting reagents, there are two recent reports that discuss local concentration fluctuations of labeling reagents.[145,1220]

Understanding the effect of the microenvironment on labeling efficiencies is crucial in determining the intrinsic structural factors that are involved in interpreting footprinting data, and these effects should not be overlooked. Some recent reports suggest that the extent of labeling by irreversible labeling reagents are not strictly driven by SASA, but by "chemical accessibility" that are a function of local or microenvironments[143,144] (the idea of "chemical accessibility" was mentioned in section 1.3.1). Only from a deep understanding can we take full advantage of footprinting results and their incorporation into protein structural modeling. The needed studies for irreversible labeling, however, are more challenging than those for HDX, whose labeling reagent is high-concentration solvent, D_2O, itself. One can reasonably assume that the effect of a microenvironment will be reagent-dependent, further complicating efforts to unravel the effects but also offering an opportunity to tailor reagents.

From a practical perspective, several issues need to be addressed in upcoming research in this field. First is the implementation of novel labeling reagents with distinct mass tags. As mentioned earlier, a good footprinter needs to provide a distinct mass tag that can be properly differentiated from naturally occurring PTMs. Because targeted labeling reagents generally introduce mass tags that do not occur naturally, protein footprinting by radical species is mostly carried out with hydroxyl radicals, which label protein by oxidation, producing a +16 Da mass shift inter alia. As oxidation often occurs not only naturally but also during post labeling sample handling and even during LC separations, it is sometimes ambiguous to distinguish the oxidations from footprinting and those from other processes. This small limitation can be overcome by development of novel labeling reagents that not only retain the advantages of widely adopted hydroxyl radicals but also "paint" the protein surfaces, affording another distinctive mass tag.

Another possible way to address this issue is to develop cleavable or isotopically encoded radical precursors as labeling reagents, an idea that has been extensively utilized in protein chemical cross-linking. Using this approach, investigators employ targeted labeling reagents that can readily accommodate isotopic encoding to shift the mass of the mass tag. These efforts will surely shorten the analysis time and increase the confidence of identification and quantification in radical-based protein footprinting. This is difficult with hydroxyl radical footprinting because at best the reagent mass can only

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

be shifted by 2 Da (^{16}O vs ^{18}O), which is insufficient to resolve for large peptides and proteins.

The second is the development of novel labeling reagents (examples were discussed in section 5) for residues that are activated with difficulty and, therefore, react sparingly. Although it is feasible to activate selective C−H bonds from an organic chemistry perspective, to achieve it under physiological conditions is considerably more challenging because the modification must be done rapidly. Labeling of Ala, Gly, and Pro whose side chains are functionalized with difficulty is likely only with free radicals, but even then, the reactivity will be low considering that proteins contain many other more reactive residues. Although specific labeling of Gly residues at the protein N-terminus by targeted labeling reagents was demonstrated recently,[1221] general modification of these residues is difficult and been realized only in one study to our knowledge.[1222] Moreover, bridging the fields of protein chemical modification and protein footprinting offers the opportunity to "reinvent the old". Protein chemical modifications have been extensively developed for many years to satisfy several biological and analytical needs. Some of the reactions and reagents that are adopted in those studies can be used to footprint proteins as well, as has been done in the past and continues today (consider DEPC, GEE, IAM, NHB, and TNM footprinting, for example).

Last is a technical comment to emphasize the need to conduct thorough evaluations when applying a labeling reagent to study a specific protein system. Technical issues that need to be faced include the compatibility of the labeling reagent and the protein, the optimum reagent dose, the labeling time to minimize double labeling. Optimization will achieve the proper normalization of labeling extents in differential experiments. Further, statistical treatments of the differences in labeling extents must be in place. These points should be considered as standard operating procedures, yet few studies to date report evaluations prior to presenting the results, perhaps not surprising given that the field is young. Specific labeling reagents themselves can perturb protein HOS, especially when large amounts of reagents are used to afford optimum protein footprinting. The installed labels can also compromise the structural integrity of the target protein. All in all, structural integrity of the target protein should be evaluated when initiating a new protein footprinting study, either by low-resolution approaches including ion mobility MS, CD, infrared spectroscopy, or by mid-to-high resolution approaches including HDX-MS and NMR. If the target protein is biologically active (e.g., an enzyme), a functional assay can be helpful. The use of small-molecule reporters or peptide dosimeters to enable normalization between different states of the protein is also critical. These efforts and results should be reported in future studies to convince the user community that the methods are robust.

7.3. Broad Perspective

From a broad view, each MS-based labeling approach will contribute to the understanding of protein HOS from a different perspective, motivating the thoughtful and creative choice of different reagents. As mentioned above, the methods of HDX, targeted labeling reagents, and radical species inform on protein HOS and their changes on different time scales. Their targeted amino acid residues on the protein are also different. Among all three, HDX is undoubtfully the most mature choice whereas radical labeling is the least developed. All three approaches are robust, as seen in the extensive demonstrations during the three decades. In terms of throughput, HDX is the highest as it requires minimal modifications of existing LC-MS platforms. Its data analysis is also the fastest owing to long-standing commercial efforts to provide HDX data analysis software, as introduced in section 2. The targeted amino acid labeling and labeling by radical species shares similar characteristics and issues; the data analysis is more complex as compared with HDX. The radical labeling often involves a special apparatus to generate radicals; therefore, it is relatively less advanced as compared with the other two. These platforms, both for software and hardware, are becoming available commercially (e.g., GenNext Technologies, Inc., LEAP HDX Platform). Spatial resolutions of these approaches are hard to compare. Irreversible labeling approaches often come with residue-level spatial resolution with commonly used fragmentation approaches (e.g., HCD or CID), whereas one can only achieve residue-level resolution with HDX by ETD/ECD fragmentation or extensive digestion (as discussed earlier). The spatial resolution of chemical cross-linking is considered to be lowest, as discussed in section 4.4, and computer modeling by utilizing footprinting data may offer highest structural resolution of a chemical approach, but a mature workflow remains to be established. Applying combinations of different approaches can provide a deeper understanding of the protein and a higher resolution structure. Recent demonstrations have shown the power of such integrated strategies.[743,773,841,842,1054,1223−1227] Broad application of integrated methods, however, remains to be explored.

Platform integration not only expands the protein footprinting approaches but also requires clever application of the various sample handling and MS techniques. For example, one can couple native MS and protein footprinting in a preliminary analysis to assist in the study design of protein ligand interactions by determining mixing ratios that push the binding equilibriums to the product complex and by exploring the binding regions. Ion mobility MS coupled with protein footprinting will illuminate protein conformational changes during and after labeling, providing a view of protein conformational changes induced by chemical labeling. Multiple ion dissociation techniques coupled with different enzymatic digestions will deliver better sequence coverage, even to the residue level, which is an important goal in MS-based protein HOS analysis.

Improved reversed-phase chromatography and adoption of other separation strategies (e.g., normal phase, capillary electrophoresis) will allow separation of isomeric peptides formed in labeling (e.g., chemically modified peptides that bear the same modification on multiple residues) and their quantification from extracted ion chromatograms and other approaches. Improved digestion strategies (e.g., mixed and immobilized enzymes) will also improve spatial resolution. Better separation will also minimize ion suppression during ionization, separate structural isomers that have the same m/z, and ultimately improve the precision of quantification and the spatial resolution of the protein HOS analysis. Enhanced separation capability, possibly even with 2D approaches, is particularly important for footprinting complex mixtures of proteins or large proteins that digest to provide many peptides. Modern protein and peptide separation approaches, including multidimensional LC,[1228,1229] size exclusion chromatography,[1230,1231] and ion-exchange chromatography,[1232] have all been coupled with MS analysis. Although most of the current

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

demonstrations are for primary-structure proteomics and native MS, they will surely be adopted for protein HOS analysis soon.

Upon obtaining the MS data, database searching and spectra identifications by software are also critical. Over the past decade, we witnessed a burst of MS software for MS/MS identification, de novo sequencing, HDX, irreversible labeling, chemical cross-linking, native MS, ion mobility MS, for proteomics, glycomics, lipidomics, and metabolomics. Although these developments largely contribute to data processing and visualization, they will continue to be important components in the "engine" of MS-based structural proteomics and ultimately lead to more automated data processing.

Lastly, to increase the spatial resolution of MS-based protein HOS studies, computer modeling is also critical. As discussed in section 6.6, pioneering efforts demonstrate a promising future in this area. Incorporating footprinting data into structural predictions in a straightforward way breaks the resolution limit of MS approaches, making MS-based protein HOS analysis a biophysical method that delivers high resolution results as does X-ray crystallography, Cryo-EM, and NMR or that, at least, complements them, even for protein complexes.[1188] A simple example is the use of footprinting to recover the information loss for flexible protein regions in an X-ray crystal structure. Thus, footprinting data can increase the confidence of protein structural predictions both on its own and in combination with other approaches.

To achieve these goals, deeper understanding of footprinting fundamentals is crucial because only from a solid understanding of fundamentals can we evaluate the data properly, especially for a single protein state rather than employing a differential experiment. Development of user-friendly software that can take footprinting data as input for structural modeling is also needed.

7.4. Concluding Remarks

The combination of advanced MS instrumentation, novel protein labeling workflows, wide ranging labeling reagents, careful experimental design, and precise data processing will continue to advance MS footprinting as a tool well beyond protein molecular weight and primary structure determination. MS as a robust and informative protein structural characterization tool has been established during the past 30 years, and we predict that MS-based approaches will continue to contribute significantly to the field of structural proteomics and the broader field of structural biology.

AUTHOR INFORMATION

Corresponding Authors

Xiaoran Roger Liu − *Department of Chemistry, Washington University in St. Louis, St. Louis, Missouri 63130, United States;* orcid.org/0000-0003-3198-1588; Email: xliu167@wustl.edu

Michael L. Gross − *Department of Chemistry, Washington University in St. Louis, St. Louis, Missouri 63130, United States;* orcid.org/0000-0003-1159-4636; Email: mgross@wustl.edu

Author

Mengru Mira Zhang − *Department of Chemistry, Washington University in St. Louis, St. Louis, Missouri 63130, United States;* orcid.org/0000-0001-9082-1272

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.chemrev.9b00815

Author Contributions

†X. R. Liu and M. M. Zhang contributed equally.

Notes

The authors declare no competing financial interest.

Biographies

Xiaoran Roger Liu received his B.E. in Applied Chemistry from East China University of Science and Technology in 2014. He then moved to The University of Akron. During his M.S. work with Dr. Toshikazu Miyoshi, he focused on structural characterization of carbon fiber precursors during thermal stabilization by solid-state NMR. He is currently working toward his Ph.D. in Washington University in St. Louis with Dr. Michael L. Gross. His research interest is to develop and apply novel mass spectrometry-based approaches in characterizing protein higher-order structures.

Mengru Mira Zhang received her B.E. degree in 2014 at East China University of Science and Technology. In 2013, she moved to Ohio and joined the Department of Polymer Science at The University of Akron under the supervision of Dr. Li Jia, focusing on organometallic synthesis and its catalytic application for novel polymers. After obtaining her M.S. degree in 2015, she decided to pursue her Ph.D. degree in Dr. Michael L. Gross' lab at Washington University in St. Louis, with interests in developing and utilizing multiple biophysical methods to address different biological questions.

Michael L. Gross is a chemist who has worked independently in mass spectrometry since 1968. He began his career at the University of Nebraska—Lincoln, where he was distinguished professor of chemistry and director of an NSF Center for Mass Spectrometry until 1994. He then moved to Washington University in St. Louis as professor of chemistry, immunology, medicine, and principal investigator of the WU NIH NIGMS Mass Spectrometry Research Resource. During his career of 50+ years, he has authored and coauthored over 650 publications and book chapters. His primary research interests have evolved to structural proteomics and MS-based biophysics.

ACKNOWLEDGMENTS

This work was supported by the National Institute of Health NIGMS grants 5P41GM103422, 1R01GM13100801, and R24GM135766 (to M.L.G.). We also like to thank Dr. Hailiang Jin and Dr. Jiawei Liu for their help in tuning the colors and enhancing the resolution for selected figures.

ABBREVIATIONS

$^1O_2$ = singlet oxygen
2D = two-dimensional
3D = three-dimensional
5-IAF = 5-idoacetamidofluorescein
6-Cl-IMP = 6-chloropurine riboside 5′-monophosphate
Aβ = amyloid beta
ADH = adipic acid dihydrazide
AFM = atomic force microscopy
Ala = alanine
aMb = apo-myoglobin
ANB-NOS = *N*-5-azido-2-nitrobenzoyloxysuccinimide
APEX = ascorbate peroxidase
Arg = arginine
Asn = asparagine
Asp = aspartic acid

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

αSyn = alpha-synuclein
azide-A-DSBSO = acid-cleavable disuccinimidyl bis-sulfoxide
β2m = β-2-microglobulin
BAMG = bis(succinimidyl)-3-azidomethylglutarate
BDC = *N*-benzyl-*N*′-3-(dimethylamino)propylcarbodiimide
BDP-NHP = *N*-hydroxyphthalamide ester of biotin aspartate proline
BHD = benzhydrazide
BME = b-mercaptoethanol
BS3 = Bis(sulfosuccinimidyl)suberate
C. elegans = *Caenorhabditis elegans*
CaM = calmodulin
CASP = critical assesment of protein structure presiction
CBDPS = cyanurbiotindimercaptopropionyl succinimide
CD = circular dichroism
CDR = complementarity determining region
chloramine T = *N*-chloro-4-methyl-benzenesulfonamide
CID = collision induced dissociation
CL-MS = covalent labeling MS
COSY = correlation spectroscopy
Cryo-EM = cryogenic electron microscopy
Cys = cysteine
DEPC = diethylpyrocarbonate
di-UB = diubiquitin
DMP = dimethyl pimelimidate
DMS = dimethyl suberimidate
DMTMM = 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride
DSBU = disuccinimidyl dibutyric urea
DSG = disuccinimidyl glutarate
DSS = disuccinimidyl suberate
DSSO = disuccinimidyl sulfoxide
DTBP = dimethyl 3,3′-dithiobispropionimidate
DTNB = 5,5′-dithiobis(2-nitrobenzoic acid)
DTSSP = 3,3′-dithiobis(sulfosuccinimidylpropionate)
DTT = Dithiothreitol
E. coli = *Escherichia coli*
e_{aq}^- = hydrated electrons
ECD = electron capture dissociation
EDC = 1-ethyl-3-(3-(dimethylamino)propyl)carbodiimide
EDD = electron detachment dissociation
EDTA = ethylenediaminetetraacetic acid
EGF = epidermal growth factor
EGFR = human epidermal growth factor receptor
eIF3 = yeast initiation factor 3
ELISA = enzyme linked immunosorbent assay
EPR = electron paramagnetic resonance
ESI = electrospray ionization
ETD = electron transfer dissociation
EThcD = electron transfer higher energy dissociation
FAB = fast atom bombardment
FDR = false discoverer rate
Fe-BABE = Fe-(*S*)-1-(*p*-bromoacetimidobenzyl)-EDTA
FMO = Fenna−Matthews−Olson
FPOP = fast photochemical oxidation of proteins
FRET = Fürster resonance energy transfer
FTICR = Fourier-transform ion cyclotron resonance
FT-IR = Fourier transform infrared spectroscopy
GA = glycinamide
GEE = glycine ethyl ester
Gln = glutamine
Glu = glutamic acid
GLUTs = human glucose transporters
Gly = glycine
γ-ray = gamma ray
H_2O^* = activated water
HBx = hepatitis B virus X protein
HDX = hydrogen−deuterium exchange
hERα = human estrogen receptor alpha
His = histidine
hMb = holo-myoglobin
HNB = 2-hydroxy-5-nitrobenzyl-bromide
HNSB = dimethyl(2-hydroxy-5-nitrobenzyl)sulfonium bromide
holo-CaM = calcium-bound calmodulin
HOS = higher-order structure
HRF = hydroxyl radical footprinting
HRP = horseradish peroxidase
hVKOR = human vitamin K epoxide reductase
IAM = iodoacetamide
ICAT = isotope-coded affinity tag
IgG1 = human Immunoglobulin G subclass 1 antibody
IL-6R = human interleukin-6 receptor α-chain
Ile = isoleucine
IMPDH = human type II inosine 5′-monophosphate dehydrogenase
iodogen = 1,3,4,6-tetra-chloro-3a,6a-diphenyl-glycouril
IRMPD = infrared multiphoton dissociation
ITC = isothermal titration calorimetry
kethoxal = 3-ethoxy-1,1-dihydroxy-2-butanone
LC = liquid chromatography
Leu = leucine
LITPOMS = ligand titration, fast photochemical oxidation of proteins and mass spectrometry
Lys = lysine
m/*z* = mass to charge ratio
mAb = monoclonal antibodies
MALDI = matrix-assisted laser desorption ionization
MCO = metal-catalyzed oxidation
Melarsen oxide = *p*-(4,6-diamino-1,3,5-triazin-2-yl)-aminophenylarsonous acid
MES = 2-(*N*-morpholino) ethanesulfonic acid
Met = methionine
MMTS = methylmethanethiosulfonate
MS = mass spectrometry
MS/MS = tandem MS
MURR1 = mouse U2af1-rs1 region1
N-AcO-AAF = *N*-acetoxy-*N*-acetyl-2-aminofluorene
NAI = *N*-acetylimidazole
NBS = *N*-bromosuccinimide
NCS = *N*-chlorosuccinimide
NEM = *N*-ethylmaleimide
NHS-ester = *N*-hydroxysuccinimide ester
NMR = nuclear magnetic resonance
NTCB = 2-nitro-5-thiocyanobenzoic acid
PAL = photoaffinity labeling
p-bromophenacyl bromide = 2-bromo-1-(4-bromophenyl)-ethenone
PD = plasma desorption
PDB = Protein Data Bank
PDH = pimelic acid dihydrazide
PF = protection factor
Phe = phenylalanine
PIR = protein interaction reporter
PLIMB = plasma induced modification of biomolecules

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

PLIMSTEX = protein−ligand interactions by mass spectrometry, titration, and H/D exchange
PLP = pyridoxal-5′-phosphate
POSE = protein oxidation of sialic acid environment
PPIs = protein−protein interactions
Pro = proline
PrP = prion protein
PSII = photosystem II
PTM = post translational modification
Q-ToF = quadropole time-of-flight
ReACT = real-time analysis for cross-linked peptides technology
RMSD = root-mean-square deviation
ROS = reactive oxygen species
SASA = solvent accessible surface area
Ser = serine
SID = surface induced dissociation
SOD = superoxide dimustase
SPR = surface plasmon resonance
SPROX = stability of proteins from rates of oxidation
sulfo-SDA = sulfosuccinimidyl 4,4′-azipentanoate
SUPREX = stability of unpurified proteins from rates of H/D exchange
TBHP = *t*-butylhydroperoxide
TECP = tris(2-carboxyethyl)phosphine
Thr = theronine
ThT = thioflaven T
TNBS = 2,4,6-trinitrobenzenesulfonic acid
TNM = tetranitromethane
ToF = time-of-flight
TPDs = 3-trifluoromethyl-3-phenyldiazirines
TPO = thyroid peroxidase
Trp = tryptophan
Tyr = tyrosine
UBA = ubiquitin associated domain of USP5
Ugi = PBS2 uracil-DNA glycosylase inhibitor
Ung = uracil-DNA glycosylase
UPS5 = ubiquitin specific protease 5
UV = ultraviolet
UVPD = ultraviolet photodissociation
Val = valine
WRK = Woodward's reagent K
XL-MS = cross-linking mass spectrometry
XO = xanthine oxidase
ZnF-UBP = Zn-finger ubiquitin-binding domain of USP5

