Biochemistry

Article

Cite This: *Biochemistry* 2019, 58, 1810−1817

pubs.acs.org/biochemistry

Using Peptide Arrays To Discover the Sequence-Specific Acetylation of the Histidine-Tyrosine Dyad

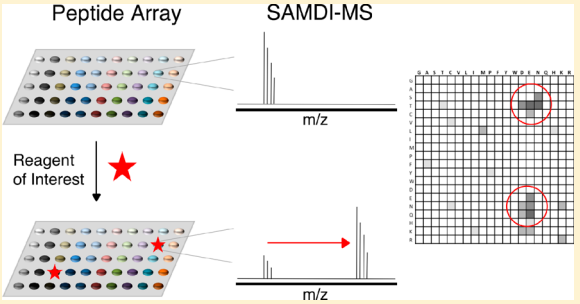
Lindsey C. Szymczak[†] and Milan Mrksich*[,†,‡]

[†]Department of Chemistry, Northwestern University, Evanston, Illinois 60208, United States

[‡]Department of Biomedical Engineering, Northwestern University, Evanston, Illinois 60208, United States

Supporting Information

ABSTRACT: Reactions that can selectively modify amino acid sequences within peptides and proteins are important for preparing protein reagents, immobilizing proteins, and making antibody-drug conjugates. The development of new reactions often begins with known chemistries and optimizes yields using a small set of peptide reactants. This article describes the use of peptide arrays and self-assembled monolayers for matrix-assisted laser desorption/ionization mass spectrometry (SAMDI-MS) to discover and characterize unanticipated sequence-selective reactions of peptides. This work reports the selective acetylation of HY (histidine-tyrosine) and YH (tyrosine-histidine) dyads when treated with acetic anhydride in aqueous conditions. More broadly, this example illustrates the benefits of using peptide arrays and a label-free analysis method to discover peptide-modifying reactions and gain mechanistic insight into their sequence specificity.



Reactions of chemical reagents with peptides and proteins are common and remain important for fundamental studies of protein function,[1] developing probes for *in vivo* imaging,[2,3] introducing motifs for purification,[4] and the development of biotherapeutics[5,6] and protein-based materials.[7,8] Important work continues to develop reactions that can selectively modify proteins, either at individual amino acids, or more selectively at peptide motifs.[9,10] However, the development of new chemistries is often validated on a small set of peptide sequences rather than across all peptide sequence variations, which limits the understanding of the specificity of the reaction. Developing sequence-selective reactions remains challenging, owing to the limited set of functional groups present in proteins and the need to use aqueous solutions near a neutral pH. The time required to isolate and characterize peptide reaction products also limits the number of peptide substrates that are analyzed. Hence, current approaches are not well-suited to identify unexpected reactivities and provide an understanding of the specificity of the reactions. In this article, we introduce the use of peptide arrays and self-assembled monolayers for matrix-assisted laser desorption/ionization (SAMDI) mass spectrometry (MS) as a method for discovering new acylation reactions on peptides.

The SAMDI-MS technique uses self-assembled monolayers (SAMs) of alkanethiolates on gold that present functional groups for immobilization, often a maleimide group, at a density of 10% against a background of tri(ethylene glycol) groups, which prevent nonspecific adsorption to the surface.[11,12] The maleimide group is used to immobilize thiol-derived molecules, including cysteine-terminated peptides. The functionalized SAMs can be treated with reagents of interest and then analyzed using matrix-assisted laser desorption/ionization (MALDI) mass spectrometry to reveal new peaks that correspond to reaction products. SAMDI is a high-throughput technique that provides a quantitative measure of reaction without the need for purification or tedious sample preparation. Because SAMDI-MS is a label-free technique, it can also be applied to a broad range of reactions and identify unanticipated products.