REFERENCES

(1) Lodish, H.; Berk, A.; Kaiser, C. A.; Krieger, M.; Scott, M. P.; Bretscher, A.; Ploegh, H.; Matsudaira, P. *Mol. Cell. Biol.*, 6th ed; W.H. Freeman: New York, 2008.
(2) Anfinsen, C. B.; Haber, E.; Sela, M.; White, F. H. The Kinetics of Formation of Native Ribonuclease During Oxidation of the Reduced Polypeptide Chain. *Proc. Natl. Acad. Sci. U. S. A.* **1961**, *47*, 1309−1314.
(3) Fermi, G.; Perutz, M. F.; Shaanan, B.; Fourme, R. The Crystal Structure of Human Deoxyhaemoglobin at 1.74 Å Resolution. *J. Mol. Biol.* **1984**, *175*, 159−174.
(4) Mathews, C. K.; van Holde, K. E.; Appling, D. R.; Anthony-Cahill, S. J. *Biochemistry*, 4th ed; Pearson: New York, 2013.
(5) Fruton, J. S. Early Theories of Protein Structure. *Ann. N. Y. Acad. Sci.* **1979**, *325*, 1−20.
(6) Sanger, F.; Tuppy, H. The Amino-Acid Sequence in the Phenylalanyl Chain of Insulin. 1. The Identification of Lower Peptides from Partial Hydrolysates. *Biochem. J.* **1951**, *49*, 463−481.
(7) Sanger, F. The Arrangement of Amino Acids in Proteins. In *Advances in Protein Chemistry*; Anson, M. L., Bailey, K., Edsall, J. T., Eds.; Academic Press: New York, 1952; Vol. 7, pp 1−67, DOI: 10.1016/S0065-3233(08)60017-0.
(8) Baker, E. N.; Hubbard, R. E. Hydrogen Bonding in Globular Proteins. *Prog. Biophys. Mol. Biol.* **1984**, *44*, 97−179.
(9) Fersht, A. R.; Shi, J.-P.; Knill-Jones, J.; Lowe, D. M.; Wilkinson, A. J.; Blow, D. M.; Brick, P.; Carter, P.; Waye, M. M. Y.; Winter, G. Hydrogen Bonding and Biological Specificity Analysed by Protein Engineering. *Nature* **1985**, *314*, 235−238.
(10) Waldburger, C. D.; Schildbach, J. F.; Sauer, R. T. Are Buried Salt Bridges Important for Protein Stability and Conformational Specificity? *Nat. Struct. Mol. Biol.* **1995**, *2*, 122−128.
(11) Kumar, S.; Nussinov, R. Salt Bridge Stability in Monomeric Proteins. *J. Mol. Biol.* **1999**, *293*, 1241−1255.
(12) Privalov, P. L.; Gill, S. J. Stability of Protein Structure and Hydrophobic Interaction. In *Advances in Protein Chemistry*; Anfinsen, C. B., Edsall, J. T., Richards, F. M., Eisenberg, D. S., Eds.; Academic Press: New York, 1988; Vol. 39, pp 191−234, DOI: 10.1016/S0065-3233(08)60377-0.
(13) Baldwin, R. L. Temperature Dependence of the Hydrophobic Interaction in Protein Folding. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 8069−8072.
(14) Burley, S.; Petsko, G. Aromatic-aromatic Interaction: A Mechanism of Protein Structure Stabilization. *Science* **1985**, *229*, 23−28.
(15) Gallivan, J. P.; Dougherty, D. A. Cation-π Interactions in Structural Biology. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 9459−9464.
(16) Roth, C. M.; Neal, B. L.; Lenhoff, A. M. Van der Waals Interactions Involving Proteins. *Biophys. J.* **1996**, *70*, 977−987.
(17) Betz, S. F. Disulfide Bonds and the Stability of Globular Proteins. *Protein Sci.* **1993**, *2*, 1551−1558.
(18) Wedemeyer, W. J.; Welker, E.; Narayan, M.; Scheraga, H. A. Disulfide Bonds and Protein Folding. *Biochemistry* **2000**, *39*, 4207−4216.
(19) Doig, A. J.; Sternberg, M. J. E. Side-chain Conformational Entropy in Protein Folding. *Protein Sci.* **1995**, *4*, 2247−2251.
(20) Frederick, K. K.; Marlow, M. S.; Valentine, K. G.; Wand, A. J. Conformational entropy in Molecular Recognition by Proteins. *Nature* **2007**, *448*, 325−329.
(21) Tzeng, S.-R.; Kalodimos, C. G. Protein Activity Regulation by Conformational Entropy. *Nature* **2012**, *488*, 236−240.
(22) Kendrew, J. C.; Bodo, G.; Dintzis, H. M.; Parrish, R. G.; Wyckoff, H.; Phillips, D. C. A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature* **1958**, *181*, 662−666.
(23) Perutz, M. F.; Rossmann, M. G.; Cullis, A. F.; Muirhead, H.; Will, G.; North, A. C. T. Structure of Hæmoglobin: A Three-Dimensional Fourier Synthesis at 5.5-Å. Resolution, Obtained by X-Ray Analysis. *Nature* **1960**, *185*, 416−422.
(24) Kuboniwa, H.; Tjandra, N.; Grzesiek, S.; Ren, H.; Klee, C. B.; Bax, A. Solution Structure of Calcium-free Calmodulin. *Nat. Struct. Mol. Biol.* **1995**, *2*, 768−776.
(25) Drenth, J. *Principles of Protein X-ray Crystallography*, 3rd ed.; Springer Science & Business Media: New York, 2007.
(26) McPherson, A. Introduction to Protein Crystallization. *Methods* **2004**, *34*, 254−265.
(27) McPherson, A.; Gavira, J. A. Introduction to Protein Crystallization. *Acta Crystallogr., Sect. F: Struct. Biol. Commun.* **2014**, *70*, 2−20.
(28) Billeter, M. Comparison of Protein Structures Determined by NMR in Solution and by X-ray Diffraction in Single Crystals. *Q. Rev. Biophys.* **1992**, *25*, 325−377.
(29) Garbuzynskiy, S. O.; Melnik, B. S.; Lobanov, M. Y.; Finkelstein, A. V.; Galzitskaya, O. V. Comparison of X-ray and NMR Structures: Is there a Systematic Difference in Residue Contacts Between X-ray- and NMR-Resolved Protein Structures? *Proteins: Struct., Funct., Genet.* **2005**, *60*, 139−147.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(30) Wüthrich, K. *NMR in Biological Research: Peptides and Proteins*; North-Holland Publishing Company: Amsterdam, 1976.
(31) Wüthrich, K. Protein Structure Determination in Solution by Nuclear Magnetic Resonance Spectroscopy. *Science* **1989**, *243*, 45−50.
(32) Wüthrich, K. The Way to NMR Structures of Proteins. *Nat. Struct. Biol.* **2001**, *8*, 923−925.
(33) Bax, A.; Grzesiek, S. Methodological Advances in Protein NMR. *Acc. Chem. Res.* **1993**, *26*, 131−138.
(34) Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G., III; Rance, M.; Skelton, N. J. *Protein NMR Spectroscopy: Principles and Practice*, 2nd ed.; Elsevier Academic Press: San Diego, 1995.
(35) Williamson, M. P.; Havel, T. F.; Wüthrich, K. Solution Conformation of Proteinase Inhibitor IIA from Bull Seminal Plasma by 1H Nuclear Magnetic Resonance and Distance Geometry. *J. Mol. Biol.* **1985**, *182*, 295−315.
(36) Wüthrich, K. NMR with Proteins and Nucleic Acids. *Europhys. News* **1986**, *17*, 11−13.
(37) Mainz, A.; Religa, T. L.; Sprangers, R.; Linser, R.; Kay, L. E.; Reif, B. NMR Spectroscopy of Soluble Protein Complexes at One Mega-Dalton and Beyond. *Angew. Chem., Int. Ed.* **2013**, *52*, 8746−8751.
(38) Ishima, R.; Torchia, D. A. Protein Dynamics from NMR. *Nat. Struct. Biol.* **2000**, *7*, 740−743.
(39) Andronesi, O. C.; Becker, S.; Seidel, K.; Heise, H.; Young, H. S.; Baldus, M. Determination of Membrane Protein Structure and Dynamics by Magic-Angle-Spinning Solid-State NMR Spectroscopy. *J. Am. Chem. Soc.* **2005**, *127*, 12965−12974.
(40) Tuttle, M. D.; Comellas, G.; Nieuwkoop, A. J.; Covell, D. J.; Berthold, D. A.; Kloepper, K. D.; Courtney, J. M.; Kim, J. K.; Barclay, A. M.; Kendall, A.; et al. Solid-state NMR Structure of a Pathogenic Fibril of Full-length Human α-synuclein. *Nat. Struct. Mol. Biol.* **2016**, *23*, 409−415.
(41) Castellani, F.; van Rossum, B.; Diehl, A.; Schubert, M.; Rehbein, K.; Oschkinat, H. Structure of a Protein Determined by Solid-state Magic-angle-spinning NMR Spectroscopy. *Nature* **2002**, *420*, 99−102.
(42) Baldus, M. Solid-State NMR Spectroscopy: Molecular Structure and Organization at the Atomic Level. *Angew. Chem., Int. Ed.* **2006**, *45*, 1186−1188.
(43) Kline, A. D.; Braun, W.; Wüthrich, K. Determination of the Complete Three-dimensional Structure of the α-amylase Inhibitor Tendamistat in Aqueous Solution by Nuclear Magnetic Resonance and Distance Geometry. *J. Mol. Biol.* **1988**, *204*, 675−724.
(44) Yee, A. A.; Savchenko, A.; Ignachenko, A.; Lukin, J.; Xu, X.; Skarina, T.; Evdokimova, E.; Liu, C. S.; Semesi, A.; Guido, V.; Edwards, A. M.; Arrowsmith, C. H. NMR and X-ray Crystallography, Complementary Tools in Structural Proteomics of Small Proteins. *J. Am. Chem. Soc.* **2005**, *127*, 16512−16517.
(45) *Protein DataBank Data Distribution by Experimental Method and Molecular Type*; RCSB, 2020; https://www.rcsb.org/stats/summary (accessed 2020-03-20).
(46) Bai, X.-c.; McMullan, G.; Scheres, S. H. W. How Cryo-EM is Revolutionizing Structural Biology. *Trends Biochem. Sci.* **2015**, *40*, 49−57.
(47) Cheng, Y. Single-Particle Cryo-EM at Crystallographic Resolution. *Cell* **2015**, *161*, 450−457.
(48) Dubochet, J.; McDowall, A. W. Vitrification of Pure Water for Electron Microscopy. *J. Microsc.* **1981**, *124*, 3−4.
(49) Knapek, E. Properties of Organic Specimens and Their Supports at 4 K under Irradiation in an Electron Microscope. *Ultramicroscopy* **1982**, *10*, 71−86.
(50) Frank, J. Averaging of Low Exposure Electron Micrographs of Non-periodic Objects. *Ultramicroscopy* **1975**, *1*, 159−162.
(51) Henderson, R.; Baldwin, J. M.; Ceska, T. A.; Zemlin, F.; Beckmann, E.; Downing, K. H. Model for the Structure of Bacteriorhodopsin based on High-Resolution Electron Cryo-Microscopy. *J. Mol. Biol.* **1990**, *213*, 899−929.
(52) Yu, X.; Jin, L.; Zhou, Z. H. 3.88 Å Structure of Cytoplasmic Polyhedrosis Virus by Cryo-electron Microscopy. *Nature* **2008**, *453*, 415−419.
(53) Zhang, X.; Jin, L.; Fang, Q.; Hui, W. H.; Zhou, Z. H. 3.3 Å Cryo-EM Structure of a Nonenveloped Virus Reveals a Priming Mechanism for Cell Entry. *Cell* **2010**, *141*, 472−482.
(54) Merk, A.; Bartesaghi, A.; Banerjee, S.; Falconieri, V.; Rao, P.; Davis, M. I.; Pragani, R.; Boxer, M. B.; Earl, L. A.; Milne, J. L. S.; et al. Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery. *Cell* **2016**, *165*, 1698−1707.
(55) Matadeen, R.; Patwardhan, A.; Gowen, B.; Orlova, E. V.; Pape, T.; Cuff, M.; Mueller, F.; Brimacombe, R.; van Heel, M. The Escherichia coli Large Ribosomal Subunit at 7.5 Å Resolution. *Structure* **1999**, *7*, 1575−1583.
(56) Jomaa, A.; Boehringer, D.; Leibundgut, M.; Ban, N. Structures of the E. coli Translating Ribosome with SRP and its Receptor and with the Translocon. *Nat. Commun.* **2016**, *7*, 10471.
(57) Du, Y.; Duc, N. M.; Rasmussen, S. G. F.; Hilger, D.; Kubiak, X.; Wang, L.; Bohon, J.; Kim, H. R.; Wegrecki, M.; Asuru, A.; et al. Assembly of a GPCR-G Protein Complex. *Cell* **2019**, *177*, 1232−1242 e11..
(58) Thompson, R. F.; Walker, M.; Siebert, C. A.; Muench, S. P.; Ranson, N. A. An Introduction to Sample Preparation and Imaging by Cryo-electron Microscopy for Structural Biology. *Methods* **2016**, *100*, 3−15.
(59) Provencher, S. W.; Gloeckner, J. Estimation of Globular Protein Secondary Structure from Circular Dichroism. *Biochemistry* **1981**, *20*, 33−37.
(60) Greenfield, N. J. Using Circular Dichroism Spectra to Estimate Protein Secondary Structure. *Nat. Protoc.* **2006**, *1*, 2876−2890.
(61) Oladepo, S. A.; Xiong, K.; Hong, Z.; Asher, S. A.; Handen, J.; Lednev, I. K. UV Resonance Raman Investigations of Peptide and Protein Structure and Dynamics. *Chem. Rev.* **2012**, *112*, 2604−2628.
(62) Jackson, M.; Mantsch, H. H. The Use and Misuse of FTIR Spectroscopy in the Determination of Protein Structure. *Crit. Rev. Biochem. Mol. Biol.* **1995**, *30*, 95−120.
(63) Kong, J.; Yu, S. Fourier Transform Infrared Spectroscopic Analysis of Protein Secondary Structures. *Acta Biochim. Biophys. Sin.* **2007**, *39*, 549−559.
(64) Noble, J. E.; Bailey, M. J. A. Quantitation of Protein. In *Methods in Enzymology, Guide to Protein Purification*, 2nd ed; Burgess, R. R., Deutscher, M. P., Eds.; Elsevier Academic Press: San Diego, 2009; Vol. *463*, pp 73−95, DOI: 10.1016/S0076-6879(09)63008-1.
(65) Eftink, M. R. Fluorescence Techniques for Studying Protein Structure. In *Methods of Biochemical Analysis*; Suelter, C. H., Ed.; John Wiley & Sons: Hoboken, NJ, 1991; Vol. *35*, pp 127−205, DOI: 10.1002/9780470110560.ch3.
(66) Royer, C. A. Probing Protein Folding and Conformational Transitions with Fluorescence. *Chem. Rev.* **2006**, *106*, 1769−1784.
(67) Piston, D. W.; Kremers, G.-J. Fluorescent protein FRET: the good, the bad and the ugly. *Trends Biochem. Sci.* **2007**, *32*, 407−414.
(68) Jachimska, B.; Wasilewska, M.; Adamczyk, Z. Characterization of Globular Protein Solutions by Dynamic Light Scattering, Electrophoretic Mobility, and Viscosity Measurements. *Langmuir* **2008**, *24*, 6866−6872.
(69) Stetefeld, J.; McKenna, S. A.; Patel, T. R. Dynamic Light Scattering: A Practical Guide and Applications in Biomedical Sciences. *Biophys. Rev.* **2016**, *8*, 409−427.
(70) Wen, J.; Arakawa, T.; Philo, J. S. Size-Exclusion Chromatography with On-Line Light-Scattering, Absorbance, and Refractive Index Detectors for Studying Proteins and Their Interactions. *Anal. Biochem.* **1996**, *240*, 155−166.
(71) Krause, F. Das magnetische Elektronenmikroskop und seine Anwendung in der Biologie. *Naturwissenschaften* **1937**, *25*, 817−825.
(72) Brenner, S.; Horne, R. W. A Negative Staining Method for High Resolution Electron Microscopy of Viruses. *Biochim. Biophys. Acta* **1959**, *34*, 103−110.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(73) Müller, D. J.; Engel, A. Atomic Force Microscopy and Spectroscopy of Native Membrane Proteins. *Nat. Protoc.* **2007**, *2*, 2191−2197.
(74) Wiseman, T.; Williston, S.; Brandts, J. F.; Lin, L.-N. Rapid Measurement of Binding Constants and Heats of Binding Using a New Titration Calorimeter. *Anal. Biochem.* **1989**, *179*, 131−137.
(75) Leavitt, S.; Freire, E. Direct Measurement of Protein Binding Energetics by Isothermal Titration Calorimetry. *Curr. Opin. Struct. Biol.* **2001**, *11*, 560−566.
(76) Johnsson, B.; LöfÅs, S.; Lindquist, G. Immobilization of Proteins to a Carboxymethyldextran-modified Gold Surface for Biospecific Interaction Analysis in Surface Plasmon Resonance Sensors. *Anal. Biochem.* **1991**, *198*, 268−277.
(77) Smith, E. A.; Thomas, W. D.; Kiessling, L. L.; Corn, R. M. Surface Plasmon Resonance Imaging Studies of Protein-Carbohydrate Interactions. *J. Am. Chem. Soc.* **2003**, *125*, 6140−6148.
(78) Baker, D.; Sali, A. Protein Structure Prediction and Structural Genomics. *Science* **2001**, *294*, 93−96.
(79) Zhang, Y. Progress and Challenges in Protein Structure Prediction. *Curr. Opin. Struct. Biol.* **2008**, *18*, 342−348.
(80) Kihara, D. *Protein Structure Prediction*, 3rd ed.; Humana Press: Totowa, NJ, 2014.
(81) Guzzo, A. V. The Influence of Amino Acid Sequence on Protein Structure. *Biophys. J.* **1965**, *5*, 809−822.
(82) Prothero, J. W. Correlation Between the Distribution of Amino Acids and Alpha Helices. *Biophys. J.* **1966**, *6*, 367−370.
(83) Chou, P. Y.; Fasman, G. D. Prediction of Protein Conformation. *Biochemistry* **1974**, *13*, 222−245.
(84) Ginalski, K. Comparative Modeling for Protein Structure Prediction. *Curr. Opin. Struct. Biol.* **2006**, *16*, 172−177.
(85) Xiang, Z. Advances in Homology Protein Structure Modeling. *Curr. Protein Pept. Sci.* **2006**, *7*, 217−227.
(86) Bowie, J. U.; Eisenberg, D. An Evolutionary Approach to Folding Small Alpha-helical Proteins that Uses Sequence Information and an Empirical Guiding Fitness Function. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 4436−4440.
(87) Lee, J.; Freddolino, P. L.; Zhang, Y. Ab Initio Protein Structure Prediction. In *From Protein Structure to Function with Bioinformatics*; Rigden, J. D., Ed.; Springer Netherlands: Dordrecht, 2017; pp 3−35, DOI: 10.1007/978-94-024-1069-3_1.
(88) Bradley, P.; Misura, K. M. S.; Baker, D. Toward High-Resolution de Novo Structure Prediction for Small Proteins. *Science* **2005**, *309*, 1868−1871.
(89) Marks, D. S.; Hopf, T. A.; Sander, C. Protein Structure Prediction from Sequence Variation. *Nat. Biotechnol.* **2012**, *30*, 1072−1080.
(90) Moult, J.; Fidelis, K.; Kryshtafovych, A.; Schwede, T.; Tramontano, A. Critical Assessment of Methods of Protein Structure Prediction (CASP)—Round XII. *Proteins: Struct., Funct., Genet.* **2018**, *86*, 7−15.
(91) Munson, M. S. B.; Field, F. H. Chemical Ionization Mass Spectrometry. I. General Introduction. *J. Am. Chem. Soc.* **1966**, *88*, 2621−2630.
(92) Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. Molecular Beams of Macroions. *J. Chem. Phys.* **1968**, *49*, 2240−2249.
(93) Winkler, H. U.; Beckey, H. D. Field Desorption Mass Spectrometry of Amino Acids. *Org. Mass Spectrom.* **1972**, *6*, 655−660.
(94) Mumma, R. O.; Vastola, F. J. Analysis of Organic Salts by Laser Ionization Mass Spectrometry. Sulfonates, Sulfates and Thiosulfates. *Org. Mass Spectrom.* **1972**, *6*, 1373−1376.
(95) Posthumus, M. A.; Kistemaker, P. G.; Meuzelaar, H. L. C.; Ten Noever de Brauw, M. C. Laser Desorption-mass Spectrometry of Polar Nonvolatile Bio-organic Molecules. *Anal. Chem.* **1978**, *50*, 985−991.
(96) Torgerson, D. F.; Skowronski, R. P.; Macfarlane, R. D. New Approach to the Mass Spectroscopy of Non-Volatile Compounds. *Biochem. Biophys. Res. Commun.* **1974**, *60*, 616−621.
(97) Benninghoven, A.; Jaspers, D.; Sichtermann, W. Secondary-ion Emission of Amino Acids. *Appl. Phys.* **1976**, *11*, 35−39.
(98) Barber, M.; Bordoli, R. S.; Sedgwick, R. D.; Tyler, A. N. Fast Atom Bombardment of Solids (F.A.B.): a New Ion Source for Mass Spectrometry. *J. Chem. Soc., Chem. Commun.* **1981**, 325−327.
(99) Barber, M.; Bordoli, R. S.; Elliott, G. J.; Sedgwick, R. D.; Tyler, A. N. Fast Atom Bombardment Mass Spectrometry. *Anal. Chem.* **1982**, *54*, 645−657.
(100) Karas, M.; Bachmann, D.; Hillenkamp, F. Influence of the Wavelength in High-Irradiance Ultraviolet Laser Desorption Mass Spectrometry of Organic Molecules. *Anal. Chem.* **1985**, *57*, 2935−2939.
(101) Karas, M.; Bahr, U.; Gießmann, U. Matrix-Assisted Laser Desorption Ionization Mass Spectrometry. *Mass Spectrom. Rev.* **1991**, *10*, 335−357.
(102) Sundqvist, B.; Kamensky, I.; HÅkansson, P.; Kjellberg, J.; Salehpour, M.; Widdiyasekera, S.; Fohlman, J.; Peterson, P. A.; Roepstorff, P. Californium-252 Plasma Desorption Time of Flight Mass Spectroscopy of Proteins. *Biomed. Mass Spectrom.* **1984**, *11*, 242−257.
(103) Sundqvist, B.; Roepstorff, P.; Fohlman, J.; Hedin, A.; Hakansson, P.; Kamensky, I.; Lindberg, M.; Salehpour, M.; Sawe, G. Molecular Weight Determinations of Proteins by Californium Plasma Desorption Mass Spectrometry. *Science* **1984**, *226*, 696−698.
(104) Barber, M.; Green, B. N.; Jennings, K. R. The Analysis of Small Proteins in the Molecular Weight Range 10−24 kDa by Magnetic Sector Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1987**, *1*, 80−83.
(105) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T. Protein and Polymer Analyses up to m/z 100 000 by Laser Ionization Time-of-flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1988**, *2*, 151−153.
(106) Karas, M.; Hillenkamp, F. Laser Desorption Ionization of Proteins with Molecular Masses Exceeding 10,000 Da. *Anal. Chem.* **1988**, *60*, 2299−2301.
(107) Fenn, J.; Mann, M.; Meng, C.; Wong, S.; Whitehouse, C. Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science* **1989**, *246*, 64−71.
(108) Whitehouse, C. M.; Dreyer, R. N.; Yamashita, M.; Fenn, J. B. Electrospray Interface for Liquid Chromatographs and Mass Spectrometers. *Anal. Chem.* **1985**, *57*, 675−679.
(109) Wilm, M. S.; Mann, M. Electrospray and Taylor-Cone theory, Dole's Beam of Macromolecules at Last? *Int. J. Mass Spectrom. Ion Processes* **1994**, *136*, 167−180.
(110) Wilm, M.; Mann, M. Analytical Properties of the Nano-electrospray Ion Source. *Anal. Chem.* **1996**, *68*, 1−8.
(111) Paul, W.; Steinwedel, H. Ein neues massenspektrometer ohne magnetfeld. *Z. Naturforsch., A: Phys. Sci.* **1953**, *8*, 448−450.
(112) Johnson, E. G.; Nier, A. O. Angular Aberrations in Sector Shaped Electromagnetic Lenses for Focusing Beams of Charged Particles. *Phys. Rev.* **1953**, *91*, 10−17.
(113) Gross, M. L. Tandem Mass Spectrometry: Multisector Magnetic Instruments. In *Methods in Enzymology*; Elsevier Academic Press: San Diego, 1990; Vol. *193*, pp 131−153, DOI: 10.1016/0076-6879(90)93414-G.
(114) Gross, M. L.; Chess, E. K.; Lyon, P. A.; Crow, F. W.; Evans, S.; Tudge, H. Triple Analyzer Mass Spectrometry for High Resolution MS/MS Studies. *Int. J. Mass Spectrom. Ion Phys.* **1982**, *42*, 243−254.
(115) McLafferty, F. W.; Todd, P. J.; McGilvery, D. C.; Baldwin, M. A. Collisional Activation and Metastable Ion Characteristics. 73. High-resolution Tandem Mass Spectrometer (MS/MS) of Increased Sensitivity and Mass Range. *J. Am. Chem. Soc.* **1980**, *102*, 3360−3363.
(116) Bateman, R. H.; Brown, J.; Lefevere, M.; Flammang, R.; Van Haverbeke, Y. Applications in Gaseous Ion and Neutral Chemistry using a Six-Sector Mass Spectrometer. *Int. J. Mass Spectrom. Ion Processes* **1992**, *115*, 205−218.
(117) Wiley, W. C.; McLaren, I. H. Time-of-Flight Mass Spectrometer with Improved Resolution. *Rev. Sci. Instrum.* **1955**, *26*, 1150−1157.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(118) Mamyrin, B.; Karataev, V.; Shmikk, D.; Zagulin, V. The Mass-reflectron, a New Nonmagnetic Time-of-flight Mass Spectrometer with High Resolution. *Zh. Eksp. Teor. Fiz.* **1973**, *64*, 82−89.
(119) Guilhaus, M.; Selby, D.; Mlynski, V. Orthogonal Acceleration Time-of-flight Mass Spectrometry. *Mass Spectrom. Rev.* **2000**, *19*, 65−107.
(120) Wolfgang, P.; Helmut, S.; Apparatus for Separating Charged Particles of Different Specific Charges. US US2939952A, 1960.
(121) Comisarow, M. B.; Marshall, A. G. Fourier Transform Ion Cyclotron Resonance Spectroscopy. *Chem. Phys. Lett.* **1974**, *25*, 282−283.
(122) Makarov, A. Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Anal. Chem.* **2000**, *72*, 1156−1162.
(123) Zubarev, R. A.; Makarov, A. Orbitrap Mass Spectrometry. *Anal. Chem.* **2013**, *85*, 5288−5296.
(124) Morris, H. R.; Paxton, T.; Dell, A.; Langhorne, J.; Berg, M.; Bordoli, R. S.; Hoyes, J.; Bateman, R. H. High Sensitivity Collisionally-activated Decomposition Tandem Mass Spectrometry on a Novel Quadrupole/Orthogonal-acceleration Time-of-flight Mass Spectrometer. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 889−896.
(125) Hu, Q.; Noll, R. J.; Li, H.; Makarov, A.; Hardman, M.; Cooks, R. G. The Orbitrap: a New Mass Spectrometer. *J. Mass Spectrom.* **2005**, *40*, 430−443.
(126) Sleno, L.; Volmer, D. A. Ion Activation Methods for Tandem Mass Spectrometry. *J. Mass Spectrom.* **2004**, *39*, 1091−1112.
(127) Jennings, K. R. Collision-induced Decompositions of Aromatic Molecular Ions. *Int. J. Mass Spectrom. Ion Phys.* **1968**, *1*, 227−235.
(128) Mitchell Wells, J.; McLuckey, S. A. Collision-Induced Dissociation (CID) of Peptides and Proteins. In *Methods in Enzymology*; Elsevier Academic Press: San Diego, 2005; Vol. *402*, pp148−185, DOI: 10.1016/S0076-6879(05)02005-7.
(129) Gross, M. L.; McCrery, D.; Crow, F.; Tomer, K. B.; Pope, M. R.; Ciuffetti, L. M.; Knoche, H. W.; Daly, J. M.; Dunkle, L. D. The Structure of the Toxin from Helminthosporium Carbonum. *Tetrahedron Lett.* **1982**, *23*, 5381−5384.
(130) Hunt, D. F.; Yates, J. R.; Shabanowitz, J.; Winston, S.; Hauer, C. R. Protein Sequencing by Tandem Mass Spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 6233−6237.
(131) Biemann, K.; Papayannopoulos, I. A. Amino Acid Sequencing of Proteins. *Acc. Chem. Res.* **1994**, *27*, 370−378.
(132) Mabud, M. A.; Dekrey, M. J.; Cooks, R. G. Surface-induced Dissociation of Molecular Ions. *Int. J. Mass Spectrom. Ion Processes* **1985**, *67*, 285−294.
(133) Chorush, R. A.; Little, D. P.; Beu, S. C.; Wood, T. D.; McLafferty, F. W. Surface-Induced Dissociation of Multiply-Protonated Proteins. *Anal. Chem.* **1995**, *67*, 1042−1046.
(134) Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. Peptide and Protein Sequence Analysis by Electron Transfer Dissociation Mass Spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 9528−9533.
(135) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. Electron Capture Dissociation of Multiply Charged Protein Cations. A Nonergodic Process. *J. Am. Chem. Soc.* **1998**, *120*, 3265−3266.
(136) Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. Electron Detachment Dissociation of Peptide Di-anions: an Electron−hole Recombination Phenomenon. *Chem. Phys. Lett.* **2001**, *342*, 299−302.
(137) Bowers, W. D.; Delbert, S. S.; Hunter, R. L.; McIver, R. T. Fragmentation of Oligopeptide Ions using Ultraviolet Laser Radiation and Fourier Transform Mass Spectrometry. *J. Am. Chem. Soc.* **1984**, *106*, 7288−7289.
(138) Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S. Complete Protein Characterization Using Top-Down Mass Spectrometry and Ultraviolet Photodissociation. *J. Am. Chem. Soc.* **2013**, *135*, 12646−12651.
(139) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. Infrared Multiphoton Dissociation of Large Multiply Charged Ions for Biomolecule Sequencing. *Anal. Chem.* **1994**, *66*, 2809−2815.
(140) Brodbelt, J. S. Ion Activation Methods for Peptides and Proteins. *Anal. Chem.* **2016**, *88*, 30−51.
(141) Sheshberadaran, H.; Payne, L. G. Protein Antigen-Monoclonal Antibody Contact Sites Investigated by Limited Proteolysis of Monoclonal Antibody-Bound Antigen: Protein ″footprinting. *Proc. Natl. Acad. Sci. U. S. A.* **1988**, *85*, 1−5.
(142) Konermann, L.; Rodriguez, A. D.; Sowole, M. A. Type 1 and Type 2 Scenarios in Hydrogen Exchange Mass Spectrometry Studies on Protein−Ligand Complexes. *Analyst* **2014**, *139*, 6078−6087.
(143) Limpikirati, P.; Pan, X.; Vachet, R. W. Covalent Labeling with Diethylpyrocarbonate: Sensitive to the Residue Microenvironment, Providing Improved Analysis of Protein Higher Order Structure by Mass Spectrometry. *Anal. Chem.* **2019**, *91*, 8516−8523.
(144) Ziemianowicz, D. S.; MacCallum, J. L.; Schriemer, D. C. Correlation between Labeling Yield and Surface Accessibility in Covalent Labeling Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2020**, *31*, 207−216.
(145) Liu, X. R.; Zhang, M. M.; Zhang, B.; Rempel, D. L.; Gross, M. L. Hydroxyl-Radical Reaction Pathways for the Fast Photochemical Oxidation of Proteins Platform As Revealed by 18O Isotopic Labeling. *Anal. Chem.* **2019**, *91*, 9238−9245.
(146) Domon, B.; Aebersold, R. Mass Spectrometry and Protein Analysis. *Science* **2006**, *312*, 212−217.
(147) Dančík, V.; Addona, T. A.; Clauser, K. R.; Vath, J. E.; Pevzner, P. A. De Novo Peptide Sequencing via Tandem Mass Spectrometry. *J. Comput. Biol.* **1999**, *6*, 327−342.
(148) Edman, P. Method for Determination of the Amino Acid Sequence in Peptides. *Acta Chem. Scand.* **1950**, *4*, 283−293.
(149) Mann, M.; Jensen, O. N. Proteomic Analysis of Post-Translational Modifications. *Nat. Biotechnol.* **2003**, *21*, 255−261.
(150) McLachlin, D. T.; Chait, B. T. Analysis of phosphorylated proteins and peptides by Mass Spectrometry. *Curr. Opin. Chem. Biol.* **2001**, *5*, 591−602.
(151) Nørregaard Jensen, O. Modification-specific Proteomics: Characterization of Post-Translational Modifications by Mass Spectrometry. *Curr. Opin. Chem. Biol.* **2004**, *8*, 33−41.
(152) Witze, E. S.; Old, W. M.; Resing, K. A.; Ahn, N. G. Mapping Protein Post-Translational Modifications with Mass Spectrometry. *Nat. Methods* **2007**, *4*, 798−806.
(153) Hao, P.; Adav, S. S.; Gallart-Palau, X.; Sze, S. K. Recent Advances in Mass Spectrometric Analysis of Protein Deamidation. *Mass Spectrom. Rev.* **2017**, *36*, 677−692.
(154) Ly, T.; Julian, R. R. Using ESI-MS to Probe Protein Structure by Site-Specific Noncovalent Attachment of 18-crown-6. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1209−1215.
(155) Loo, J. A. Studying Noncovalent Protein Complexes by Electrospray Ionization Mass Spectrometry. *Mass Spectrom. Rev.* **1997**, *16*, 1−23.
(156) Hernández, H.; Robinson, C. V. Determining the Stoichiometry and Interactions of Macromolecular Assemblies from Mass Spectrometry. *Nat. Protoc.* **2007**, *2*, 715−726.
(157) Benesch, J. L. P.; Ruotolo, B. T.; Simmons, D. A.; Robinson, C. V. Protein Complexes in the Gas Phase: Technology for Structural Genomics and Proteomics. *Chem. Rev.* **2007**, *107*, 3544−3567.
(158) Laganowsky, A.; Reading, E.; Hopper, J. T. S.; Robinson, C. V. Mass Spectrometry of Intact Membrane Protein Complexes. *Nat. Protoc.* **2013**, *8*, 639−651.
(159) Boeri Erba, E.; Petosa, C. The Emerging Role of Native Mass Spectrometry in Characterizing the Structure and Dynamics of Macromolecular Complexes. *Protein Sci.* **2015**, *24*, 1176−1192.
(160) Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S.-J.; Robinson, C. V. Ion Mobility−Mass Spectrometry Analysis of Large Protein Complexes. *Nat. Protoc.* **2008**, *3*, 1139−1152.
(161) Uetrecht, C.; Rose, R. J.; van Duijn, E.; Lorenzen, K.; Heck, A. J. R. Ion Mobility Mass Spectrometry of Proteins and Protein Assemblies. *Chem. Soc. Rev.* **2010**, *39*, 1633−1655.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(162) Zhong, Y.; Hyung, S.-J.; Ruotolo, B. T. Ion Mobility−Mass Spectrometry for Structural Proteomics. *Expert Rev. Proteomics* **2012**, *9*, 47−58.
(163) Sinz, A. Chemical Cross-Linking and Mass Spectrometry to Map Three-Dimensional Protein Structures and Protein−Protein Interactions. *Mass Spectrom. Rev.* **2006**, *25*, 663−682.
(164) Leitner, A. Cross-linking and Other Structural Proteomics Techniques: How Chemistry is Enabling Mass Spectrometry Applications in Structural Biology. *Chem. Sci.* **2016**, *7*, 4792−4803.
(165) Leitner, A.; Walzthoeni, T.; Kahraman, A.; Herzog, F.; Rinner, O.; Beck, M.; Aebersold, R. Probing Native Protein Structures by Chemical Cross-linking, Mass Spectrometry, and Bioinformatics. *Mol. Cell. Proteomics* **2010**, *9*, 1634−1649.
(166) Galas, D. J.; Schmitz, A. DNAase Footprinting a Simple Method for the Detection of Protein-DNA Binding Specificity. *Nucleic Acids Res.* **1978**, *5*, 3157−3170.
(167) Fontana, A.; de Laureto, P. P.; Spolaore, B.; Frare, E.; Picotti, P.; Zambonin, M. Probing Protein Structure by Limited Proteolysis. *Acta Biochim. Polym. Eng. Ed.* **2019**, *51*, 299−322.
(168) Schopper, S.; Kahraman, A.; Leuenberger, P.; Feng, Y.; Piazza, I.; Müller, O.; Boersema, P. J.; Picotti, P. Measuring protein structural changes on a Proteome-Wide Scale Using Limited Proteolysis-Coupled Mass Spectrometry. *Nat. Protoc.* **2017**, *12*, 2391−2410.
(169) Ermácora, M. R.; Delfino, J. M.; Cuenoud, B.; Schepartz, A.; Fox, R. O. Conformation-dependent Cleavage of Staphylococcal Nuclease with a Disulfide-Linked Iron Chelate. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 6383−6387.
(170) Konermann, L.; Pan, J.; Liu, Y.-H. Hydrogen Exchange Mass Spectrometry for Studying Protein Structure and Dynamics. *Chem. Soc. Rev.* **2011**, *40*, 1224−1234.
(171) Wold, F. Vivo Chemical Modification of Proteins (Post-Translational Modification). *Annu. Rev. Biochem.* **1981**, *50*, 783−814.
(172) Krishna, R. G.; Wold, F. Post-Translational Modifications of Proteins. In *Methods in Protein Sequence Analysis*; Imahori, K., Sakiyama, F., Eds.; Springer: Boston, 1993; pp 167−172, DOI: 10.1007/978-1-4899-1603-7_21.
(173) Deribe, Y. L.; Pawson, T.; Dikic, I. Post-Translational Modifications in Signal Integration. *Nat. Struct. Mol. Biol.* **2010**, *17*, 666−672.
(174) Mendoza, V. L.; Vachet, R. W. Probing Protein Structure by Amino Acid-Specific Covalent Labeling and Mass Spectrometry. *Mass Spectrom. Rev.* **2009**, *28*, 785−815.
(175) Xu, G.; Chance, M. R. Hydroxyl Radical-Mediated Modification of Proteins as Probes for Structural Proteomics. *Chem. Rev.* **2007**, *107*, 3514−3543.
(176) Chait, B. T. Mass Spectrometry: Bottom-Up or Top-Down? *Science* **2006**, *314*, 65−66.
(177) Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W. Top Down versus Bottom Up Protein Characterization by Tandem High-Resolution Mass Spectrometry. *J. Am. Chem. Soc.* **1999**, *121*, 806−812.
(178) Zhang, Y.; Fonslow, B. R.; Shan, B.; Baek, M.-C.; Yates, J. R. Protein Analysis by Shotgun/Bottom-up Proteomics. *Chem. Rev.* **2013**, *113*, 2343−2394.
(179) Kelleher, N. L. Peer Reviewed: Top-Down Proteomics. *Anal. Chem.* **2004**, *76*, 196A−203A.
(180) Siuti, N.; Kelleher, N. L. Decoding Protein Modifications using Top-Down Mass Spectrometry. *Nat. Methods* **2007**, *4*, 817−821.
(181) Toby, T. K.; Fornelli, L.; Kelleher, N. L. Progress in Top-Down Proteomics and the Analysis of Proteoforms. *Annu. Rev. Anal. Chem.* **2016**, *9*, 499−519.
(182) Cui, W.; Rohrs, H. W.; Gross, M. L. Top-down Mass Spectrometry: Recent Developments, Applications and Perspectives. *Analyst* **2011**, *136*, 3854−3864.
(183) Compton, P. D.; Zamdborg, L.; Thomas, P. M.; Kelleher, N. L. On the Scalability and Requirements of Whole Protein Mass Spectrometry. *Anal. Chem.* **2011**, *83*, 6868−6874.
(184) Kruppa, G. H.; Schoeniger, J.; Young, M. M. A Top Down Approach to Protein Structural Studies using Chemical Cross-Linking and Fourier Transform Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 155−162.
(185) Novak, P.; Kruppa, G. H.; Young, M. M.; Schoeniger, J. A Top-down Method for the Determination of Residue-Specific Solvent Accessibility in Proteins. *J. Mass Spectrom.* **2004**, *39*, 322−328.
(186) Chen, J.; Cui, W.; Giblin, D.; Gross, M. L. New Protein Footprinting: Fast Photochemical Iodination Combined with Top-Down and Bottom-Up Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 1306−1318.
(187) Hvidt, A.; Linderstrøm-Lang, K. Exchange of Hydrogen Atoms in Insulin with Deuterium Atoms in Aqueous Solutions. *Biochim. Biophys. Acta* **1954**, *14*, 574−575.
(188) Englander, S. W.; Mayne, L.; Bai, Y.; Sosnick, T. R. Hydrogen Exchange: The Modern Legacy of Linderstrøm-Lang. *Protein Sci.* **1997**, *6*, 1101−1109.
(189) Linderstrøm-Lang, K. The pH-dependence of the Deuterium Exchange of Insulin. *Biochim. Biophys. Acta* **1955**, *18*, 308.
(190) Berger, A.; Linderstrøm-Lang, K. Deuterium Exchange of Poly-*DL*-Alanine in Aqueous Solution. *Arch. Biochem. Biophys.* **1957**, *69*, 106−118.
(191) Linderstrøm-Lang, K. Deuterium Exchange and Protein Structure. In *Symposium on Protein Structure*; Neuberger, A., Ed; Methuen & Co. Ltd.: London, 1958; pp 23−34.
(192) Hvidt, A.; Nielsen, S. O. Hydrogen Exchange in Proteins. In *Advances in Protein Chemistry*; Anfinsen, C. B., Anson, M. L., Edsall, J. T., Richards, F. M., Eds.; Academic Press: New York, 1966; Vol. *21*, pp 287−386, DOI: 10.1016/S0065-3233(08)60129-1.
(193) Hvidt, A.; Johansen, G.; Linderstrøm-Lang, K. Deuterium and 18O exchange. In *The Composition, Structure and Reactivity of Proteins*; Alexander, P., Block, R. J., Eds.; Pergamon: Oxford, 1960; pp 101−130, DOI: 10.1016/B978-0-08-009425-0.50010-8.
(194) Englander, S. W. Hydrogen Exchange and Mass Spectrometry: A Historical Perspective. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1481−1489.
(195) Englander, J. J.; Calhoun, D. B.; Englander, S. W. Measurement and Calibration of Peptide Group Hydrogen-Deuterium Exchange by Ultraviolet Spectrophotometry. *Anal. Biochem.* **1979**, *92*, 517−524.
(196) Nielsen, S. O. Hydrogen-deuterium Exchange in N-methylacetamide. *Biochim. Biophys. Acta* **1960**, *37*, 146−147.
(197) Blout, E. R.; De Loze, C.; Asadourian, A. The Deuterium Exchange of Water-soluble Polypeptides and Proteins as Measured by Infrared Spectroscopy. *J. Am. Chem. Soc.* **1961**, *83*, 1895−1900.
(198) Kossiakoff, A. A. Protein Dynamics Investigated by the Neutron Diffraction−Hydrogen Exchange Technique. *Nature* **1982**, *296*, 713−721.
(199) Saunders, M.; Wishnia, A.; Kirkwood, J. G. The Nuclear Magnetic Resonance Spectrum of Ribonuclease1. *J. Am. Chem. Soc.* **1957**, *79*, 3289−3290.
(200) Saunders, M.; Wishnia, A. Nuclear Magnetic Resonance Spectra of Proteins. *Ann. N. Y. Acad. Sci.* **1958**, *70*, 870−874.
(201) Wishnia, A.; Saunders, M. The Nature of the Slowly Exchanging Protons of Ribonuclease. *J. Am. Chem. Soc.* **1962**, *84*, 4235−4239.
(202) Glickson, J. D.; McDonald, C. C.; Phillips, W. D. Assignment of Tryptophan Indole NH Proton Resonances of Lysozyme. *Biochem. Biophys. Res. Commun.* **1969**, *35*, 492−498.
(203) Wagner, G.; Wüthrich, K. Amide Proton Exchange and Surface Conformation of the Basic Pancreatic Trypsin Inhibitor in Solution: Studies with Two-Dimensional Nuclear Magnetic Resonance. *J. Mol. Biol.* **1982**, *160*, 343−361.
(204) Schanda, P.; Brutscher, B. Very Fast Two-Dimensional NMR Spectroscopy for Real-Time Investigation of Dynamic Events in Proteins on the Time Scale of Seconds. *J. Am. Chem. Soc.* **2005**, *127*, 8014−8015.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(205) Englander, S. W.; Kallenbach, N. R. Hydrogen Exchange and Structural Dynamics of Proteins and Nucleic Acids. *Q. Rev. Biophys.* **1983**, *16*, 521−655.
(206) Englander, S. W.; Sosnick, T. R.; Englander, J. J.; Mayne, L. Mechanisms and Uses of Hydrogen Exchange. *Curr. Opin. Struct. Biol.* **1996**, *6*, 18−23.
(207) Paterson, Y.; Englander, S.; Roder, H. An Antibody Binding Site on Cytochrome C Defined by Hydrogen Exchange and Two-Dimensional NMR. *Science* **1990**, *249*, 755−759.
(208) Katta, V.; Chait, B. T.; Carr, S. Conformational Changes in Proteins Probed by Hydrogen-Exchange Electrospray-Ionization Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1991**, *5*, 214−217.
(209) Zhang, Z.; Smith, D. L. Determination of Amide Hydrogen Exchange by Mass Spectrometry: A New Tool for Protein Structure Elucidation. *Protein Sci.* **1993**, *2*, 522−531.
(210) Englander, S. W.; Downer, N. W.; Teitelbaum, H. Hydrogen Exchange. *Annu. Rev. Biochem.* **1972**, *41*, 903−924.
(211) Bai, Y.; Milne, J. S.; Mayne, L.; Englander, S. W. Primary Structure Effects on Peptide Group Hydrogen Exchange. *Proteins: Struct., Funct., Genet.* **1993**, *17*, 75−86.
(212) Smith, D. L.; Deng, Y.; Zhang, Z. Probing the Non-covalent Structure of Proteins by Amide Hydrogen Exchange and Mass Spectrometry. *J. Mass Spectrom.* **1997**, *32*, 135−146.
(213) Pauling, L.; Corey, R. B.; Branson, H. R. The Structure of proteins: Two Hydrogen-Bonded Helical Configurations of the Polypeptide Chain. *Proc. Natl. Acad. Sci. U. S. A.* **1951**, *37*, 205−211.
(214) Konermann, L.; Tong, X.; Pan, Y. Protein Structure and Dynamics Studied by Mass Spectrometry: H/D Exchange, Hydroxyl Radical Labeling, and Related Approaches. *J. Mass Spectrom.* **2008**, *43*, 1021−1036.
(215) Chalmers, M. J.; Busby, S. A.; Pascal, B. D.; West, G. M.; Griffin, P. R. Differential Hydrogen/Deuterium Exchange Mass Spectrometry Analysis of Protein−Ligand Interactions. *Expert Rev. Proteomics* **2011**, *8*, 43−59.
(216) Masson, G. R.; Burke, J. E.; Ahn, N. G.; Anand, G. S.; Borchers, C.; Brier, S.; Bou-Assaf, G. M.; Engen, J. R.; Englander, S. W.; et al. Recommendations for Performing, Interpreting and Reporting Hydrogen Deuterium Exchange Mass Spectrometry (HDX-MS) Experiments. *Nat. Methods* **2019**, *16*, 595−602.
(217) Rosa, J. J.; Richards, F. M. An Experimental Procedure for Increasing the Structural Resolution of Chemical Hydrogen-Exchange Measurements on Proteins: Application to Ribonuclease S Peptide. *J. Mol. Biol.* **1979**, *133*, 399−416.
(218) Englander, J. J.; Rogero, J. R.; Englander, S. W. Protein hydrogen exchange studied by the Fragment Separation Method. *Anal. Biochem.* **1985**, *147*, 234−244.
(219) Jørgensen, T. J. D.; GÅrdsvoll, H.; Ploug, M.; Roepstorff, P. Intramolecular Migration of Amide Hydrogens in Protonated Peptides upon Collisional Activation. *J. Am. Chem. Soc.* **2005**, *127*, 2785−2793.
(220) Jørgensen, T. J. D.; Bache, N.; Roepstorff, P.; GÅrdsvoll, H.; Ploug, M. Collisional Activation by MALDI Tandem Time-of-flight Mass Spectrometry Induces Intramolecular Migration of Amide Hydrogens in Protonated Peptides. *Mol. Cell. Proteomics* **2005**, *4*, 1910−1919.
(221) Rand, K. D.; Adams, C. M.; Zubarev, R. A.; Jørgensen, T. J. D. Electron Capture Dissociation Proceeds with a Low Degree of Intramolecular Migration of Peptide Amide Hydrogens. *J. Am. Chem. Soc.* **2008**, *130*, 1341−1349.
(222) Zehl, M.; Rand, K. D.; Jensen, O. N.; Jørgensen, T. J. D. Electron Transfer Dissociation Facilitates the Measurement of Deuterium Incorporation into Selectively Labeled Peptides with Single Residue Resolution. *J. Am. Chem. Soc.* **2008**, *130*, 17453−17459.
(223) Pan, J.; Han, J.; Borchers, C. H.; Konermann, L. Hydrogen/Deuterium Exchange Mass Spectrometry with Top-Down Electron Capture Dissociation for Characterizing Structural Transitions of a 17 kDa Protein. *J. Am. Chem. Soc.* **2009**, *131*, 12801−12808.
(224) Hamuro, Y.; E, S. Y. Determination of Backbone Amide Hydrogen Exchange Rates of Cytochrome c Using Partially Scrambled Electron Transfer Dissociation Data. *J. Am. Soc. Mass Spectrom.* **2018**, *29*, 989−1001.
(225) Kan, Z.-Y.; Walters, B. T.; Mayne, L.; Englander, S. W. Protein Hydrogen Exchange at Residue Resolution by Proteolytic Fragmentation Mass Spectrometry Analysis. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 16438−16443.
(226) Hamuro, Y.; Zhang, T. High-Resolution HDX-MS of Cytochrome c Using Pepsin/Fungal Protease Type XIII Mixed Bed Column. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 227−234.
(227) López-Ferrer, D.; Petritis, K.; Lourette, N. M.; Clowers, B.; Hixson, K. K.; Heibeck, T.; Prior, D. C.; Paša-Tolić, L.; Camp, D. G.; et al. On-line Digestion System for Protein Characterization and Proteome Analysis. *Anal. Chem.* **2008**, *80*, 8930−8936.
(228) Jones, L. M.; Zhang, H.; Vidavsky, I.; Gross, M. L. Online, High-Pressure Digestion System for Protein Characterization by Hydrogen/Deuterium Exchange and Mass Spectrometry. *Anal. Chem.* **2010**, *82*, 1171−1174.
(229) Rand, K. D.; Zehl, M.; Jensen, O. N.; Jørgensen, T. J. D. Protein Hydrogen Exchange Measured at Single-Residue Resolution by Electron Transfer Dissociation Mass Spectrometry. *Anal. Chem.* **2009**, *81*, 5577−5584.
(230) Landgraf, R. R.; Chalmers, M. J.; Griffin, P. R. Automated Hydrogen/Deuterium Exchange Electron Transfer Dissociation High Resolution Mass Spectrometry Measured at Single-Amide Resolution. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 301−309.
(231) Huang, R. Y.-C.; Krystek, S. R., Jr; Felix, N.; Graziano, R. F.; Srinivasan, M.; Pashine, A.; Chen, G. Hydrogen/Deuterium Exchange Mass Spectrometry and Computational Modeling Reveal a Discontinuous Epitope of an Antibody/TL1A Interaction. *MAbs* **2018**, *10*, 95−103.
(232) Gessner, C.; Steinchen, W.; Bédard, S.; Skinner, J. J.; Woods, V. L.; Walsh, T. J.; Bange, G.; Pantazatos, D. P. Computational Method Allowing Hydrogen-Deuterium Exchange Mass Spectrometry at Single Amide Resolution. *Sci. Rep.* **2017**, *7*, 3789.
(233) Saltzberg, D. J.; Broughton, H. B.; Pellarin, R.; Chalmers, M. J.; Espada, A.; Dodge, J. A.; Pascal, B. D.; Griffin, P. R.; Humblet, C.; Sali, A. A Residue-Resolved Bayesian Approach to Quantitative Interpretation of Hydrogen−Deuterium Exchange from Mass Spectrometry: Application to Characterizing Protein−Ligand Interactions. *J. Phys. Chem. B* **2017**, *121*, 3493−3501.
(234) Jansson, E. T.; Lai, Y.-H.; Santiago, J. G.; Zare, R. N. Rapid Hydrogen−Deuterium Exchange in Liquid Droplets. *J. Am. Chem. Soc.* **2017**, *139*, 6851−6854.
(235) Rand, K. D.; Pringle, S. D.; Murphy, J. P.; Fadgen, K. E.; Brown, J.; Engen, J. R. Gas-Phase Hydrogen/Deuterium Exchange in a Traveling Wave Ion Guide for the Examination of Protein Conformations. *Anal. Chem.* **2009**, *81*, 10019−10028.
(236) Mistarz, U. H.; Brown, J. M.; Haselmann, K. F.; Rand, K. D. Probing the Binding Interfaces of Protein Complexes Using Gas-Phase H/D Exchange Mass Spectrometry. *Structure* **2016**, *24*, 310−318.
(237) Uppal, S. S.; Beasley, S. E.; Scian, M.; Guttman, M. Gas-Phase Hydrogen/Deuterium Exchange for Distinguishing Isomeric Carbohydrate Ions. *Anal. Chem.* **2017**, *89*, 4737−4742.
(238) Powell, K. D.; Ghaemmaghami, S.; Wang, M. Z.; Ma, L.; Oas, T. G.; Fitzgerald, M. C. A General Mass Spectrometry-Based Assay for the Quantitation of Protein−Ligand Binding Interactions in Solution. *J. Am. Chem. Soc.* **2002**, *124*, 10256−10257.
(239) Zhu, M. M.; Rempel, D. L.; Du, Z.; Gross, M. L. Quantification of Protein−Ligand Interactions by Mass Spectrometry, Titration, and H/D Exchange: PLIMSTEX. *J. Am. Chem. Soc.* **2003**, *125*, 5252−5253.
(240) Wang, H.; Rempel, D. L.; Giblin, D.; Frieden, C.; Gross, M. L. Peptide-Level Interactions between Proteins and Small-Molecule Drug Candidates by Two Hydrogen−Deuterium Exchange MS-Based Methods: The Example of Apolipoprotein E3. *Anal. Chem.* **2017**, *89*, 10687−10695.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(241) Rey, M.; Mrázek, H.; Pompach, P.; Novák, P.; Pelosi, L.; Brandolin, G.; Forest, E.; Havlíček, V.; Man, P. Effective Removal of Nonionic Detergents in Protein Mass Spectrometry, Hydrogen/Deuterium Exchange, and Proteomics. *Anal. Chem.* **2010**, *82*, 5107−5116.
(242) Anderson, K. W.; Gallagher, E. S.; Hudgens, J. W. Automated Removal of Phospholipids from Membrane Proteins for H/D Exchange Mass Spectrometry Workflows. *Anal. Chem.* **2018**, *90*, 6409−6412.
(243) Möller, I. R.; Slivacka, M.; Hausner, J.; Nielsen, A. K.; Pospíšilová, E.; Merkle, P. S.; Lišková, R.; Polák, M.; Loland, C. J.; Kádek, A.; et al. Improving the Sequence Coverage of Integral Membrane Proteins during Hydrogen/Deuterium Exchange Mass Spectrometry Experiments. *Anal. Chem.* **2019**, *91*, 10970−10978.
(244) Jensen, P. F.; Comamala, G.; Trelle, M. B.; Madsen, J. B.; Jørgensen, T. J. D.; Rand, K. D. Removal of N-Linked Glycosylations at Acidic pH by PNGase A Facilitates Hydrogen/Deuterium Exchange Mass Spectrometry Analysis of N-Linked Glycoproteins. *Anal. Chem.* **2016**, *88*, 12479−12488.
(245) Hamuro, Y.; Coales, S. J.; Southern, M. R.; Nemeth-Cawley, J. F.; Stranz, D. D.; Griffin, P. R. Rapid Analysis of Protein Structure and Dynamics by Hydrogen/Deuterium Exchange Mass Spectrometry. *J. Biomol. Technol.* **2003**, *14*, 171−182.
(246) Weis, D. D.; Engen, J. R.; Kass, I. J. Semi-Automated Data Processing of Hydrogen Exchange Mass Spectra using HX-Express. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1700−1703.
(247) Pascal, B. D.; Willis, S.; Lauer, J. L.; Landgraf, R. R.; West, G. M.; Marciano, D.; Novick, S.; Goswami, D.; Chalmers, M. J.; Griffin, P. R. HDX Workbench: Software for the Analysis of H/D Exchange MS Data. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 1512−1521.
(248) Palmblad, M.; Buijs, J.; HÅkansson, P. Automatic Analysis of Hydrogen/Deuterium Exchange Mass Spectra of Peptides and Proteins Using Calculations of Isotopic Distributions. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 1153−1162.
(249) Zhang, Z. Large-Scale Identification and Quantification of Covalent Modifications in Therapeutic Proteins. *Anal. Chem.* **2009**, *81*, 8354−8364.
(250) Zhang, Z.; Zhang, A.; Xiao, G. Improved Protein Hydrogen/Deuterium Exchange Mass Spectrometry Platform with Fully Automated Data Processing. *Anal. Chem.* **2012**, *84*, 4942−4949.
(251) Pascal, B. D.; Chalmers, M. J.; Busby, S. A.; Griffin, P. R. HD desktop: An Integrated Platform for the Analysis and Visualization of H/D Exchange Data. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 601−610.
(252) Lou, X.; Kirchner, M.; Renard, B. Y.; Köthe, U.; Boppel, S.; Graf, C.; Lee, C.-T.; Steen, J. A. J.; Steen, H.; Mayer, M. P.; et al. Deuteration Distribution Estimation with Improved Sequence Coverage for HX/MS Experiments. *Bioinformatics* **2010**, *26*, 1535−1541.
(253) Kan, Z.-Y.; Mayne, L.; Sevugan Chetty, P.; Englander, S. W. ExMS: Data Analysis for HX-MS Experiments. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 1906.
(254) Liu, S.; Liu, L.; Uzuner, U.; Zhou, X.; Gu, M.; Shi, W.; Zhang, Y.; Dai, S. Y.; Yuan, J. S. HDX-Analyzer: A Novel Package for Statistical Analysis of Protein Structure Dynamics. *BMC Bioinf.* **2011**, *12*, S43−S44.
(255) Miller, D. E.; Prasannan, C. B.; Villar, M. T.; Fenton, A. W.; Artigues, A. HDXFinder: Automated Analysis and Data Reporting of Deuterium/Hydrogen Exchange Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 425−429.
(256) Rey, M.; Sarpe, V.; Burns, K. M.; Buse, J.; Baker, C. A. H.; van Dijk, M.; Wordeman, L.; Bonvin, A. M. J. J.; Schriemer, D. C. Mass Spec Studio for Integrative Structural Biology. *Structure* **2014**, *22*, 1538−1548.
(257) Kavan, D.; Man, P. MSTools—Web Based Application for Visualization and Presentation of HXMS Data. *Int. J. Mass Spectrom.* **2011**, *302*, 53−58.
(258) Hourdel, V.; Volant, S.; O'Brien, D. P.; Chenal, A.; Chamot-Rooke, J.; Dillies, M.-A.; Brier, S. MEMHDX: An Interactive Tool to Expedite the Statistical Validation and Visualization of Large HDX-MS Datasets. *Bioinformatics* **2016**, *32*, 3413−3419.
(259) Lau, A. M. C.; Ahdash, Z.; Martens, C.; Politis, A. Deuteros: Software for Rapid Analysis and Visualization of Data from Differential Hydrogen Deuterium Exchange-Mass Spectrometry. *Bioinformatics* **2019**, *35*, 3171−3173.
(260) Bouyssié, D.; Lesne, J.; Locard-Paulet, M.; Albigot, R.; Burlet-Schiltz, O.; Marcoux, J. HDX-Viewer: Interactive 3D Visualization of Hydrogen−Deuterium Exchange Data. *Bioinformatics* **2019**, *35*, 5331−5333.
(261) Pirrone, G. F.; Iacob, R. E.; Engen, J. R. Applications of Hydrogen/Deuterium Exchange MS from 2012 to 2014. *Anal. Chem.* **2015**, *87*, 99−118.
(262) Engen, J. R.; Wales, T. E. Analytical Aspects of Hydrogen Exchange Mass Spectrometry. *Annu. Rev. Anal. Chem.* **2015**, *8*, 127−148.
(263) Marciano, D. P.; Dharmarajan, V.; Griffin, P. R. HDX-MS Guided Drug Discovery: Small Molecules and Biopharmaceuticals. *Curr. Opin. Struct. Biol.* **2014**, *28*, 105−111.
(264) Wei, H.; Mo, J.; Tao, L.; Russell, R. J.; Tymiak, A. A.; Chen, G.; Iacob, R. E.; Engen, J. R. Hydrogen/Deuterium Exchange Mass Spectrometry for Probing Higher Order Structure of Protein Therapeutics: Methodology and Applications. *Drug Discovery Today* **2014**, *19*, 95−102.
(265) Deng, B.; Lento, C.; Wilson, D. J. Hydrogen Deuterium Exchange Mass Spectrometry in Biopharmaceutical Discovery and Development − A review. *Anal. Chim. Acta* **2016**, *940*, 8−20.
(266) Masson, G. R.; Jenkins, M. L.; Burke, J. E. An Overview of Hydrogen Deuterium Exchange Mass Spectrometry (HDX-MS) in Drug Discovery. *Expert Opin. Drug Discovery* **2017**, *12*, 981−994.
(267) Englander, S. W.; Mayne, L.; Kan, Z.-Y.; Hu, W. Protein Folding—How and Why: By Hydrogen Exchange, Fragment Separation, and Mass Spectrometry. *Annu. Rev. Biophys.* **2016**, *45*, 135−152.
(268) Georgescauld, F.; Wales, T. E.; Engen, J. R. Hydrogen Deuterium Exchange Mass Spectrometry Applied to Chaperones and Chaperone-Assisted Protein Folding. *Expert Rev. Proteomics* **2019**, *16*, 613−625.
(269) Gallagher, E. S.; Hudgens, J. W. Mapping Protein−Ligand Interactions with Proteolytic Fragmentation, Hydrogen/Deuterium Exchange-Mass Spectrometry. In *Methods in Enzymology*; Kelman, Z., Ed.; Elsevier Academic Press: San Diego, 2016; Vol. *566*, pp 357−404, DOI: 10.1016/bs.mie.2015.08.010.
(270) Percy, A. J.; Rey, M.; Burns, K. M.; Schriemer, D. C. Probing Protein Interactions with Hydrogen/Deuterium Exchange and Mass Spectrometry—A Review. *Anal. Chim. Acta* **2012**, *721*, 7−21.
(271) Harrison, R. A.; Engen, J. R. Conformational Insight into Multi-Protein Signaling Assemblies by Hydrogen−Deuterium Exchange Mass Spectrometry. *Curr. Opin. Struct. Biol.* **2016**, *41*, 187−193.
(272) Engen, J. R. Analysis of Protein Conformation and Dynamics by Hydrogen/Deuterium Exchange MS. *Anal. Chem.* **2009**, *81*, 7870−7875.
(273) Kochert, B. A.; Iacob, R. E.; Wales, T. E.; Makriyannis, A.; Engen, J. R. Hydrogen-Deuterium Exchange Mass Spectrometry to Study Protein Complexes. In *Protein Complex Assembly: Methods and Protocols*; Marsh, J. A., Ed.; Springer: New York, 2018; pp 153−171, DOI: 10.1007/978-1-4939-7759-8_10.
(274) Englander, J. J.; Del Mar, C.; Li, W.; Englander, S. W.; Kim, J. S.; Stranz, D. D.; Hamuro, Y.; Woods, V. L. Protein Structure Change Studied by Hydrogen-Deuterium Exchange, Functional Labeling, and Mass Spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 7057−7062.
(275) Vadas, O.; Jenkins, M. L.; Dornan, G. L.; Burke, J. E. Using Hydrogen−Deuterium Exchange Mass Spectrometry to Examine Protein−Membrane Interactions. In *Methods in Enzymology*; Gelb, M. H., Ed.; Elsevier Academic Press: San Diego, 2017; Vol. *583*, pp 143−172, DOI: 10.1016/bs.mie.2016.09.008.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(276) Rand, K. D.; Zehl, M.; Jørgensen, T. J. D. Measuring the Hydrogen/Deuterium Exchange of Proteins at High Spatial Resolution by Mass Spectrometry: Overcoming Gas-Phase Hydrogen/Deuterium Scrambling. *Acc. Chem. Res.* **2014**, *47*, 3018−3027.
(277) Baslé, E.; Joubert, N.; Pucheault, M. Protein Chemical Modification on Endogenous Amino Acids. *Chem. Biol.* **2010**, *17*, 213−227.
(278) Lundblad, R. L. *Chemical Reagents for Protein Modification*, 4th ed; CRC Press: Boca Raton, 2014.
(279) Spicer, C. D.; Davis, B. G. Selective Chemical Protein Modification. *Nat. Commun.* **2014**, *5*, 4740.
(280) Boutureira, O.; Bernardes, G. J. L. Advances in Chemical Protein Modification. *Chem. Rev.* **2015**, *115*, 2174−2195.
(281) Leitner, A. A. Review of the Role of Chemical Modification Methods in Contemporary Mass Spectrometry-Based Proteomics Research. *Anal. Chim. Acta* **2018**, *1000*, 2−19.
(282) Bauer, R. A. Covalent Inhibitors in Drug Discovery: from Accidental Discoveries to Avoided Liabilities and Designed Therapies. *Drug Discovery Today* **2015**, *20*, 1061−1073.
(283) Lonsdale, R.; Ward, R. A. Structure-based Design of Targeted Covalent Inhibitors. *Chem. Soc. Rev.* **2018**, *47*, 3816−3830.
(284) Mukherjee, H.; Grimster, N. P. Beyond Cysteine: Recent Developments in the Area of Targeted Covalent Inhibition. *Curr. Opin. Chem. Biol.* **2018**, *44*, 30−38.
(285) Tamura, T.; Hamachi, I. Chemistry for Covalent Modification of Endogenous/Native Proteins: From Test Tubes to Complex Biological Systems. *J. Am. Chem. Soc.* **2019**, *141*, 2782−2799.
(286) Giles, N. M.; Giles, G. I.; Jacob, C. Multiple Roles of Cysteine in Biocatalysis. *Biochem. Biophys. Res. Commun.* **2003**, *300*, 1−4.
(287) Giron, P.; Dayon, L.; Sanchez, J.-C. Cysteine Tagging for MS-based Proteomics. *Mass Spectrom. Rev.* **2011**, *30*, 366−395.
(288) Wu, X.; Liang, H.; O'Hara, K. A.; Yalowich, J. C.; Hasinoff, B. B. Thiol-Modulated Mechanisms of the Cytotoxicity of Thimerosal and Inhibition of DNA Topoisomerase II*α*. *Chem. Res. Toxicol.* **2008**, *21*, 483−493.
(289) Smolka, M. B.; Zhou, H.; Purkayastha, S.; Aebersold, R. Optimization of the Isotope-Coded Affinity Tag-Labeling Procedure for Quantitative Proteome Analysis. *Anal. Biochem.* **2001**, *297*, 25−31.
(290) Smythe, C. V. The Reaction of Iodoacetate and of Iodoacetamide with Various Sulfhydryl Groups, with urease and with yeast preparations. *J. Biol. Chem.* **1936**, *114*, 601−612.
(291) Anson, M. L. The Reactions of Iodine and Iodoacetamide with Native Egg Albumin. *J. Gen. Physiol.* **1940**, *23*, 321−331.
(292) Schroeder, E. F.; Woodward, G. E.; Platt, M. E. The Relation of Sulfhydryl to Inhibition of Yeast Fermentation by Iodoacetic Acid. *J. Biol. Chem.* **1933**, *101*, 133−144.
(293) Pasquarello, C.; Sanchez, J.-C.; Hochstrasser, D. F.; Corthals, G. L. N-t-butyliodoacetamide and iodoacetanilide: Two New Cysteine Alkylating Reagents for Relative Quantitation of Proteins. *Rapid Commun. Mass Spectrom.* **2004**, *18*, 117−127.
(294) Gundlach, H. G.; Stein, W. H.; Moore, S. The Nature of the Amino Acid Residues Involved in the Inactivation of Ribonuclease by Iodoacetate. *J. Biol. Chem.* **1959**, *234*, 1754−1760.
(295) Heinrikson, R. L.; Stein, W. H.; Crestfield, A. M.; Moore, S. The Reactivities of the Histidine Residues at the Active Site of Ribonuclease toward Halo Acids of Different Structures. *J. Biol. Chem.* **1965**, *240*, 2921−2934.
(296) Gurd, F. R. N. Carboxymethylation. In *Methods in Enzymology*; Hirs, C. H. W., Ed.; Academic Press: New York, 1967; Vol. *11*, pp 532−541, DOI: 10.1016/S0076-6879(67)11064-1.
(297) Gundlach, H. G.; Moore, S.; Stein, W. H. The Reaction of Iodoacetate with Methionine. *J. Biol. Chem.* **1959**, *234*, 1761−1764.
(298) Whitehurst, C. B.; Soderblom, E. J.; West, M. L.; Hernandez, R.; Goshe, M. B.; Brown, D. T. Location and Role of Free Cysteinyl Residues in the Sindbis Virus E1 and E2 Glycoproteins. *J. Virol.* **2007**, *81*, 6231−6240.
(299) Britto, P. J.; Knipling, L.; Wolff, J. The Local Electrostatic Environment Determines Cysteine Reactivity of Tubulin. *J. Biol. Chem.* **2002**, *277*, 29018−29027.
(300) Korman, S.; Clarke, H. T. Carboxymethylamino Acids and Peptides. *J. Biol. Chem.* **1956**, *221*, 113−132.
(301) Dahl, K. H.; McKinley-McKee, J. S. The Reactivity of Affinity Labels: A kinetic study of the Reaction of Alkyl Halides with Thiolate Anions—a Model Reaction for Protein Alkylation. *Bioorg. Chem.* **1981**, *10*, 329−341.
(302) Plapp, B. V. Application of Affinity Labeling for Studying Structure and Function Of Enzymes. In *Methods in Enzymology*; Purich, D. L., Ed.; Academic Press: New York, 1982; Vol. *87*, pp 469−499, DOI: 10.1016/S0076-6879(82)87027-4.
(303) Gregory, J. D. The Stability of N-Ethylmaleimide and its Reaction with Sulfhydryl Groups. *J. Am. Chem. Soc.* **1955**, *77*, 3922−3923.
(304) Leslie, J. Spectral Shifts in the Reaction of N-ethylmaleimide with Proteins. *Anal. Biochem.* **1965**, *10*, 162−167.
(305) Gorin, G.; Martic, P. A.; Doughty, G. Kinetics of the Reaction of N-ethylmaleimide with Cysteine and Some Congeners. *Arch. Biochem. Biophys.* **1966**, *115*, 593−597.
(306) Smyth, D. G.; Nagamatsu, A.; Fruton, J. S. Some Reactions of N-Ethylmaleimide1. *J. Am. Chem. Soc.* **1960**, *82*, 4600−4604.
(307) Bednar, R. A. Reactivity and pH Dependence of Thiol Conjugation to N-ethylmaleimide: Detection of a Conformational Change in Chalcone Isomerase. *Biochemistry* **1990**, *29*, 3684−3690.
(308) Apuy, J. L.; Park, Z.-Y.; Swartz, P. D.; Dangott, L. J.; Russell, D. H.; Baldwin, T. O. Pulsed-Alkylation Mass Spectrometry for the Study of Protein Folding and Dynamics: Development and Application to the Study of a Folding/Unfolding Intermediate of Bacterial Luciferase. *Biochemistry* **2001**, *40*, 15153−15163.
(309) Tedaldi, L. M.; Smith, M. E. B.; Nathani, R. I.; Baker, J. R. Bromomaleimides: New Reagents for the Selective and Reversible Modification of Cysteine. *Chem. Commun.* **2009**, 6583−6585.
(310) Smith, M. E. B.; Schumacher, F. F.; Ryan, C. P.; Tedaldi, L. M.; Papaioannou, D.; Waksman, G.; Caddick, S.; Baker, J. R. Protein Modification, Bioconjugation, and Disulfide Bridging Using Bromomaleimides. *J. Am. Chem. Soc.* **2010**, *132*, 1960−1965.
(311) Ryan, C. P.; Smith, M. E. B.; Schumacher, F. F.; Grohmann, D.; Papaioannou, D.; Waksman, G.; Werner, F.; Baker, J. R.; Caddick, S. Tunable Reagents for Multi-Functional Bioconjugation: Reversible or Permanent Chemical Modification of Proteins and Peptides by Control of Maleimide Hydrolysis. *Chem. Commun.* **2011**, *47*, 5452−5454.
(312) Antonino, L. C.; Straub, K.; Wu, J. C. Probing the Active Site of Human IMP Dehydrogenase Using Halogenated Purine Riboside 5′-Monophosphates and Covalent Modification Reagents. *Biochemistry* **1994**, *33*, 1760−1765.
(313) Toutchkine, A.; Kraynov, V.; Hahn, K. Solvent-Sensitive Dyes to Report Protein Conformational Changes in Living Cells. *J. Am. Chem. Soc.* **2003**, *125*, 4132−4145.
(314) Garrett, S. C.; Hodgson, L.; Rybin, A.; Toutchkine, A.; Hahn, K. M.; Lawrence, D. S.; Bresnick, A. R. A Biosensor of S100A4Metastasis Factor Activation: Inhibitor Screening and Cellular Activation Dynamics. *Biochemistry* **2008**, *47*, 986−996.
(315) Embaby, A. M.; Schoffelen, S.; Kofoed, C.; Meldal, M.; Diness, F. Rational Tuning of Fluorobenzene Probes for Cysteine-Selective Protein Modification. *Angew. Chem.* **2018**, *130*, 8154−8158.
(316) Chumsae, C.; Gaza-Bulseco, G.; Liu, H. Identification and Localization of Unpaired Cysteine Residues in Monoclonal Antibodies by Fluorescence Labeling and Mass Spectrometry. *Anal. Chem.* **2009**, *81*, 6449−6457.
(317) Everett, E. A.; Falick, A. M.; Reich, N. O. Identification of a Critical Cysteine in EcoRI DNA Methyltransferase by Mass Spectrometry. *J. Biol. Chem.* **1990**, *265*, 17713−17719.
(318) Hess, D. T.; Matsumoto, A.; Kim, S.-O.; Marshall, H. E.; Stamler, J. S. Protein S-nitrosylation: Purview and Parameters. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 150−166.
(319) Torta, F.; Usuelli, V.; Malgaroli, A.; Bachi, A. Proteomic Analysis of Protein S-nitrosylation. *Proteomics* **2008**, *8*, 4484−4494.
(320) Smith, S.; Kyte, J. Assessment of the Number of Free Cysteines and Isolation and Identification of Cysteine-Containing

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

Peptides from Acetylcholine Receptor. Appendix: Reversible Modification of Cysteine with Cyanogen Bromide. *Biochemistry* **1989**, *28*, 3481−3482.
(321) Xu, K.; Zhang, Y.; Tang, B.; Laskin, J.; Roach, P. J.; Chen, H. Study of Highly Selective and Efficient Thiol Derivatization Using Selenium Reagents by Mass Spectrometry. *Anal. Chem.* **2010**, *82*, 6926−6932.
(322) Wang, Z.; Zhang, Y.; Zhang, H.; Harrington, P. B.; Chen, H. Fast and Selective Modification of Thiol Proteins/Peptides by N-(Phenylseleno)Phthalimide. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 520−529.
(323) Ellman, G. L. Tissue Sulfhydryl Groups. *Arch. Biochem. Biophys.* **1959**, *82*, 70−77.
(324) Grassetti, D. R.; Murray, J. F. Determination of Sulfhydryl Groups with 2,2'- or 4,4'-Dithiodipyridine. *Arch. Biochem. Biophys.* **1967**, *119*, 41−49.
(325) Friedman, M. Application of the S-Pyridylethylation Reaction to the Elucidation of the Structures and Functions of Proteins. *J. Protein Chem.* **2001**, *20*, 431−453.
(326) Kleinova, M.; Belgacem, O.; Pock, K.; Rizzi, A.; Buchacher, A.; Allmaier, G. Characterization of Cysteinylation of pharmaceutical-Grade Human Serum Albumin by Electrospray Ionization Mass Spectrometry and Low-Energy Collision-Induced Dissociation Tandem Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 2965−2973.
(327) LoPachin, R. M.; Gavin, T.; Geohagen, B. C.; Das, S. Neurotoxic Mechanisms of Electrophilic Type-2 Alkenes: Soft−Soft Interactions Described by Quantum Mechanical Parameters. *Toxicol. Sci.* **2007**, *98*, 561−570.
(328) Zhuang, J.; Zhao, B.; Meng, X.; Schiffman, J.; Perry, S. L.; Vachet, R.; Thayumanavan, S. Programmable Chemical Switch based on Triggerable Michael Acceptors. *Chem. Sci.* **2020**, *11*, 2103−2111.
(329) Farnworth, A. J. The Reactivity of the Cystine Linkages in Wool Towards Reducing Agents. *Biochem. J.* **1955**, *60*, 626−635.
(330) Lindley, H.; Cranston, R. W. The Reactivity of the Disulphide Bonds of Wool. *Biochem. J.* **1974**, *139*, 515−523.
(331) Hiskey, R. G.; Harpp, D. N. Chemistry of Aliphatic Disulfides: VII. Cyanide Cleavage in the Presence of Thiocyanates. *J. Am. Chem. Soc.* **1964**, *86*, 2014−2018.
(332) Neumann, H.; Smith, R. A. Cleavage of the Disulfide Bonds of Cystine and Oxidized Glutathione by Phosphorothioate. *Arch. Biochem. Biophys.* **1967**, *122*, 354−361.
(333) Lukesh, J. C.; Palte, M. J.; Raines, R. T. A Potent, Versatile Disulfide-Reducing Agent from Aspartic Acid. *J. Am. Chem. Soc.* **2012**, *134*, 4057−4059.
(334) Singh, R.; Lamoureux, G. V.; Lees, W. J.; Whitesides, G. M. Reagents for Rapid Reduction of Disulfide Bonds. In *Methods in Enzymology*; Parker, L., Ed.; Academic Press: New York, 1995; Vol. *251*, pp 167−173, DOI: 10.1016/0076-6879(95)51119-9.
(335) Duane Brown, W. Reduction of Protein Disulfide Bonds by Sodium Borohydride. *Biochim. Biophys. Acta* **1960**, *44*, 365−367.
(336) Simpson, S. D.; Young, L. Biochemical Studies of Toxic Agents. 1. Experiments with Radioactive 2:3-Dimercaptopropanol (British Anti-Lewisite). *Biochem. J.* **1950**, *46*, 634−640.
(337) Evans, R. M.; Fraser, J. B.; Owen, L. N. 61. Dithiols. Part III. Derivatives of polyhydric alcohols. *J. Chem. Soc.* **1949**, 248−255.
(338) Cleland, W. W. Dithiothreitol, a New Protective Reagent for SH Groups*. *Biochemistry* **1964**, *3*, 480−482.
(339) Stevens, R.; Stevens, L.; Price, N. C. The Stabilities of Various Thiol Compounds used in Protein Purifications. *Biochem. Educ.* **1983**, *11*, 70−71.
(340) Iyer, K. S.; Klee, W. A. Direct Spectrophotometric Measurement of the Rate of Reduction of Disulfide Bonds: The Reactivity of the Disulfide Bonds of Bovine α-Lactalbumin. *J. Biol. Chem.* **1973**, *248*, 707−710.
(341) Burns, J. A.; Butler, J. C.; Moran, J.; Whitesides, G. M. Selective Reduction of Disulfides by Tris(2-Carboxyethyl)Phosphine. *J. Org. Chem.* **1991**, *56*, 2648−2650.
(342) Vâlcu, C.-M.; Schlink, K. Reduction of Proteins During Sample Preparation and Two-Dimensional Gel Electrophoresis of Woody Plant Samples. *Proteomics* **2006**, *6*, 1599−1605.
(343) Hong, R.; Nisonoff, A. Relative Labilities of the Two Types of Interchain Disulfide Bond of Rabbit γG-Immunoglobulin. *J. Biol. Chem.* **1965**, *240*, 3883−3891.
(344) Humphreys, D. P.; Heywood, S. P.; Henry, A.; Ait-Lhadj, L.; Antoniw, P.; Palframan, R.; Greenslade, K. J.; Carrington, B.; Reeks, D. G.; Bowering, L. C.; et al. Alternative Antibody Fab' Fragment PEGylation Strategies: Combination of Strong Reducing Agents, Disruption of the Interchain Disulphide Bond and Disulphide Engineering. *Protein Eng., Des. Sel.* **2007**, *20*, 227−234.
(345) Kraj, A.; Brouwer, H.-J.; Reinhoud, N.; Chervet, J.-P. A Novel Electrochemical Method for Efficient Reduction of Disulfide Bonds in Peptides and Proteins Prior to MS Detection. *Anal. Bioanal. Chem.* **2013**, *405*, 9311−9320.
(346) Mysling, S.; Salbo, R.; Ploug, M.; Jørgensen, T. J. D. Electrochemical Reduction of Disulfide-Containing Proteins for Hydrogen/Deuterium Exchange Monitored by Mass Spectrometry. *Anal. Chem.* **2014**, *86*, 340−345.
(347) Palego, L.; Betti, L.; Rossi, A.; Giannaccini, G. Tryptophan Biochemistry: Structural, Nutritional, Metabolic, and Medical Aspects in Humans. *J. Amino Acids* **2016**, *2016*, 1−13.
(348) Painter, R. G.; Sage, H. J.; Tanford, C. Contributions of Heavy and Light Chains of Rabbit Immunoglobulin G to Antibody Activity. I. Binding Studies on Isolated Heavy and Light Chains. *Biochemistry* **1972**, *11*, 1327−1337.
(349) Schibli, D. J.; Epand, R. F.; Vogel, H. J.; Epand, R. M. Tryptophan-Rich Antimicrobial Peptides: Comparative Properties and Membrane Interactions. *Biochem. Cell Biol.* **2002**, *80*, 667−677.
(350) Al-Abdul-Wahid, M. S.; DeMill, C. M.; Serwin, M. B.; Prosser, R. S.; Stewart, B. A. Effect of Juxtamembrane Tryptophans on the Immersion Depth of Synaptobrevin, an Integral Vesicle Membrane Protein. *Biochim. Biophys. Acta, Biomembr.* **2012**, *1818*, 2994−2999.
(351) Herbst, D. A.; Boll, B.; Zocher, G.; Stehle, T.; Heide, L. Structural Basis of the Interaction of MbtH-like Proteins, Putative Regulators of Nonribosomal Peptide Biosynthesis, with Adenylating Enzymes. *J. Biol. Chem.* **2013**, *288*, 1991−2003.
(352) Chen, Y.; Barkley, M. D. Toward Understanding Tryptophan Fluorescence in Proteins. *Biochemistry* **1998**, *37*, 9976−9982.
(353) Vivian, J. T.; Callis, P. R. Mechanisms of Tryptophan Fluorescence Shifts in Proteins. *Biophys. J.* **2001**, *80*, 2093−2109.
(354) MEYBECK, A.; WINDLE, J. J. An EPR Study of Peptides after UV Irradiation. *Photochem. Photobiol.* **1969**, *10*, 1−12.
(355) SANTUS, R.; GROSSWEINER, L. I. Primary Products in the Flash Photolysis of Tryptophan*. *Photochem. Photobiol.* **1972**, *15*, 101−105.
(356) Ananthaswamy, H. N.; Eisenstark, A. Near-UV-induced Breaks in Phage DNA: Sensitization by Hydrogen Peroxide (a Tryptophan Photoproduct). *Photochem. Photobiol.* **1976**, *24*, 439−442.
(357) Szajdzinska-Pietek, E.; Bednarek, J.; Plonka, A. Electron Spin Resonance Studies on Tryptophan Photolysis in Frozen Micellar Systems of Anionic Surfactants. *J. Photochem. Photobiol., A* **1993**, *75*, 131−136.
(358) Angiolillo, P. J.; Vanderkooi, J. M. Hydrogen Atoms Are Produced When Tryptophan within a Protein Is Irradiated with Ultraviolet Light. *Photochem. Photobiol.* **1996**, *64*, 492−495.
(359) Barnes, S.; Shonsey, E. M.; Eliuk, S. M.; Stella, D.; Barrett, K.; Srivastava, Om P.; Kim, H.; Renfrow, M. B. High-Resolution Mass Spectrometry Analysis of Protein Oxidations and Resultant Loss of Function. *Biochem. Soc. Trans.* **2008**, *36*, 1037−1044.
(360) Freinbichler, W.; Colivicchi, M. A.; Stefanini, C.; Bianchi, L.; Ballini, C.; Misini, B.; Weinberger, P.; Linert, W.; Varešlija, D.; et al. Highly Reactive Oxygen Species: Detection, Formation, and Possible Functions. *Cell. Mol. Life Sci.* **2011**, *68*, 2067−2079.
(361) Previero, A.; Coletti-Previero, M. A.; Cavadore, J.-C. A Reversible Chemical Modification of the Tryptophan Residue. *Biochim. Biophys. Acta, Protein Struct.* **1967**, *147*, 453−461.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(362) Foettinger, A.; Melmer, M.; Leitner, A.; Lindner, W. Reaction of the Indole Group with Malondialdehyde: Application for the Derivatization of Tryptophan Residues in Peptides. *Bioconjugate Chem.* **2007**, *18*, 1678−1683.
(363) Sturm, M.; Leitner, A.; Lindner, W. Development of an Indole-Based Chemically Cleavable Linker Concept for Immobilizing Bait Compounds for Protein Pull-Down Experiments. *Bioconjugate Chem.* **2011**, *22*, 211−217.
(364) Koshland, D. E.; Karkhanis, Y. D.; Latham, H. G. An Environmentally-Sensitive Reagent with Selectivity for the Tryptophan Residue in Proteins. *J. Am. Chem. Soc.* **1964**, *86*, 1448−1450.
(365) Horton, H. R.; Koshland, D. E. A Highly Reactive Colored Reagent with Selectivity for the Tryptophan Residue in Proteins. 2-Hydroxy-5-nitrobenzyl Bromide1. *J. Am. Chem. Soc.* **1965**, *87*, 1126−1132.
(366) Loudon, G. M.; Koshland, D. E. The Chemistry of a Reporter Group: 2-Hydroxy-5-nitrobenzyl Bromide. *J. Biol. Chem.* **1970**, *245*, 2247−2254.
(367) Horton, H. R.; Tucker, W. P. Dimethyl(2-hydroxy-5-nitrobenzyl)sulfonium Salts: Water-Soluble Environmentally Sensitive Protein Reagents. *J. Biol. Chem.* **1970**, *245*, 3397−3401.
(368) Horton, H. R.; Kelly, H.; Koshland, D. E. Environmentally Sensitive Protein Reagents: 2-methoxy-5-nitrobenzyl bromide. *J. Biol. Chem.* **1965**, *240*, 722−724.
(369) Robert Horton, H.; Young, G. 2-Acetoxy-5-nitrobenzyl chloride: A Reagent Designed to Introduce a Reporter Group near the Active Site of Chymotrypsin. *Biochim. Biophys. Acta, Protein Struct.* **1969**, *194*, 272−278.
(370) Strohalm, M.; Kodíček, M.; Pechar, M. Tryptophan Modification by 2-hydroxy-5-nitrobenzyl Bromide Studied by MALDI-TOF Mass Spectrometry. *Biochem. Biophys. Res. Commun.* **2003**, *312*, 811−816.
(371) Strohalm, M.; Šantrůček, J.; Hynek, R.; Kodíček, M. Analysis of Tryptophan Surface Accessibility in Proteins by MALDI-TOF Mass Spectrometry. *Biochem. Biophys. Res. Commun.* **2004**, *323*, 1134−1138.
(372) Peyser, Y. M.; Muhlrad, A.; Werber, M. M. Tryptophan-130 is the most reactive Tryptophan Residue in Rabbit Skeletal Myosin Subfragment-1. *FEBS Lett.* **1990**, *259*, 346−348.
(373) Xu, Y.; Strickland, E. C.; Fitzgerald, M. C. Thermodynamic Analysis of Protein Folding and Stability Using a Tryptophan Modification Protocol. *Anal. Chem.* **2014**, *86*, 7041−7048.
(374) Borotto, N. B.; Zhang, Z.; Dong, J.; Burant, B.; Vachet, R. W. Increased β-Sheet Dynamics and D−E Loop Repositioning Are Necessary for Cu(II)-Induced Amyloid Formation by β-2-Microglobulin. *Biochemistry* **2017**, *56*, 1095−1104.
(375) Patchornik, A.; Lawson, W. B.; Witkop, B. Selective Cleavage of Peptide Bonds. ii. The Tryptophyl Peptide Bond and the Cleavage of Glucagon1. *J. Am. Chem. Soc.* **1958**, *80*, 4747−4748.
(376) Spande, T. F.; Witkop, B. Determination of the Tryptophan Content of Proteins with N-Bromosuccinimide. In *Methods in Enzymology*; Hirs, C. H. W., Ed.; Academic Press: New York, 1967; Vol. *11*, pp 498−506, DOI: 10.1016/S0076-6879(67)11060-4
(377) Takahashi, T.; Hiramoto, S.; Wato, S.-I.; Nishimoto, T.; Wada, Y.; Nagai, K.; Yamaguchi, H. Identification of Essential Amino Acid Residues of an α-Amylase Inhibitor from Phaseolus vulgaris White Kidney Beans1. *J. Biochem.* **1999**, *126*, 838−844.
(378) Padrines, M.; Rabaud, M.; Bieth, J. G. Oxidized α1-Proteinase-Formula Inhibitor: A Fast-Acting Inhibitor of Human Pancreatic Elastase. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1992**, *1118*, 174−178.
(379) Lischwe, M. A.; Sung, M. T. Use of N-Chlorosuccinimide/Urea for the Selective Cleavage of Tryptophanyl Peptide Bonds in Proteins. Cytochrome C. *J. Biol. Chem.* **1977**, *252*, 4976−4980.
(380) Smith, B. J. Chemical Cleavage of Proteins at Tryptophan Residues. In *The Protein Protocols Handbook*; Walker, J. M., Ed.; Humana Press: Totowa, NJ, 1996; pp 375−380, DOI: 10.1007/978-1-60327-259-9_64.
(381) Scoffone, E.; Fontana, A.; Rocchi, R. Sulfenyl Halides as Modifying Reagents for Polypeptides and Proteins. I. Modification of tryptophan residues. *Biochemistry* **1968**, *7*, 971−979.
(382) Wilchek, M.; Miron, T. The Conversion of Tryptophan to 2-Thioltryptophan in Peptides and Proteins. *Biochem. Biophys. Res. Commun.* **1972**, *47*, 1015−1020.
(383) Zhang, J.-G.; Reid, G. E.; Moritz, R. L.; Ward, L. D.; Simpson, R. J. Specific Covalent Modification of the Tryptophan Residues in Murine Interleukin-6. *Eur. J. Biochem.* **1993**, *217*, 53−59.
(384) Zhang, Z.; Ostanin, K.; Van Etten, R. L. Covalent Modification and Site-Directed Mutagenesis of an Active Site Tryptophan of Human Prostatic Acid Phosphatase. *Acta Biochim. Polym. Eng. Ed.* **1997**, *44*, 659−672.
(385) Spande, T. F.; Green, N. M.; Witkop, B. The Reactivity toward N-Bromosuccinimide of Tryptophan in Enzymes, Zymogens, and Inhibited Enzymes*. *Biochemistry* **1966**, *5*, 1926−1933.
(386) Inokuchi, N.; Takahashi, T.; Yoshimoto, A.; Irie, M. N-Bromosuccinimide Oxidation of a Glucoamylase from Aspergillus saitoi. *J. Biochem.* **1982**, *91*, 1661−1668.
(387) Ohnishi, M.; Kawagishi, T.; Hiromi, K. Stopped-Flow Chemical Modification with N-Bromosuccinimide: A Good Probe for Changes in the Microenvironment of the Trp 62 Residue of Chicken Egg White Lysozyme. *Arch. Biochem. Biophys.* **1989**, *272*, 46−51.
(388) Jiang, Y.; Ma, D. Regioselective Acylation at the 5- or 6-Position of l-Tryptophan Derivatives. *Tetrahedron Lett.* **2002**, *43*, 7013−7015.
(389) Hawkins, C. L.; Pattison, D. I.; Stanley, N. R.; Davies, M. J. Tryptophan Residues are Targets in Hypothiocyanous Acid-Mediated Protein Oxidation. *Biochem. J.* **2008**, *416*, 441−452.
(390) Edwards, R. A.; Jickling, G.; Turner, R. J. The Light-Induced Reactions of Tryptophan with Halocompounds. *Photochem. Photobiol.* **2002**, *75*, 362−368.
(391) Rall, S. C.; Cole, R. D. Amino Acid Sequence and Sequence Variability of the Amino-terminal Regions of Lysine-rich Histones. *J. Biol. Chem.* **1971**, *246*, 7175−7190.
(392) Bogan, A. A.; Thorn, K. S. Anatomy of Hot Spots in Protein Interfaces. *J. Mol. Biol.* **1998**, *280*, 1−9.
(393) Abello, N.; Kerstjens, H. A. M.; Postma, D. S.; Bischoff, R. Protein Tyrosine Nitration: Selectivity, Physicochemical and Biological Consequences, Denitration, and Proteomics Methods for the Identification of Tyrosine-Nitrated Proteins. *J. Proteome Res.* **2009**, *8*, 3222−3238.
(394) Castro, L.; Demicheli, V.; Tórtora, V.; Radi, R. Mitochondrial Protein Tyrosine Nitration. *Free Radical Res.* **2011**, *45*, 37−52.
(395) Aslan, M.; Dogan, S. Proteomic Detection of Nitroproteins as Potential Biomarkers for Cardiovascular Disease. *J. Proteomics* **2011**, *74*, 2274−2288.
(396) Radi, R. Protein Tyrosine Nitration: Biochemical Mechanisms and Structural Basis of Functional Effects. *Acc. Chem. Res.* **2013**, *46*, 550−559.
(397) Monigatti, F.; Hekking, B.; Steen, H. Protein Sulfation Analysis—A Primer. *Biochim. Biophys. Acta, Proteins Proteomics* **2006**, *1764*, 1904−1913.
(398) Stone, M. J.; Chuang, S.; Hou, X.; Shoham, M.; Zhu, J. Z. Tyrosine Sulfation: An Increasingly Recognised Post-Translational Modification of Secreted Proteins. *New Biotechnol.* **2009**, *25*, 299−317.
(399) Machida, K.; Mayer, B. J.; Nollau, P. Profiling the Global Tyrosine Phosphorylation State. *Mol. Cell. Proteomics* **2003**, *2*, 215−233.
(400) Leitner, A.; Sturm, M.; Lindner, W. Tools for Analyzing the Phosphoproteome and other Phosphorylated Biomolecules: A review. *Anal. Chim. Acta* **2011**, *703*, 19−30.
(401) Dushek, O.; Goyette, J.; van der Merwe, P. A. Non-Catalytic Tyrosine-Phosphorylated Receptors. *Immunol. Rev.* **2012**, *250*, 258−276.
(402) Hunter, T. Why Nature Chose Phosphate to Modify Proteins. *Philos. Trans. R. Soc., B* **2012**, *367*, 2513−2516.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(403) Johnson, T. B.; Kohmann, E. F. Studies on Nitrated Proteins: I. the Determination of the Structure of Nitrotyrosine. *J. Am. Chem. Soc.* **1915**, *37*, 1863−1884.
(404) Wormall, A. The Immunological Specificity of Chemically Altered Proteins: Halogenated and Nitrated Proteins. *J. Exp. Med.* **1930**, *51*, 295−317.
(405) Riordan, J. F.; Sokolovsky, M.; Vallee, B. L. Tetranitromethane. A Reagent for the Nitration of Tyrosine and Tyrosyl Residues of Proteins. *J. Am. Chem. Soc.* **1966**, *88*, 4104−4105.
(406) Sokolovsky, M.; Harell, D.; Riordan, J. F. Reaction of Tetranitromethane with Sulfhydryl Groups in Proteins. *Biochemistry* **1969**, *8*, 4740−4745.
(407) Riordan, J. F.; Vallee, B. L. Nitration with Tetranitromethane. In *Methods in Enzymology*; Hirs, C. H. W., Timasheff, S. N., Eds.; Academic Press: New York, 1972; Vol. *25*, pp 515−521, DOI: 10.1016/S0076-6879(72)25048-0.
(408) Rabani, J.; Mulac, W. A.; Matheson, M. S. The Pulse Radiolysis of Aqueous Tetranitromethane. I. Rate Constants and the Extinction Coefficient of e_{aq}^-. II. Oxygenated Solutions. *J. Phys. Chem.* **1965**, *69*, 53−70.
(409) Cuatrecasas, P.; Fuchs, S.; Anfinsen, C. B. The Tyrosyl Residues at the Active Site of Staphylococcal Nuclease: Modifications by Tetranitromethane. *J. Biol. Chem.* **1968**, *243*, 4787−4798.
(410) Palamalai, V.; Miyagi, M. Mechanism of Glyceraldehyde-3-Phosphate Dehydrogenase Inactivation by Tyrosine Nitration. *Protein Sci.* **2010**, *19*, 255−262.
(411) Doyle, R. J.; Bello, J.; Roholt, O. A. Probable Protein Crosslinking with Tetranitromethane. Biochim. Biophys. *Biochim. Biophys. Acta, Protein Struct.* **1968**, *160*, 274−276.
(412) Boesel, R. W.; Carpenter, F. H. Crosslinking during the Nitration of Bovine Insulin with Tetranitromethane. *Biochem. Biophys. Res. Commun.* **1970**, *38*, 678−682.
(413) Kunkel, G. R.; Mehrabian, M.; Martinson, H. G. Contact-Site Cross-Linking Agents. *Mol. Cell. Biochem.* **1981**, *34*, 3−13.
(414) Cutfield, S. M.; Dodson, G. G.; Ronco, N.; Cutfield, J. F. Preparation and Activity of Nitrated Insulin Dimer. *Int. J. Pept. Protein Res.* **1986**, *27*, 335−343.
(415) Ploug, M.; Rahbek-Nielsen, H.; Ellis, V.; Roepstorff, P.; Dano, K. Chemical Modification of the Urokinase-Type Plasminogen Activator and Its Receptor Using Tetranitromethane. Evidence for the Involvement of Specific Tyrosine Residues in Both Molecules during Receptor-Ligand Interaction. *Biochemistry* **1995**, *34*, 12524−12534.
(416) Šantrůček, J.; Strohalm, M.; Kadlčík, V.; Hynek, R.; Kodíček, M. Tyrosine Residues Modification Studied by MALDI-TOF Mass Spectrometry. *Biochem. Biophys. Res. Commun.* **2004**, *323*, 1151−1156.
(417) Greenacre, S. A. B.; Ischiropoulos, H. Tyrosine Nitration: Localisation, Quantification, Consequences for Protein Function and Signal Transduction. *Free Radical Res.* **2001**, *34*, 541−581.
(418) Ischiropoulos, H. Biological Selectivity and Functional Aspects of Protein Tyrosine Nitration. *Biochem. Biophys. Res. Commun.* **2003**, *305*, 776−783.
(419) Lennon, C. W.; Cox, H. D.; Hennelly, S. P.; Chelmo, S. J.; McGuirl, M. A. Probing Structural Differences in Prion Protein Isoforms by Tyrosine Nitration. *Biochemistry* **2007**, *46*, 4850−4860.
(420) Matters, D.; Cooper, H. J.; McDonnell, L.; Iniesta, J.; Heptinstall, J.; Derrick, P.; Walton, D.; Peterson, I. Mass spectrometry in Demonstrating the Site-Specific Nitration of Hen Egg White Lysozyme by an Improved Electrochemical Method. *Anal. Biochem.* **2006**, *356*, 171−181.
(421) Qiao, L.; Lu, Y.; Liu, B.; Girault, H. H. Copper-Catalyzed Tyrosine Nitration. *J. Am. Chem. Soc.* **2011**, *133*, 19823−19831.
(422) Chapman, E. M.; Skanse, B. N.; Evans, R. D. Treatment of Hyperthyroidism with Radioactive Iodine. *Radiology* **1948**, *51*, 558−563.
(423) Banks, W. A.; Niehoff, M. L.; Ponzio, N. M.; Erickson, M. A.; Zalcman, S. S. Pharmacokinetics and Modeling of Immune Cell Trafficking: quantifying differential influences of Target Tissues Versus Lymphocytes in SJL and Lipopolysaccharide-Treated Mice. *J. Neuroinflammation* **2012**, *9*, 231.
(424) Hunter, W. M.; Greenwood, F. C. Preparation of Iodine-131 Labelled Human Growth Hormone of High Specific Activity. *Nature* **1962**, *194*, 495−496.
(425) Greenwood, F.; Hunter, W.; Glover, J. The Preparation of 131I-Labelled Human Growth Hormone of High Specific Radioactivity. *Biochem. J.* **1963**, *89*, 114−123.
(426) McConahey, P. J.; Dixon, F. J. Radioiodination of Proteins by the use of the Chloramine-T Method. In *Methods in Enzymology*; Vunakis, H. V., Langone, J. J., Eds.; Academic Press: New York, 1980; Vol. *70*, pp 210−213, DOI: 10.1016/S0076-6879(80)70050-2.
(427) McFarlane, A. S. Efficient Trace-labelling of Proteins with Iodine. *Nature* **1958**, *182*, 53−54.
(428) Angeles Contreras, M.; Bale, W. F.; Spar, I. L. Iodine Monochloride (ICI) Iodination Techniques. In *Methods in Enzymology*; Langone, J. J., Vunakis, H. V., Eds.; Academic Press: New York, 1983; Vol. *92*, pp 277−292, DOI: 10.1016/0076-6879(83)92025-6.
(429) Salacinski, P. R. P.; McLean, C.; Sykes, J. E. C.; Clement-Jones, V. V.; Lowry, P. J. Iodination of Proteins, Glycoproteins, and Peptides using a Solid-Phase Oxidizing Agent, 1,3,4,6-Tetrachloro-3α,6α-Diphenyl Glycoluril (Iodogen). *Anal. Biochem.* **1981**, *117*, 136−146.
(430) Marchalonis, J. J. An Enzymic Method for the Trace Iodination of Immunoglobulins and other Proteins. *Biochem. J.* **1969**, *113*, 299−305.
(431) Liou, Y.-C.; Davies, P. L.; Jia, Z. Crystallization and Preliminary X-Ray Analysis of Insect Antifreeze Protein from the Beetle Tenebrio Molitor. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2000**, *56*, 354−356.
(432) Koshland, M. E.; Englberger, F. M.; Erwin, M. J.; Gaddone, S. M. Modification of Amino Acid Residues in Anti-p-azobenzenearsonic Acid Antibody during Extensive Iodination. *J. Biol. Chem.* **1963**, *238*, 1343−1348.
(433) Hobba, G. D.; Forbes, B. E.; Parkinson, E. J.; Francis, G. L.; Wallace, J. C. The Insulin-like Growth Factor (IGF) Binding Site of Bovine Insulin-like Growth Factor Binding Protein-2 (bIGFBP-2) Probed by Iodination. *J. Biol. Chem.* **1996**, *271*, 30529−30536.
(434) Simpson, R. T.; Riordan, J. F.; Vallee, B. L. Functional Tyrosyl Residues in the Active Center of Bovine Pancreatic Carboxypeptidase A. *Biochemistry* **1963**, *2*, 616−622.
(435) Riordan, J. F.; Wacker, W. E. C.; Vallee, B. L. N-Acetylimidazole: A Reagent for Determination of "Free" Tyrosyl Residues of Proteins. *Biochemistry* **1965**, *4*, 1758−1765.
(436) Houston, L. L.; Walsh, K. Transient Inactivation of Trypsin by Mild Acetylation with N-acetylimidazole. *Biochemistry* **1970**, *9*, 156−166.
(437) Riordan, J. F.; Vallee, B. L. O-Acetyltyrosine. In *Methods in Enzymology*; Hirs, C. H. W., Timasheff, S. N., Eds.; Academic Press: New York, 1972; Vol. *25*, pp 500−506, DOI: 10.1016/S0076-6879(72)25046-7.
(438) Koltun, W. L.; Dexter, R. N.; Clark, R. E.; Gurd, F. R. N. Coördination Complexes and Catalytic Properties of Proteins and Related Substances. I. Effect of Cupric and Zinc Ions on the Hydrolysis of p-Nitrophenyl Acetate by Imidazole1. *J. Am. Chem. Soc.* **1958**, *80*, 4188−4194.
(439) Fife, T. H.; Natarajan, R.; Werner, M. H. Effect of the Leaving Group in the Hydrolysis of N-acylimidazoles. The Hydroxide Ion, Water, and General-Base Catalyzed Hydrolysis of N-acyl-4(5)-Nitroimidazoles. *J. Org. Chem.* **1987**, *52*, 740−746.
(440) Zappacosta, F.; Ingallinella, P.; Scaloni, A.; Pessi, A.; Bianchi, E.; Sollazzo, M.; Tramontano, A.; Marino, G.; Pucci, P. Surface topology of Minibody by Selective Chemical Modifications and Mass Spectrometry. *Protein Sci.* **1997**, *6*, 1901−1909.
(441) Irie, M.; Miyasaka, T.; Arakawa, K. 3-Acetoxy-1-acetyl-5-methylpyrazole, a Novel Acetylating Reagent of Protein. *J. Biochem.* **1972**, *72*, 65−72.
(442) Adackaparayil, M.; Smith, J. H. Preparation and Reactivity of a New Spin Label Reagent. *J. Org. Chem.* **1977**, *42*, 1655−1656.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(443) Fraenkel-Conrat, H.; Colloms, M. Reactivity of Tobacco Mosaic Virus and its Protein toward Acetic Anhydride. *Biochemistry* **1967**, *6*, 2740−2745.
(444) Kurihara, K.; Horinishi, H.; Shibata, K. Reactions of Cyanuric Halides with Proteins I. Bound Tyrosine Residues of Insulin and Lysozyme as Identified with Cyanuric Fluoride. *Biochim. Biophys. Acta* **1963**, *74*, 678−687.
(445) Gorbunoff, M. J. Cyanuration. In *Methods in Enzymology*; Hirs, C. H. W., Timasheff, S. N., Eds.; Academic Press: New York, 1972; Vol. *25*, pp 506−514, DOI: 10.1016/S0076-6879(72)25047-9.
(446) Riordan, J. F.; Vallee, B. L. Diazonium Salts as Specific Reagents and Probes of Protein Conformation. In *Methods in Enzymology*; Hirs, C. H. W., Timasheff, S. N., Eds.; Academic Press: New York, 1972; Vol. *25*, pp 521−531, DOI: 10.1016/S0076-6879(72)25049-2.
(447) Jones, M. W.; Mantovani, G.; Blindauer, C. A.; Ryan, S. M.; Wang, X.; Brayden, D. J.; Haddleton, D. M. Direct Peptide Bioconjugation/PEGylation at Tyrosine with Linear and Branched Polymeric Diazonium Salts. *J. Am. Chem. Soc.* **2012**, *134*, 7406−7413.
(448) Qu, N.; Li, F.; Shao, B.; Shao, J.; Zhai, G.; Wang, F.; Zhu, B.-Z. The Unexpected and Exceptionally Facile Chemical Modification of the Phenolic Hydroxyl Group of Tyrosine by Polyhalogenated Quinones under Physiological Conditions. *Chem. Res. Toxicol.* **2016**, *29*, 1699−1705.
(449) Liao, T. H.; Ting, R. S.; Yeung, J. E. Reactivity of Tyrosine in bovine pancreatic Deoxyribonuclease with p-Nitrobenzenesulfonyl Fluoride. *J. Biol. Chem.* **1982**, *257*, 5637−5644.
(450) Gitlin, G.; Bayer, E. A.; Wilchek, M. Studies on the Biotin-Binding Sites of Avidin and Streptavidin. Tyrosine Residues are Involved in the Binding Site. *Biochem. J.* **1990**, *269*, 527−530.
(451) Murachi, T.; Inagami, T.; Yasui, M. Evidence for Alkylphosphorylation of Tyrosyl Residues of Stem Bromelain by Diisopropylphosphorofluoridate. *Biochemistry* **1965**, *4*, 2815−2825.
(452) Means, G. E.; Wu, H.-L. The Reactive Tyrosine Residue of Human Serum Albumin: Characterization of its Reaction with Diisopropylfluorophosphate. *Arch. Biochem. Biophys.* **1979**, *194*, 526−530.
(453) Awad-Elkarim, A.; Means, G. E. The Reactivity of p-Nitrophenyl Acetate with Serum Albumins. *Comp. Biochem. Physiol. B* **1988**, *91*, 267−272.
(454) Sato, S.; Nakamura, K.; Nakamura, H. Tyrosine-Specific Chemical Modification with in Situ Hemin-Activated Luminol Derivatives. *ACS Chem. Biol.* **2015**, *10*, 2633−2640.
(455) Fraenkel-Conrat, H.; Olcott, H. S. Esterification of Proteins with Alcohols of Low Molecular Weight. *J. Biol. Chem.* **1945**, *161*, 259−268.
(456) Erlanger, B. F.; Vratsanos, S. M.; Wassermann, N.; Cooper, A. G. A Chemical Investigation of the Active Center of Pepsin. *Biochem. Biophys. Res. Commun.* **1966**, *23*, 243−245.
(457) Gross, E.; Morell, J. L. Evidence for an Active Carboxyl Group in Pepsin. *J. Biol. Chem.* **1966**, *241*, 3638−3639.
(458) Kemp, D. S.; Chien, S. W. A New Peptide Coupling Reagent. *J. Am. Chem. Soc.* **1967**, *89*, 2743−2745.
(459) Chibnall, A. C.; Mangan, J. L.; Rees, M. W. Studies on the Amide and C-terminal Residues in Proteins. 3. The Esterification of Proteins. *Biochem. J.* **1958**, *68*, 114−118.
(460) Doscher, M. S.; Wilcox, P. E. Chemical Derivatives of α-Chymotrypsinogen: IV. A Comparison of the Reactions of α-Chymotrypsinogen and of Simple Carboxylic Acids with Diazoacetamide. *J. Biol. Chem.* **1961**, *236*, 1328−1337.
(461) Delpierre, G. R.; Fruton, J. S. Inactivation of Pepsin by Diphenyldiazomethane. *Proc. Natl. Acad. Sci. U. S. A.* **1965**, *54*, 1161−1167.
(462) Rajagopalan, T. G.; Stein, W. H.; Moore, S. The Inactivation of Pepsin by Diazoacetylnorleucine Methyl Ester. *J. Biol. Chem.* **1966**, *241*, 4295−4297.
(463) Mix, K. A.; Raines, R. T. Optimized Diazo Scaffold for Protein Esterification. *Org. Lett.* **2015**, *17*, 2358−2361.
(464) Woodward, R. B.; Olofson, R. A.; Mayer, H. A New Synthesis of Peptides. *J. Am. Chem. Soc.* **1961**, *83*, 1010−1012.
(465) Bodlaender, P.; Feinstein, G.; Shaw, E. Use of Isoxazolium Salts for Carboxyl Group Modification in Proteins. Trypsin. *Biochemistry* **1969**, *8*, 4941−4949.
(466) Feinstein, G.; Bodlaender, P.; Shaw, E. Modification of essential carboxylic Acid Side Chains of Trypsin. *Biochemistry* **1969**, *8*, 4949−4955.
(467) Komissarov, A. A.; Romanova, D. V.; Debabov, V. G. Complete Inactivation of Escherichia coli Uridine Phosphorylase by Modification of Asp5 with Woodward's Reagent K. *J. Biol. Chem.* **1995**, *270*, 10050−10055.
(468) Martín-Gago, P.; Fansa, E. K.; Winzker, M.; Murarka, S.; Janning, P.; Schultz-Fademrecht, C.; Baumann, M.; Wittinghofer, A.; Waldmann, H. Covalent Protein Labeling at Glutamic Acids. *Cell Chem. Biol.* **2017**, *24*, 589−597 e585..
(469) Sheehan, J. C.; Hlavka, J. J. The Use of Water-Soluble and Basic Carbodiimides in Peptide Synthesis. *J. Org. Chem.* **1956**, *21*, 439−441.
(470) Franzblau, C.; Gallop, P. M.; Seifter, S. The Presence in Collagen of γ-Glutamyl Peptide Linkages. *Biopolymers* **1963**, *1*, 79−97.
(471) Goodfriend, T. L.; Levine, L.; Fasman, G. D. Antibodies to Bradykinin and Angiotensin: A Use of Carbodiimides in Immunology. *Science* **1964**, *144*, 1344−1346.
(472) Riehm, J. P.; Scheraga, H. A. Structural Studies of Ribonuclease. XXI. The Reaction between Ribonuclease and a Water-Soluble Carbodiimide. *Biochemistry* **1966**, *5*, 99−115.
(473) Hoare, D. G.; Koshland, D. E. A Procedure for the Selective Modification of Carboxyl Groups in Proteins. *J. Am. Chem. Soc.* **1966**, *88*, 2057−2058.
(474) Hoare, D. G.; Koshland, D. E. A Method for the Quantitative Modification and Estimation of Carboxylic Acid Groups in Proteins. *J. Biol. Chem.* **1967**, *242*, 2447−2453.
(475) Carraway, K. L.; Koshland, D. E. Carbodiimide modification of proteins. In *Methods in Enzymology*; Hirs, C. H. W., Timasheff, S. N., Eds.; Academic Press: New York, 1972; Vol. *25*, pp 616−623, DOI: 10.1016/S0076-6879(72)25060-1.
(476) Akashi, S.; Niitsu, U.; Yuji, R.; Ide, H.; Hirayama, K. Investigation of the Interaction between Enzyme and Inhibitor by the Combination of Chemical Modification, Electrospray Ionization Mass Spectrometry and Frit-Fast Atom Bombardment liquid Chromatography/Mass Spectrometry. *Biol. Mass Spectrom.* **1993**, *22*, 124−132.
(477) Sanderson, R. J.; Mosbaugh, D. W. Identification of Specific Carboxyl Groups on Uracil-DNA Glycosylase Inhibitor Protein That Are Required for Activity. *J. Biol. Chem.* **1996**, *271*, 29170−29181.
(478) Wen, J.; Zhang, H.; Gross, M. L.; Blankenship, R. E. Membrane Orientation of the FMO Antenna Protein from Chlorobaculum Tepidum as Determined by Mass Spectrometry-Based Footprinting. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 6134−6139.
(479) Zhang, H.; Liu, H.; Blankenship, R. E.; Gross, M. L. Isotope-Encoded Carboxyl Group Footprinting for Mass Spectrometry-Based Protein Conformational Studies. *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 178−181.
(480) Guo, C.; Cheng, M.; Gross, M. L. Protein-Metal-Ion Interactions Studied by Mass Spectrometry-Based Footprinting with Isotope-Encoded Benzhydrazide. *Anal. Chem.* **2019**, *91*, 1416−1423.
(481) Prell, J. S.; O'Brien, J. T.; Steill, J. D.; Oomens, J.; Williams, E. R. Structures of Protonated Dipeptides: The Role of Arginine in Stabilizing Salt Bridges. *J. Am. Chem. Soc.* **2009**, *131*, 11442−11449.
(482) Donald, J. E.; Kulp, D. W.; DeGrado, W. F. Salt bridges: Geometrically Specific, Designable Interactions. *Proteins: Struct., Funct., Genet.* **2011**, *79*, 898−915.
(483) Migliori, V.; Phalke, S.; Bezzi, M.; Guccione, E. Arginine/Lysine−Methyl/Methyl Switches: Biochemical Role of Histone Arginine Methylation in Transcriptional Regulation. *Epigenomics* **2010**, *2*, 119−137.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(484) Boisvert, F.-M.; Chénard, C. A.; Richard, S. Protein Interfaces in Signaling Regulated by Arginine Methylation. *Sci. Signaling* **2005**, *2005*, re2.
(485) Fitch, C. A.; Platzer, G.; Okon, M.; Garcia-Moreno E, B.; McIntosh, L. P. Arginine: Its pKa Value Revisited. *Protein Sci.* **2015**, *24*, 752−761.
(486) Takahashi, K. The Reaction of Phenylglyoxal with Arginine Residues in Proteins. *J. Biol. Chem.* **1968**, *243*, 6171−6179.
(487) Riordan, J. F. Arginyl Residues and Anion Binding Sites in Proteins. *Mol. Cell. Biochem.* **1979**, *26*, 71−92.
(488) Cheung, S.-T.; Fonda, M. L. Reaction of Phenylglyoxal with Arginine. The Effect of Buffers and pH. *Biochem. Biophys. Res. Commun.* **1979**, *90*, 940−947.
(489) Eun, H. M.; Miles, E. W. Reaction of Phenylglyoxal and {p-Hydroxyphenyl} Glyoxal with Arginines and Cysteines in the. Alpha. Subunit of Tryptophan Synthase. *Biochemistry* **1984**, *23*, 6484−6491.
(490) Borders, C. L.; Pearson, L. J.; McLaughlin, A. E.; Gustafson, M. E.; Vasiloff, J.; An, F. Y.; Morgan, D. J. 4-Hydroxy-3-Nitrophenylglyoxal. A Chromophoric Reagent for Arginyl Residues in Proteins. *Biochim. Biophys. Acta Enzymol.* **1979**, *568*, 491−495.
(491) Yamasaki, R. B.; Shimer, D. A.; Feeney, R. E. Colorimetric Determination of Arginine Residues in Proteins by p-Nitrophenylglyoxal. *Anal. Biochem.* **1981**, *111*, 220−226.
(492) Politz, S. M.; Noller, H. F.; McWhirter, P. D. Ribonucleic Acid-Protein Cross-Linking in Escherichia Coli Ribosomes: (4-Azidophenyl) Glyoxal, a Novel Heterobifunctional Reagent. *Biochemistry* **1981**, *20*, 372−378.
(493) Krell, T.; Pitt, A. R.; Coggins, J. R. The Use of Electrospray Mass Spectrometry to Identify an Essential Arginine Residue in Type II dehydroquinases. *FEBS Lett.* **1995**, *360*, 93−96.
(494) Wood, T. D.; Guan, Z.; Borders, C. L.; Chen, L. H.; Kenyon, G. L.; McLafferty, F. W. Creatine Kinase: Essential Arginine Residues at the Nucleotide Binding Site Identified by Chemical Modification and High-Resolution Tandem Mass Spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 3362−3365.
(495) Yankeelov, J. A.; Mitchell, C. D.; Crawford, T. H. Simple Trimerization of 2,3-Butanedione Yielding a Selective Reagent for the Modification of Arginine in Proteins. *J. Am. Chem. Soc.* **1968**, *90*, 1664−1666.
(496) Riordan, J. F. Functional Arginyl Residues in Carboxypeptidase A. Modification with Butanedione. *Biochemistry* **1973**, *12*, 3915−3923.
(497) Leitner, A.; Lindner, W. Probing of Arginine Residues in Peptides and Proteins Using Selective Tagging and Electrospray Ionization Mass Spectrometry. *J. Mass Spectrom.* **2003**, *38*, 891−899.
(498) Leitner, A.; Lindner, W. Effects of an Arginine-Selective Tagging Procedure on the Fragmentation Behavior of Peptides Studied by Electrospray Ionization Tandem Mass Spectrometry (ESI-MS/MS). *Anal. Chim. Acta* **2005**, *528*, 165−173.
(499) Leitner, A.; Lindner, W. Functional Probing of Arginine Residues in Proteins Using Mass Spectrometry and an Arginine-Specific Covalent Tagging Concept. *Anal. Chem.* **2005**, *77*, 4481−4488.
(500) Leitner, A.; Amon, S.; Rizzi, A.; Lindner, W. Use of the Arginine-Specific Butanedione/Phenylboronic Acid Tag for Analysis of peptides and Protein Digests Using Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 1321−1330.
(501) Stensland, M.; Holm, A.; Kiehne, A.; Fleckenstein, B. Targeted Analysis of Protein Citrullination using Chemical Modification and Tandem Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 2754−2762.
(502) De Ceuleneer, M.; De Wit, V.; Van Steendam, K.; Van Nieuwerburgh, F.; Tilleman, K.; Deforce, D. Modification of Citrulline Residues with 2,3-butanedione Facilitates Their Detection by Liquid Chromatography/Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2011**, *25*, 1536−1542.
(503) Toi, K.; Bynum, E.; Norris, E.; Itano, H. A. Studies on the Chemical Modification of Arginine: I. The Reaction of 1,2-Cyclohexanedione with Arginine and Arginyl Residues of Proteins. *J. Biol. Chem.* **1967**, *242*, 1036−1043.