The combination of peptide arrays and SAMDI-MS offers an efficient approach for the discovery of sequence-specific reactions; a single array, for example, can be used to assess the reactivity of a reagent with hundreds to thousands of peptide sequences.[13,14] For example, we recently demonstrated the use of peptide arrays and SAMDI-MS to identify optimal substrates for glycosyl transferases.[15,16] Here, we describe the use of two peptide arrays with SAMDI-MS to discover and characterize the sequence-selective reaction of acetic anhydride with the histidine-tyrosine dyad. We find that acetic anhydride has preferential reactivity with peptides containing both histidine and tyrosine, and we investigate the mechanism behind this selective reactivity. We show that this reaction proceeds with initial acetylation on histidine, followed by acetyl transfer to tyrosine. We further characterize this reactivity with a second peptide library to examine the effects of amino acids neighboring the HY dyad. Finally, we examine the reactivities of two additional acylating electrophiles, as well as the spatial effects of this acetyl transfer. The combination of

Received: January 9, 2019
Revised: February 18, 2019
Published: February 28, 2019

DOI: 10.1021/acs.biochem.9b00022
Biochemistry 2019, 58, 1810−1817

peptide arrays and SAMDI-MS provides a powerful platform for reaction discovery on peptides.

MATERIAL AND METHODS

Peptide Library Synthesis. Peptide libraries, Ac-GXZRGC and Ac-GXYHZGRC, were synthesized using standard Fmoc-based solid-phase peptide synthesis[17] on SynPhase polyamide lanterns with Fmoc-Rink amide linkers purchased from Mimotopes. Synthesis was performed in 96-well filter plates purchased from Arctic White with the use of a 96-well plate vacuum manifold purchased from Millipore. Fmoc- and side chain-protected amino acids and *N*-acetylglycine were purchased from either Anaspec or Sigma-Aldrich. Coupling reagents, Pybop and *N*-methylmorpholine (NMM), as well as dimethylformamide (DMF) were purchased from Fisher or Sigma-Aldrich.

For each library, lanterns were swelled for 1 h in DMF and then separated into four 96-well filter plates. The lanterns were Fmoc-deprotected with a solution of 20% piperidine in DMF for 20 min. Using vacuum filtration, the wells were drained and the lanterns were rinsed with DMF. The first coupling reaction was performed for 20 min with Fmoc-Cys(Trt)−OH, Pybop, and NMM in 250 μL of DMF at a molar excess of 8:8:16, respectively, to the molar loading capacity of the lanterns. The wells were then drained, and the lanterns were rinsed with DMF. Fmoc-deprotection, washing, and coupling reactions with respective amino acids were repeated until the last amino acid (*N*-acetylglycine) was coupled. After the last coupling reaction, the lanterns were rinsed with DMF and dichloromethane and dried under a vacuum for 1 h. Lanterns were then transferred into new 96-well plates. Amino acid side chain deprotection and peptide cleavage from the lanterns was performed in tandem for 2 h using a solution of 2.5% triethylsilane, 2.5% H_2O, and 95% trifluoroacetic acid (TFA). The lanterns were then discarded, and the cleavage solution was evaporated under nitrogen. The peptides were resuspended in 0.1% TFA in H_2O and lyophilized. The peptides were resuspended again in 0.1% TFA in H_2O to a final concentration of 500 μM.

Additional Peptide Synthesis. Sixteen additional peptides, Ac-GYHGRG, Ac-GKGRG, Ac-GAGRC, NH2-GAGRG, Ac-GHGRG, Ac-GYGRG, Ac-GAGRG, Ac-GYHGC, Ac-GY^{OMe}HGC (O-methylated), Ac-GHGC, Ac-GYGC, Ac-GYHRGC, Ac-GYGHRGC, Ac-GYGGHRGC, Ac-GYGGGHRGC, and Ac-GYGGGGHRGC, were synthesized using standard Fmoc-based solid-phase peptide synthesis on Fmoc-Rink Amide MBHA resin purchased from Anaspec. The O-methylated Fmoc-protected tyrosine was purchased from Advanced Chem Tech. The resin was swelled for 30 min in DMF. Fmoc-deprotection and amino acid coupling reactions were performed as described for the peptide library synthesis. The peptides were cleaved, and the amino acid side chains were deprotected in a solution of 2.5% triethylsilane, 2.5% H_2O, and 95% TFA for 2 h. The cleaved peptides were filtered to remove the resin, and the peptides were precipitated with ethyl ether. The precipitated peptides were resuspended in 0.1% TFA in H_2O and lyophilized. They were then resuspended in 0.1% TFA to a final concentration of 500 μM.