(504) Patthy, L.; Smith, E. L. Reversible modification of arginine residues. Application to Sequence Studies by Restriction of Tryptic Hydrolysis to Lysine Residues. *J. Biol. Chem.* **1975**, *250*, 557−564.
(505) Patthy, L.; Smith, E. L. Identification of Functional Arginine Residues in Ribonuclease A and Lysozyme. *J. Biol. Chem.* **1975**, *250*, 565−569.
(506) Suckau, D.; Mak, M.; Przybylski, M. Protein Surface Topology-Probing by Selective Chemical Modification and Mass Spectrometric Peptide Mapping. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 5630−5634.
(507) Litt, M. Structural Studies on Transfer Ribonucleic Acid. I. Labeling of Exposed Guanine Sites in Yeast Phenylalanine Transfer Ribonucleic Acid with Kethoxal. *Biochemistry* **1969**, *8*, 3249−3253.
(508) Quarrier, S.; Martin, J. S.; Davis-Neulander, L.; Beauregard, A.; Laederach, A. Evaluation of the Information Content of RNA Structure Mapping Data for Secondary Structure Prediction. *RNA* **2010**, *16*, 1108−1117.
(509) Delihas, N.; Zorn, G. A.; Strobel, E. The Reaction of Escherichia Coli Ribosomes with Kethoxal. *Biochimie* **1974**, *55*, 1227−1234.
(510) Iijima, H.; Patrzyc, H.; Bello, J. Modification of Amino Acids and Bovine Pancreatic Ribonuclease a by Kethoxal. *Biochim. Biophys. Acta, Protein Struct.* **1977**, *491*, 305−316.
(511) Akinsiku, O. T.; Yu, E. T.; Fabris, D. Mass Spectrometric Investigation of Protein Alkylation by the RNA Footprinting Probe Kethoxal. *J. Mass Spectrom.* **2005**, *40*, 1372−1381.
(512) Itano, H. A.; Gottlieb, A. J. Blocking of Tryptic Cleavage of Arginyl Bonds by the Chemical Modification of the Guanido Group with Benzil. *Biochem. Biophys. Res. Commun.* **1963**, *12*, 405−408.
(513) Foettinger, A.; Leitner, A.; Lindner, W. Derivatisation of Arginine Residues with Malondialdehyde for the Analysis of Peptides and Protein Digests by LC-ESI-MS/MS. *J. Mass Spectrom.* **2006**, *41*, 623−632.
(514) Leitner, A.; Foettinger, A.; Lindner, W. Improving Fragmentation of Poorly Fragmenting Peptides and Phosphopeptides during Collision-Induced Dissociation by Malondialdehyde Modification of Arginine Residues. *J. Mass Spectrom.* **2007**, *42*, 950−959.
(515) Onofrejova, L.; Leitner, A.; Lindner, W. Malondialdehyde Tagging Improves the Analysis of Arginine Oligomers and Arginine-Containing Dendrimers by HPLC-MS. *J. Sep. Sci.* **2008**, *31*, 499−506.
(516) Gao, Y.; Wang, Y. Site-Selective Modifications of Arginine Residues in Human Hemoglobin Induced by Methylglyoxal. *Biochemistry* **2006**, *45*, 15654−15660.
(517) Klöpfer, A.; Spanneberg, R.; Glomb, M. A. Formation of Arginine Modifications in a Model System of Nα-tert-Butoxycarbonyl (Boc)-Arginine with Methylglyoxal. *J. Agric. Food Chem.* **2011**, *59*, 394−401.
(518) Takahashi, K. Specific Modification of Arginine Residues in Proteins with Ninhydrin. *J. Biochem.* **1976**, *80*, 1173−1176.
(519) Boppana, V. K.; Rhodes, G. R. High-performance Liquid Chromatographic Determination of an Arginine-Containing Octapeptide Antagonist of Vasopressin in Human Plasma by Means of a Selective Post-Column Reaction with Fluorescence Detection. *J. Chromatogr. A* **1990**, *507*, 79−84.
(520) Schneider, F. Histidine in Enzyme Active Centers. *Angew. Chem., Int. Ed. Engl.* **1978**, *17*, 583−592.
(521) Hoffman, K. W.; Romei, M. G.; Londergan, C. H. A New Raman Spectroscopic Probe of Both the Protonation State and Noncovalent Interactions of Histidine Residues. *J. Phys. Chem. A* **2013**, *117*, 5987−5996.
(522) Gromiha, M. M.; Selvaraj, S. Inter-Residue Interactions in Protein Folding and Stability. *Prog. Biophys. Mol. Biol.* **2004**, *86*, 235−277.
(523) Krupenko, S. A.; Vlasov, A. P.; Wagner, C. On the Role of Conserved Histidine 106 in 10-Formyltetrahydrofolate Dehydrogenase Catalysis: Connection between Hydrolase and Dehydrogenase Mechanisms. *J. Biol. Chem.* **2001**, *276*, 24030−24037.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(524) Barnard, E. A.; Stein, W. D. The Histidine Residue in the Active Centre of Ribonuclease: I. A Specific Reaction with Bromoacetic Acid. *J. Mol. Biol.* **1959**, *1*, 339−349.
(525) Nigen, A. M.; Keim, P.; Marshall, R. C.; Morrow, J. S.; Gurd, F. R. N. Carbon 13 Nuclear Magnetic Resonance Spectroscopy of Myoglobins and Ribonuclease A Carboxymethylated with Enriched [2−13C]Bromoacetate. *J. Biol. Chem.* **1972**, *247*, 4100−4102.
(526) Plapp, B. V. Mechanisms of Carboxymethylation of Bovine Pancreatic Nucleases by Haloacetates and Tosylglycolate. *J. Biol. Chem.* **1973**, *248*, 4896−4900.
(527) Wieghardt, T.; Goren, H. J. The Reactivity of Imidazole Nitrogens in Histidine to Alkylation. *Bioorg. Chem.* **1975**, *4*, 30−40.
(528) Lennette, E. P.; Plapp, B. V. Kinetics of Carboxymethylation of Histidine Hydantoin. *Biochemistry* **1979**, *18*, 3933−3938.
(529) Malinowski, D. P.; Fridovich, I. Subunit Association and Side-Chain Reactivities of Bovine Erythrocyte Superoxide Dismutase in Denaturing Solvents. *Biochemistry* **1979**, *18*, 5055−5060.
(530) Postnikova, G. B.; Moiseeva, S. A.; Shekhovtsova, E. A. The Main Role of Inner Histidines in the Molecular Mechanism of Myoglobin Oxidation Catalyzed by Copper Compounds. *Inorg. Chem.* **2010**, *49*, 1347−1354.
(531) Lin, M. C.; Stein, W. H.; Moore, S. Further Studies on the Alkylation of the Histidine Residues at the Active Site of Pancreatic Ribonuclease. *J. Biol. Chem.* **1968**, *243*, 6167−6170.
(532) Pincus, M. R.; Hummel, C. F.; Brandt-rauf, P. W.; Carty, R. P. Enthalpic and Entropic Determinants for the Specificity of Alkylation of the Histidine-12 Residue of Ribonuclease A by Four Bromoacetamido Nucleoside Affinity Labels and Bromoacetamide. *Int. J. Pept. Protein Res.* **1990**, *36*, 56−66.
(533) Anderton, B. H.; Rabin, B. R. Alkylation Studies on a Reactive Histidine in Pig Heart Malate Dehydrogenase. *Eur. J. Biochem.* **1970**, *15*, 568−573.
(534) Nakagawa, Y.; Bender, M. L. Modification of Alpha-Chymotrypsin by Methyl p-Nitrobenzenesulfonate. *J. Am. Chem. Soc.* **1969**, *91*, 1566−1567.
(535) Nakagawa, Y.; Bender, M. L. Methylation of Histidine-57 in α-Chymotrypsin by Methyl ρ-Nitrobenzenesulfonate. New Approach to Enzyme Modification. *Biochemistry* **1970**, *9*, 259−267.
(536) Volwerk, J. J.; Pieterson, W. A.; De Haas, G. H. Phospholipase A2 and its Zymogen from Porcine Pancreas. VI. Histidine at the Active Site of Phospholipase A2. *Biochemistry* **1974**, *13*, 1446−1454.
(537) Roberts, M. F.; Deems, R. A.; Mincey, T. C.; Dennis, E. A. Chemical Modification of the Histidine Residue in Phospholipase A2 (Naja naja naja). A Case of Half-Site Reactivity. *J. Biol. Chem.* **1977**, *252*, 2405−2411.
(538) Marchi-Salvador, D. P.; Fernandes, C. A. H.; Silveira, L. B.; Soares, A. M.; Fontes, M. R. M. Crystal Structure of a Phospholipase A2 Homolog Complexed with p-Bromophenacyl Bromide Reveals Important Structural Changes Associated with the Inhibition of Myotoxic Activity. *Biochim. Biophys. Acta, Proteins Proteomics* **2009**, *1794*, 1583−1590.
(539) Blacklow, B.; Escoubas, P.; Nicholson, G. M. Characterisation of the Heterotrimeric Presynaptic Phospholipase A2 Neurotoxin Complex from the Venom of the Common Death Adder (Acanthophis antarcticus). *Biochem. Pharmacol.* **2010**, *80*, 277−287.
(540) Hullán, L.; Szontágh, T.; Turtóczky, I.; Fedorcsák, I. The Inactivation of Trypsin by Diethyl Pyrocarbonate. *Acta Chem. Scand.* **1965**, *19*, 2440−2441.
(541) Fedorcsak, I.; Ehrenberg, L. Effects of Diethyl Pyrocarbonate and Methyl Methanesulfonate on Nucleic Acids and Nucleases. *Acta Chem. Scand.* **1966**, *20*, 107−112.
(542) Rosen, C.-G.; Fedorcsák, I. Studies on the Action of Diethyl Pyrocarbonate on Proteins. *Biochim. Biophys. Acta, Gen. Subj.* **1966**, *130*, 401−405.
(543) Holbrook, J. J.; Ingram, V. A. Ionic Properties of an Essential Histidine Residue in Pig Heart Lactate Dehydrogenase. *Biochem. J.* **1973**, *131*, 729−738.
(544) Mendoza, V. L.; Vachet, R. W. Protein Surface Mapping Using Diethylpyrocarbonate with Mass Spectrometric Detection. *Anal. Chem.* **2008**, *80*, 2895−2904.
(545) Hnízda, A.; Šantrůček, J.; Šanda, M.; Strohalm, M.; Kodíček, M. Reactivity of Histidine and Lysine Side-Chains with Diethylpyrocarbonate — A Method to Identify Surface Exposed Residues in Proteins. *J. Biochem. Biophys. Methods* **2008**, *70*, 1091−1097.
(546) Murad, A. H. G.; Toth, G. Effect of Diethylpyr-Ocarbonate on Proteins. I. Reaction of Diethylpyrocarbonate with Amino Acids. *Acta Biochim. Biophys. Acad. Sci. Hung.* **1967**, *2*, 19−29.
(547) Miles, E. W. Modification of Histidyl Residues in Proteins by Diethylpyrocarbonate. In *Methods in Enzymology*; Hirs, C. H. W., Timasheff, S. N., Eds.; Academic Press: New York, 1977; Vol. *47*, pp 431−442, DOI: 10.1016/0076-6879(77)47043-5.
(548) Melchior, W. B.; Fahrney, D. Ethoxyformylation of Proteins. Reaction of Ethoxyformic Anhydride with Alpha-Chymotrypsin, Pepsin, and Pancreatic Ribonuclease at pH 4. *Biochemistry* **1970**, *9*, 251−258.
(549) Kalkum, M.; Przybylski, M.; Glocker, M. O. Structure Characterization of Functional Histidine Residues and Carbethoxylated Derivatives in Peptides and Proteins by Mass Spectrometry. *Bioconjugate Chem.* **1998**, *9*, 226−235.
(550) Foti, S.; Marletta, D.; Saletti, R.; Petrone, G.; Daolio, S. Fast-atom Bombardment Mass Spectrometry of Peptide Derivatives with Diethylpyrocarbonate. *Rapid Commun. Mass Spectrom.* **1991**, *5*, 336−339.
(551) Glocker, M. O.; Kalkum, M.; Yamamoto, R.; Schreurs, J. Selective Biochemical Modification of Functional Residues in Recombinant Human Macrophage Colony-Stimulating Factor β (rhM-CSF β): Identification by Mass Spectrometry. *Biochemistry* **1996**, *35*, 14625−14633.
(552) Dage, J. L.; Sun, H.; Halsall, H. B. Determination of Diethylpyrocarbonate-Modified Amino Acid Residues in α1-Acid Glycoprotein by High-Performance Liquid Chromatography Electrospray Ionization−Mass Spectrometry and Matrix-Assisted Laser Desorption/Ionization Time-of-Flight−Mass Spectrometry. *Anal. Biochem.* **1998**, *257*, 176−185.
(553) Srikanth, R.; Mendoza, V. L.; Bridgewater, J. D.; Zhang, G.; Vachet, R. W. Copper Binding to β-2-Microglobulin and Its Pre-Amyloid Oligomers. *Biochemistry* **2009**, *48*, 9871−9881.
(554) Mendoza, V. L.; Antwi, K.; Barón-Rodríguez, M. A.; Blanco, C.; Vachet, R. W. Structure of the Preamyloid Dimer of β-2-Microglobulin from Covalent Labeling and Mass Spectrometry. *Biochemistry* **2010**, *49*, 1522−1532.
(555) Zhou, Y.; Vachet, R. W. Diethylpyrocarbonate Labeling for the Structural Analysis of Proteins: Label Scrambling in Solution and How to Avoid It. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 899−907.
(556) Zhou, Y.; Vachet, R. W. Increased Protein Structural Resolution from Diethylpyrocarbonate-based Covalent Labeling and Mass Spectrometric Detection. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 708−717.
(557) Joshi, P. N.; Rai, V. Single-Site Labeling of Histidine in Proteins, On-Demand Reversibility, and Traceless Metal-Free Protein Purification. *Chem. Commun.* **2019**, *55*, 1100−1103.
(558) Riordan, J. F.; Vallee, B. L. Acetylation. In *Methods in Enzymology*; Hirs, C. H. W., Ed.; Academic Press: New York, 1967; Vol. *11*, pp 565−570, DOI: 10.1016/S0076-6879(67)11068-9.
(559) Scholten, A.; Visser, N. F. C.; van den Heuvel, R. H. H.; Heck, A. J. R. Analysis of Protein-Protein Interaction Surfaces using a Combination of Efficient Lysine Acetylation and Nanolc-MALDI-MS/MS Applied to the E9:Im9 Bacteriotoxin—Immunity Protein Complex. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 983−994.
(560) Miyazaki, K.; Tsugita, A. C-Terminal Sequencing Method for Proteins in Polyacrylamide Gel by the Reaction of Acetic Anhydride. *Proteomics* **2006**, *6*, 2026−2033.
(561) Fraenkel-Conrat, H. Methods for Investigating the Essential Groups for Enzyme Activity. In *Methods in Enzymology*; Colowick, S. P., Kaplan, N. O., Eds.; Academic Press: New York, 1957; Vol. *4*, pp 247−269, DOI: 10.1016/0076-6879(57)04059-8.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(562) Ohguro, H.; Palczewski, K.; Walsh, K. A.; Johnson, R. S. Topographic Study of Arrestin using Differential Chemical Modifications and Hydrogen/Deuterium Exchange. *Protein Sci.* **1994**, *3*, 2428−2434.
(563) Bennett, K. L.; Smith, S. V.; Lambrecht, R. M.; Truscott, R. J. W.; Sheil, M. M. Rapid Characterization of Chemically-Modified Proteins by Electrospray Mass Spectrometry. *Bioconjugate Chem.* **1996**, *7*, 16−22.
(564) Fligge, T. A.; Kast, J.; Bruns, K.; Przybylski, M. Direct Monitoring of Protein−Chemical Reactions Utilising Nanoelectrospray Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 112−118.
(565) Bothner, B.; Schneemann, A.; Marshall, D.; Reddy, V.; Johnson, J. E.; Siuzdak, G. Crystallographically Identical Virus Capsids Display Different Properties in Solution. *Nat. Struct. Biol.* **1999**, *6*, 114−116.
(566) Izumi, S.; Kaneko, H.; Yamazaki, T.; Hirata, T.; Kominami, S. Membrane Topology of Guinea Pig Cytochrome P450 17α Revealed by a Combination of Chemical Modifications and Mass Spectrometry. *Biochemistry* **2003**, *42*, 14663−14669.
(567) Smith, C. M.; Gafken, P. R.; Zhang, Z.; Gottschling, D. E.; Smith, J. B.; Smith, D. L. Mass Spectrometric Quantification of Acetylation at Specific Lysines within the Amino-Terminal Tail of Histone H4. *Anal. Biochem.* **2003**, *316*, 23−33.
(568) D'Ambrosio, C.; Talamo, F.; Vitale, R. M.; Amodeo, P.; Tell, G.; Ferrara, L.; Scaloni, A. Probing the Dimeric Structure of Porcine Aminoacylase 1 by Mass Spectrometric and Modeling Procedures. *Biochemistry* **2003**, *42*, 4430−4443.
(569) Turner, B. T.; Sabo, T. M.; Wilding, D.; Maurer, M. C. Mapping of Factor XIII Solvent Accessibility as a Function of Activation State using Chemical Modification Methods. *Biochemistry* **2004**, *43*, 9755−9765.
(570) Gong, B.; Ramos, A.; Vázquez-Fernández, E.; Silva, C. J.; Alonso, J.; Liu, Z.; Requena, J. R. Probing Structural Differences between PrPC and PrPSc by Surface Nitration and Acetylation: Evidence of Conformational Change in the C-Terminus. *Biochemistry* **2011**, *50*, 4963−4972.
(571) Steiner, R. F.; Albaugh, S.; Fenselau, C.; Murphy, C.; Vestling, M. A Mass Spectrometry Method for Mapping the Interface Topography of Interacting Proteins, Illustrated by the Melittin-Calmodulin System. *Anal. Biochem.* **1991**, *196*, 120−125.
(572) Klotz, I. M. Succinylation. In *Methods in Enzymology*; Hirs, C. H. W., Ed.; Academic Press: New York, 1967; Vol. *11*, pp 576−580, DOI: 10.1016/S0076-6879(67)11070-7.
(573) Meighen, E. A.; Nicoli, M. Z.; Hastings, J. W. Hybridization of bacterial luciferase with a Variant Produced by Chemical Modification. *Biochemistry* **1971**, *10*, 4062−4068.
(574) Nagel, G. M.; Schachman, H. Cooperative Interactions in Hybrids of Aspartate Transcarbamylase containing Succinylated Regulatory Polypeptide Chains. *Biochemistry* **1975**, *14*, 3195−3203.
(575) Rossi, A.; Menezes, L. C.; Pudles, J. Yeast Hexokinase A. *Eur. J. Biochem.* **1975**, *59*, 423−432.
(576) Schwenke, K. D.; Zirwer, D.; Gast, K.; Görnitz, E.; Linow, K.-J.; Gueguen, J. Changes of the Oligomeric Structure of Legumin from Pea (Pisum sativum L.) after Succinylation. *Eur. J. Biochem.* **1990**, *194*, 621−627.
(577) Klapper, M. H.; Klotz, I. M. Acylation with Dicarboxylic Acid Anhydrides. In *Methods in Enzymology*; Hirs, C. H. W., Timasheff, S. N., Eds.; Academic Press: New York, 1972; Vol. *25*, pp 531−536, DOI: 10.1016/S0076-6879(72)25050-9.
(578) Przybylski, M.; Glocker, M. O.; Nestel, U.; Schnaible, V.; Blüggel, M.; Diederichs, K.; Weckesser, J.; Schad, M.; Schmid, A.; et al. X-ray Crystallographic and Mass Spectrometric Structure Determination and Functional Characterization of Succinylated Porin from Rhodobacter Capsulatus: Implications for Ion Selectivity and Single-Channel Conductance. *Protein Sci.* **1996**, *5*, 1477−1489.
(579) Gould, A. R.; Norton, R. S. Chemical Modification of Cationic Groups in the Polypeptide Cardiac Stimulant Anthopleurin-A. *Toxicon* **1995**, *33*, 187−199.
(580) Becker, L.; McLeod, R. S.; Marcovina, S. M.; Yao, Z.; Koschinsky, M. L. Identification of a Critical Lysine Residue in Apolipoprotein B-100 That Mediates Noncovalent Interaction with Apolipoprotein(a). *J. Biol. Chem.* **2001**, *276*, 36155−36162.
(581) Ehrhard, B.; Misselwitz, R.; Welfle, K.; Hausdorf, G.; Glaser, R. W.; Schneider-Mergener, J.; Welfle, H. Chemical Modification of Recombinant HIV-1 Capsid Protein p24 Leads to the Release of a Hidden Epitope Prior to Changes of the Overall Folding of the Protein. *Biochemistry* **1996**, *35*, 9097−9105.
(582) Paetzel, M.; Strynadka, N. C. J.; Tschantz, W. R.; Casareno, R.; Bullinger, P. R.; Dalbey, R. E. Use of Site-directed Chemical Modification to Study an Essential Lysine in Escherichia coli Leader Peptidase. *J. Biol. Chem.* **1997**, *272*, 9994−10003.
(583) Neurath, A. R.; Debnath, A. K.; Strick, N.; Li, Y.-Y.; Lin, K.; Jiang, S. Blocking of CD4 Cell Receptors for the Human Immunodeficiency Virus Type 1 (HIV-1) by Chemically Modified Bovine Milk Proteins: Potential for AIDS Prophylaxis. *J. Mol. Recognit.* **1995**, *8*, 304−316.
(584) O'Brien, A. M.; Smith, A. T.; Ó'Fágáin, C. Effects of Phthalic Anhydride Modification on Horseradish Peroxidase Stability and Activity. *Biotechnol. Bioeng.* **2003**, *81*, 233−240.
(585) Lindh, C. H.; Jönsson, B. A. G. Human Hemoglobin Adducts Following Exposure to Hexahydrophthalic Anhydride and Methylhexahydrophthalic Anhydride. *Toxicol. Appl. Pharmacol.* **1998**, *153*, 152−160.
(586) Kristiansson, M. H.; Jönsson, B. A. G.; Lindh, C. H. Mass Spectrometric Characterization of Human Hemoglobin Adducts Formed in Vitro by Hexahydrophthalic Anhydride. *Chem. Res. Toxicol.* **2002**, *15*, 562−569.
(587) Jonsson, B. A. G.; Wishnok, J. S.; Skipper, P. L.; Stillwell, W. G.; Tannenbaum, S. R. Lysine Adducts Between Methyltetrahydrophthalic Anhydride and Collagen in Guinea Pig Lung. *Toxicol. Appl. Pharmacol.* **1995**, *135*, 156−162.
(588) Swart, P. J.; Kuipers, E. M.; Smit, C.; Van Der Strate, B. W. A.; Harmsen, M. C.; Meijer, D. K. F. Lactoferrin. In *Advances in Lactoferrin Research*; Spik, G., Legrand, D., Mazurier, J., Pierce, A., Perraudin, J.-P., Eds.; Springer: Boston, 1998; pp 205−213, DOI: 10.1007/978-1-4757-9068-9_24.
(589) Alcalde, M.; Plou, F. J.; Teresa Martín, M.; Valdés, I.; Méndez, E.; Ballesteros, A. Succinylation of Cyclodextrin Glycosyltransferase from Thermoanaerobacter sp. 501 Enhances its Transferase Activity using Starch as Donor. *J. Biotechnol.* **2001**, *86*, 71−80.
(590) Swart, P. J.; Kuipers, M. E.; Smit, C.; Pauwels, R.; deBéthune, M. P.; de Clercq, E.; Meijer, D. K.; Huisman, J. G. Antiviral Effects of Milk Proteins: Acylation Results in Polyanionic Compounds with Potent Activity against Human Immunodeficiency Virus Types 1 and 2 in Vitro. *AIDS Res. Hum. Retroviruses* **1996**, *12*, 769−775.
(591) Pool, C. T.; Thompson, T. E. Methods for Dual, Site-Specific Derivatization of Bovine Pancreatic Trypsin Inhibitor: Trypsin Protection of Lysine-15 and Attachment of Fatty Acids or Hydrophobic Peptides at the N-Terminus. *Bioconjugate Chem.* **1999**, *10*, 221−230.
(592) Makoff, A. J.; Malcolm, A. D. B. Properties of Methyl Acetimidate and its use as a Protein-Modifying Reagent. *Biochem. J.* **1981**, *193*, 245−249.
(593) Sekiguchi, T.; Oshiro, S.; Goingo, E. M.; Nosoh, Y. Chemical Modification of ϵ-Amino Groups in Glutamine Synthetase from Bacillus Stearothermophilus with Ethyl Acetimidate. *J. Biochem.* **1979**, *85*, 75−78.
(594) Zoltobrocki, M.; Kim, J. C.; Plapp, B. V. Activity of Liver Alcohol Dehydrogenase with Various Substituents on the Amino Groups. *Biochemistry* **1974**, *13*, 899−903.
(595) Thumm, M.; Hoenes, J.; Pfleiderer, G. S-methylthioacetimidate is a New Reagent for the Amidination of Proteins at Low pH. *Biochim. Biophys. Acta, Gen. Subj.* **1987**, *923*, 263−267.
(596) Rack, M. Effects of Chemical Modification of Amino Groups by Two Different Imidoesters on Voltage-Clamped Nerve Fibres of the Frog. *Pfluegers Arch.* **1985**, *404*, 126−130.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(597) Jaffee, E. G.; Lauber, M. A.; Running, W. E.; Reilly, J. P. In Vitro and In Vivo Chemical Labeling of Ribosomal Proteins: A Quantitative Comparison. *Anal. Chem.* **2012**, *84*, 9355−9361.
(598) Whiteley, N. M.; Berg, H. C. Amidination of the Outer and Inner Surfaces of the Human Erythrocyte Membrane. *J. Mol. Biol.* **1974**, *87*, 541−561.
(599) Janecki, D. J.; Beardsley, R. L.; Reilly, J. P. Probing Protein Tertiary Structure with Amidination. *Anal. Chem.* **2005**, *77*, 7274−7281.
(600) Brown, D. T.; Kent, S. B. H. Formation of Non-Amidine Products in the Chemical Modification of Horse Liver Alcohol Dehydrogenase with Imido Esters. *Biochem. Biophys. Res. Commun.* **1975**, *67*, 133−138.
(601) Blumberg, S.; Vallee, B. L. Superactivation of Thermolysin by Acylation with Amino Acid N-Hydroxysuccinimide Esters. *Biochemistry* **1975**, *14*, 2410−2419.
(602) Staros, J. V. N-hydroxysulfosuccinimide Active Esters: Bis(N-hydroxysulfosuccinimide) Esters of Two Dicarboxylic Acids are Hydrophilic, Membrane-Impermeant, Protein Cross-Linkers. *Biochemistry* **1982**, *21*, 3950−3955.
(603) Staros, J. V. Membrane-Impermeant Crosslinking Reagents: Probes of the Structure and Dynamics of Membrane Proteins. *Acc. Chem. Res.* **1988**, *21*, 435−441.
(604) Kalkhof, S.; Sinz, A. Chances and Pitfalls of Chemical Cross-Linking with Amine-Reactive N-Hydroxysuccinimide Esters. *Anal. Bioanal. Chem.* **2008**, *392*, 305−312.
(605) Wilchek, M.; Bayer, E. A. Introduction to Avidin-Biotin Technology. In *Methods in Enzymology*; Wilchek, M., Bayer, E. A., Eds.; Academic Press: New York, 1990; Vol. *184*, pp 5−13, DOI: 10.1016/0076-6879(90)84256-G.
(606) Bayer, E. A.; Wilchek, M. Protein Biotinylation. In *Methods in Enzymology*; Wilchek, M., Bayer, E. A., Eds.; Academic Press: New York, 1990; Vol. *184*, pp 138−160, DOI: 10.1016/0076-6879(90)84268-L.
(607) Liu, Y.; Kvaratskhelia, M.; Hess, S.; Qu, Y.; Zou, Y. Modulation of Replication Protein A Function by Its Hyperphosphorylation-induced Conformational Change Involving DNA Binding Domain B. *J. Biol. Chem.* **2005**, *280*, 32775−32783.
(608) Knock, S. L.; Miller, B. T.; Blankenship, J. E.; Nagle, G. T.; Smith, J. S.; Kurosky, A. N-Acylation of Aplysia Egg-Laying Hormone with Biotin. Characterization of Bioactive and Inactive Derivatives. *J. Biol. Chem.* **1991**, *266*, 24413−24419.
(609) Shell, S. M.; Hess, S.; Kvaratskhelia, M.; Zou, Y. Mass Spectrometric Identification of Lysines Involved in the Interaction of Human Replication Protein A with Single-Stranded DNA. *Biochemistry* **2005**, *44*, 971−978.
(610) Gabant, G.; Augier, J.; Armengaud, J. Assessment of solvent residues accessibility using Three Sulfo-NHS-Biotin Reagents in Parallel: Application to Footprint Changes of a Methyltransferase upon Binding Its Substrate. *J. Mass Spectrom.* **2008**, *43*, 360−370.
(611) Dreger, M.; Leung, B. W.; Brownlee, G. G.; Deng, T. A Quantitative Strategy to Detect Changes in Accessibility of Protein Regions to Chemical Modification on Heterodimerization. *Protein Sci.* **2009**, *18*, 1448−1458.
(612) Ori, A.; Free, P.; Courty, J.; Wilkinson, M. C.; Fernig, D. G. Identification of Heparin-Binding Sites in Proteins by Selective Labeling. *Mol. Cell. Proteomics* **2009**, *8*, 2256−2265.
(613) Bolton, A. E.; Hunter, W. M. The Labelling of Proteins to high specific radioactivities by Conjugation to a 125I-Containing Acylating Agent. Application to the Radioimmunoassay. *Biochem. J.* **1973**, *133*, 529−538.
(614) Harmon, C. M.; Luce, P.; Beth, A. H.; Abumrad, N. A. Labeling of Adipocyte Membranes by Sulfo-N-Succinimidyl Derivatives of Long-Chain Fatty Acids: Inhibition of Fatty Acid Transport. *J. Membr. Biol.* **1991**, *121*, 261−268.
(615) Okuyama, T.; Satake, K. On the Preparation and Properties of 2, 4, 6-Trinitrophenyl-Amino Acids and-Peptides. *J. Biochem.* **1960**, *47*, 454−466.
(616) Habeeb, A. F. S. A. Determination of Free Amino Groups in Proteins by Trinitrobenzenesulfonic Acid. *Anal. Biochem.* **1966**, *14*, 328−336.
(617) Fields, R. The Rapid Determination of Amino Groups with TNBS. In *Methods in Enzymology*; Hirs, C. H. W., Timasheff, S. N., Eds.; Academic Press: New York, 1972; Vol. *25*, pp 464−468, DOI: 10.1016/S0076-6879(72)25042-X.
(618) Yang, C. C.; Chang, L. S. Studies on the Status of Lysine Residues in Phospholipase A2 from Naja naja atra (Taiwan cobra) Snake Venom. *Biochem. J.* **1989**, *262*, 855−860.
(619) Komatsu, H.; Emoto, Y.; Tawada, K. Half-Stoichiometric Trinitrophenylation of Myosin Subfragment 1 in the Presence of Pyrophosphate or Adenosine Diphosphate. *J. Biol. Chem.* **1993**, *268*, 7799−7808.
(620) Chang, L.-s.; Lin, S.-r.; Chang, C.-c. Probing Calcium Ion-Induced Conformational Changes of Taiwan Cobra Phospholipase A2 by Trinitrophenylation of Lysine Residues. *J. Protein Chem.* **1997**, *16*, 51−57.
(621) Gevaert, K.; Goethals, M.; Martens, L.; Van Damme, J.; Staes, A.; Thomas, G. R.; Vandekerckhove, J. Exploring Proteomes and analyzing protein Processing by Mass Spectrometric Identification of Sorted N-Terminal Peptides. *Nat. Biotechnol.* **2003**, *21*, 566−569.
(622) Kluger, R.; Tsui, W.-C. Methyl Acetyl Phosphate. A Small Anionic Acetylating Agent. *J. Org. Chem.* **1980**, *45*, 2723−2724.
(623) Ueno, H.; Pospischil, M. A.; Manning, J. M. Methyl Acetyl Phosphate as a Covalent Probe for Anion-Binding Sites in Human and Bovine Hemoglobins. *J. Biol. Chem.* **1989**, *264*, 12344−12351.
(624) Xu, A. S. L.; Labotka, R. J.; London, R. E. Acetylation of Human Hemoglobin by Methyl Acetylphosphate: Evidence of Broad Regio-Selectivity Revealed by NMR Studies. *J. Biol. Chem.* **1999**, *274*, 26629−26632.
(625) Raibekas, A. A.; Bures, E. J.; Siska, C. C.; Kohno, T.; Latypov, R. F.; Kerwin, B. A. Anion Binding and Controlled Aggregation of Human Interleukin-1 Receptor Antagonist. *Biochemistry* **2005**, *44*, 9871−9879.
(626) Anderson, B. M.; Anderson, C. D.; Churchich, J. E. Inhibition of Glutamic Dehydrogenase by Pyridoxal 5′-Phosphate. *Biochemistry* **1966**, *5*, 2893−2900.
(627) Cole, S. C. J.; Yon, R. J. Active-Site-Directed Inactivation of Wheat-Germ Aspartate Transcarbamoylase by Pyridoxal 5′-Phosphate. *Biochem. J.* **1987**, *248*, 403−408.
(628) Xiao, G.-S.; Zhou, J.-M. Conformational Changes at the Active Site of Bovine Pancreatic Rnase A at Low Concentrations of Guanidine Hydrochloride Probed by Pyridoxal 5′-Phosphate. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1996**, *1294*, 1−7.
(629) Ahmed, S. A.; McPhie, P.; Miles, E. W. A Thermally Induced Reversible Conformational Transition of the Tryptophan Synthase Subunit Probed by the Spectroscopic Properties of Pyridoxal Phosphate and by Enzymatic Activity. *J. Biol. Chem.* **1996**, *271*, 8612−8617.
(630) Brautigan, D. L.; Ferguson-Miller, S.; Margoliash, E. Definition of Cytochrome C Binding Domains by Chemical Modification. I. Reaction with 4-Chloro-3,5-Dinitrobenzoate and Chromatographic Separation of Singly Substituted Derivatives. *J. Biol. Chem.* **1978**, *253*, 130−139.
(631) Stark, G. R. Modification of Proteins with Cyanate. In *Methods in Enzymology*; Hirs, C. H. W., Timasheff, S. N., Eds.; Academic Press: New York, 1972; Vol. *25*, pp 579−584, DOI: 10.1016/S0076-6879(72)25057-1.
(632) Means, G. E.; Feeney, R. E. Reductive Alkylation of Proteins. *Anal. Biochem.* **1995**, *224*, 1−16.
(633) Rayment, I. Reductive Alkylation of Lysine Residues to Alter Crystallization Properties of Proteins. In *Methods in Enzymology*; Carter, C. W., Jr., Ed.; Academic Press: New York, 1997; Vol. *276*, pp 171−179, DOI: 10.1016/S0076-6879(97)76058-0.
(634) Day, J. F.; Thorpe, S. R.; Baynes, J. W. Nonenzymatically Glucosylated Albumin. In Vitro Preparation and Isolation from Normal Human Serum. *J. Biol. Chem.* **1979**, *254*, 595−597.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(635) Gurd, F. R. N. Carboxymethylation. In *Methods in Enzymology*; Hirs, C. H. W., Timasheff, S. N., Eds.; Academic Press: New York, 1972; Vol. *25*, pp 424−438, DOI: 10.1016/S0076-6879(72)25038-8.
(636) Jones, W. C., Jr; Rothgeb, T. M.; Gurd, F. R. Specific Enrichment with Carbon-13 of the Methionine Methyl Groups of Sperm Whale Myoglobin. *J. Am. Chem. Soc.* **1975**, *97*, 3875−3877.
(637) Degen, J.; Kyte, J. The Purification of Peptides which Contain Methionine Residues. *Anal. Biochem.* **1978**, *89*, 529−539.
(638) Houghten, R. A.; Glaser, C. B.; Li, C. H. Human Somatotropin: Reaction with Hydrogen Peroxide. *Arch. Biochem. Biophys.* **1977**, *178*, 350−355.
(639) Mizzer, J. P.; Thorpe, C. An Essential Methionine in Pig Kidney General Acyl-Coa Dehydrogenase. *Biochemistry* **1980**, *19*, 5500−5504.
(640) Gross, E. The Cyanogen Bromide Reaction. In *Methods in Enzymology*; Hirs, C. H. W., Ed.; Academic Press, 1967; Vol. *11*, pp 238−255, DOI: 10.1016/S0076-6879(67)11029-X.
(641) Moerman, P. P.; Sergeant, K.; Debyser, G.; Devreese, B.; Samyn, B. A New Chemical Approach to Differentiate Carboxy Terminal Peptide Fragments in Cyanogen Bromide Digests of Proteins. *J. Proteomics* **2010**, *73*, 1454−1460.
(642) McLoughlin, S. M.; Mazur, M. T.; Miller, L. M.; Yin, J.; Liu, F.; Walsh, C. T.; Kelleher, N. L. Chemoenzymatic Approaches for Streamlined Detection of Active Site Modifications on Thiotemplate Assembly Lines Using Mass Spectrometry. *Biochemistry* **2005**, *44*, 14159−14169.
(643) Reid, G. E.; Roberts, K. D.; Simpson, R. J.; O'Hair, R. A. J. Selective Identification and Quantitative Analysis of Methionine Containing Peptides by Charge Derivatization and Tandem Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1131−1150.
(644) Savige, W. E.; Fontana, A. Interconversion of Methionine and Methionine Sulfoxide. In *Methods in Enzymology*; Hirs, C. H. W., Timasheff, S. N., Eds.; Academic Press, 1977; Vol. *47*, pp 453−459, DOI: 10.1016/0076-6879(77)47045-9.
(645) Levine, R. L.; Mosoni, L.; Berlett, B. S.; Stadtman, E. R. Methionine Residues as Endogenous Antioxidants in Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 15036−15040.
(646) West, G. M.; Tang, L.; Fitzgerald, M. C. Thermodynamic Analysis of Protein Stability and Ligand Binding Using a Chemical Modification- and Mass Spectrometry-Based Strategy. *Anal. Chem.* **2008**, *80*, 4175−4185.
(647) Steen, P. V. d.; Rudd, P. M.; Dwek, R. A.; Opdenakker, G. Concepts and Principles of O-Linked Glycosylation. *Crit. Rev. Biochem. Mol. Biol.* **1998**, *33*, 151−208.
(648) Prescher, J. A.; Bertozzi, C. R. Chemical Technologies for Probing Glycans. *Cell* **2006**, *126*, 851−854.
(649) North, S. J.; Hitchen, P. G.; Haslam, S. M.; Dell, A. Mass Spectrometry in the Analysis of N-linked and O-linked Glycans. *Curr. Opin. Struct. Biol.* **2009**, *19*, 498−506.
(650) Cohen, P. The Origins of Protein Phosphorylation. *Nat. Cell Biol.* **2002**, *4*, E127−E130.
(651) Lockridge, O.; Xue, W.; Gaydess, A.; Grigoryan, H.; Ding, S.-J.; Schopfer, L. M.; Hinrichs, S. H.; Masson, P. Pseudo-Esterase Activity of Human Albumin: Slow Turnover on Tyrosine 411 and Stable Acetylation of 82 Residues including 59 Lysines. *J. Biol. Chem.* **2008**, *283*, 22582−22590.
(652) Means, G. E.; Bender, M. L. Acetylation of Human Serum Albumin by p-Nitrophenyl Acetate. *Biochemistry* **1975**, *14*, 4989−4994.
(653) Jabusch, J. R.; Deutsch, H. F. Localization of Lysines Acetylated in Ubiquitin Reacted with p-Nitrophenyl Acetate. *Arch. Biochem. Biophys.* **1985**, *238*, 170−177.
(654) Chelius, D.; Shaler, T. A. Capture of Peptides with N-Terminal Serine and Threonine: A Sequence-Specific Chemical Method for Peptide Mixture Simplification. *Bioconjugate Chem.* **2003**, *14*, 205−211.
(655) Loo, R. R. O.; Loo, J. A. Matrix-Assisted Laser Desorption/Ionization-Mass Spectrometry of Hydrophobic Proteins in Mixtures Using Formic Acid, Perfluorooctanoic Acid, and Sorbitol. *Anal. Chem.* **2007**, *79*, 1115−1125.
(656) Robinson, N. E. Protein Deamidation. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 5283−5288.
(657) Won, J.-I.; Meagher, R. J.; Barron, A. E. Characterization of Glutamine Deamidation in a Long, Repetitive Protein Polymer via Bioconjugate Capillary Electrophoresis. *Biomacromolecules* **2004**, *5*, 618−627.
(658) Lomant, A. J.; Fairbanks, G. Chemical Probes of Extended Biological Structures: Synthesis and Properties of the Cleavable Protein Cross-Linking Reagent [35S] Dithiobis (Succinimidyl Propionate). *J. Mol. Biol.* **1976**, *104*, 243−261.
(659) Cuatrecasas, P.; Anfinsen, C. B. Affinity Chromatography. In *Methods in Enzymology*; Jakoby, W. B., Ed.; Academic Press: New York, 1971; Vol. *22*, pp 345−378, DOI: 10.1016/0076-6879(71)22033-4.
(660) *Thermo Scientific Crosslinking Technical Handbook*; Thermo Scientific: Waltham, MA, 2012; https://tools.thermofisher.com/content/sfs/brochures/1602163-Crosslinking-Reagents-Handbook.pdf (accessed 2020-03-20).
(661) Hermanson, G. T. *Bioconjugate Techniques*, 3rd ed.; Elsevier Academic Press: San Diego, 2013.
(662) Iacobucci, C.; Piotrowski, C.; Aebersold, R.; Amaral, B. C.; Andrews, P.; Bernfur, K.; Borchers, C.; Brodie, N. I.; Bruce, J. E.; Cao, Y.; et al. First Community-Wide, Comparative Cross-Linking Mass Spectrometry Study. *Anal. Chem.* **2019**, *91*, 6953−6961.
(663) Bragg, P.; Hou, C. Subunit Composition, Function, and Spatial Arrangement in the Ca^{2+}-and Mg^{2+}-Activated Adenosine Triphosphatases of Escherichia Coli and Salmonella Typhimurium. *Arch. Biochem. Biophys.* **1975**, *167*, 311−321.
(664) Merkley, E. D.; Rysavy, S.; Kahraman, A.; Hafen, R. P.; Daggett, V.; Adkins, J. N. Distance Restraints from Crosslinking Mass Spectrometry: Mining a Molecular Dynamics Simulation Database to Evaluate Lysine−Lysine Distances. *Protein Sci.* **2014**, *23*, 747−759.
(665) Yu, C.; Huang, L. Cross-Linking Mass Spectrometry: an Emerging Technology for Interactomics and Structural Biology. *Anal. Chem.* **2018**, *90*, 144−165.
(666) Kao, A.; Chiu, C.-l.; Vellucci, D.; Yang, Y.; Patel, V. R.; Guan, S.; Randall, A.; Baldi, P.; Rychnovsky, S. D.; Huang, L. Development of a novel cross-linking strategy for fast and accurate identification of Cross-Linked Peptides of Protein Complexes. *Mol. Cell. Proteomics* **2011**, *10*, M110.002212.
(667) Petrotchenko, E. V.; Serpa, J. J.; Borchers, C. H. An Isotopically Coded CID-Cleavable Biotinylated Cross-Linker for Structural Proteomics. *Mol. Cell. Proteomics* **2011**, *10*, M110.001420.
(668) Müller, M. Q.; Dreiocker, F.; Ihling, C. H.; Schäfer, M.; Sinz, A. Cleavable Cross-Linker for Protein Structure Analysis: Reliable Identification of Cross-Linking Products by Tandem MS. *Anal. Chem.* **2010**, *82*, 6958−6968.
(669) Weisbrod, C. R.; Chavez, J. D.; Eng, J. K.; Yang, L.; Zheng, C.; Bruce, J. E. Vivo Protein Interaction Network Identified with a Novel Real-Time Cross-Linked Peptide Identification Strategy. *J. Proteom. Res.* **2013**, *12*, 1569−1579.
(670) Hartman, F. C.; Wold, F. Bifunctional Reagents. Cross-Linking of Pancreatic Ribonuclease with a Diimido Ester1. *J. Am. Chem. Soc.* **1966**, *88*, 3890−3891.
(671) Hunter, M.; Ludwig, M. The reaction of Imidoesters with Proteins and Related Small Molecules. *J. Am. Chem. Soc.* **1962**, *84*, 3491−3504.
(672) Browne, D. T.; Kent, S. B. Formation of Non-Amidine Products in the Reaction of Primary Amines with Imido Esters. *Biochem. Biophys. Res. Commun.* **1975**, *67*, 126−132.
(673) Grabarek, Z.; Gergely, J. Zero-Length Crosslinking Procedure with the Use of Active Esters. *Anal. Biochem.* **1990**, *185*, 131−135.
(674) Novak, P.; Kruppa, G. H. Intra-molecular cross-linking of acidic residues for Protein Structure Studies. *Eur. J. Mass Spectrom.* **2008**, *14*, 355−365.
(675) Leitner, A.; Joachimiak, L. A.; Unverdorben, P.; Walzthoeni, T.; Frydman, J.; Förster, F.; Aebersold, R. Chemical Cross-Linking/