Preparation of SAMDI Peptide Arrays. Peptide arrays were prepared as described previously.[16,18−20] Briefly, 384 gold spots were evaporated onto titanium-coated steel plates using a Thermionics E-beam evaporator. The plates were soaked in a 1 mM total disulfide monolayer solution of 0.8 mM tri(ethylene glycol) disulfide and 0.2 mM tri(ethylene glycol)-maleimide disulfide in ethanol. The alkanethiolate monolayer self-assembles onto the gold surfaces and presents a functional maleimide group against a background of tri(ethylene glycol).

Peptides were neutralized by dilution in 50 mM Tris pH 7.5 to a concentration of 50 μM and pipetted onto the gold spots using Tecan robotics (one peptide per gold monolayer). The peptides were incubated on the surface in a humidified chamber at room temperature for 1 h for immobilization. Peptide immobilization occurs through 1,4-Michael addition of the cysteine-thiol side chain to the maleimide. The arrays were then washed with H_2O and then ethanol, dried under nitrogen, and stored in vacuum sealed bags at 4 °C until ready for use.

SAMDI Peptide Array Reactions. Acetic anhydride, propionic anhydride, and m-dPEG-NHS ester were purchased from Sigma-Aldrich. Four prepared SAMDI Ac-GXZRGC peptide arrays were treated with 100 mM acetic anhydride in H_2O at room temperature for 1 h. Two additional prepared Ac-GXZRGC peptide arrays were treated with either 100 mM propionic anhydride or 100 mM m-dPEG-NHS ester in H_2O at room temperature for 1 h. After the reaction completed, the plates were rinsed with H_2O and then ethanol and dried under nitrogen. Ten SAMDI Ac-GXYHZGRC peptide arrays were prepared; five were treated with 100 mM acetic anhydride, and five were treated with 0.1 mM acetic anhydride in H_2O at room temperature for 1 h. Two additional prepared Ac-GXYHZGRC peptide arrays were treated with either 0.1 mM propionic anhydride or 0.1 mM m-dPEG-NHS ester in H_2O for 1 h at room temperature. The plates were then rinsed with H_2O and then ethanol and dried under nitrogen. All arrays were analyzed by SAMDI-MS. The heat maps shown in Figure 2 and Figure 5 portray average values of acetylation on each peptide.

In-Solution Peptide Reactions. Seven peptides, Ac-GHYGRG, Ac-GKGRG, Ac-GAGRC, NH2-GAGRG, Ac-GYGRG, Ac-GHGRG, and Ac-GAGRG, were treated with acetic anhydride in solution. The in-solution reactions consisted of 100 μM of peptide with 100 mM acetic anhydride in H_2O for 1 h at room temperature. The reactions were then flash frozen in liquid nitrogen and lyophilized to remove the acetic anhydride. The peptides were then resuspended in H_2O. The resuspended peptide solutions were mixed at a ratio of 1:9 with a matrix solution consisting of 10 mg/mL α-cyano-4-hydroxycinnamic acid in 50% acetonitrile, 50% H_2O, and 0.1% TFA and spotted onto a MALDI plate. The peptides were analyzed using reflector positive mode on an AB Sciex 5800 MALDI TOF/TOF. Spectra were collected from 900 laser shots with a stage velocity of 2500 μm/sec. The proportions of acetylation are plotted in Figure 3, and the spectra are shown in Figure S2.

Timed SAMDI Reactions on Individual Peptides. Ac-GYHGC, Ac-GY^{OMe}HGC (O-methylated), Ac-GHGC, and Ac-GYGC peptides were immobilized (as described above) in a quintuplet on 11 SAMDI plates. Each plate was treated with 100 mM acetic anhydride in H_2O for 0, 2, 4, 6, 8, 10, 12, 18, 20, 25, or 30 min. Each plate was rinsed with water and then ethanol, dried under nitrogen at the completion of its time point, and then analyzed using SAMDI-MS. The averages and standard deviations were calculated from five replicates of each time point. The averages and standard deviations are plotted in Figure 4b.

SAMDI Acetyltyrosine Hydrolysis Experiments. Forty micro-SAMDI plates were prepared, Ac-GYHGC was immo-

DOI: 10.1021/acs.biochem.9b00022
Biochemistry 2019, 58, 1810−1817

bilized (as described above) in a quintuplet on 20 plates, and Ac-GYGC was immobilized in a quintuplet on each of the remaining 20 plates. The 20 Ac-GYHGC-immobilized plates were chemically acetylated by three treatments of 100 mM acetic anhydride in acetonitrile for 1 h at room temperature. The 20 Ac-GYGC-immobilized plates were chemically acetylated by three treatments of 100 mM acetic anhydride and 50 mM triethylamine in acetonitrile for 1 h at room temperature. All plates were then rinsed with ethanol, water, and then ethanol and dried under nitrogen. Ten of the Ac-GY^{ac}HGC-immobilized plates, and ten of the Ac-GY^{ac}GC-immobilized plates were soaked in H_2O for 0, 6, 12, 24, 36, 48, 60, 72, 84, or 96 h. The 10 remaining Ac-GY^{ac}HGC-immobilized plates and the 10 remaining Ac-GY^{ac}GC-immobilized plates were soaked in PBS pH 7.4 buffer for 0, 6, 12, 24, 36, 48, 60, 72, 84, or 96 h. The plates were rinsed with water and then ethanol, dried under nitrogen, and analyzed by SAMDI-MS. The average remaining amount of acetylation was measured at each time point, and the pseudo-first-order rate plot is shown in Figure S4.

SAMDI Intramolecular Distance Experiments. To examine the effects of increased intramolecular distance between the histidine and tyrosine, five peptides (Ac-GYHRGC, Ac-GYGHRGC, Ac-GYGGHRGC, Ac-GYGGGHRGC, and Ac-GYGGGGHRGC) were all immobilized in triplicate onto seven SAMDI plates. Each plate was treated with 100 mM acetic anhydride in H_2O at room temperature for either 0, 4, 8, 12, 16, 20, or 25 min and rinsed with water and then ethanol and dried under nitrogen. Each plate was analyzed by SAMDI-MS. The proportion of acetylation on each peptide at each time point was averaged and plotted in Figure 6.

SAMDI Intermolecular Distance Experiments. To examine the effects of increased intermolecular distance between the histidine and tyrosine, six gold patterned plates were soaked in a 1 mM total disulfide monolayer solution with different ratios of tri(ethylene glycol) disulfide and tri(ethylene glycol)-maleimide disulfide in ethanol to present maleimide at a density of 1%, 5%, 10%, 15%, 20%, or 25%. Peptides Ac-GYGC and Ac-GHGC were coimmobilized in triplicate on each plate. As a control, peptide Ac-GYGC alone was also spotted in triplicate on each plate. All plates were treated with 100 mM acetic anhydride for 30 min a room temperature, rinsed with water and then ethanol, and dried under nitrogen. The plates were analyzed using SAMDI-MS, and the proportion of acetylation on the Ac-GYGC peptide (when coimmobilized with Ac-GHGC and when immobilized independently) was plotted in Figure 7.