Mass Spectrometry Targeting Acidic Residues in Proteins and Protein Complexes. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 9455−9460.
(676) Zhang, X.; Wang, J.-H.; Tan, D.; Li, Q.; Li, M.; Gong, Z.; Tang, C.; Liu, Z.; Dong, M.-Q.; Lei, X. Carboxylate-Selective Chemical Cross-Linkers for Mass Spectrometric Analysis of Protein Structures. *Anal. Chem.* **2018**, *90*, 1195−1201.
(677) Smyth, D. G.; Blumenfeld, O.; Konigsberg, W. Reactions of N-Ethylmaleimide with Peptides and Amino Acids. *Biochem. J.* **1964**, *91*, 589.
(678) Brewer, C. F.; Riehm, J. P. Evidence for Possible Nonspecific Reactions Between N-Ethylmaleimide and Proteins. *Anal. Biochem.* **1967**, *18*, 248−255.
(679) Gilchrist, T. L.; Rees, C. W. *Carbenes, Nitrenes and Arynes*; Springer: New York, 1969.
(680) Brunner, J. New Photolabeling and Crosslinking Methods. *Annu. Rev. Biochem.* **1993**, *62*, 483−514.
(681) Suchanek, M.; Radzikowska, A.; Thiele, C. Photo-Leucine and Photo-Methionine Allow Identification of Protein-Protein Interactions in Living Cells. *Nat. Methods* **2005**, *2*, 261.
(682) Lewis, R. V.; Roberts, M. F.; Dennis, E. A.; Allison, W. S. Photoactivated Heterobifunctional Cross-Linking Reagents which Demonstrate the Aggregation State of Phospholipase A2. *Biochemistry* **1977**, *16*, 5650−5654.
(683) Li, X.; Li, Z.; Xie, B.; Sharp, J. S. Improved Identification and Relative Quantification of Sites of Peptide and Protein Oxidation for Hydroxyl Radical Footprinting. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 1767−1776.
(684) Dudev, T.; Lim, C. Competition among Metal Ions for Protein Binding Sites: Determinants of Metal Ion Selectivity in Proteins. *Chem. Rev.* **2014**, *114*, 538−556.
(685) Tainer, J. A.; Roberts, V. A.; Getzoff, E. D. Protein metal-binding sites. *Curr. Opin. Biotechnol.* **1992**, *3*, 378−387.
(686) Bowman, S. E. J.; Bridwell-Rabb, J.; Drennan, C. L. Metalloprotein Crystallography: More than a Structure. *Acc. Chem. Res.* **2016**, *49*, 695−702.
(687) Bertini, I.; Gray, H. B.; Stiefel, E. I.; Valentine, J. S. *Biological Inorganic Chemistry: Structure and Reactivity*; University Science Books: Sausalito, 2007.
(688) Gregory, D. S.; Martin, A. C. R.; Cheetham, J. C.; Rees, A. R. The Prediction and Characterization of Metal Binding Sites in Proteins. *Protein Eng., Des. Sel.* **1993**, *6*, 29−35.
(689) Xu, H.; Xu, D. C.; Wang, Y. Natural Indices for the Chemical Hardness/Softness of Metal Cations and Ligands. *ACS Omega* **2017**, *2*, 7185−7193.
(690) Carlton, D. D.; Schug, K. A. A Review on the Interrogation of Peptide−Metal Interactions using Electrospray Ionization-Mass Spectrometry. *Anal. Chim. Acta* **2011**, *686*, 19−39.
(691) Sung, Y.-h.; Rospigliosi, C.; Eliezer, D. NMR Mapping of Copper Binding Sites in Alpha-Synuclein. *Biochim. Biophys. Acta, Proteins Proteomics* **2006**, *1764*, 5−12.
(692) Zhang, Q.; Dai, X.; Cong, Y.; Zhang, J.; Chen, D.-H.; Dougherty, M. T.; Wang, J.; Ludtke, S. J.; Schmid, M. F.; Chiu, W. Cryo-EM Structure of a Molluscan Hemocyanin Suggests Its Allosteric Mechanism. *Structure* **2013**, *21*, 604−613.
(693) Carter, K. P.; Young, A. M.; Palmer, A. E. Fluorescent Sensors for Measuring Metal Ions in Living Systems. *Chem. Rev.* **2014**, *114*, 4564−4601.
(694) Huang, R. Y.; Rempel, D. L.; Gross, M. L. HD Exchange and PLIMSTEX Determine the Affinities and Order of Binding of Ca2+ with Troponin C. *Biochemistry* **2011**, *50*, 5426−5435.
(695) Liu, X. R.; Zhang, M. M.; Rempel, D. L.; Gross, M. L. A Single Approach Reveals the Composite Conformational Changes, Order of Binding, and Affinities for Calcium Binding to Calmodulin. *Anal. Chem.* **2019**, *91*, 5508−5512.
(696) Zhang, H.; Gau, B. C.; Jones, L. M.; Vidavsky, I.; Gross, M. L. Fast Photochemical Oxidation of Proteins for Comparing Structures of Protein-Ligand Complexes: The Calmodulin-Peptide Model System. *Anal. Chem.* **2011**, *83*, 311−318.
(697) Qin, K.; Yang, Y.; Mastrangelo, P.; Westaway, D. Mapping Cu(II) Binding Sites in Prion Proteins by Diethyl Pyrocarbonate Modification and Matrix-assisted Laser Desorption Ionization-Time of Flight (MALDI-TOF) Mass Spectrometric Footprinting. *J. Biol. Chem.* **2002**, *277*, 1981−1990.
(698) Burns, C. S.; Aronoff-Spencer, E.; Legname, G.; Prusiner, S. B.; Antholine, W. E.; Gerfen, G. J.; Peisach, J.; Millhauser, G. L. Copper Coordination in the Full-Length, Recombinant Prion Protein. *Biochemistry* **2003**, *42*, 6794−6803.
(699) Qin, K.; Coomaraswamy, J.; Mastrangelo, P.; Yang, Y.; Lugowski, S.; Petromilli, C.; Prusiner, S. B.; Fraser, P. E.; Goldberg, J. M.; Chakrabartty, A.; et al. The PrP-like Protein Doppel Binds Copper. *J. Biol. Chem.* **2003**, *278*, 8888−8896.
(700) Narindrasorasak, S.; Kulkarni, P.; Deschamps, P.; She, Y.-M.; Sarkar, B. Characterization and Copper Binding Properties of Human COMMD1 (MURR1). *Biochemistry* **2007**, *46*, 3116−3128.
(701) Zhao, H.; Waite, J. H. Proteins in Load-Bearing Junctions: The Histidine-Rich Metal-Binding Protein of Mussel Byssus. *Biochemistry* **2006**, *45*, 14223−14231.
(702) Binolfi, A.; Lamberto, G. R.; Duran, R.; Quintanar, L.; Bertoncini, C. W.; Souza, J. M.; Cerveñansky, C.; Zweckstetter, M.; Griesinger, C.; Fernández, C. O. Site-Specific Interactions of Cu(II) with α and β-Synuclein: Bridging the Molecular Gap between Metal Binding and Aggregation. *J. Am. Chem. Soc.* **2008**, *130*, 11801−11812.
(703) Karmakar, S.; Das, K. P. Interaction of Cu+2 with α-Crystallin: A Biophysical and Mass Spectrometric Study. *Protein Pept. Lett.* **2018**, *25*, 275−284.
(704) Ginotra, Y. P.; Kulkarni, P. P. Solution Structure of Physiological Cu(His)2: Novel Considerations into Imidazole Coordination. *Inorg. Chem.* **2009**, *48*, 7000−7002.
(705) Miller, J.; McLachlan, A. D.; Klug, A. Repetitive Zinc-Binding Domains in the Protein Transcription Factor IIIA from Xenopus Oocytes. *EMBO J.* **1985**, *4*, 1609−1614.
(706) Gonzalez de Peredo, A.; Saint-Pierre, C.; Adrait, A.; Jacquamet, L.; Latour, J.-M.; Michaud-Soret, I.; Forest, E. Identification of the Two Zinc-Bound Cysteines in the Ferric Uptake Regulation Protein from Escherichia coli: Chemical Modification and Mass Spectrometry Analysis. *Biochemistry* **1999**, *38*, 8582−8589.
(707) Apuy, J. L.; Chen, X.; Russell, D. H.; Baldwin, T. O.; Giedroc, D. P. Ratiometric Pulsed Alkylation/Mass Spectrometry of the Cysteine Pairs in Individual Zinc Fingers of MRE-Binding Transcription Factor-1 (MTF-1) as a Probe of Zinc Chelate Stability. *Biochemistry* **2001**, *40*, 15164−15175.
(708) Atsriku, C.; Scott, G. K.; Benz, C. C.; Baldwin, M. A. Reactivity of Zinc Finger Cysteines: Chemical Modifications within Labile Zinc Fingers in Estrogen Receptor. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 2017−2026.
(709) Larabee, J. L.; Hocker, J. R.; Hanas, J. S. Mechanisms of Aurothiomalate−Cys2His2 Zinc Finger Interactions. *Chem. Res. Toxicol.* **2005**, *18*, 1943−1954.
(710) Larabee, J. L.; Hocker, J. R.; Hanas, J. S. Cys Redox Reactions and Metal Binding of a Cys2His2 Zinc Finger. *Arch. Biochem. Biophys.* **2005**, *434*, 139−149.
(711) Walter, E. D.; Stevens, D. J.; Visconte, M. P.; Millhauser, G. L. The Prion Protein is a Combined Zinc and Copper Binding Protein: Zn2+ Alters the Distribution of Cu2+ Coordination Modes. *J. Am. Chem. Soc.* **2007**, *129*, 15440−15441.
(712) Karmakar, S.; Das, K. P. Identification of Histidine Residues Involved in Zn2+ Binding to αA- and αB-Crystallin by Chemical Modification and MALDI TOF Mass Spectrometry. *Protein J.* **2012**, *31*, 623−640.
(713) Chen, S.-H.; Russell, W. K.; Russell, D. H. Combining Chemical Labeling, Bottom-Up and Top-Down Ion-Mobility Mass Spectrometry To Identify Metal-Binding Sites of Partially Metalated Metallothionein. *Anal. Chem.* **2013**, *85*, 3229−3237.
(714) Puljung, M. C.; Zagotta, W. N. Labeling of Specific Cysteines in Proteins Using Reversible Metal Protection. *Biophys. J.* **2011**, *100*, 2513−2521.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(715) Ramakrishnan, D.; Xing, W.; Beran, R. K.; Chemuru, S.; Rohrs, H.; Niedziela-Majka, A.; Marchand, B.; Mehra, U.; Zábranský, A.; Doležal, M.; et al. Hepatitis B Virus X Protein Function Requires Zinc Binding. *J. Virol.* **2019**, *93*, e00250.
(716) Clapham, D. E. Calcium Signaling. *Cell* **1995**, *80*, 259−268.
(717) Clapham, D. E. Calcium Signaling. *Cell* **2007**, *131*, 1047−1058.
(718) Gifford, J. L.; Walsh, M. P.; Vogel, H. J. Structures and Metal-Ion-Binding Properties of the Ca^{2+}-Binding Helix−Loop−Helix EF-Hand Motifs. *Biochem. J.* **2007**, *405*, 199−221.
(719) Zhang, H.; Wen, J.; Huang, R. Y. C.; Blankenship, R. E.; Gross, M. L. Mass Spectrometry-Based Carboxyl Footprinting of Proteins: Method Evaluation. *Int. J. Mass Spectrom.* **2012**, *312*, 78−86.
(720) Narumi, R.; Yamamoto, T.; Inoue, A.; Arata, T. Substrate-Induced Conformational Changes in Sarcoplasmic Reticulum Ca2+-ATPase Probed by Surface Modification using Diethylpyrocarbonate with Mass Spectrometry. *FEBS Lett.* **2012**, *586*, 3172−3178.
(721) Senguen, F. T.; Grabarek, Z. X-ray Structures of Magnesium and Manganese Complexes with the N-Terminal Domain of Calmodulin: Insights into the Mechanism and Specificity of Metal Ion Binding to an EF-Hand. *Biochemistry* **2012**, *51*, 6182−6194.
(722) Gorman, J. J.; Wallis, T. P.; Pitt, J. J. Protein Disulfide Bond Determination by Mass Spectrometry. *Mass Spectrom. Rev.* **2002**, *21*, 183−216.
(723) Xia, Y.; Cooks, R. G. Plasma Induced Oxidative Cleavage of Disulfide Bonds in Polypeptides during Nanoelectrospray Ionization. *Anal. Chem.* **2010**, *82*, 2856−2864.
(724) Wang, Y.; Lu, Q.; Wu, S.-L.; Karger, B. L.; Hancock, W. S. Characterization and Comparison of Disulfide Linkages and Scrambling Patterns in Therapeutic Monoclonal Antibodies: Using LC-MS with Electron Transfer Dissociation. *Anal. Chem.* **2011**, *83*, 3133−3140.
(725) Liu, F.; van Breukelen, B.; Heck, A. J. R. Facilitating Protein Disulfide Mapping by a Combination of Pepsin Digestion, Electron Transfer Higher Energy Dissociation (EThcD), and a Dedicated Search Algorithm SlinkS. *Mol. Cell. Proteomics* **2014**, *13*, 2776−2786.
(726) Agarwal, A.; Diedrich, J. K.; Julian, R. R. Direct Elucidation of Disulfide Bond Partners Using Ultraviolet Photodissociation Mass Spectrometry. *Anal. Chem.* **2011**, *83*, 6455−6458.
(727) Quick, M. M.; Crittenden, C. M.; Rosenberg, J. A.; Brodbelt, J. S. Characterization of Disulfide Linkages in Proteins by 193 nm Ultraviolet Photodissociation (UVPD) Mass Spectrometry. *Anal. Chem.* **2018**, *90*, 8523−8530.
(728) Zhang, Y.; Dewald, H. D.; Chen, H. Online Mass Spectrometric Analysis of Proteins/Peptides Following Electrolytic Cleavage of Disulfide Bonds. *J. Proteom. Res.* **2011**, *10*, 1293−1304.
(729) Echterbille, J.; Quinton, L.; Gilles, N.; De Pauw, E. Ion Mobility Mass Spectrometry as a Potential Tool To Assign Disulfide Bonds Arrangements in Peptides with Multiple Disulfide Bridges. *Anal. Chem.* **2013**, *85*, 4405−4413.
(730) Choi, S.; Jeong, J.; Na, S.; Lee, H. S.; Kim, H.-Y.; Lee, K.-J.; Paek, E. New Algorithm for the Identification of Intact Disulfide Linkages Based on Fragmentation Characteristics in Tandem Mass Spectra. *J. Proteom. Res.* **2010**, *9*, 626−635.
(731) Huang, S. Y.; Chen, S. F.; Chen, C. H.; Huang, H. W.; Wu, W. G.; Sung, W. C. Global Disulfide Bond Profiling for Crude Snake Venom Using Dimethyl Labeling Coupled with Mass Spectrometry and RADAR Algorithm. *Anal. Chem.* **2014**, *86*, 8742−8750.
(732) Morris, H. R.; Pucci, P. A New Method for Rapid Assignment of S-S Bridges in Proteins. *Biochem. Biophys. Res. Commun.* **1985**, *126*, 1122−1128.
(733) Smith, D. L.; Zhou, Z. Strategies for Locating Disulfide Bonds in Proteins. In *Methods in Enzymology*; McCloskey, J. A., Ed; Academic Press: New York, 1990; Vol. *193*, pp 374−389, DOI: 10.1016/0076-6879(90)93428-N.
(734) Zaluzec, E. J.; Gage, D. A.; Watson, J. T. Quantitative Assessment of Cysteine and Cystine in Peptides and Proteins Following Organomercurial Derivatization and Analysis by Matrix-Assisted Laser Desorption Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 359−366.
(735) Zaluzec, E. J.; Gage, D. A.; Watson, J. T. Matrix-Assisted Laser Desorption Ionization Mass Spectrometry: Applications in Peptide and Protein Characterization. *Protein Expression Purif.* **1995**, *6*, 109−123.
(736) Wu, J.; Gage, D. A.; Watson, J. T. A Strategy to Locate Cysteine Residues in Proteins by Specific Chemical Cleavage Followed by Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry. *Anal. Biochem.* **1996**, *235*, 161−174.
(737) Jacobson, G. R.; Schaffer, M. H.; Stark, G. R.; Vanaman, T. C. Specific Chemical Cleavage in High Yield at the Amino Peptide Bonds of Cysteine and Cystine Residues. *J. Biol. Chem.* **1973**, *248*, 6583−6591.
(738) Barbirz, S.; Jakob, U.; Glocker, M. O. Mass Spectrometry Unravels Disulfide Bond Formation as the Mechanism That Activates a Molecular Chaperone. *J. Biol. Chem.* **2000**, *275*, 18759−18766.
(739) Abo, M.; Li, C.; Weerapana, E. Isotopically-Labeled Iodoacetamide-Alkyne Probes for Quantitative Cysteine-Reactivity Profiling. *Mol. Pharmaceutics* **2018**, *15*, 743−749.
(740) Weerapana, E.; Wang, C.; Simon, G. M.; Richter, F.; Khare, S.; Dillon, M. B. D.; Bachovchin, D. A.; Mowen, K.; Baker, D.; Cravatt, B. F. Quantitative Reactivity Profiling Predicts Functional Cysteines in Proteomes. *Nature* **2010**, *468*, 790−795.
(741) Chalker, J. M.; Bernardes, G. J. L.; Lin, Y. A.; Davis, B. G. Chemical Modification of Proteins at Cysteine: Opportunities in Chemistry and Biology. *Chem. - Asian J.* **2009**, *4*, 630−640.
(742) Glocker, M. O.; Arbogast, B.; Schreurs, J.; Deinzer, M. L. Assignment of the Inter- and Intramolecular Disulfide Linkages in Recombinant Human Macrophage Colony Stimulating Factor using Fast Atom Bombardment Mass Spectrometry. *Biochemistry* **1993**, *32*, 482−488.
(743) Li, K. S.; Chen, G.; Mo, J.; Huang, R. Y. C.; Deyanova, E. G.; Beno, B. R.; O'Neil, S. R.; Tymiak, A. A.; Gross, M. L. Orthogonal Mass Spectrometry-Based Footprinting for Epitope Mapping and Structural Characterization: The IL-6 Receptor upon Binding of Protein Therapeutics. *Anal. Chem.* **2017**, *89*, 7742−7749.
(744) Kaur, P.; Tomechko, S.; Kiselar, J.; Shi, W.; Deperalta, G.; Wecksler, A. T.; Gokulrangan, G.; Ling, V.; Chance, M. R. Characterizing Monoclonal Antibody Structure by Carbodiimide/GEE Footprinting. *mAbs* **2014**, *6*, 1486−1499.
(745) Kaur, P.; Kiselar, J.; Shi, W.; Yang, S.; Chance, M. R. Covalent Labeling Techniques for Characterizing Higher Order Structure of Monoclonal Antibodies. In *State-of-the-Art and Emerging Technologies for Therapeutic Monoclonal Antibody Characterization Vol. 3. Defining the Next Generation of Analytical and Biophysical Techniques*; Schiel, J. E., Davis, D. L., Borisov, O. V., Eds; American Chemical Society: Washington, DC, 2015; Vol. *1202*, pp 45−73 DOI: 10.1021/bk-2015-1202.ch003.
(746) Wecksler, A. T.; Kalo, M. S.; Deperalta, G. Mapping of Fab-1:VEGF Interface Using Carboxyl Group Footprinting Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 2077−2080.
(747) Kaur, P.; Tomechko, S. E.; Kiselar, J.; Shi, W.; Deperalta, G.; Wecksler, A. T.; Gokulrangan, G.; Ling, V.; Chance, M. R. Characterizing Monoclonal Antibody Structure by Carboxyl Group Footprinting. *mAbs* **2015**, *7*, 540−552.
(748) Pan, L. Y.; Salas-Solano, O.; Valliere-Douglass, J. F. Localized Conformational Interrogation of Antibody and Antibody-Drug Conjugates by Site-Specific Carboxyl Group Footprinting. *mAbs* **2017**, *9*, 307−318.
(749) Underbakke, E. S.; Zhu, Y.; Kiessling, L. L. Protein Footprinting in a Complex Milieu: Identifying the Interaction Surfaces of the Chemotaxis Adaptor Protein CheW. *J. Mol. Biol.* **2011**, *409*, 483−495.
(750) Underbakke, E. S.; Zhu, Y.; Kiessling, L. L. Isotope-Coded Affinity Tags with Tunable Reactivities for Protein Footprinting. *Angew. Chem., Int. Ed.* **2008**, *47*, 9677−9680.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(751) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **1999**, *17*, 994−999.
(752) Aebersold, R.; Goodlett, D. R. Mass Spectrometry in Proteomics. *Chem. Rev.* **2001**, *101*, 269−296.
(753) Mendoza, V. L.; Barón-Rodríguez, M. A.; Blanco, C.; Vachet, R. W. Structural Insights into the Pre-Amyloid Tetramer of β-2-Microglobulin from Covalent Labeling and Mass Spectrometry. *Biochemistry* **2011**, *50*, 6711−6722.
(754) Liu, T.; Marcinko, T. M.; Kiefer, P. A.; Vachet, R. W. Using Covalent Labeling and Mass Spectrometry To Study Protein Binding Sites of Amyloid Inhibiting Molecules. *Anal. Chem.* **2017**, *89*, 11583−11591.
(755) Shrivastava, S.; Nuffer, J. H.; Siegel, R. W.; Dordick, J. S. Position-Specific Chemical Modification and Quantitative Proteomics Disclose Protein Orientation Adsorbed on Silica Nanoparticles. *Nano Lett.* **2012**, *12*, 1583−1587.
(756) Thyparambil, A. A.; Wei, Y.; Wu, Y.; Latour, R. A. Determination of Orientation and Adsorption-Induced Changes in the Tertiary Structure of Proteins on Material Surfaces by Chemical Modification and Peptide Mapping. *Acta Biomater.* **2014**, *10*, 2404−2414.
(757) Tollefson, E. J.; Allen, C. R.; Chong, G.; Zhang, X.; Rozanov, N. D.; Bautista, A.; Cerda, J. J.; Pedersen, J. A.; Murphy, C. J.; Carlson, E. E.; et al. Preferential Binding of Cytochrome c to Anionic Ligand-Coated Gold Nanoparticles: A Complementary Computational and Experimental Approach. *ACS Nano* **2019**, *13*, 6856−6866.
(758) Yang, B.; Wu, Y.-J.; Zhu, M.; Fan, S.-B.; Lin, J.; Zhang, K.; Li, S.; Chi, H.; Li, Y.-X.; Chen, H.-F.; et al. Identification of Cross-Linked Peptides from Complex Samples. *Nat. Methods* **2012**, *9*, 904−906.
(759) Rinner, O.; Seebacher, J.; Walzthoeni, T.; Mueller, L. N.; Beck, M.; Schmidt, A.; Mueller, M.; Aebersold, R. Identification of Cross-Linked Peptides from Large Sequence Databases. *Nat. Methods* **2008**, *5*, 315−318.
(760) Liu, F.; Rijkers, D. T.; Post, H.; Heck, A. J. Proteome-Wide Profiling of Protein Assemblies by Cross-Linking Mass Spectrometry. *Nat. Methods* **2015**, *12*, 1179−1184.
(761) Götze, M.; Pettelkau, J.; Schaks, S.; Bosse, K.; Ihling, C. H.; Krauth, F.; Fritzsche, R.; Kühn, U.; Sinz, A. StavroX—a Software for Analyzing Crosslinked Products in Protein Interaction Studies. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 76−87.
(762) Young, M. M.; Tang, N.; Hempel, J. C.; Oshiro, C. M.; Taylor, E. W.; Kuntz, I. D.; Gibson, B. W.; Dollinger, G. High Throughput Protein Fold Identification by using experimental constraints Derived from Intramolecular Cross-Links and Mass Spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 5802−5806.
(763) Alber, F.; Dokudovskaya, S.; Veenhoff, L. M.; Zhang, W.; Kipper, J.; Devos, D.; Suprapto, A.; Karni-Schmidt, O.; Williams, R.; Chait, B. T.; Sali, A.; Rout, M. P. The Molecular Architecture of the Nuclear Pore Complex. *Nature* **2007**, *450*, 695−701.
(764) Shi, Y.; Fernandez-Martinez, J.; Tjioe, E.; Pellarin, R.; Kim, S. J.; Williams, R.; Schneidman-Duhovny, D.; Sali, A.; Rout, M. P.; Chait, B. T. Structural Characterization by Cross-Linking Reveals the Detailed Architecture of a Coatomer-Related Heptameric Module from the Nuclear Pore Complex. *Mol. Cell. Proteomics* **2014**, *13*, 2927−2943.
(765) Erzberger, J. P.; Stengel, F.; Pellarin, R.; Zhang, S.; Schaefer, T.; Aylett, C. H.; Cimermančič, P.; Boehringer, D.; Sali, A.; Aebersold, R.; Ban, N. Molecular Architecture of the 40S· eIF1· eIF3 Translation Initiation Complex. *Cell* **2014**, *158*, 1123−1135.
(766) Herzog, F.; Kahraman, A.; Boehringer, D.; Mak, R.; Bracher, A.; Walzthoeni, T.; Leitner, A.; Beck, M.; Hartl, F.-U.; Ban, N.; et al. Structural Probing of a Protein Phosphatase 2A Network by Chemical Cross-Linking and Mass Spectrometry. *Science* **2012**, *337*, 1348−1352.
(767) Hofmann, T.; Fischer, A. W.; Meiler, J.; Kalkhof, S. Protein Structure Prediction Guided by Crosslinking Restraints−A Systematic Evaluation of the Impact of the Crosslinking Spacer Length. *Methods* **2015**, *89*, 79−90.
(768) Gutierrez, C. B.; Yu, C.; Novitsky, E. J.; Huszagh, A. S.; Rychnovsky, S. D.; Huang, L. Developing an Acidic Residue Reactive and Sulfoxide-Containing MS-Cleavable Homobifunctional Cross-Linker for Probing Protein−Protein Interactions. *Anal. Chem.* **2016**, *88*, 8315−8322.
(769) Belsom, A.; Mudd, G.; Giese, S.; Auer, M.; Rappsilber, J. Complementary Benzophenone Cross-Linking/Mass Spectrometry Photochemistry. *Anal. Chem.* **2017**, *89*, 5319−5324.
(770) Belsom, A.; Schneider, M.; Fischer, L.; Brock, O.; Rappsilber, J. Serum Albumin Domain Structures in Human Blood Serum by Mass Spectrometry and Computational Biology. *Mol. Cell. Proteomics* **2016**, *15*, 1105−1116.
(771) Leitner, A.; Faini, M.; Stengel, F.; Aebersold, R. Crosslinking and Mass Spectrometry: An Integrated Technology to Understand the Structure and Function of Molecular Machines. *Trends Biochem. Sci.* **2016**, *41*, 20−32.
(772) Schneider, M.; Belsom, A.; Rappsilber, J. Protein Tertiary Structure by Crosslinking/Mass Spectrometry. *Trends Biochem. Sci.* **2018**, *43*, 157−169.
(773) Zhang, M. M.; Beno, B. R.; Huang, R. Y. C.; Adhikari, J.; Deyanova, E. G.; Li, J.; Chen, G.; Gross, M. L. An Integrated Approach for Determining a Protein-Protein Binding Interface in Solution and an Evaluation of HDX Kinetics for Adjudicating Candidate Docking Models. *Anal. Chem.* **2019**, *91*, 15709−15717.
(774) Komolov, K. E.; Du, Y.; Duc, N. M.; Betz, R. M.; Rodrigues, J. P. G. L. M.; Leib, R. D.; Patra, D.; Skiniotis, G.; Adams, C. M.; Dror, R. O.; et al. Structural and Functional Analysis of a β2-Adrenergic Receptor Complex with GRK5. *Cell* **2017**, *169*, 407−421.
(775) Lin, S.-J.; Chen, Y.-F.; Hsu, K.-C.; Chen, Y.-L.; Ko, T.-P.; Lo, C.-F.; Wang, H.-C.; Wang, H.-C. Structural Insights to the Heterotetrameric Interaction between the Vibrio parahaemolyticus PirAvp and PirBvp Toxins and Activation of the Cry-Like Pore-Forming Domain. *Toxins* **2019**, *11*, 233.
(776) Xu, H.; Hsu, P.-H.; Zhang, L.; Tsai, M.-D.; Freitas, M. A. Database Search Algorithm for Identification of Intact Cross-Links in Proteins and Peptides Using Tandem Mass Spectrometry. *J. Proteom. Res.* **2010**, *9*, 3384−3393.
(777) Liu, F.; Lössl, P.; Scheltema, R.; Viner, R.; Heck, A. J. R. Optimized Fragmentation Schemes and Data Analysis Strategies for Proteome-Wide Cross-Link Identification. *Nat. Commun.* **2017**, *8*, 15473.
(778) Tan, D.; Li, Q.; Zhang, M.-J.; Liu, C.; Ma, C.; Zhang, P.; Ding, Y.-H.; Fan, S.-B.; Tao, L.; Yang, B.; et al. Trifunctional Cross-Linker for Mapping Protein-Protein Interaction Networks and Comparing Protein Conformational States. *eLife* **2016**, *5*, No. e12509.
(779) Yu, C.; Huang, L. Cross-Linking Mass Spectrometry: An Emerging Technology for Interactomics and Structural Biology. *Anal. Chem.* **2018**, *90*, 144−165.
(780) Konermann, L.; Simmons, D. A. Protein-Folding Kinetics and Mechanisms Studied by Pulse-Labeling and Mass Spectrometry. *Mass Spectrom. Rev.* **2003**, *22*, 1−26.
(781) Happersberger, H. P.; Stapleton, J.; Cowgill, C.; Glocker, M. O. Characterization of the Folding Pathway of Recombinant Human Macrophage-Colony Stimulating-Factor β (rhM-CSF β) by Bis-Cysteinyl Modification and Mass Spectrometry. *Proteins: Struct., Funct., Genet.* **1998**, *33*, 50−62.
(782) Smith, D. J.; Kenyon, G. L. Nonessentiality of the Active Sulfhydryl Group of Rabbit Muscle Creatine Kinase. *J. Biol. Chem.* **1974**, *249*, 3317−3318.
(783) Feng, Z.; Butler, M. C.; Alam, S. L.; Loh, S. N. On the Nature of Conformational Openings: Native and Unfolded-State Hydrogen and Thiol-Disulfide Exchange Studies of Ferric Aquomyoglobin. *J. Mol. Biol.* **2001**, *314*, 153−166.
(784) Ha, J.-H.; Loh, S. N. Changes in Side Chain Packing during Apomyoglobin Folding Characterized by Pulsed Thiol-Disulfide Exchange. *Nat. Struct. Mol. Biol.* **1998**, *5*, 730−737.
(785) Feng, Z.; Ha, J.-H.; Loh, S. N. Identifying the Site of Initial Tertiary Structure Disruption during Apomyoglobin Unfolding. *Biochemistry* **1999**, *38*, 14433−14439.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(786) Jha, S. K.; Udgaonkar, J. B. Exploring the Cooperativity of the Fast Folding Reaction of a Small Protein Using Pulsed Thiol Labeling and Mass Spectrometry. *J. Biol. Chem.* **2007**, *282*, 37479−37491.
(787) Han, H.; McLuckey, S. A. Selective Covalent Bond Formation in Polypeptide Ions via Gas-Phase Ion/Ion Reaction Chemistry. *J. Am. Chem. Soc.* **2009**, *131*, 12884−12885.
(788) Mentinova, M.; McLuckey, S. A. Covalent Modification of Gaseous Peptide Ions with N-Hydroxysuccinimide Ester Reagent Ions. *J. Am. Chem. Soc.* **2010**, *132*, 18248−18257.
(789) Peng, Z.; McGee, W. M.; Bu, J.; Barefoot, N. Z.; McLuckey, S. A. Gas Phase Reactivity of Carboxylates with N-Hydroxysuccinimide Esters. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 174−180.
(790) Peng, Z.; Pilo, A. L.; Luongo, C. A.; McLuckey, S. A. Gas-Phase Amidation of Carboxylic Acids with Woodward's Reagent K Ions. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 1686−1694.
(791) Mourão, M. A.; Hakim, J. B.; Schnell, S. Connecting the Dots: The Effects of Macromolecular Crowding on Cell Physiology. *Biophys. J.* **2014**, *107*, 2761−2766.
(792) Guin, D.; Gruebele, M. Weak Chemical Interactions That Drive Protein Evolution: Crowding, Sticking, and Quinary Structure in Folding and Function. *Chem. Rev.* **2019**, *119*, 10691−10717.
(793) Xue, L.; Karpenko, I. A.; Hiblot, J.; Johnsson, K. Imaging and Manipulating Proteins in Live Cells through Covalent Labeling. *Nat. Chem. Biol.* **2015**, *11*, 917−923.
(794) Chen, I.; Ting, A. Y. Site-Specific Labeling of Proteins with Small Molecules in Live Cells. *Curr. Opin. Biotechnol.* **2005**, *16*, 35−40.
(795) Jing, C.; Cornish, V. W. Chemical Tags for Labeling Proteins Inside Living Cells. *Acc. Chem. Res.* **2011**, *44*, 784−792.
(796) Griffin, B. A.; Adams, S. R.; Tsien, R. Y. Specific Covalent Labeling of Recombinant Protein Molecules Inside Live Cells. *Science* **1998**, *281*, 269−272.
(797) Marks, K. M.; Braun, P. D.; Nolan, G. P. A General Approach for Chemical Labeling and Rapid, Spatially Controlled Protein Inactivation. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 9982−9987.
(798) Takaoka, Y.; Ojida, A.; Hamachi, I. Protein Organic Chemistry and Applications for Labeling and Engineering in Live-Cell Systems. *Angew. Chem., Int. Ed.* **2013**, *52*, 4088−4106.
(799) Szymański, W.; Beierle, J. M.; Kistemaker, H. A. V.; Velema, W. A.; Feringa, B. L. Reversible Photocontrol of Biological Systems by the Incorporation of Molecular Photoswitches. *Chem. Rev.* **2013**, *113*, 6114−6178.
(800) Angel, T. E.; Aryal, U. K.; Hengel, S. M.; Baker, E. S.; Kelly, R. T.; Robinson, E. W.; Smith, R. D. Mass Spectrometry-Based Proteomics: Existing Capabilities and Future Directions. *Chem. Soc. Rev.* **2012**, *41*, 3912−3928.
(801) Sechi, S.; Oda, Y. Quantitative Proteomics using Mass Spectrometry. *Curr. Opin. Chem. Biol.* **2003**, *7*, 70−77.
(802) Birkemeyer, C.; Luedemann, A.; Wagner, C.; Erban, A.; Kopka, J. Metabolome Analysis: The Potential of In Vivo Labeling with Stable Isotopes for Metabolite Profiling. *Trends Biotechnol.* **2005**, *23*, 28−33.
(803) Selenko, P.; Wagner, G. Looking into Live Cells with In-Cell NMR Spectroscopy. *J. Struct. Biol.* **2007**, *158*, 244−253.
(804) Sakakibara, D.; Sasaki, A.; Ikeya, T.; Hamatsu, J.; Hanashima, T.; Mishima, M.; Yoshimasu, M.; Hayashi, N.; Mikawa, T.; Wälchli, M.; et al. Protein Structure Determination in Living Cells by In-Cell NMR Spectroscopy. *Nature* **2009**, *458*, 102−105.
(805) Sinz, A. The Advancement of Chemical Cross-Linking and Mass Spectrometry for Structural Proteomics: From Single Proteins to Protein Interaction Networks. *Expert Rev. Proteomics* **2014**, *11*, 733−743.
(806) Chavez, J. D.; Bruce, J. E. Chemical Cross-Linking with Mass Spectrometry: A Tool for Systems Structural Biology. *Curr. Opin. Chem. Biol.* **2019**, *48*, 8−18.
(807) Zhang, H.; Tang, X.; Munske, G. R.; Tolic, N.; Anderson, G. A.; Bruce, J. E. Identification of Protein-Protein Interactions and Topologies in Living Cells with Chemical Cross-Linking and Mass Spectrometry. *Mol. Cell. Proteomics* **2009**, *8*, 409−420.
(808) Chavez, J. D.; Weisbrod, C. R.; Zheng, C.; Eng, J. K.; Bruce, J. E. Protein Interactions, Post-Translational Modifications and Topologies in Human Cells. *Mol. Cell. Proteomics* **2013**, *12*, 1451−1467.
(809) Zheng, C.; Yang, L.; Hoopmann, M. R.; Eng, J. K.; Tang, X.; Weisbrod, C. R.; Bruce, J. E. Cross-linking Measurements of In Vivo Protein Complex Topologies. *Mol. Cell. Proteomics* **2011**, *10*, M110.006841.
(810) Navare, A. T.; Chavez, J. D.; Zheng, C.; Weisbrod, C. R.; Eng, J. K.; Siehnel, R.; Singh, P. K.; Manoil, C.; Bruce, J. E. Probing the Protein Interaction Network of Pseudomonas Aeruginosa Cells by Chemical Cross-Linking Mass Spectrometry. *Structure* **2015**, *23*, 762−773.
(811) Wu, X.; Chavez, J. D.; Schweppe, D. K.; Zheng, C.; Weisbrod, C. R.; Eng, J. K.; Murali, A.; Lee, S. A.; Ramage, E.; Gallagher, L. A.; et al. In Vivo Protein Interaction Network Analysis Reveals Porin-Localized Antibiotic Inactivation in Acinetobacter Baumannii Strain AB5075. *Nat. Commun.* **2016**, *7*, 13414.
(812) Zhong, X.; Navare, A. T.; Chavez, J. D.; Eng, J. K.; Schweppe, D. K.; Bruce, J. E. Large-Scale and Targeted Quantitative Cross-Linking MS using Isotope-Labeled Protein Interaction Reporter (PIR) Cross-Linkers. *J. Proteom. Res.* **2017**, *16*, 720−727.
(813) de Jong, L.; de Koning, E. A.; Roseboom, W.; Buncherd, H.; Wanner, M. J.; Dapic, I.; Jansen, P. J.; van Maarseveen, J. H.; Corthals, G. L.; Lewis, P. J.; Hamoen, L. W.; de Koster, C. G. In-Culture Cross-Linking of Bacterial Cells Reveals Large-Scale Dynamic Protein−Protein Interactions at the Peptide Level. *J. Proteom. Res.* **2017**, *16*, 2457−2471.
(814) Kaake, R. M.; Wang, X.; Burke, A.; Yu, C.; Kandur, W.; Yang, Y.; Novtisky, E. J.; Second, T.; Duan, J.; Kao, A.; et al. A New In Vivo Cross-Linking Mass Spectrometry Platform to Define Protein−Protein Interactions in Living Cells. *Mol. Cell. Proteomics* **2014**, *13*, 3533−3543.
(815) Liu, F.; Lössl, P.; Rabbitts, B. M.; Balaban, R. S.; Heck, A. J. The Interactome of Intact Mitochondria by Cross-Linking Mass Spectrometry Provides Evidence for Coexisting Respiratory Supercomplexes. *Mol. Cell. Proteomics* **2018**, *17*, 216−232.
(816) Liu, H.; Zhang, H.; Niedzwiedzki, D. M.; Prado, M.; He, G.; Gross, M. L.; Blankenship, R. E. Phycobilisomes Supply Excitations to Both Photosystems in a Megacomplex in Cyanobacteria. *Science* **2013**, *342*, 1104−1107.
(817) Liu, H.; Weisz, D. A.; Zhang, M. M.; Cheng, M.; Zhang, B.; Zhang, H.; Gerstenecker, G. S.; Pakrasi, H. B.; Gross, M. L.; Blankenship, R. E. Phycobilisomes Harbor FNR_L in Cyanobacteria. *mBio* **2019**, *10*, e00669-19.
(818) Liu, F.; Heck, A. J. Interrogating the architecture of protein assemblies and Protein Interaction Networks by Cross-Linking Mass Spectrometry. *Curr. Opin. Struct. Biol.* **2015**, *35*, 100−108.
(819) Shen, G.; Cui, W.; Zhang, H.; Zhou, F.; Huang, W.; Liu, Q.; Yang, Y.; Li, S.; Bowman, G. R.; Sadler, J. E.; et al. Warfarin Traps Human Vitamin K Epoxide Reductase in an Intermediate State during Electron Transfer. *Nat. Struct. Mol. Biol.* **2017**, *24*, 69−76.
(820) Calabrese, A. N.; Radford, S. E. Mass spectrometry-enabled structural biology of Membrane Proteins. *Methods* **2018**, *147*, 187−205.
(821) Bolla, J. R.; Agasid, M. T.; Mehmood, S.; Robinson, C. V. Membrane Protein−Lipid Interactions Probed Using Mass Spectrometry. *Annu. Rev. Biochem.* **2019**, *88*, 85−111.
(822) Leney, A. C.; Heck, A. J. R. Native Mass Spectrometry: What is in the Name? *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 5−13.
(823) Shen, G.; Li, S.; Cui, W.; Liu, S.; Yang, Y.; Gross, M.; Li, W. Membrane Protein Structure in Live Cells: Methodology for Studying Drug Interaction by Mass Spectrometry-Based Footprinting. *Biochemistry* **2018**, *57*, 286−294.
(824) Li, K. Mass Spectrometry-based Strategies for Protein Characterization: Amyloid Formation, Protein-Ligand Interactions and Structures *of Membrane Proteins in Live Cells*. Ph.D. Dissertation. Washington University in St. Louis: St. Louis, MO, 2018.
(825) Rouillon, C.; Zhou, M.; Zhang, J.; Politis, A.; Beilsten-Edmands, V.; Cannone, G.; Graham, S.; Robinson, C. V.; Spagnolo,