SAMDI Mass Spectrometry. A matrix solution of 10 mg/mL 2,4,6-trihydroxyacetophenone in 0.1% TFA in acetonitrile was applied to all treated peptide array plates. The arrays were analyzed in reflector positive mode on an AB Sciex 5800 MALDI TOF/TOF. Spectra were collected from 900 laser shots with a stage velocity of 3000 μm/sec.

Data Analysis. All spectra were analyzed by integrating the area under the curve of the peptide product and substrate peaks. The proportion of acylation on each peptide was calculated by dividing the area under the curve (AUC) of the acylated peptide product divided by the AUC of both the acylated product and unreacted substrate. All replicates were averaged in presented results. The peptide library results are displayed as heat maps, which are a grayscale graphical representation of the average proportion of acetylation of each peptide in the array, where black represents complete reactivity. Each box represents an individual peptide in the array with the X position amino acid on the vertical axis and the Z position amino acid on the horizontal axis.

RESULTS AND DISCUSSION

Acetic anhydride is known to acetylate amines, including the side chain of lysine residues, under basic conditions. We asked whether there were certain peptide sequences that would react with the acylating reagent under neutral conditions and therefore aimed to map the reactivity of acetic anhydride with a peptide array encompassing all combinations of amino acid dyads. We synthesized a peptide library consisting of 361 peptides with sequences of Ac-GXZRGC, where X and Z are variable and represent all 19 natural amino acids (excluding cysteine). The peptides were capped at their N-termini to prevent reaction at this site and also included an arginine to increase ionization efficiency for MALDI-MS. The cysteine residue provides the thiol functional group for immobilization to SAMs that present the corresponding maleimide group. We prepared a peptide array by immobilizing the peptides onto an array plate having 384 self-assembled monolayers on gold presenting maleimide groups at a density of 10% against a background of tri(ethylene glycol) groups, as described previously.[18] We then treated the peptide array with an aqueous solution of acetic anhydride (100 mM) for 1 h at room temperature, after which the plate was rinsed, treated with matrix, and analyzed by SAMDI-MS to provide a mass spectrum for each spot (Figure 1).

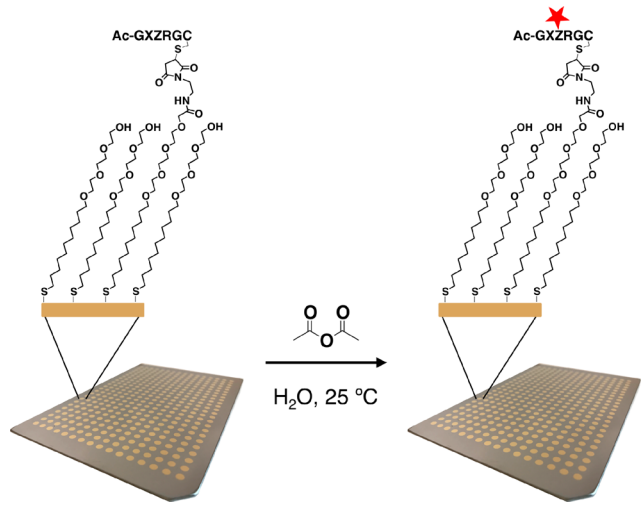


Figure 1. The use of peptide arrays to discover peptide-modifying reactions and to characterize their sequence specificity. A key benefit of the approach is the use of SAMDI-MS as a label-free method to identify and quantitate reaction products. Here, a peptide array based on the sequence Ac-GXZRGC, where X and Z represent all natural amino acids except for cysteine, is treated with acetic anhydride and analyzed.

The mass spectra showed that the masses of the majority of the peptides were unchanged following treatment with acetic anhydride and therefore did not react with the anhydride. Surprisingly, only two peptides in the array underwent efficient acetylation, as shown by an increase in mass of 42 Da. These peptides, Ac-GHYRGC and Ac-GYHRGC, both contain histidine and tyrosine in adjacent positions (Figure 2). We

DOI: 10.1021/acs.biochem.9b00022
Biochemistry 2019, 58, 1810−1817

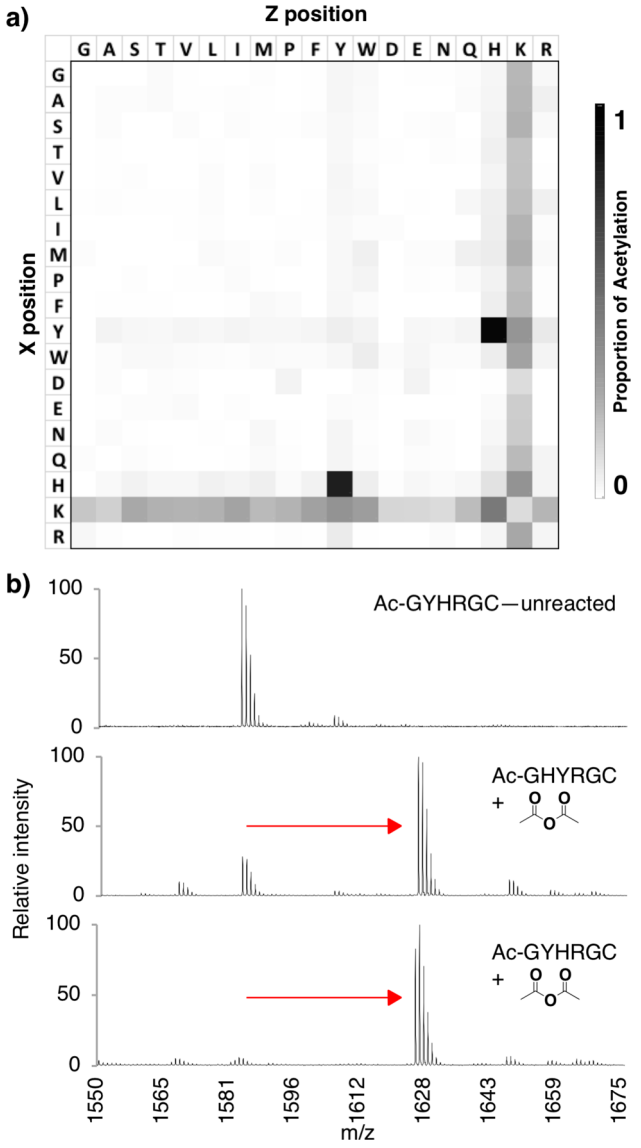


Figure 2. (a) Heat map showing the average extent of acetylation from four replicates of Ac-GXZRGC peptides with acetic anhydride, where each square represents a peptide having X and Z residues as denoted on the vertical and horizontal axes, respectively. The extent of acetylation is denoted in a grayscale, with black representing a complete reaction. Acetic anhydride selectively reacted with peptides containing both histidine and tyrosine in the X and Z positions. The proportion of acetylated peptide was calculated by dividing the area under the curve (AUC) of the peak for the acetylated product by the total AUC of the peaks for the unreacted substrate and acetylated product. (b) Spectra for Ac-GHYRGC and Ac-GYHRGC before and after reaction with acetic anhydride. The top spectrum shows Ac-GYHRGC before treatment with acetic anhydride. (The spectrum of unreacted Ac-GHYRGC is similar and not shown.) The middle and bottom spectra show that, after treatment with acetic anhydride, the mass peaks of Ac-GHYRGC and Ac-GYHRGC are shifted 42 Da corresponding to acetylation.

repeated this experiment four times to confirm the relative reaction yields, and we found that Ac-GHYRGC and Ac-GYHRGC were consistently acetylated with average values of 90% with a standard deviation of 2% and 97% with a standard deviation of 2%, respectively.