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

L.; White, M. F. Structure of the CRISPR Interference Complex CSM Reveals Key Similarities with Cascade. *Mol. Cell* **2013**, *52*, 124−134.
(826) Bui, K. H.; von Appen, A.; DiGuilio, A. L.; Ori, A.; Sparks, L.; Mackmull, M.-T.; Bock, T.; Hagen, W.; Andrés-Pons, A.; Glavy, J. S.; Beck, M. Integrated Structural Analysis of the Human Nuclear Pore Complex Scaffold. *Cell* **2013**, *155*, 1233−1243.
(827) Lasker, K.; Förster, F.; Bohn, S.; Walzthoeni, T.; Villa, E.; Unverdorben, P.; Beck, F.; Aebersold, R.; Sali, A.; Baumeister, W. Molecular Architecture of the 26S Proteasome Holocomplex Determined by an Integrative Approach. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 1380−1387.
(828) Greber, B. J.; Boehringer, D.; Leitner, A.; Bieri, P.; Voigts-Hoffmann, F.; Erzberger, J. P.; Leibundgut, M.; Aebersold, R.; Ban, N. Architecture of the Large Subunit of the Mammalian Mitochondrial Ribosome. *Nature* **2014**, *505*, 515−519.
(829) Kar, U. K.; Simonian, M.; Whitelegge, J. P. Integral Membrane Proteins: Bottom-Up, Top-Down and Structural Proteomics. *Expert Rev. Proteomics* **2017**, *14*, 715−723.
(830) Weisz, D. A.; Gross, M. L.; Pakrasi, H. B. The Use of Advanced Mass Spectrometry to Dissect the Life-Cycle of Photosystem II. *Front. Plant Sci.* **2016**, *7*, 617.
(831) Hurt, E.; Beck, M. Towards Understanding Nuclear Pore Complex Architecture and Dynamics in the Age of Integrative Structural Analysis. *Curr. Opin. Cell Biol.* **2015**, *34*, 31−38.
(832) Sinz, A.; Arlt, C.; Chorev, D.; Sharon, M. Chemical Cross-Linking and Native Mass Spectrometry: A Fruitful Combination for Structural Biology. *Protein Sci.* **2015**, *24*, 1193−1209.
(833) Politis, A.; Schmidt, C.; Tjioe, E.; Sandercock, A. M.; Lasker, K.; Gordiyenko, Y.; Russel, D.; Sali, A.; Robinson, C. V. Topological Models of Heteromeric Protein Assemblies from Mass Spectrometry: Application to the Yeast Eif3: Eif5 Complex. *Chem. Biol.* **2015**, *22*, 117−128.
(834) Robinson, P. J.; Trnka, M. J.; Bushnell, D. A.; Davis, R. E.; Mattei, P.-J.; Burlingame, A. L.; Kornberg, R. D. Structure of a Complete Mediator-RNA Polymerase II Pre-Initiation Complex. *Cell* **2016**, *166*, 1411.
(835) Schmidt, C.; Macpherson, J. A.; Lau, A. M.; Tan, K. W.; Fraternali, F.; Politis, A. Surface Accessibility and Dynamics of Macromolecular Assemblies Probed by Covalent Labeling Mass Spectrometry and Integrative Modeling. *Anal. Chem.* **2017**, *89*, 1459−1468.
(836) Liu, W. H.; Roemer, S. C.; Zhou, Y.; Shen, Z.-J.; Dennehey, B. K.; Balsbaugh, J. L.; Liddle, J. C.; Nemkov, T.; Ahn, N. G.; Hansen, K. C.; Tyler, J. K.; Churchill, M. E. A. The Cac1 Subunit of Histone Chaperone CAF-1 Organizes CAF-1-H3/H4 Architecture and Tetramerizes Histones. *eLife* **2016**, *5*, No. e18023.
(837) Mysling, S.; Kristensen, K. K.; Larsson, M.; Beigneux, A. P.; GÅrdsvoll, H.; Fong, L. G.; Bensadouen, A.; Jørgensen, T. J.; Young, S. G.; Ploug, M. The Acidic Domain of the Endothelial Membrane Protein GPIHBP1 Stabilizes Lipoprotein Lipase Activity by Preventing Unfolding of its Catalytic Domain. *eLife* **2016**, *5*, No. e12095.
(838) Zanphorlin, L. M.; Lima, T. B.; Wong, M. J.; Balbuena, T. S.; Minetti, C. A.; Remeta, D. P.; Young, J. C.; Barbosa, L. R.; Gozzo, F. C.; Ramos, C. H. Heat shock protein 90 kDa (Hsp90) has a Second Functional Interaction Site with the Mitochondrial Import Receptor Tom70. *J. Biol. Chem.* **2016**, *291*, 18620−18631.
(839) Aprahamian, M. L.; Lindert, S. Utility of Covalent Labeling Mass Spectrometry Data in Protein Structure Prediction with Rosetta. *J. Chem. Theory Comput.* **2019**, *15*, 3410−3424.
(840) Kaltashov, I. A.; Bobst, C. E.; Abzalimov, R. R.; Berkowitz, S. A.; Houde, D. Conformation and Dynamics of Biopharmaceuticals: Transition of Mass Spectrometry-Based Tools from Academe to Industry. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 323−337.
(841) Gau, B.; Garai, K.; Frieden, C.; Gross, M. L. Mass Spectrometry-Based Protein Footprinting Characterizes the Structures of Oligomeric Apolipoprotein E2, E3, and E4. *Biochemistry* **2011**, *50*, 8117−8126.
(842) Liu, T.; Limpikirati, P.; Vachet, R. W. Synergistic Structural Information from Covalent Labeling and Hydrogen−Deuterium Exchange Mass Spectrometry for Protein−Ligand Interactions. *Anal. Chem.* **2019**, *91*, 15248−15254.
(843) Hager-Braun, C.; Tomer, K. B. Characterization of the Tertiary Structure of Soluble CD4 Bound to Glycosylated Full-Length HIVgp120 by Chemical Modification of Arginine Residues and Mass Spectrometric Analysis. *Biochemistry* **2002**, *41*, 1759−1766.
(844) Hambly, D. M.; Gross, M. L. Laser Flash Photolysis of Hydrogen Peroxide to Oxidize Protein Solvent-Accessible Residues on the Microsecond Timescale. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 2057−2063.
(845) Murphy, M. P. How Mitochondria Produce Reactive Oxygen Species. *Biochem. J.* **2009**, *417*, 1−13.
(846) Apel, K.; Hirt, H. Reactive oxygen species: Metabolism, Oxidative Stress, and Signal Transduction. *Annu. Rev. Plant Biol.* **2004**, *55*, 373−399.
(847) D'Autréaux, B.; Toledano, M. B. ROS as Signalling Molecules: Mechanisms that Generate Specificity in ROS Homeostasis. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 813−824.
(848) Takamoto, K.; Chance, M. R. Radiolytic Protein Footprinting with Mass Spectrometry to Probe the Structure of Macromolecular Complexes. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 251−276.
(849) Buxton, G. V.; Greenstock, C. L.; Helman, W. P.; Ross, A. B. Critical Review of Rate Constants for Reactions of Hydrated Electrons, Hydrogen Atoms and Hydroxyl Radicals (·OH/·O− in Aqueous Solution. *J. Phys. Chem. Ref. Data* **1988**, *17*, 513−886.
(850) Garrison, W. M. Reaction Mechanisms in the Radiolysis of Peptides, Polypeptides, and Proteins. *Chem. Rev.* **1987**, *87*, 381−398.
(851) Fenton, H. J. H. LXXIII.—Oxidation of Tartaric Acid in Presence of Iron. *J. Chem. Soc., Trans.* **1894**, *65*, 899−910.
(852) Masarwa, A.; Rachmilovich-Calis, S.; Meyerstein, N.; Meyerstein, D. Oxidation of Organic Substrates in Aerated Aqueous Solutions by the Fenton Reagent. *Coord. Chem. Rev.* **2005**, *249*, 1937−1943.
(853) Kryatov, S. V.; Rybak-Akimova, E. V.; Schindler, S. Kinetics and Mechanisms of Formation and Reactivity of Non-Heme Iron Oxygen Intermediates. *Chem. Rev.* **2005**, *105*, 2175−2226.
(854) Goldstein, S.; Meyerstein, D.; Czapski, G. The Fenton Reagents. *Free Radical Biol. Med.* **1993**, *15*, 435−445.
(855) Sawyer, D. T.; Sobkowiak, A.; Matsushita, T. Metal [MLx; M = Fe, Cu, Co, Mn]/Hydroperoxide-Induced Activation of Dioxygen for the Oxygenation of Hydrocarbons: Oxygenated Fenton Chemistry. *Acc. Chem. Res.* **1996**, *29*, 409−416.
(856) Haber, F.; Weiss, J. Über die Katalyse des Hydroperoxydes. *Naturwissenschaften* **1932**, *20*, 948−950.
(857) Haber, F.; Weiss, J.; Pope, W. J. The Catalytic Decomposition of Hydrogen Peroxide by Iron Salts. *Proc. R. Soc. London A Math. Phys. Sci.* **1934**, *147*, 332−351.
(858) Barb, W. G.; Baxendale, J. H.; George, P.; Hargrave, K. R. Reactions of Ferrous and Ferric Ions with Hydrogen Peroxide. Part I.—The Ferrous Ion Reaction. *Trans. Faraday Soc.* **1951**, *47*, 462−500.
(859) Barb, W. G.; Baxendale, J. H.; George, P.; Hargrave, K. R. Reactions of Ferrous and Ferric Ions with Hydrogen Peroxide. Part II.—The Ferric Ion Reaction. *Trans. Faraday Soc.* **1951**, *47*, 591−616.
(860) Tullius, T.; Dombroski, B. Iron(II) EDTA used to Measure the Helical Twist Along any DNA Molecule. *Science* **1985**, *230*, 679−681.
(861) Tullius, T. D.; Dombroski, B. A. Hydroxyl Radical "Footprinting": High-Resolution Information about DNA-Protein Contacts and Application to Lambda Repressor and Cro Protein. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 5469−5473.
(862) Pogozelski, W. K.; McNeese, T. J.; Tullius, T. D. What Species Is Responsible for Strand Scission in the Reaction of [FeIIEDTA]2- and H2O2 with DNA? *J. Am. Chem. Soc.* **1995**, *117*, 6428−6433.
(863) Baron, C. P.; Refsgaard, H. H. F.; Skibsted, L. H.; Andersen, M. L. Oxidation of Bovine Serum Albumin Initiated by the Fenton

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

Reaction—Effect of EDTA, Tert-Butylhydroperoxide and Tetrahydrofuran. *Free Radical Res.* **2006**, *40*, 409−417.
(864) Tullius, T. D.; Greenbaum, J. A. Mapping Nucleic Acid Structure by Hydroxyl Radical Cleavage. *Curr. Opin. Chem. Biol.* **2005**, *9*, 127−134.
(865) Heyduk, E.; Heyduk, T. Mapping Protein Domains Involved in Macromolecular Interactions: A Novel Protein Footprinting Approach. *Biochemistry* **1994**, *33*, 9643−9650.
(866) Sharp, J. S.; Becker, J. M.; Hettich, R. L. Protein Surface Mapping by Chemical Oxidation: Structural Analysis by Mass Spectrometry. *Anal. Biochem.* **2003**, *313*, 216−225.
(867) Monroe, E. B.; Heien, M. L. Electrochemical Generation of Hydroxyl Radicals for Examining Protein Structure. *Anal. Chem.* **2013**, *85*, 6185−6189.
(868) Rana, T. M.; Meares, C. F. Specific Cleavage of a Protein by an Attached Iron Chelate. *J. Am. Chem. Soc.* **1990**, *112*, 2457−2458.
(869) Rana, T. M.; Meares, C. F. Iron Chelate Mediated Proteolysis: Protein Structure Dependence. *J. Am. Chem. Soc.* **1991**, *113*, 1859−1861.
(870) Rana, T. M.; Meares, C. F. Transfer of Oxygen from an Artificial Protease to Peptide Carbon During Proteolysis. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 10578−10582.
(871) Heilek, G. M.; Noller, H. F. Site-Directed Hydroxyl Radical Probing of the rRNA Neighborhood of Ribosomal Protein S5. *Science* **1996**, *272*, 1659−1662.
(872) Lancaster, L.; Kiel, M. C.; Kaji, A.; Noller, H. F. Orientation of Ribosome Recycling Factor in the Ribosome from Directed Hydroxyl Radical Probing. *Cell* **2002**, *111*, 129−140.
(873) Duval, M.; Marenna, A.; Chevalier, C.; Marzi, S. Site-Directed Chemical Probing to Map Transient RNA/Protein Interactions. *Methods* **2017**, *117*, 48−58.
(874) Cheal, S. M.; Ng, M.; Barrios, B.; Miao, Z.; Kalani, A. K.; Meares, C. F. Mapping Protein−Protein Interactions by Localized Oxidation: Consequences of the Reach of Hydroxyl Radical. *Biochemistry* **2009**, *48*, 4577−4586.
(875) Li, Q.; Xie, Y.; Xu, G.; Lebrilla, C. B. Identification of Potential sialic acid binding Proteins on Cell Membranes by Proximity Chemical Labeling. *Chem. Sci.* **2019**, *10*, 6199−6209.
(876) Zhu, Y.; Guo, T.; Park, J. E.; Li, X.; Meng, W.; Datta, A.; Bern, M.; Lim, S. K.; Sze, S. K. Elucidating in Vivo Structural Dynamics in Integral Membrane Protein by Hydroxyl Radical Footprinting. *Mol. Cell. Proteomics* **2009**, *8*, 1999−2010.
(877) Petit, A.; Mwale, F.; Tkaczyk, C.; Antoniou, J.; Zukor, D. J.; Huk, O. L. Induction of Protein Oxidation by Cobalt and Chromium Ions in Human U937 Macrophages. *Biomaterials* **2005**, *26*, 4416−4422.
(878) Bridgewater, J. D.; Lim, J.; Vachet, R. W. Transition Metal−Peptide Binding Studied by Metal-Catalyzed Oxidation Reactions and Mass Spectrometry. *Anal. Chem.* **2006**, *78*, 2432−2438.
(879) Lim, J.; Vachet, R. W. Development of a Methodology Based on Metal-Catalyzed Oxidation Reactions and Mass Spectrometry To Determine the Metal Binding Sites in Copper Metalloproteins. *Anal. Chem.* **2003**, *75*, 1164−1172.
(880) Lim, J.; Vachet, R. W. Using Mass Spectrometry To Study Copper−Protein Binding under Native and Non-Native Conditions: *β*-2-Microglobulin. *Anal. Chem.* **2004**, *76*, 3498−3504.
(881) Gau, B. C.; Sharp, J. S.; Rempel, D. L.; Gross, M. L. Fast Photochemical Oxidation of Protein Footprints Faster than Protein Unfolding. *Anal. Chem.* **2009**, *81*, 6563−6571.
(882) Shcherbakova, I.; Mitra, S.; Beer, R. H.; Brenowitz, M. Fast Fenton Footprinting: a Laboratory-Based Method for the Time-Resolved Analysis of DNA, RNA and Proteins. *Nucleic Acids Res.* **2006**, *34*, No. e48.
(883) Sclavi, B.; Sullivan, M.; Chance, M. R.; Brenowitz, M.; Woodson, S. A. RNA Folding at Millisecond Intervals by Synchrotron Hydroxyl Radical Footprinting. *Science* **1998**, *279*, 1940−1943.
(884) Sclavi, B.; Woodson, S.; Sullivan, M.; Chance, M. R.; Brenowitz, M. Time-Resolved Synchrotron X-Ray "Footprinting", a New Approach to the Study of Nucleic Acid Structure and Function: Application to Protein-DNA Interactions and RNA Folding. *J. Mol. Biol.* **1997**, *266*, 144−159.
(885) Chance, M. R.; Sclavi, B.; Woodson, S. A.; Brenowitz, M. Examining the Conformational Dynamics of Macromolecules with Time-Resolved Synchrotron X-Ray 'Footprinting'. *Structure* **1997**, *5*, 865−869.
(886) Sclavi, B.; Woodson, S.; Sullivan, M.; Chance, M.; Brenowitz, M. Following the Folding of RNA with Time-Resolved Synchrotron X-Ray Footprinting. In *Methods in Enzymology*; Ackers, G. K., Johnson, M. L., Eds.; Academic Press: New York, 1998; Vol. *295*, pp 379−402, DOI: 10.1016/S0076-6879(98)95050-9.
(887) Maleknia, S. D.; Brenowitz, M.; Chance, M. R. Millisecond Radiolytic Modification of Peptides by Synchrotron X-Rays Identified by Mass Spectrometry. *Anal. Chem.* **1999**, *71*, 3965−3973.
(888) Simic, M. G. Pulse Radiolysis in Study of Oxygen Radicals. In *Methods in Enzymology*; Packer, L., Glazer, A. N., Eds.; Academic Press: New York, 1990; Vol. *186*, pp 89−100, DOI: 10.1016/0076-6879(90)86094-C.
(889) Gupta, S.; Sullivan, M.; Toomey, J.; Kiselar, J.; Chance, M. R. The Beamline X28C of the Center for Synchrotron Biosciences: A National Resource for Biomolecular Structure and Dynamics Experiments Using Synchrotron Footprinting. *J. Synchrotron Radiat.* **2007**, *14*, 233−243.
(890) Asuru, A.; Farquhar, E. R.; Sullivan, M.; Abel, D.; Toomey, J.; Chance, M. R.; Bohon, J. The XFP (17-BM) Beamline for X-Ray Footprinting at NSLS-II. *J. Synchrotron Radiat.* **2019**, *26*, 1388−1399.
(891) Weeks, J. L.; Matheson, M. S. The Primary Quantum Yield of Hydrogen Peroxide Decomposition1. *J. Am. Chem. Soc.* **1956**, *78*, 1273−1278.
(892) Dainton, F. S. The Primary Quantum Yield in the Photolysis of Hydrogen Peroxide at 3130 Å. and the Primary Radical Yield in the X- and *γ*-Radiolysis of Water. *J. Am. Chem. Soc.* **1956**, *78*, 1278−1279.
(893) Sharp, J. S.; Becker, J. M.; Hettich, R. L. Analysis of Protein Solvent Accessible Surfaces by Photochemical Oxidation and Mass Spectrometry. *Anal. Chem.* **2004**, *76*, 672−683.
(894) Aye, T. T.; Low, T. Y.; Sze, S. K. Nanosecond Laser-Induced Photochemical Oxidation Method for Protein Surface Mapping with Mass Spectrometry. *Anal. Chem.* **2005**, *77*, 5814−5822.
(895) Hambly, D. M.; Gross, M. L. Cold Chemical Oxidation of Proteins. *Anal. Chem.* **2009**, *81*, 7235−7242.
(896) Li, K. S.; Shi, L.; Gross, M. L. Mass Spectrometry-Based Fast Photochemical Oxidation of Proteins (FPOP) for Higher Order Structure Characterization. *Acc. Chem. Res.* **2018**, *51*, 736−744.
(897) Zhang, B.; Cheng, M.; Rempel, D.; Gross, M. L. Implementing Fast Photochemical Oxidation of Proteins (FPOP) as a Footprinting Approach to Solve Diverse Problems in Structural Biology. *Methods* **2018**, *144*, 94−103.
(898) Chen, J.; Rempel, D. L.; Gross, M. L. Temperature Jump and Fast Photochemical Oxidation Probe Submillisecond Protein Folding. *J. Am. Chem. Soc.* **2010**, *132*, 15502−15504.
(899) Chen, J.; Rempel, D. L.; Gau, B. C.; Gross, M. L. Fast Photochemical Oxidation of Proteins and Mass Spectrometry Follow Submillisecond Protein Folding at the Amino-Acid Level. *J. Am. Chem. Soc.* **2012**, *134*, 18724−18731.
(900) Zhang, Y.; Rempel, D. L.; Zhang, H.; Gross, M. L. An Improved Fast Photochemical Oxidation of Proteins (FPOP) Platform for Protein Therapeutics. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 526−529.
(901) Vahidi, S.; Konermann, L. Probing the Time Scale of FPOP (Fast Photochemical Oxidation of Proteins): Radical Reactions Extend Over Tens of Milliseconds. *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 1156−1164.
(902) Sheng, Y.; Capri, J.; Waring, A.; Valentine, J. S.; Whitelegge, J. Exposure of Solvent-Inaccessible Regions in the Amyloidogenic Protein Human SOD1 Determined by Hydroxyl Radical Footprinting. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 218−226.
(903) Maleknia, S. D.; Chance, M. R.; Downard, K. M. Electrospray-Assisted Modification of Proteins: a Radical Probe of Protein Structure. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 2352−2358.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(904) Maleknia, S. D.; Wong, J. W. H.; Downard, K. M. Photochemical and Electrophysical Production of Radicals on Millisecond Timescales to Probe the Structure, Dynamics and Interactions of Proteins. *Photochem. Photobiol. Sci.* **2004**, *3*, 741−748.
(905) Maleknia, S. D.; Downard, K. M. Radical Approaches to Probe Protein Structure, Folding, and Interactions by Mass Spectrometry. *Mass Spectrom. Rev.* **2001**, *20*, 388−401.
(906) Maleknia, S. D.; Downard, K. M. Advances in Radical Probe Mass Spectrometry for Protein Footprinting in Chemical Biology Applications. *Chem. Soc. Rev.* **2014**, *43*, 3244−3258.
(907) Thomas, M. C.; Mitchell, T. W.; Blanksby, S. J. Ozonolysis of Phospholipid Double Bonds during Electrospray Ionization: A New Tool for Structure Determination. *J. Am. Chem. Soc.* **2006**, *128*, 58−59.
(908) Wong, J. W. H.; Maleknia, S. D.; Downard, K. M. Study of the Ribonuclease−S-Protein−Peptide Complex Using a Radical Probe and Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2003**, *75*, 1557−1563.
(909) Wong, J. W. H.; Maleknia, S. D.; Downard, K. M. Hydroxyl Radical Probe of the Calmodulin-Melittin Complex Interface by Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 225−233.
(910) Maleknia, S. D.; Downard, K. M. On-Plate Deposition of Oxidized Proteins to Facilitate Protein Footprinting Studies by Radical Probe Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2012**, *26*, 2311−2318.
(911) Lin, H.; Kitova, E. N.; Johnson, M. A.; Eugenio, L.; Ng, K. K. S.; Klassen, J. S. Electrospray Ionization-Induced Protein Unfolding. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 2122−2131.
(912) Minkoff, B. B.; Blatz, J. M.; Choudhury, F. A.; Benjamin, D.; Shohet, J. L.; Sussman, M. R. Plasma-Generated OH Radical Production for Analyzing Three-Dimensional Structure in Protein Therapeutics. *Sci. Rep.* **2017**, *7*, 12946.
(913) LaVerne, J. A. OH Radicals and Oxidizing Products in the Gamma Radiolysis of Water. *Radiat. Res.* **2000**, *153*, 196−200.
(914) Hayes, J. J.; Kam, L.; Tullius, T. D. Footprinting Protein-DNA Complexes with γ-rays. In *Methods in Enzymology*; Packer, L., Glazer, A. N., Eds.; Academic Press: New York, 1990; Vol. *186*, pp 545−549, DOI: 10.1016/0076-6879(90)86148-O.
(915) Nukuna, B. N.; Sun, G.; Anderson, V. E. Hydroxyl Radical Oxidation of Cytochrome C by Aerobic Radiolysis. *Free Radical Biol. Med.* **2004**, *37*, 1203−1213.
(916) Xu, G.; Kiselar, J.; He, Q.; Chance, M. R. Secondary Reactions and Strategies To Improve Quantitative Protein Footprinting. *Anal. Chem.* **2005**, *77*, 3029−3037.
(917) Xu, G.; Chance, M. R. Radiolytic Modification and Reactivity of Amino Acid Residues Serving as Structural Probes for Protein Footprinting. *Anal. Chem.* **2005**, *77*, 4549−4555.
(918) Sharp, J. S.; Tomer, K. B. Effects of Anion Proximity in Peptide Primary Sequence on the Rate and Mechanism of Leucine Oxidation. *Anal. Chem.* **2006**, *78*, 4885−4893.
(919) Tong, X.; Wren, J. C.; Konermann, L. Effects of Protein Concentration on the Extent of γ-Ray-Mediated Oxidative Labeling Studied by Electrospray Mass Spectrometry. *Anal. Chem.* **2007**, *79*, 6376−6382.
(920) Sharp, J. S.; Tomer, K. B. Analysis of the Oxidative Damage-Induced Conformational Changes of Apo- and Holocalmodulin by Dose-Dependent Protein Oxidative Surface Mapping. *Biophys. J.* **2007**, *92*, 1682−1692.
(921) Tong, X.; Wren, J. C.; Konermann, L. γ-Ray-Mediated Oxidative Labeling for Detecting Protein Conformational Changes by Electrospray Mass Spectrometry. *Anal. Chem.* **2008**, *80*, 2222−2231.
(922) Fang, X.; Mark, G.; von Sonntag. OH Radical Formation by Ultrasound in Aqueous Solutions Part I: The Chemistry Underlying the Terephthalate Dosimeter. *Ultrason. Sonochem.* **1996**, *3*, 57−63.
(923) Mark, G.; Tauber, A.; Laupert, R.; Schuchmann, H.-P.; Schulz, D.; Mues, A.; von Sonntag, C. OH-Radical Formation by Ultrasound in Aqueous Solution − Part II: Terephthalate and Fricke Dosimetry and the Influence of Various Conditions on the Sonolytic Yield. *Ultrason. Sonochem.* **1998**, *5*, 41−52.
(924) Spotheim-Maurizot, M.; Charlier, M.; Sabattier, R. DNA Radiolysis by Fast Neutrons. *Int. J. Radiat. Biol.* **1990**, *57*, 301−313.
(925) Spotheim-Maurizot, M.; Franchet, J.; Sabattier, R.; Charlier, M. DNA Radiolysis by Fast Neutrons. II. Oxygen, Thiols and Ionic Strength Effects. *Int. J. Radiat. Biol.* **1991**, *59*, 1313−1324.
(926) Franchet-Beuzit, J.; Spotheim-Maurizot, M.; Sabattier, R.; Blazy-Baudras, B.; Charlier, M. Radiolytic Footprinting. .Beta. Rays, .Gamma. Photons, and Fast Neutrons Probe DNA-Protein Interactions. *Biochemistry* **1993**, *32*, 2104−2110.
(927) King, P. A.; Jamison, E.; Strahs, D.; Anderson, V. E.; Brenowitz, M. 'Footprinting' Proteins on DNA with Peroxonitrous Acid. *Nucleic Acids Res.* **1993**, *21*, 2473−2478.
(928) Coddington, J. W.; Hurst, J. K.; Lymar, S. V. Hydroxyl Radical Formation during Peroxynitrous Acid Decomposition. *J. Am. Chem. Soc.* **1999**, *121*, 2438−2443.
(929) Watson, C.; Janik, I.; Zhuang, T.; Charvátová, O.; Woods, R. J.; Sharp, J. S. Pulsed Electron Beam Water Radiolysis for Submicrosecond Hydroxyl Radical Protein Footprinting. *Anal. Chem.* **2009**, *81*, 2496−2505.
(930) Chaulk, S. G.; Pezacki, J. P.; MacMillan, A. M. Studies of RNA Cleavage by Photolysis of N-Hydroxypyridine-2(1H)-thione. A New Photochemical Footprinting Method. *Biochemistry* **2000**, *39*, 10448−10453.
(931) Kelly, J. J.; Chernov, B. K.; Tovstanovsky, I.; Mirzabekov, A. D.; Bavykin, S. G. Radical-Generating Coordination Complexes as Tools for Rapid and Effective Fragmentation and Fluorescent Labeling of Nucleic Acids for Microchip Hybridization. *Anal. Biochem.* **2002**, *311*, 103−118.
(932) McClintock, C. S.; Hettich, R. L. Experimental Approach to Controllably Vary Protein Oxidation While Minimizing Electrode Adsorption for Boron-Doped Diamond Electrochemical Surface Mapping Applications. *Anal. Chem.* **2013**, *85*, 213−219.
(933) Xu, G.; Takamoto, K.; Chance, M. R. Radiolytic Modification of Basic Amino Acid Residues in Peptides: Probes for Examining Protein−Protein Interactions. *Anal. Chem.* **2003**, *75*, 6995−7007.
(934) Xu, G.; Chance, M. R. Radiolytic Modification of Acidic Amino Acid Residues in Peptides: Probes for Examining Protein−Protein Interactions. *Anal. Chem.* **2004**, *76*, 1213−1221.
(935) Xu, G.; Chance, M. R. Radiolytic Modification of Sulfur-Containing Amino Acid Residues in Model Peptides: Fundamental Studies for Protein Footprinting. *Anal. Chem.* **2005**, *77*, 2437−2449.
(936) Karmakar, T.; Balasubramanian, S. Elucidating the interaction of H2O2 with polar amino acids − Quantum chemical calculations. *Chem. Phys. Lett.* **2014**, *613*, 5−9.
(937) Davies, M.; Dean, R. *Radical-Mediated Protein Oxidation: From Chemistry to Medicine*. Oxford University Press: New York, 1997.
(938) Xiao, R.; Liu, K.; Bai, L.; Minakata, D.; Seo, Y.; Göktaş, R. K.; Dionysiou, D. D.; Tang, C.-J.; Wei, Z.; Spinney, R. Inactivation of pathogenic microorganisms by sulfate radical: Present and future. *Chem. Eng. J.* **2019**, *371*, 222−232.
(939) Davies, M. J. Protein oxidation and peroxidation. *Biochem. J.* **2016**, *473*, 805−825.
(940) Fancy, D. A.; Kodadek, T. Chemistry for the analysis of protein−protein interactions: rapid and efficient cross-linking triggered by long wavelength light. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 6020−6024.
(941) Borsarelli, C. D.; Falomir-Lockhart, L. J.; Ostatna, V.; Fauerbach, J. A.; Hsiao, H.-H.; Urlaub, H.; Paleček, E.; Jares-Erijman, E. A.; Jovin, T. M. Biophysical properties and cellular toxicity of covalent crosslinked oligomers of α-synuclein formed by photo-induced side-chain tyrosyl radicals. *Free Radical Biol. Med.* **2012**, *53*, 1004−1015.
(942) Rayshell, M.; Ross, J.; Werbin, H. Evidence that N-acetoxy-N-acetyl-2-aminofluorene crosslinks DNA to protein by a free radical mechanism. *Carcinogenesis* **1983**, *4*, 501−507.
(943) Bridgewater, J. D.; Vachet, R. W. Using microwave-assisted metal-catalyzed oxidation reactions and mass spectrometry to increase

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

the rate at which the copper-binding sites of a protein are determined. *Anal. Chem.* **2005**, *77*, 4649−4653.
(944) Bridgewater, J. D.; Lim, J.; Vachet, R. W. Using Metal-Catalyzed Oxidation Reactions and Mass Spectrometry to Identify Amino Acid Residues within 10 Å of the Metal in Cu-Binding Proteins. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1552−1559.
(945) Bridgewater, J. D.; Vachet, R. W. Metal-Catalyzed Oxidation Reactions and Mass Spectrometry: The Roles of Ascorbate and Different Oxidizing Agents in Determining Cu−Protein-Binding Sites. *Anal. Biochem.* **2005**, *341*, 122−130.
(946) Gau, B. C.; Chen, H.; Zhang, Y.; Gross, M. L. Sulfate Radical Anion as a New Reagent for Fast Photochemical Oxidation of Proteins. *Anal. Chem.* **2010**, *82*, 7821−7827.
(947) Davies, M. J.; Gilbert, B. C.; McCleland, C. W.; Thomas, C. B.; Young, J. E.s.r. Evidence for the Multiplicity of Side-Chain Oxidation Pathways in the Acid-Catalysed Decomposition of Substituted Hydroxycyclohexadienyl Radicals. *J. Chem. Soc., Chem. Commun.* **1984**, 966−967.
(948) Davies, M. J.; Gilbert, B. C.; Norman, R. O. C. Electron Spin Resonance. Part 67. Oxidation of Aliphatic Sulphides and Sulfoxides by the Sulphate Radical Anion ($SO4^{\cdot-}$) and of Aliphatic Radicals by the Peroxydisulphate Anion ($S2O8^{2-}$). *J. Chem. Soc., Perkin Trans. 2* **1984**, *2*, 503−509.
(949) Rustgi, S. M.; Riesz, P. An E.S.R. and Spin-Trapping Study of the Reactions of the $S^{\cdot}O4^{-}$ Radical with Protein and Nucleic Acid Constituents. *Int. J. Radiat. Biol. Relat. Stud. Phys., Chem. Med.* **1978**, *34*, 301−316.
(950) Lu, Y.; Zhang, H.; Niedzwiedzki, D. M.; Jiang, J.; Blankenship, R. E.; Gross, M. L. Fast Photochemical Oxidation of Proteins Maps the Topology of Intrinsic Membrane Proteins: Light-Harvesting Complex 2 in a Nanodisc. *Anal. Chem.* **2016**, *88*, 8827−8834.
(951) Medinas, D. B.; Cerchiaro, G.; Trindade, D. F.; Augusto, O. The Carbonate Radical And Related Oxidants Derived from Bicarbonate Buffer. *IUBMB Life* **2007**, *59*, 255−262.
(952) Sankarapandi, S.; Zweier, J. L. Bicarbonate Is Required for the Peroxidase Function of Cu,Zn-Superoxide Dismutase at Physiological pH. *J. Biol. Chem.* **1999**, *274*, 1226−1232.
(953) Hodgson, E. K.; Fridovich, I. The Mechanism of the Activity-Dependent Luminescence of Xanthine Oxidase. *Arch. Biochem. Biophys.* **1976**, *172*, 202−205.
(954) Augusto, O.; Bonini, M. G.; Amanso, A. M.; Linares, E.; Santos, C. C.; De Menezes, S. L. Nitrogen Dioxide and Carbonate Radical Anion: Two Emerging Radicals in Biology. *Free Radical Biol. Med.* **2002**, *32*, 841−859.
(955) Radi, R. Nitric Oxide, Oxidants, and Protein Tyrosine Nitration. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 4003−4008.
(956) Paviani, V.; Queiroz, R. F.; Marques, E. F.; Di Mascio, P.; Augusto, O. Production of Lysozyme and Lysozyme-Superoxide Dismutase Dimers Bound by a Ditryptophan Cross-Link in Carbonate Radical-Treated Lysozyme. *Free Radical Biol. Med.* **2015**, *89*, 72−82.
(957) Gebicka, L.; Didik, J.; Gebicki, J. Reactions of Heme Proteins with Carbonate Radical Anion. *Res. Chem. Intermed.* **2009**, *35*, 401−409.
(958) Shafirovich, V.; Dourandin, A.; Huang, W. D.; Geacintov, N. E. The Carbonate Radical is a Site-Selective Oxidizing Agent of Guanine in Double-Stranded Oligonucleotides. *J. Biol. Chem.* **2001**, *276*, 24621−24626.
(959) Chen, S. N.; Hoffman, M. Z. Rate Constants for the Reaction of the Carbonate Radical with Compounds of Biochemical Interest in Neutral Aqueous Solution. *Radiat. Res.* **1973**, *56*, 40−47.
(960) Medinas, D. B.; Gozzo, F. C.; Santos, L. F.; Iglesias, A. H.; Augusto, O. A Ditryptophan Cross-Link is Responsible for the Covalent Dimerization of Human Superoxide Dismutase 1 during its Bicarbonate-Dependent Peroxidase Activity. *Free Radical Biol. Med.* **2010**, *49*, 1046−1053.
(961) Coelho, F. R.; Iqbal, A.; Linares, E.; Silva, D. F.; Lima, F. S.; Cuccovia, I. M.; Augusto, O. Oxidation of the Tryptophan 32 Residue of Human Superoxide Dismutase 1 caused by its Bicarbonate-Dependent Peroxidase Activity Triggers the Non-Amyloid Aggregation of the Enzyme. *J. Biol. Chem.* **2014**, *289*, 30690−30701.
(962) Zhang, M. M.; Rempel, D. L.; Gross, M. L. A Fast Photochemical Oxidation of Proteins (FPOP) Platform for Free-Radical Reactions: The Carbonate Radical Anion with Peptides and Proteins. *Free Radical Biol. Med.* **2019**, *131*, 126−132.
(963) Chen, S.-n.; Hoffman, M. Z. Rate Constants for the Reaction of the Carbonate Radical with Compounds of Biochemical Interest in Neutral Aqueous Solution. *Radiat. Res.* **1973**, *56*, 40−47.
(964) Huie, R. E.; Clifton, C. L.; Neta, P. Electron-Transfer Reaction-Rates and Equilibria of the Carbonate and Sulfate Radical-Anions. *Radiat. Phys. Chem.* **1991**, *38*, 477−481.
(965) Paulsen, S. R. 3.3-Dialkyl-Diazacyclopropen-(1). *Angew. Chem.* **1960**, *72*, 781−782.
(966) Smith, R. A.; Knowles, J. R. Aryldiazirines. Potential Reagents for Photolabeling of Biological Receptor Sites. *J. Am. Chem. Soc.* **1973**, *95*, 5072−5073.
(967) Wofsy, L.; Metzger, H.; Singer, S. Affinity Labeling—a General Method for Labeling the Active Sites of Antibody and Enzyme Molecules. *Biochemistry* **1962**, *1*, 1031−1039.
(968) Ruoho, A. E.; Kiefer, H.; Roeder, P. E.; Singer, S. J. The Mechanism of Photoaffinity Labeling. *Proc. Natl. Acad. Sci. U. S. A.* **1973**, *70*, 2567−2571.
(969) Halloran, M. W.; Lumb, J.-P. Recent Applications of Diazirines in Chemical Proteomics. *Chem. - Eur. J.* **2019**, *25*, 4885−4898.
(970) Nicholas, R. A. Purification of the Membrane-Spanning Tryptic Peptides of the. Alpha. Polypeptide from (Sodium and Potassium Ion)-Activated Adenosine Triphosphatase Labeled with 1-Tritiospiro [Adamantane-4, 1'-Diazirine]. *Biochemistry* **1984**, *23*, 888−898.
(971) Farley, R. A. Identification of Hydrophobic Regions of the Calcium-Transport ATPase from Sarcoplasmic Reticulum after Photochemical Labeling with Adamantane Diazirine. *Int. J. Biochem.* **1983**, *15*, 1423−1427.
(972) Xi, J.; Liu, R.; Rossi, M. J.; Yang, J.; Loll, P. J.; Dailey, W. P.; Eckenhoff, R. G. Photoactive Analogues of the Haloether Anesthetics Provide High-Resolution Features from Low-Affinity Interactions. *ACS Chem. Biol.* **2006**, *1*, 377−384.
(973) Gomes, A. F.; Gozzo, F. C. Chemical Cross-Linking with a Diazirine Photoactivatable Cross-Linker Investigated by MALDI-and ESI-MS/MS. *J. Mass Spectrom.* **2010**, *45*, 892−899.
(974) Richards, F. M.; Lamed, R.; Wynn, R.; Patel, D.; Olack, G. Methylene as a Possible Universal Footprinting Reagent that will Include Hydrophobic Surface Areas: Overview and Feasibility: Properties of Diazirine as a Precursor. *Protein Sci.* **2000**, *9*, 2506−2517.
(975) Jumper, C. C.; Schriemer, D. C. Mass Spectrometry of Laser-Initiated Carbene Reactions for Protein Topographic Analysis. *Anal. Chem.* **2011**, *83*, 2913−2920.
(976) Jumper, C. C.; Bomgarden, R.; Rogers, J.; Etienne, C.; Schriemer, D. C. High-Resolution Mapping of Carbene-Based Protein Footprints. *Anal. Chem.* **2012**, *84*, 4411−4418.
(977) Ziemianowicz, D. S.; Bomgarden, R.; Etienne, C.; Schriemer, D. C. Amino Acid Insertion Frequencies Arising from Photoproducts Generated Using Aliphatic Diazirines. *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 2011−2021.
(978) Blencowe, A.; Hayes, W. Development and Application of Diazirines in Biological and Synthetic Macromolecular Systems. *Soft Matter* **2005**, *1*, 178−205.
(979) Hu, Y.; Guo, T.; Ye, X.; Li, Q.; Guo, M.; Liu, H.; Wu, Z. Dye Adsorption by Resins: Effect of Ionic Strength on Hydrophobic and Electrostatic Interactions. *Chem. Eng. J.* **2013**, *228*, 392−397.
(980) Christen, R. P.; Nomikos, S. I.; Smith, E. T. Probing Protein Electrostatic Interactions through Temperature/Reduction Potential Profiles. *JBIC, J. Biol. Inorg. Chem.* **1996**, *1*, 515−522.
(981) Zhang, B.; Rempel, D. L.; Gross, M. L. Protein Footprinting by Carbenes on a Fast Photochemical Oxidation of Proteins (FPOP) Platform. *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 552−555.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(982) Manzi, L.; Barrow, A. S.; Scott, D.; Layfield, R.; Wright, T. G.; Moses, J. E.; Oldham, N. J. Carbene Footprinting Accurately Maps Binding Sites in Protein−Ligand and Protein−Protein Interactions. *Nat. Commun.* **2016**, *7*, 13288.
(983) Irikura, K. K.; Goddard, W., III; Beauchamp, J. Singlet-Triplet Gaps in Substituted Carbenes CXY (X, Y= H, Fluoro, Chloro, Bromo, Iodo, Silyl). *J. Am. Chem. Soc.* **1992**, *114*, 48−51.
(984) *CRC Handbook of Organic Photochemistry and Photobiology*; Horspool, W. M., Lenci, F., Eds.; CRC Press: Boca Raton, FL, 2003; Vols. *1* and *2*.
(985) Das, J. Aliphatic Diazirines as Photoaffinity Probes for Proteins: Recent Developments. *Chem. Rev.* **2011**, *111*, 4405−4417.
(986) Bourissou, D.; Guerret, O.; Gabbai, F. P.; Bertrand, G. Stable Carbenes. *Chem. Rev.* **2000**, *100*, 39−92.
(987) Richards, C. A., Jr; Kim, S.-J.; Yamaguchi, Y.; Schaefer, H. F., III Dimethylcarbene: A Singlet Ground State? *J. Am. Chem. Soc.* **1995**, *117*, 10104−10107.
(988) Bertrand, G. *Carbene Chemistry: From Fleeting Intermediates to Powerful Reagents*; CRC Press: Boca Raton, FL, 2002.
(989) Chan, K. K. J.; O'Hagan, D. The Rare Fluorinated Natural Products and Biotechnological Prospects for Fluorine Enzymology. In *Methods in Enzymology*; Hopwood, D. A., Ed.; Elsevier Academic Press: San Diego, 2012; Vol. *516*, pp 219−235, DOI: 10.1016/B978-0-12-394291-3.00003-4.
(990) O'Hagan, D. Understanding Organofluorine Chemistry. An Introduction to the C−F Bond. *Chem. Soc. Rev.* **2008**, *37*, 308−319.
(991) Shannon, R. D. Revised Effective Ionic Radii and Systematic Studies of Interatomic Distances in Halides and Chalcogenides. *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr.* **1976**, *32*, 751−767.
(992) Müller, K.; Faeh, C.; Diederich, F. Fluorine in Pharmaceuticals: Looking Beyond Intuition. *Science* **2007**, *317*, 1881−1886.
(993) Brunner, J.; Senn, H.; Richards, F. 3-Trifluoromethyl-3-Phenyldiazirine. A New Carbene Generating Group for Photolabeling Reagents. *J. Biol. Chem.* **1980**, *255*, 3313−3318.
(994) Emeleus, H.; Haszeldine, R. 623. Organometallic Fluorine Compounds. Part I. The Synthesis of Trifluoromethyl and Pentafluoroethyl Mercurials. *J. Chem. Soc.* **1949**, *0*, 2948−2952.
(995) Naumann, D.; Kischkewitz, J. Trifluormethylierungsreaktionen von Te(CF3)2 mit halogenbenzolen und methylbenzolen. *J. Fluorine Chem.* **1990**, *47*, 283−299.
(996) Lane, S. I.; Oexler, E. V.; Staricco, E. H. The Gas-Phase Reaction of Cf3 Radicals with C6f5h in the Temperature-Range 373−658-K. *Int. J. Chem. Kinet.* **1991**, *23*, 361−367.
(997) Charpentier, J.; Früh, N.; Togni, A. Electrophilic Trifluoromethylation by Use of Hypervalent Iodine Reagents. *Chem. Rev.* **2015**, *115*, 650−682.
(998) Teruo, U.; Sumi, I. Power-Variable Trifluoromethylating Agents, (Trifluoromethyl)Dibenzothio- and -Selenophenium Salt System. *Tetrahedron Lett.* **1990**, *31*, 3579−3582.
(999) Mizuta, S.; Verhoog, S.; Engle, K. M.; Khotavivattana, T.; O'duill, M.; Wheelhouse, K.; Rassias, G.; Médebielle, M.; Gouverneur, V. r. Catalytic Hydrotrifluoromethylation Of Unactivated Alkenes. *J. Am. Chem. Soc.* **2013**, *135*, 2505−2508.
(1000) Mizuta, S.; Engle, K. M.; Verhoog, S.; Galicia-López, O.; O'Duill, M.; Médebielle, M.; Wheelhouse, K.; Rassias, G.; Thompson, A. L.; Gouverneur, V. Trifluoromethylation of Allylsilanes Under Photoredox Catalysis. *Org. Lett.* **2013**, *15*, 1250−1253.
(1001) Chaabouni, S.; Simonet, F.; François, A.; Abid, S.; Galaup, C.; Chassaing, S. 3-Trifluoromethylated Coumarins and Carbostyrils by Radical Trifluoromethylation of Ortho-Functionalized Cinnamic Esters. *Eur. J. Org. Chem.* **2017**, *2017*, 271−277.
(1002) Langlois, B. R.; Laurent, E.; Roidot, N. Trifluoromethylation of Aromatic Compounds with Sodium Trifluoromethanesulfinate Under Oxidative Conditions. *Tetrahedron Lett.* **1991**, *32*, 7525−7528.
(1003) Ji, Y.; Brueckl, T.; Baxter, R. D.; Fujiwara, Y.; Seiple, I. B.; Su, S.; Blackmond, D. G.; Baran, P. S. Innate CH Trifluoromethylation of Heterocycles. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 14411−14415.
(1004) Fujiwara, Y.; Dixon, J. A.; O'Hara, F.; Funder, E. D.; Dixon, D. D.; Rodriguez, R. A.; Baxter, R. D.; Herle, B.; Sach, N.; Collins, M. R.; Ishihara, Y.; Baran, P. S. Practical and Innate Carbon−Hydrogen Functionalization of Heterocycles. *Nature* **2012**, *492*, 95−99.
(1005) Fennewald, J. C.; Lipshutz, B. H. Trifluoromethylation of Heterocycles in Water at Room Temperature. *Green Chem.* **2014**, *16*, 1097−1100.
(1006) Cheng, M.; Zhang, B.; Cui, W.; Gross, M. L. Laser-Initiated Radical Trifluoromethylation of Peptides and Proteins: Application to Mass-Spectrometry-Based Protein Footprinting. *Angew. Chem., Int. Ed.* **2017**, *56*, 14007−14010.
(1007) Winkler, R. Iodine—a Potential Antioxidant and the Role of Iodine/Iodide in Health and Disease. *Nat. Sci.* **2015**, *7*, 548−557.
(1008) Cavalieri, R. R. Iodine Metabolism and Thyroid Physiology: Current Concepts. *Thyroid* **1997**, *7*, 177−181.
(1009) Soundarrajan, M.; Kopp, P. A. Thyroid Hormone Biosynthesis and Physiology. In *Thyroid Disease and Reproduction*; Eaton, J. L., Ed.; Springer: Berlin, 2019; pp 1−17, DOI: 10.1007/978-3-319-99079-8_1.
(1010) Kessler, J.; Obinger, C.; Eales, G. Factors influencing the study of Peroxidase-Generated Iodine Species and Implications for Thyroglobulin Synthesis. *Thyroid* **2008**, *18*, 769−774.
(1011) Rousset, B.; Dupuy, C.; Miot, F.; Dumont, J. Thyroid Hormone Synthesis And Secretion. In *Endotext*; Feingold, K. R., Anawalt, B., Boyce, A., Ed.; MDText. com, Inc., 2015; https://www.endotext.org/Chapter/Chapter-2-thyroid-hormone-synthesis-and-secretion/ (accessed 2020-03-20).
(1012) Verma, S.; Pradeep Kumar, G.; Laloraya, M.; Singh, A. Activation of Iodine into a Free-Radical Intermediate by Superoxide: a physiologically significant Step in the Iodination of Tyrosine. *Biochem. Biophys. Res. Commun.* **1990**, *170*, 1026−1034.
(1013) Goldsmith, S. J. Radioimmunoassay: Review of Basic Principles. In *Seminars in Nuclear Medicine*; Freeman, L. M., Donald Blaufox, M., Eds.; Elsevier: Amsterdam, 1975; pp 125−152, DOI: 10.1016/S0001-2998(75)80028-6.
(1014) Béhé, M.; Gotthardt, M.; Behr, T. M. Radioiodination of Proteins and Peptides. In *Cell Biology: A Laboratory Handbook*, 3rd Ed.; Celis, J. E., Ed.; Academic Press: Burlington, 2006; pp 149−154, DOI: 10.1016/B978-012164730-8/50205-7.
(1015) Caldwell, H. D.; Kromhout, J.; Schachter, J. Purification and Partial Characterization of the Major Outer Membrane Protein of Chlamydia Trachomatis. *Infect. Immun.* **1981**, *31*, 1161−1176.
(1016) Hatch, T. P.; Vance, D.; Al-Hossainy, E. Identification of a Major Envelope Protein in Chlamydia Spp. *J. Bacteriol.* **1981**, *146*, 426−429.
(1017) Rupp, W. D.; Prusoff, W. H. Photochemistry of iodouracil. II Effects of Sulfur Compounds, Ethanol and Oxygen. *Biochem. Biophys. Res. Commun.* **1965**, *18*, 158−164.
(1018) Watanabe, T.; Bando, T.; Xu, Y.; Tashiro, R.; Sugiyama, H. Efficient Generation of 2'-Deoxyuridin-5-yl at 5'-(G/C) AAXUXU-3'(X= Br, I) Sequences in Duplex DNA under UV Irradiation. *J. Am. Chem. Soc.* **2005**, *127*, 44−45.
(1019) Xu, Y.; Sugiyama, H. Photochemical Approach to Probing Different DNA Structures. *Angew. Chem., Int. Ed.* **2006**, *45*, 1354−1362.
(1020) Dai, X.; Song, D.; Liu, K.; Su, H. Photoinduced C—I bond Homolysis of 5-Iodouracil: A Singlet Predissociation Pathway. *J. Chem. Phys.* **2017**, *146*, 025103.
(1021) Ly, T.; Julian, R. R. Residue-Specific Radical-Directed Dissociation of Whole Proteins in the Gas Phase. *J. Am. Chem. Soc.* **2008**, *130*, 351−358.
(1022) Ly, T.; Julian, R. R. Elucidating the Tertiary Structure of Protein Ions in vacuo with Site Specific Photoinitiated Radical Reactions. *J. Am. Chem. Soc.* **2010**, *132*, 8602−8609.
(1023) Tao, Y.; Quebbemann, N. R.; Julian, R. R. Discriminating D-Amino Acid-Containing Peptide Epimers by Radical-Directed Dissociation Mass Spectrometry. *Anal. Chem.* **2012**, *84*, 6814−6820.
(1024) Ranka, K.; Zhao, N.; Yu, L.; Stanton, J. F.; Polfer, N. C. Radical Rearrangement Chemistry in Ultraviolet Photodissociation of

Iodotyrosine Systems: Insights from Metastable Dissociation, Infrared Ion Spectroscopy, and Reaction Pathway Calculations. *J. Am. Soc. Mass Spectrom.* **2018**, *29*, 1791−1801.
(1025) Li, C. H. Kinetics of Reactions Between Iodine and Histidine. *J. Am. Chem. Soc.* **1944**, *66*, 225−227.
(1026) Li, K. S.; Rempel, D. L.; Gross, M. L. Conformational-Sensitive Fast Photochemical Oxidation of Proteins and Mass Spectrometry Characterize Amyloid Beta 1−42 Aggregation. *J. Am. Chem. Soc.* **2016**, *138*, 12090−12098.
(1027) Köhler, G.; Milstein, C. Continuous Cultures of Fused Cells Secreting Antibody of Predefined Specificity. *Nature* **1975**, *256*, 495−497.
(1028) Coons, A. H.; Creech, H. J.; Jones, R. N. Immunological Properties of an Antibody Containing a Fluorescent Group. *Exp. Biol. Med.* **1941**, *47*, 200−202.
(1029) Engvall, E.; Perlmann, P. Enzyme-Linked Immunosorbent Assay, Elisa. III. Quantitation of Specific Antibodies by Enzyme-Labeled Anti-Immunoglobulin in Antigen-Coated Tubes. *J. Immunol.* **1972**, *109*, 129−135.
(1030) Burnette, W. N. Western Blotting": Electrophoretic Transfer of Proteins from Sodium Dodecyl Sulfate-Polyacrylamide Gels to Unmodified Nitrocellulose and Radiographic Detection with Antibody and Radioiodinated Protein A. *Anal. Biochem.* **1981**, *112*, 195−203.
(1031) Adams, G. P.; Weiner, L. M. Monoclonal Antibody Therapy of Cancer. *Nat. Biotechnol.* **2005**, *23*, 1147−1157.
(1032) Scott, A. M.; Wolchok, J. D.; Old, L. J. Antibody Therapy of Cancer. *Nat. Rev. Cancer* **2012**, *12*, 278−287.
(1033) Wang, W.; Singh, S.; Zeng, D. L.; King, K.; Nema, S. Antibody Structure, Instability, and Formulation. *J. Pharm. Sci.* **2007**, *96*, 1−26.
(1034) Tonegawa, S. Somatic Generation of Antibody Diversity. *Nature* **1983**, *302*, 575−581.
(1035) Van Regenmortel, M. H. V. Mapping Epitope Structure and Activity: From One-Dimensional Prediction to Four-Dimensional Description of Antigenic Specificity. *Methods* **1996**, *9*, 465−472.
(1036) Barlow, D. J.; Edwards, M. S.; Thornton, J. M. Continuous and Discontinuous Protein Antigenic Determinants. *Nature* **1986**, *322*, 747−748.
(1037) Amit, A.; Mariuzza, R.; Phillips, S.; Poljak, R. Three-Dimensional Structure of an Antigen-Antibody Complex at 2.8 A Resolution. *Science* **1986**, *233*, 747−753.
(1038) Mayer, M.; Meyer, B. Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angew. Chem., Int. Ed.* **1999**, *38*, 1784−1788.
(1039) Cunningham, B.; Wells, J. High-Resolution Epitope Mapping of hGH-Receptor Interactions by Alanine-Scanning Mutagenesis. *Science* **1989**, *244*, 1081−1085.
(1040) Geysen, H. M.; Meloen, R. H.; Barteling, S. J. Use of Peptide Synthesis to Probe Viral Antigens for Epitopes to a Resolution of a Single Amino Acid. *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81*, 3998−4002.
(1041) Chen, J.; Marechal, V.; Levine, A. J. Mapping of the p53 and mdm-2 Interaction Domains. *Mol. Cell. Biol.* **1993**, *13*, 4107−4114.
(1042) Suckau, D.; Köhl, J.; Karwath, G.; Schneider, K.; Casaretto, M.; Bitter-Suermann, D.; Przybylski, M. Molecular Epitope Identification by Limited Proteolysis of an Immobilized Antigen-Antibody Complex and Mass Spectrometric Peptide Mapping. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*, 9848−9852.
(1043) Zhao, Y.; Chait, B. T. Protein Epitope Mapping by Mass Spectrometry. *Anal. Chem.* **1994**, *66*, 3723−3726.
(1044) Opuni, K. F. M.; Al-Majdoub, M.; Yefremova, Y.; El-Kased, R. F.; Koy, C.; Glocker, M. O. Mass Spectrometric Epitope Mapping. *Mass Spectrom. Rev.* **2018**, *37*, 229−241.
(1045) Macht, M.; Fiedler, W.; Kürzinger, K.; Przybylski, M. Mass Spectrometric Mapping of Protein Epitope Structures of Myocardial Infarct Markers Myoglobin and Troponin T. *Biochemistry* **1996**, *35*, 15633−15639.
(1046) Al-Majdoub, M.; Opuni, K. F. M.; Koy, C.; Glocker, M. O. Facile Fabrication and Instant Application of Miniaturized Antibody-Decorated Affinity Columns for Higher-Order Structure and Functional Characterization of TRIM21 Epitope Peptides. *Anal. Chem.* **2013**, *85*, 10479−10487.
(1047) Baerga-Ortiz, A.; Hughes, C. A.; Mandell, J. G.; Komives, E. A. Epitope mapping of a Monoclonal Antibody Against Human Thrombin by H/D-Exchange Mass Spectrometry Reveals Selection of a Diverse Sequence in a Highly Conserved Protein. *Protein Sci.* **2002**, *11*, 1300−1308.
(1048) Zhang, H.-M.; Kazazic, S.; Schaub, T. M.; Tipton, J. D.; Emmett, M. R.; Marshall, A. G. Enhanced Digestion Efficiency, Peptide Ionization Efficiency, and Sequence Resolution for Protein Hydrogen/Deuterium Exchange Monitored by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **2008**, *80*, 9034−9041.
(1049) Johnson, D. J. D.; Adams, Ty E.; Li, W.; Huntington, J. A. Crystal Structure of Wild-Type Human Thrombin in the Na^+-Free State. *Biochem. J.* **2005**, *392*, 21−28.
(1050) Jones, L. M.; Sperry, J. B.; Carroll, J. A.; Gross, M. L. Fast Photochemical Oxidation of Proteins for Epitope Mapping. *Anal. Chem.* **2011**, *83*, 7657−7661.
(1051) Yan, Y.; Chen, G.; Wei, H.; Huang, R. Y.-C.; Mo, J.; Rempel, D. L.; Tymiak, A. A.; Gross, M. L. Fast Photochemical Oxidation of Proteins (FPOP) Maps the Epitope of EGFR Binding to Adnectin. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 2084−2092.
(1052) Zhang, Y.; Wecksler, A. T.; Molina, P.; Deperalta, G.; Gross, M. L. Mapping the Binding Interface of VEGF and a Monoclonal Antibody Fab-1 Fragment with Fast Photochemical Oxidation of Proteins (FPOP) and Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 850−858.
(1053) Li, X.; Grant, O. C.; Ito, K.; Wallace, A.; Wang, S.; Zhao, P.; Wells, L.; Lu, S.; Woods, R. J.; Sharp, J. S. Structural Analysis of the Glycosylated Intact HIV-1 gp120−b12 Antibody Complex Using Hydroxyl Radical Protein Footprinting. *Biochemistry* **2017**, *56*, 957−970.
(1054) Li, J.; Wei, H.; Krystek, S. R.; Bond, D.; Brender, T. M.; Cohen, D.; Feiner, J.; Hamacher, N.; Harshman, J.; et al. Mapping the Energetic Epitope of an Antibody/Interleukin-23 Interaction with Hydrogen/Deuterium Exchange, Fast Photochemical Oxidation of Proteins Mass Spectrometry, and Alanine Shave Mutagenesis. *Anal. Chem.* **2017**, *89*, 2250−2258.
(1055) Xie, B.; Sharp, J. S. Hydroxyl Radical Dosimetry for High Flux Hydroxyl Radical Protein Footprinting Applications Using a Simple Optical Detection Method. *Anal. Chem.* **2015**, *87*, 10719−10723.
(1056) Niu, B.; Mackness, B. C.; Rempel, D. L.; Zhang, H.; Cui, W.; Matthews, C. R.; Zitzewitz, J. A.; Gross, M. L. Incorporation of a Reporter Peptide in FPOP Compensates for Adventitious Scavengers and Permits Time-Dependent Measurements. *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 389−392.
(1057) Levinthal, C. Are There Pathways for Protein Folding? *J. Chim. Phys. Phys.-Chim. Biol.* **1968**, *65*, 44−45.
(1058) Anfinsen, C. B. Principles that Govern the Folding of Protein Chains. *Science* **1973**, *181*, 223−230.
(1059) Dill, K. A. Dominant Forces in Protein Folding. *Biochemistry* **1990**, *29*, 7133−7155.
(1060) Leopold, P. E.; Montal, M.; Onuchic, J. N. Protein Folding Funnels: A Kinetic Approach to the Sequence-Structure Relationship. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 8721−8725.
(1061) Dill, K. A.; Chan, H. S. From Levinthal to Pathways to Funnels. *Nat. Struct. Mol. Biol.* **1997**, *4*, 10−19.
(1062) Moult, J. A. decade of CASP: Progress, Bottlenecks and Prognosis in Protein Structure Prediction. *Curr. Opin. Struct. Biol.* **2005**, *15*, 285−289.
(1063) Dobson, C. M. Protein Folding and Misfolding. *Nature* **2003**, *426*, 884−890.
(1064) Dill, K. A.; MacCallum, J. L. The Protein-Folding Problem, 50 Years On. *Science* **2012**, *338*, 1042−1046.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(1065) Englander, S. W.; Mayne, L. The Nature of Protein Folding Pathways. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 15873−15880.
(1066) Roder, H.; Maki, K.; Cheng, H. Early Events in Protein Folding Explored by Rapid Mixing Methods. *Chem. Rev.* **2006**, *106*, 1836−1861.
(1067) Ballew, R. M.; Sabelko, J.; Gruebele, M. Direct Observation of Fast Protein Folding: The Initial Collapse of Apomyoglobin. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 5759−5764.
(1068) Gruebele, M.; Sabelko, J.; Ballew, R.; Ervin, J. Laser Temperature Jump Induced Protein Refolding. *Acc. Chem. Res.* **1998**, *31*, 699−707.
(1069) Jones, C. M.; Henry, E. R.; Hu, Y.; Chan, C. K.; Luck, S. D.; Bhuyan, A.; Roder, H.; Hofrichter, J.; Eaton, W. A. Fast Events in Protein Folding Initiated by Nanosecond Laser Photolysis. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 11860−11864.
(1070) Chan, C.-K.; Hu, Y.; Takahashi, S.; Rousseau, D. L.; Eaton, W. A.; Hofrichter, J. Submillisecond Protein Folding Kinetics Studied by Ultrarapid Mixing. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94*, 1779−1784.
(1071) Jeong, B.-S.; Dyer, R. B. Proton Transport Mechanism of M2 Proton Channel Studied by Laser-Induced pH Jump. *J. Am. Chem. Soc.* **2017**, *139*, 6621−6628.
(1072) Charlier, C.; Alderson, T. R.; Courtney, J. M.; Ying, J.; Anfinrud, P.; Bax, A. Study of Protein Folding under Native Conditions by Rapidly Switching the Hydrostatic Pressure Inside an NMR Sample Cell. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, E4169−E4178.
(1073) Eaton, W. A.; Muñoz, V.; Thompson, P. A.; Henry, E. R.; Hofrichter, J. Kinetics and Dynamics of Loops, α-Helices, β-Hairpins, and Fast-Folding Proteins. *Acc. Chem. Res.* **1998**, *31*, 745−753.
(1074) Mayor, U.; Johnson, C. M.; Daggett, V.; Fersht, A. R. Protein Folding and Unfolding in Microseconds to Nanoseconds by Experiment and Simulation. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 13518−13522.
(1075) Jennings, P.; Wright, P. Formation of a Molten Globule Intermediate Early in the Kinetic Folding Pathway of Apomyoglobin. *Science* **1993**, *262*, 892−896.
(1076) Barth, A. Infrared Spectroscopy of Proteins. *Biochim. Biophys. Acta, Bioenerg.* **2007**, *1767*, 1073−1101.
(1077) Udgaonkar, J. B.; Baldwin, R. L. NMR Evidence for an early Framework Intermediate on the Folding Pathway of Ribonuclease A. *Nature* **1988**, *335*, 694−699.
(1078) Roder, H.; Elöve, G. A.; Englander, S. W. Structural Characterization of Folding Intermediates in Cytochrome C by H-Exchange Labelling and Proton NMR. *Nature* **1988**, *335*, 700−704.
(1079) Dyer, R. B.; Gai, F.; Woodruff, W. H.; Gilmanshin, R.; Callender, R. H. Infrared Studies of Fast Events in Protein Folding. *Acc. Chem. Res.* **1998**, *31*, 709−716.
(1080) Deniz, A. A.; Laurence, T. A.; Beligere, G. S.; Dahan, M.; Martin, A. B.; Chemla, D. S.; Dawson, P. E.; Schultz, P. G.; Weiss, S. Single-Molecule Protein Folding: Diffusion Fluorescence Resonance Energy Transfer Studies of the Denaturation of Chymotrypsin Inhibitor 2. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 5179−5184.
(1081) Itzhaki, L. S.; Otzen, D. E.; Fersht, A. R. The Structure of the Transition State for Folding of Chymotrypsin Inhibitor 2 Analysed by Protein Engineering Methods: Evidence for a Nucleation-condensation Mechanism for Protein Folding. *J. Mol. Biol.* **1995**, *254*, 260−288.
(1082) Chowdhury, S. K.; Katta, V.; Chait, B. T. Probing Conformational Changes in Proteins by Mass Spectrometry. *J. Am. Chem. Soc.* **1990**, *112*, 9012−9013.
(1083) Simmons, D. A.; Konermann, L. Characterization of Transient Protein Folding Intermediates during Myoglobin Reconstitution by Time-Resolved Electrospray Mass Spectrometry with On-Line Isotopic Pulse Labeling. *Biochemistry* **2002**, *41*, 1906−1914.
(1084) Katta, V.; Chait, B. T. Hydrogen/Deuterium Exchange Electrospray Ionization Mass Spectrometry: A Method for Probing Protein Conformational Changes in Solution. *J. Am. Chem. Soc.* **1993**, *115*, 6317−6321.
(1085) Miranker, A.; Robinson, C.; Radford, S.; Aplin, R.; Dobson, C. Detection of Transient Protein Folding Populations by Mass Spectrometry. *Science* **1993**, *262*, 896−900.
(1086) Kragelund, B. B.; Robinson, C. V.; Knudsen, J.; Dobson, C. M.; Poulsen, F. M. Folding of a Four-Helix Bundle: Studies of Acyl-Coenzyme A Binding Protein. *Biochemistry* **1995**, *34*, 7217−7224.
(1087) West, G. M.; Tucker, C. L.; Xu, T.; Park, S. K.; Han, X.; Yates, J. R.; Fitzgerald, M. C. Quantitative Proteomics Approach for Identifying Protein−Drug Interactions in Complex Mixtures using Protein Stability Measurements. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 9078−9082.
(1088) Jin, L.; Wang, D.; Gooden, D. M.; Ball, C. H.; Fitzgerald, M. C. Targeted Mass Spectrometry-Based Approach for Protein−Ligand Binding Analyses in Complex Biological Mixtures Using a Phenacyl Bromide Modification Strategy. *Anal. Chem.* **2016**, *88*, 10987−10993.
(1089) Strickland, E. C.; Geer, M. A.; Tran, D. T.; Adhikari, J.; West, G. M.; DeArmond, P. D.; Xu, Y.; Fitzgerald, M. C. Thermodynamic Analysis of Protein-Ligand Binding Interactions in Complex Biological Mixtures using the Stability of Proteins from Rates of Oxidation. *Nat. Protoc.* **2013**, *8*, 148−161.
(1090) Geer, M. A.; Fitzgerald, M. C. Characterization of the Saccharomyces Cerevisiae ATP-Interactome using the iTRAQ-SPROX Technique. *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 233−243.
(1091) Adhikari, J.; West, G. M.; Fitzgerald, M. C. Global Analysis of Protein Folding Thermodynamics for Disease State Characterization. *J. Proteom. Res.* **2015**, *14*, 2287−2297.
(1092) Liu, F.; Meng, H.; Fitzgerald, M. C. Large-Scale Analysis of Breast Cancer-Related Conformational Changes in Proteins Using SILAC-SPROX. *J. Proteom. Res.* **2017**, *16*, 3277−3286.
(1093) Maleknia, S. D.; Downard, K. M. Unfolding of Apomyoglobin Helices by Synchrotron Radiolysis and Mass Spectrometry. *Eur. J. Biochem.* **2001**, *268*, 5578−5588.
(1094) Poor, T. A.; Jones, L. M.; Sood, A.; Leser, G. P.; Plasencia, M. D.; Rempel, D. L.; Jardetzky, T. S.; Woods, R. J.; Gross, M. L.; Lamb, R. A. Probing the Paramyxovirus Fusion (F) Protein-Refolding Event from Pre- to Postfusion by Oxidative Footprinting. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, E2596−E2605.
(1095) Calabrese, A. N.; Ault, J. R.; Radford, S. E.; Ashcroft, A. E. Using Hydroxyl Radical Footprinting to Explore the Free Energy Landscape of Protein Folding. *Methods* **2015**, *89*, 38−44.
(1096) Stocks, B. B.; Konermann, L. Structural Characterization of Short-Lived Protein Unfolding Intermediates by Laser-Induced Oxidative Labeling and Mass Spectrometry. *Anal. Chem.* **2009**, *81*, 20−27.
(1097) Stocks, B. B.; Konermann, L. Time-Dependent Changes in Side-Chain Solvent Accessibility during Cytochrome c Folding Probed by Pulsed Oxidative Labeling and Mass Spectrometry. *J. Mol. Biol.* **2010**, *398*, 362−373.
(1098) Stocks, B. B.; Rezvanpour, A.; Shaw, G. S.; Konermann, L. Temporal Development of Protein Structure during S100A11 Folding and Dimerization Probed by Oxidative Labeling and Mass Spectrometry. *J. Mol. Biol.* **2011**, *409*, 669−679.
(1099) Pan, Y.; Brown, L.; Konermann, L. Kinetic Folding Mechanism of an Integral Membrane Protein Examined by Pulsed Oxidative Labeling and Mass Spectrometry. *J. Mol. Biol.* **2011**, *410*, 146−158.
(1100) Vahidi, S.; Stocks, B. B.; Liaghati-Mobarhan, Y.; Konermann, L. Submillisecond Protein Folding Events Monitored by Rapid Mixing and Mass Spectrometry-Based Oxidative Labeling. *Anal. Chem.* **2013**, *85*, 8618−8625.
(1101) Wu, L.; Lapidus, L. J. Combining Ultrarapid Mixing with Photochemical Oxidation to Probe Protein Folding. *Anal. Chem.* **2013**, *85*, 4920−4924.
(1102) Mortensen, D. N.; Williams, E. R. Theta-Glass Capillaries in Electrospray Ionization: Rapid Mixing and Short Droplet Lifetimes. *Anal. Chem.* **2014**, *86*, 9315−9321.
(1103) Mortensen, D. N.; Williams, E. R. Ultrafast (1 μs) Mixing and Fast Protein Folding in Nanodrops Monitored by Mass Spectrometry. *J. Am. Chem. Soc.* **2016**, *138*, 3453−3460.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(1104) Braak, H.; Tredici, K. D.; Rüb, U.; de Vos, R. A. I.; Jansen Steur, E. N. H.; Braak, E. Staging of Brain Pathology Related to Sporadic Parkinson's disease. *Neurobiol. Aging* **2003**, *24*, 197−211.
(1105) Goedert, M.; Spillantini, M. G.; Del Tredici, K.; Braak, H. 100 Years of Lewy Pathology. *Nat. Rev. Neurol.* **2013**, *9*, 13−24.
(1106) Lewy, F. H. Paralysis agitans. I. Pathologische anatomie. *Handbuch der Neurologie* **1912**, *3*, 920−958.
(1107) Berchtold, N. C.; Cotman, C. W. Evolution in the Conceptualization of Dementia and Alzheimer's Disease: Greco-Roman Period to the 1960s. *Neurobiol. Aging* **1998**, *19*, 173−189.
(1108) Hamley, I. W. The Amyloid Beta Peptide: A Chemist's Perspective. Role in Alzheimer's and Fibrillization. *Chem. Rev.* **2012**, *112*, 5147−5192.
(1109) Vos, T.; Allen, C.; Arora, M.; Barber, R. M.; Bhutta, Z. A.; Brown, A.; Carter, A.; Casey, D. C.; Charlson, F. J.; Chen, A. Z.; et al. Global, Regional, and National Incidence, Prevalence, and Years Lived with Disability for 310 Diseases and Injuries, 1990−2015: A Systematic Analysis for the Global Burden of Disease Study 2015. *Lancet* **2016**, *388*, 1545−1602.
(1110) Hardy, J.; Selkoe, D. J. The Amyloid Hypothesis of Alzheimer's Disease: Progress and Problems on the Road to Therapeutics. *Science* **2002**, *297*, 353−356.
(1111) Chiti, F.; Dobson, C. M. Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* **2006**, *75*, 333−366.
(1112) Lührs, T.; Ritter, C.; Adrian, M.; Riek-Loher, D.; Bohrmann, B.; Döbeli, H.; Schubert, D.; Riek, R. 3D Structure of Alzheimer's Amyloid-β(1−42) Fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 17342−17347.
(1113) Petkova, A. T.; Ishii, Y.; Balbach, J. J.; Antzutkin, O. N.; Leapman, R. D.; Delaglio, F.; Tycko, R. A structural model for Alzheimer's β-amyloid Fibrils Based on Experimental Constraints from Solid State NMR. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 16742−16747.
(1114) Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J. W.; McFarlane, H. T.; et al. Atomic Structures of Amyloid Cross-β Spines Reveal Varied Steric Zippers. *Nature* **2007**, *447*, 453−457.
(1115) Miller, Y.; Ma, B.; Tsai, C.-J.; Nussinov, R. Hollow Core of Alzheimer's Aβ42 Amyloid Observed by Cryoem is Relevant at Physiological pH. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 14128−14133.
(1116) Gremer, L.; Schölzel, D.; Schenk, C.; Reinartz, E.; Labahn, J.; Ravelli, R. B. G.; Tusche, M.; Lopez-Iglesias, C.; Hoyer, W.; Heise, H.; et al. Fibril Structure of Amyloid-β(1−42) by Cryo−Electron Microscopy. *Science* **2017**, *358*, 116−119.
(1117) Economou, N. J.; Giammona, M. J.; Do, T. D.; Zheng, X.; Teplow, D. B.; Buratto, S. K.; Bowers, M. T. Amyloid β-Protein Assembly and Alzheimer's Disease: Dodecamers of Aβ42, but Not of Aβ40, Seed Fibril Formation. *J. Am. Chem. Soc.* **2016**, *138*, 1772−1775.
(1118) Levine, H., III Thioflavine T interaction with synthetic Alzheimer's Disease β-Amyloid Peptides: Detection of Amyloid Aggregation in Solution. *Protein Sci.* **1993**, *2*, 404−410.
(1119) Biancalana, M.; Koide, S. Molecular Mechanism of Thioflavin-T Binding to Amyloid Fibrils. *Biochim. Biophys. Acta, Proteins Proteomics* **2010**, *1804*, 1405−1412.
(1120) Groenning, M. Binding Mode of Thioflavin T and other Molecular Probes in the Context of Amyloid Fibrils—Current Status. *J. Chem. Biol.* **2010**, *3*, 1−18.
(1121) Wolfe, L. S.; Calabrese, M. F.; Nath, A.; Blaho, D. V.; Miranker, A. D.; Xiong, Y. Protein-induced Photophysical Changes to the Amyloid Indicator Dye Thioflavin T. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 16863−16868.
(1122) Stoeckli, M.; Staab, D.; Staufenbiel, M.; Wiederhold, K.-H.; Signor, L. Molecular Imaging of Amyloid β Peptides in Mouse Brain Sections using Mass Spectrometry. *Anal. Biochem.* **2002**, *311*, 33−39.
(1123) Bernstein, S. L.; Dupuis, N. F.; Lazo, N. D.; Wyttenbach, T.; Condron, M. M.; Bitan, G.; Teplow, D. B.; Shea, J.-E.; Ruotolo, B. T.; Robinson, C. V.; et al. Amyloid-β Protein Oligomerization and the Importance of Tetramers and Dodecamers in the Aetiology of Alzheimer's Disease. *Nat. Chem.* **2009**, *1*, 326−331.
(1124) Kheterpal, I.; Zhou, S.; Cook, K. D.; Wetzel, R. Aβ Amyloid Fibrils Possess a Core Structure Highly Resistant to Hydrogen Exchange. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 13597−13601.
(1125) Kheterpal, I.; Lashuel, H. A.; Hartley, D. M.; Walz, T.; Lansbury, P. T.; Wetzel, R. Aβ Protofibrils Possess a Stable Core Structure Resistant to Hydrogen Exchange. *Biochemistry* **2003**, *42*, 14092−14098.
(1126) Zhang, Y.; Rempel, D. L.; Zhang, J.; Sharma, A. K.; Mirica, L. M.; Gross, M. L. Pulsed Hydrogen−Deuterium Exchange Mass Spectrometry Probes Conformational Changes in Amyloid Beta (Aβ) Peptide Aggregation. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 14604−14609.
(1127) Pan, J.; Han, J.; Borchers, C. H.; Konermann, L. Conformer-Specific Hydrogen Exchange Analysis of Aβ(1−42) Oligomers by Top-Down Electron Capture Dissociation Mass Spectrometry. *Anal. Chem.* **2011**, *83*, 5386−5393.
(1128) Pan, J.; Han, J.; Borchers, C. H.; Konermann, L. Structure and Dynamics of Small Soluble Aβ(1−40) Oligomers Studied by Top-Down Hydrogen Exchange Mass Spectrometry. *Biochemistry* **2012**, *51*, 3694−3703.
(1129) Stoeckli, M.; Knochenmuss, R.; McCombie, G.; Mueller, D.; Rohner, T.; Staab, D.; Wiederhold, K. H. MALDI MS Imaging of Amyloid. In *Methods in Enzymology*; Kheterpal, I., Wetzel, R., Eds.; Elsevier Academic Press: San Diego, 2006; Vol. *412*, pp 94−106, DOI: 10.1016/S0076-6879(06)12007-8.
(1130) Grasso, G. The Use of Mass Spectrometry to Study Amyloid-β Peptides. *Mass Spectrom. Rev.* **2011**, *30*, 347−365.
(1131) Hoffmann, W.; von Helden, G.; Pagel, K. Ion Mobility-Mass Spectrometry and Orthogonal Gas-Phase Techniques to Study Amyloid Formation and Inhibition. *Curr. Opin. Struct. Biol.* **2017**, *46*, 7−15.
(1132) Klinger, A. L.; Kiselar, J.; Ilchenko, S.; Komatsu, H.; Chance, M. R.; Axelsen, P. H. A Synchrotron-Based Hydroxyl Radical Footprinting Analysis of Amyloid Fibrils and Prefibrillar Intermediates with Residue-Specific Resolution. *Biochemistry* **2014**, *53*, 7724−7734.
(1133) Williams, M., Daviter, T. *Protein−Ligand Interactions Methods and Applications*; Humana Press: New York, 2013.
(1134) Rossi, A. M.; Taylor, C. W. Analysis of Protein-Ligand Interactions by Fluorescence Polarization. *Nat. Protoc.* **2011**, *6*, 365−387.
(1135) Yan, Y.; Marriott, G. Analysis of Protein Interactions Using Fluorescence Technologies. *Curr. Opin. Chem. Biol.* **2003**, *7*, 635−640.
(1136) Paolini, S.; Tanfani, F.; Fini, C.; Bertoli, E.; Pelosi, P. Porcine Odorant-Binding Protein: Structural Stability and Ligand Affinities Measured by Fourier-Transform Infrared Spectroscopy and Fluorescence Spectroscopy. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1999**, *1431*, 179−188.
(1137) Meyer, B.; Peters, T. NMR Spectroscopy Techniques for Screening and Identifying Ligand Binding to Protein Receptors. *Angew. Chem., Int. Ed.* **2003**, *42*, 864−890.
(1138) Willets, K. A.; Van Duyne, R. P. Localized Surface Plasmon Resonance Spectroscopy and Sensing. *Annu. Rev. Phys. Chem.* **2007**, *58*, 267−297.
(1139) Kitova, E. N.; El-Hawiet, A.; Schnier, P. D.; Klassen, J. S. Reliable Determinations of Protein−Ligand Interactions by Direct ESI-MS Measurements. Are We There Yet? *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 431−441.
(1140) Powell, K. D.; Fitzgerald, M. C. Measurements of Protein Stability by H/D Exchange and Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry Using Picomoles of Material. *Anal. Chem.* **2001**, *73*, 3300−3304.
(1141) Liu, X. R.; Zhang, M. M.; Rempel, D. L.; Gross, M. L. Protein-Ligand Interaction by Ligand Titration, Fast Photochemical

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

Oxidation of Proteins and Mass Spectrometry: LITPOMS. *J. Am. Soc. Mass Spectrom.* **2019**, *30*, 213−217.
(1142) Katta, V.; Chait, B. T. Observation of the Heme-Globin Complex in Native Myoglobin by Electrospray-Ionization Mass Spectrometry. *J. Am. Chem. Soc.* **1991**, *113*, 8534−8535.
(1143) Ganem, B.; Li, Y. T.; Henion, J. D. Detection of Noncovalent Receptor-Ligand Complexes by Mass Spectrometry. *J. Am. Chem. Soc.* **1991**, *113*, 6294−6296.
(1144) Daniel, J. M.; Friess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R. Quantitative Determination of Noncovalent Binding Interactions Using Soft Ionization Mass Spectrometry. *Int. J. Mass Spectrom.* **2002**, *216*, 1−27.
(1145) Loo, R. R. O.; Goodlett, D. R.; Smith, R. D.; Loo, J. A. Observation of a Noncovalent Ribonuclease S-Protein/S-Peptide Complex by Electrospray Ionization Mass Spectrometry. *J. Am. Chem. Soc.* **1993**, *115*, 4391−4392.
(1146) Ayed, A.; Krutchinsky, A. N.; Ens, W.; Standing, K. G.; Duckworth, H. W. Quantitative Evaluation of Protein−Protein and Ligand−Protein Equilibria of a Large Allosteric Enzyme by Electrospray Ionization Time-Of-Flight Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 339−344.
(1147) Liu, J.; Konermann, L. Protein−Protein Binding Affinities in Solution Determined by Electrospray Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 408−417.
(1148) Loo, J. A.; Hu, P.; McConnell, P.; Mueller, W. T.; Sawyer, T. K.; Thanabal, V. A Study of Src SH2 Domain Protein-Phosphopeptide Binding Interactions by Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 234−243.
(1149) Greig, M. J.; Gaus, H.; Cummins, L. L.; Sasmor, H.; Griffey, R. H. Measurement of Macromolecular Binding Using Electrospray Mass Spectrometry. Determination of Dissociation Constants for Oligonucleotide: Serum Albumin Complexes. *J. Am. Chem. Soc.* **1995**, *117*, 10765−10766.
(1150) Jecklin, M. C.; Schauer, S.; Dumelin, C. E.; Zenobi, R. Label-Free Determination of Protein−Ligand Binding Constants using Mass Spectrometry and Validation using Surface Plasmon Resonance and Isothermal Titration Calorimetry. *J. Mol. Recognit.* **2009**, *22*, 319−329.
(1151) Jørgensen, T. J. D.; Roepstorff, P.; Heck, A. J. R. Direct Determination of Solution Binding Constants for Noncovalent Complexes between Bacterial Cell Wall Peptide Analogues and Vancomycin Group Antibiotics by Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **1998**, *70*, 4427−4432.
(1152) Griffey, R. H.; Sannes-Lowery, K. A.; Drader, J. J.; Mohan, V.; Swayze, E. E.; Hofstadler, S. A. Characterization of Low-Affinity Complexes between RNA and Small Molecules Using Electrospray Ionization Mass Spectrometry. *J. Am. Chem. Soc.* **2000**, *122*, 9933−9938.
(1153) Wang, W.; Kitova, E. N.; Klassen, J. S. Influence of Solution and Gas Phase Processes on Protein−Carbohydrate Binding Affinities Determined by Nanoelectrospray Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Anal. Chem.* **2003**, *75*, 4945−4955.
(1154) El-Hawiet, A.; Kitova, E. N.; Liu, L.; Klassen, J. S. Quantifying Labile Protein−Ligand Interactions Using Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1893−1899.
(1155) Wang, Y.; Park, H.; Lin, H.; Kitova, E. N.; Klassen, J. S. Multipronged ESI−MS Approach for Studying Glycan-Binding Protein Interactions with Glycoproteins. *Anal. Chem.* **2019**, *91*, 2140−2147.
(1156) Nguyen, G. T. H.; Tran, T. N.; Podgorski, M. N.; Bell, S. G.; Supuran, C. T.; Donald, W. A. Nanoscale Ion Emitters in Native Mass Spectrometry for Measuring Ligand−Protein Binding Affinities. *ACS Cent. Sci.* **2019**, *5*, 308−318.
(1157) Ghaemmaghami, S.; Fitzgerald, M. C.; Oas, T. G. A Quantitative, High-Throughput Screen for Protein Stability. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 8296−8301.
(1158) Tang, L.; Hopper, E. D.; Tong, Y.; Sadowsky, J. D.; Peterson, K. J.; Gellman, S. H.; Fitzgerald, M. C. H/D Exchange- and Mass Spectrometry-Based Strategy for the Thermodynamic Analysis of Protein−Ligand Binding. *Anal. Chem.* **2007**, *79*, 5869−5877.
(1159) Liu, X. R.; Rempel, D. L.; Gross, M. L. Composite Conformational Changes of Signaling Proteins upon Ligand Binding Revealed by a Single Approach: Calcium-Calmodulin Study. *Anal. Chem.* **2019**, *91*, 12560−12567.
(1160) Zhao, L.; Lai, L.; Zhang, Z. How Calcium Ion Binding Induces the Conformational Transition of the Calmodulin N-Terminal Domain—an Atomic Level Characterization. *Phys. Chem. Chem. Phys.* **2019**, *21*, 19795−19804.
(1161) Ottinger, L. M.; Tullius, T. D. High-Resolution in Vivo Footprinting of a Protein−DNA Complex Using γ-Radiation. *J. Am. Chem. Soc.* **2000**, *122*, 5901−5902.
(1162) Adilakshmi, T.; Soper, S. F. C.; Woodson, S. A. Structural Analysis of RNA in Living Cells by In Vivo Synchrotron X-Ray Footprinting. In *Methods in Enzymology*; Herschlag, D., Ed.; Elsevier Academic Press: San Diego, 2009; Vol. *468*, pp 239−258, DOI: 10.1016/S0076-6879(09)68012-5.
(1163) Clatterbuck Soper, S. F.; Dator, R. P.; Limbach, P. A.; Woodson, S. A. In Vivo X-Ray Footprinting of Pre-30S Ribosomes Reveals Chaperone-Dependent Remodeling of Late Assembly Intermediates. *Mol. Cell* **2013**, *52*, 506−516.
(1164) Adilakshmi, T.; Lease, R. A.; Woodson, S. A. Hydroxyl Radical Footprinting in vivo: Mapping Macromolecular Structures with Synchrotron Radiation. *Nucleic Acids Res.* **2006**, *34*, No. e64.
(1165) Hulscher, R. M.; Bohon, J.; Rappé, M. C.; Gupta, S.; D'Mello, R.; Sullivan, M.; Ralston, C. Y.; Chance, M. R.; Woodson, S. A. Probing the structure of ribosome assembly intermediates in vivo using DMS and hydroxyl radical footprinting. *Methods* **2016**, *103*, 49−56.
(1166) Espino, J. A.; Mali, V. S.; Jones, L. M. In Cell Footprinting Coupled with Mass Spectrometry for the Structural Analysis of Proteins in Live Cells. *Anal. Chem.* **2015**, *87*, 7971−7978.
(1167) Rinas, A.; Mali, V. S.; Espino, J. A.; Jones, L. M. Development of a Microflow System for In-Cell Footprinting Coupled with Mass Spectrometry. *Anal. Chem.* **2016**, *88*, 10052−10058.
(1168) Espino, J. A.; Jones, L. M. Illuminating Biological Interactions with in Vivo Protein Footprinting. *Anal. Chem.* **2019**, *91*, 6577−6584.
(1169) Zhu, Y.; Serra, A.; Guo, T.; Park, J. E.; Zhong, Q.; Sze, S. K. Application of Nanosecond Laser Photolysis Protein Footprinting to Study EGFR Activation by EGF in Cells. *J. Proteom. Res.* **2017**, *16*, 2282−2293.
(1170) Rhee, H.-W.; Zou, P.; Udeshi, N. D.; Martell, J. D.; Mootha, V. K.; Carr, S. A.; Ting, A. Y. Proteomic Mapping of Mitochondria in Living Cells via Spatially Restricted Enzymatic Tagging. *Science* **2013**, *339*, 1328−1331.
(1171) Han, S.; Li, J.; Ting, A. Y. Proximity Labeling: Spatially Resolved Proteomic Mapping for Neurobiology. *Curr. Opin. Neurobiol.* **2018**, *50*, 17−23.
(1172) Mortensen, A.; Skibsted, L. H. Importance of Carotenoid Structure in Radical-Scavenging Reactions. *J. Agric. Food Chem.* **1997**, *45*, 2970−2977.
(1173) Mayer, G.; Bendayan, M. Biotinyl−Tyramide: A Novel Approach for Electron Microscopic Immunocytochemistry. *J. Histochem. Cytochem.* **1997**, *45*, 1449−1454.
(1174) Lam, S. S.; Martell, J. D.; Kamer, K. J.; Deerinck, T. J.; Ellisman, M. H.; Mootha, V. K.; Ting, A. Y. Directed Evolution of APEX2 for Electron Microscopy and Proximity Labeling. *Nat. Methods* **2015**, *12*, 51−54.
(1175) Loh, K. H.; Stawski, P. S.; Draycott, A. S.; Udeshi, N. D.; Lehrman, E. K.; Wilton, D. K.; Svinkina, T.; Deerinck, T. J.; Ellisman, M. H.; Stevens, B.; Carr, S. A.; Ting, A. Y. Proteomic Analysis of Unbounded Cellular Compartments: Synaptic Clefts. *Cell* **2016**, *166*, 1295.
(1176) Lee, S.-Y.; Kang, M.-G.; Shin, S.; Kwak, C.; Kwon, T.; Seo, J. K.; Kim, J.-S.; Rhee, H.-W. Architecture Mapping of the Inner

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

Mitochondrial Membrane Proteome by Chemical Tools in Live Cells. *J. Am. Chem. Soc.* **2017**, *139*, 3651−3662.
(1177) Shen, Y.; Lange, O.; Delaglio, F.; Rossi, P.; Aramini, J. M.; Liu, G.; Eletsky, A.; Wu, Y.; Singarapu, K. K.; Lemak, A.; et al. Consistent Blind Protein Structure Generation from NMR Chemical Shift Data. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 4685−4690.
(1178) Huang, W.; Ravikumar, K. M.; Parisien, M.; Yang, S. Theoretical Modeling of Multiprotein Complexes by iSPOT: Integration of Small-Angle X-Ray Scattering, Hydroxyl Radical Footprinting, and Computational Docking. *J. Struct. Biol.* **2016**, *196*, 340−349.
(1179) DiMaio, F.; Tyka, M. D.; Baker, M. L.; Chiu, W.; Baker, D. Refinement of Protein Structures into Low-Resolution Density Maps Using Rosetta. *J. Mol. Biol.* **2009**, *392*, 181−190.
(1180) DiMaio, F.; Song, Y.; Li, X.; Brunner, M. J.; Xu, C.; Conticello, V.; Egelman, E.; Marlovits, T. C.; Cheng, Y.; Baker, D. Atomic-Accuracy Models from 4.5-Å Cryo-Electron Microscopy Data with Density-Guided Iterative Local Refinement. *Nat. Methods* **2015**, *12*, 361−365.
(1181) Alexander, N. S.; Stein, R. A.; Koteiche, H. A.; Kaufmann, K. W.; McHaourab, H. S.; Meiler, J. RosettaEPR: Rotamer Library for Spin Label Structure and Dynamics. *PLoS One* **2013**, *8*, No. e72851.
(1182) Walzthoeni, T.; Joachimiak, L. A.; Rosenberger, G.; Röst, H. L.; Malmström, L.; Leitner, A.; Frydman, J.; Aebersold, R. xTract: Software for Characterizing Conformational Changes of Protein Complexes by Quantitative Cross-Linking Mass Spectrometry. *Nat. Methods* **2015**, *12*, 1185−1190.
(1183) Gerega, S. K.; Downard, K. M. PROXIMO—a New Docking Algorithm to Model Protein Complexes using Data from Radical Probe Mass Spectrometry (RP-MS). *Bioinformatics* **2006**, *22*, 1702−1709.
(1184) Huang, W.; Ravikumar, K. M.; Chance, M. R.; Yang, S. Quantitative Mapping of Protein Structure by Hydroxyl Radical Footprinting-Mediated Structural Mass Spectrometry: A Protection Factor Analysis. *Biophys. J.* **2015**, *108*, 107−115.
(1185) Huang, W.; Peng, Y.; Kiselar, J.; Zhao, X.; Albaqami, A.; Mendez, D.; Chen, Y.; Chakravarthy, S.; Gupta, S.; Ralston, C.; Kao, H.-Y.; Chance, M. R.; Yang, S. Multidomain architecture of Estrogen Receptor Reveals Interfacial Cross-Talk between its DNA-Binding and Ligand-Binding Domains. *Nat. Commun.* **2018**, *9*, 3520.
(1186) Xie, B.; Sood, A.; Woods, R. J.; Sharp, J. S. Quantitative Protein Topography Measurements by High Resolution Hydroxyl Radical Protein Footprinting Enable Accurate Molecular Model Selection. *Sci. Rep.* **2017**, *7*, 4552.
(1187) Aprahamian, M. L.; Chea, E. E.; Jones, L. M.; Lindert, S. Rosetta Protein Structure Prediction from Hydroxyl Radical Protein Footprinting Mass Spectrometry Data. *Anal. Chem.* **2018**, *90*, 7721−7729.
(1188) Seffernick, J. T.; Harvey, S. R.; Wysocki, V. H.; Lindert, S. Predicting Protein Complex Structure from Surface-Induced Dissociation Mass Spectrometry Data. *ACS Cent. Sci.* **2019**, *5*, 1330−1341.
(1189) Nosaka, Y.; Nosaka, A. Y. Generation and Detection of Reactive Oxygen Species in Photocatalysis. *Chem. Rev.* **2017**, *117*, 11302−11336.
(1190) Weisz, D. A.; Gross, M. L.; Pakrasi, H. B. Reactive Oxygen Species Leave a Damage Trail That Reveals Water Channels in Photosystem II. *Sci. Adv.* **2017**, *3*, No. eaao3013.
(1191) Loll, B.; Kern, J.; Saenger, W.; Zouni, A.; Biesiadka, J. Towards Complete Cofactor Arrangement in the 3.0 Å Resolution Structure of Photosystem II. *Nature* **2005**, *438*, 1040−1044.
(1192) Chaplin, M. Do We Underestimate the importance of Water in Cell Biology? *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 861−866.
(1193) Ball, P. Water as an Active Constituent in Cell Biology. *Chem. Rev.* **2008**, *108*, 74−108.
(1194) Jiang, J.-S.; Brünger, A. T. Protein Hydration Observed by X-ray Diffraction: Solvation Properties of Penicillopepsin and Neuraminidase Crystal Structures. *J. Mol. Biol.* **1994**, *243*, 100−115.
(1195) Persson, E.; Halle, B. Nanosecond to Microsecond Protein Dynamics Probed by Magnetic Relaxation Dispersion of Buried Water Molecules. *J. Am. Chem. Soc.* **2008**, *130*, 1774−1787.
(1196) Gupta, S.; D'Mello, R.; Chance, M. R. Structure and Dynamics of Protein Waters Revealed by Radiolysis and Mass Spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 14882−14887.
(1197) Angel, T. E.; Chance, M. R.; Palczewski, K. Conserved waters mediate structural and Functional Activation of Family A (Rhodopsin-Like) G Protein-Coupled Receptors. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 8555−8560.
(1198) Orban, T.; Gupta, S.; Palczewski, K.; Chance, M. R. Visualizing Water Molecules in Transmembrane Proteins Using Radiolytic Labeling Methods. *Biochemistry* **2010**, *49*, 827−834.
(1199) Kleene, R.; Schachner, M. Glycans and Neural Cell Interactions. *Nat. Rev. Neurosci.* **2004**, *5*, 195−208.
(1200) Blixt, O.; Head, S.; Mondala, T.; Scanlan, C.; Huflejt, M. E.; Alvarez, R.; Bryan, M. C.; Fazio, F.; Calarese, D.; et al. Printed Covalent Glycan Array for Ligand Profiling of Diverse Glycan Binding Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 17033−17038.
(1201) Bavro, V. N.; Gupta, S.; Ralston, C. Oxidative Footprinting in the Study of Structure and Function of Membrane Proteins: Current State and Perspectives. *Biochem. Soc. Trans.* **2015**, *43*, 983−994.
(1202) Mummadisetti, M. P.; Frankel, L. K.; Bellamy, H. D.; Sallans, L.; Goettert, J. S.; Brylinski, M.; Limbach, P. A.; Bricker, T. M. Use Of Protein Cross-Linking and Radiolytic Footprinting to Elucidate PsbP and Psbq Interactions within Higher Plant Photosystem II. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 16178−16183.
(1203) Shum, W.-K.; Maleknia, S. D.; Downard, K. M. Onset of Oxidative Damage in α-Crystallin by Radical Probe Mass Spectrometry. *Anal. Biochem.* **2005**, *344*, 247−256.
(1204) Hart, K. M.; Ho, C. M. W.; Dutta, S.; Gross, M. L.; Bowman, G. R. Modelling Proteins' Hidden Conformations to Predict Antibiotic Resistance. *Nat. Commun.* **2016**, *7*, 12965.
(1205) Loizos, N.; Darst, S. A. Mapping Protein-Ligand Interactions by Footprinting, a Radical Idea. *Structure* **1998**, *6*, 691−695.
(1206) Rogstad, S.; Faustino, A.; Ruth, A.; Keire, D.; Boyne, M.; Park, J. A Retrospective Evaluation of the Use of Mass Spectrometry in FDA Biologics License Applications. *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 786−794.
(1207) Chalmers, M. J.; Busby, S. A.; Pascal, B. D.; He, Y.; Hendrickson, C. L.; Marshall, A. G.; Griffin, P. R. Probing Protein Ligand Interactions by Automated Hydrogen/Deuterium Exchange Mass Spectrometry. *Anal. Chem.* **2006**, *78*, 1005−1014.
(1208) Iacob, R. E.; Murphy, J. P., III; Engen, J. R. Ion Mobility Adds an Additional Dimension to Mass Spectrometric Analysis of Solution-Phase Hydrogen/Deuterium Exchange. *Rapid Commun. Mass Spectrom.* **2008**, *22*, 2898−2904.
(1209) Nishimura, C.; Dyson, H. J.; Wright, P. E. Enhanced Picture of Protein-Folding Intermediates using Organic Solvents in H/D Exchange and Quench-Flow Experiments. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 4765−4770.
(1210) Zhang, Z. Retention Time Alignment of LC/MS Data by a Divide-and-Conquer Algorithm. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 764−772.
(1211) Claesen, J.; Burzykowski, T. Computational Methods and Challenges in Hydrogen/Deuterium Exchange Mass Spectrometry. *Mass Spectrom. Rev.* **2017**, *36*, 649−667.
(1212) Denisov, I. G.; Sligar, S. G. Nanodiscs in Membrane Biochemistry and Biophysics. *Chem. Rev.* **2017**, *117*, 4669−4713.
(1213) Popot, J.-L. Amphipols, Nanodiscs, and Fluorinated Surfactants: Three Nonconventional Approaches to Studying Membrane Proteins in Aqueous Solutions. *Annu. Rev. Biochem.* **2010**, *79*, 737−775.
(1214) Manzi, L.; Barrow, A. S.; Hopper, J. T. S.; Kaminska, R.; Kleanthous, C.; Robinson, C. V.; Moses, J. E.; Oldham, N. J. Carbene Footprinting Reveals Binding Interfaces of a Multimeric Membrane-Spanning Protein. *Angew. Chem., Int. Ed.* **2017**, *56*, 14873−14877.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454

(1215) Cecil, R.; McPhee, J. R. A Kinetic Study of the Reactions on some Disulphides with Sodium Sulphite. *Biochem. J.* **1955**, *60*, 496−506.
(1216) Heitmann, P. A Model for Sulfhydryl Groups in Proteins. Hydrophobic Interactions of the Cysteine Side Chain in Micelles. *Eur. J. Biochem.* **1968**, *3*, 346−350.
(1217) Means, G. E.; Congdon, W. I.; Bender, M. L. Reactions of 2,4,6-Trinitrobenzenesulfonate Ion with Amines and Hydroxide Ion. *Biochemistry* **1972**, *11*, 3564−3571.
(1218) Schirch, L.; Slagel, S.; Barra, D.; Martini, F.; Bossa, F. Evidence for a sulfhydryl group at the Active Site of Serine Transhydroxymethylase. *J. Biol. Chem.* **1980**, *255*, 2986−2989.
(1219) Mennella, C.; Visciano, M.; Napolitano, A.; Del Castillo, M. D.; Fogliano, V. Glycation of Lysine-Containing Dipeptides. *J. Pept. Sci.* **2006**, *12*, 291−296.
(1220) Cornwell, O.; Bond, N. J.; Radford, S. E.; Ashcroft, A. E. Long-Range Conformational Changes in Monoclonal Antibodies Revealed Using FPOP-LC-MS/MS. *Anal. Chem.* **2019**, *91*, 15163−15170.
(1221) Purushottam, L.; Adusumalli, S. R.; Singh, U.; Unnikrishnan, V. B.; Rawale, D. G.; Gujrati, M.; Mishra, R. K.; Rai, V. Single-Site Glycine-Specific Labeling of Proteins. *Nat. Commun.* **2019**, *10*, 2539.
(1222) Osberger, T. J.; Rogness, D. C.; Kohrt, J. T.; Stepan, A. F.; White, M. C. Oxidative Diversification of Amino Acids and Peptides by Small-Molecule Iron Catalysis. *Nature* **2016**, *537*, 214−219.
(1223) Huang, R. Y. C.; Garai, K.; Frieden, C.; Gross, M. L. Hydrogen/Deuterium Exchange and Electron-Transfer Dissociation Mass Spectrometry Determine the Interface and Dynamics of Apolipoprotein E Oligomerization. *Biochemistry* **2011**, *50*, 9273−9282.
(1224) Pan, Y.; Piyadasa, H.; O'Neil, J. D.; Konermann, L. Conformational Dynamics of a Membrane Transport Protein Probed by H/D Exchange and Covalent Labeling: The Glycerol Facilitator. *J. Mol. Biol.* **2012**, *416*, 400−413.
(1225) Baud, A.; Gonnet, F.; Salard, I.; Le Mignon, M.; Giuliani, A.; Mercère, P.; Sclavi, B.; Daniel, R. Probing the Solution Structure of Factor H using Hydroxyl Radical Protein Footprinting and Cross-Linking. *Biochem. J.* **2016**, *473*, 1805−1819.
(1226) Lin, M.; Krawitz, D.; Callahan, M. D.; Deperalta, G.; Wecksler, A. T. Characterization of ELISA Antibody-Antigen Interaction using Footprinting-Mass Spectrometry and Negative Staining Transmission Electron Microscopy. *J. Am. Soc. Mass Spectrom.* **2018**, *29*, 961−971.
(1227) Shi, L.; Liu, T.; Gross, M. L.; Huang, Y. Recognition of Human IgG1 by Fcγ Receptors: Structural Insights from Hydrogen−Deuterium Exchange and Fast Photochemical Oxidation of Proteins Coupled with Mass Spectrometry. *Biochemistry* **2019**, *58*, 1074−1080.
(1228) Delahunty, C.; Yates, J. R., Iii Protein Identification using 2D-LC-MS/MS. *Methods* **2005**, *35*, 248−255.
(1229) Doneanu, C.; Xenopoulos, A.; Fadgen, K.; Murphy, J.; Skilton, S. J.; Prentice, H.; Stapels, M.; Chen, W. Analysis of Host-Cell Proteins in Biotherapeutic Proteins by Comprehensive Online Two-Dimensional Liquid Chromatography/Mass Spectrometry. *mAbs* **2012**, *4*, 24−44.
(1230) Garbis, S. D.; Roumeliotis, T. I.; Tyritzis, S. I.; Zorpas, K. M.; Pavlakis, K.; Constantinides, C. A. A Novel Multidimensional Protein Identification Technology Approach Combining Protein Size Exclusion Prefractionation, Peptide Zwitterion−Ion Hydrophilic Interaction Chromatography, and Nano-Ultraperformance RP Chromatography/nESI-MS2 for the in-Depth Analysis of the Serum Proteome and Phosphoproteome: Application to Clinical Sera Derived from Humans with Benign Prostate Hyperplasia. *Anal. Chem.* **2011**, *83*, 708−718.
(1231) Bouvier, E. S. P.; Koza, S. M. Advances in Size-Exclusion Separations of Proteins and Polymers by UHPLC. *TrAC, Trends Anal. Chem.* **2014**, *63*, 85−94.
(1232) Muneeruddin, K.; Nazzaro, M.; Kaltashov, I. A. Characterization of Intact Protein Conjugates and Biopharmaceuticals Using Ion-Exchange Chromatography with Online Detection by Native Electrospray Ionization Mass Spectrometry and Top-Down Tandem Mass Spectrometry. *Anal. Chem.* **2015**, *87*, 10138−10145.

https://dx.doi.org/10.1021/acs.chemrev.9b00815
Chem. Rev. 2020, 120, 4355−4454