The mass spectra also revealed low yields for acetylation of peptides containing lysine residues, which is expected at a neutral pH because there is still a small fraction of side chain amino groups that are not protonated (and therefore nucleophilic). Under basic conditions, in contrast, the acetylation of lysine amino groups with anhydrides is efficient and has been used for sample preparation in bottom-up proteomics.[21] The heat map also demonstrates that SAMDI has the quantitative resolution to compare reactivities for different sequences. The average yield and standard deviation for acetylation of each peptide are shown in Figure S1.

We next confirm that the acetylation of the HY and YH dyads was not influenced by a proximal monolayer and are also observed for solution-phase reactions. Additionally, we repeated the reaction on peptides having a noncapped N-terminal amino group as well as a free cysteine residue to compare relative reactivities. We performed solution-phase reactions with four peptides, each containing only a single site that can react with the anhydride, Ac-GYHGRG, Ac-GKGRG, Ac-GAGRC, and NH_2-GAGRG, and three peptides that served as negative controls, Ac-GYGRG, Ac-GHGRG, and Ac-GAGRG. We treated solutions of each peptide (100 μM) with acetic anhydride in water (100 mM) for 1 h and then analyzed the reactions using MALDI-MS. Figure 3 shows the

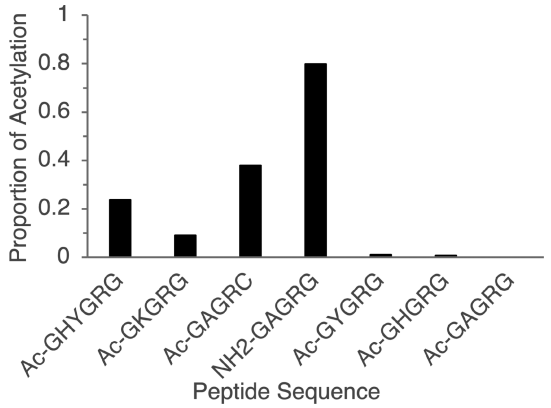


Figure 3. Solution-phase acetylation of peptides containing the HY dyad, lysine, cysteine, or a noncapped N-terminus. Each peptide (100 uM) was treated with acetic anhydride (100 mM) for 1 h, and the reactions were analyzed using MALDI mass spectrometry. This plot shows the yield for acetylation of each peptide under these conditions.

yield for each reaction, and the spectra are shown in Figure S2. The reactivities of the peptides in solution are consistent with our observations of the reactions directly on the peptide array, though the yields for the former were generally lower. The higher yields for the immobilized peptides likely reflect the different stoichiometry in the solid-phase reaction, where the anhydride is present in far greater excess relative to the immobilized peptide, which is particularly important since hydrolysis of the anhydride competes with acetylation. The HY dyad is acetylated in solution, and peptides containing only histidine and tyrosine are not significantly acetylated, eliminating the possibility that the monolayer plays a role in the reactivity observed with the peptide array. We also observed acetylation of the lysine amino group at a level of about 1/3 of that observed for the peptide containing the HY dyad, which was similar to what we observed in reactions with the peptide array. As anticipated, cysteine and the noncapped N-terminus are more reactive than the HY dyad. The observed reactivity of acetic anhydride with the peptides containing

DOI: 10.1021/acs.biochem.9b00022
Biochemistry 2019, 58, 1810−1817

cysteine and a noncapped N-terminus also agrees with previous reports.[22,23]

We also characterized reactivity of the Ac-GYHGC peptide with other electrophilic acylating reagents, including acetyl-CoA, propionic anhydride, and m-dPEG-NHS ester. While we did not observe reaction with acetyl-CoA, we did observe acylation of the peptide with propionic anhydride and with the NHS ester; we reason that a lower electrophilicity of the acetyl-CoA resulted in poor reactivity under these conditions. We also treated the Ac-GXZRGC peptide array with propionic anhydride (Figure S3a) and m-dPEG-NHS ester (Figure S3b) and found that both reagents gave a similar selectivity for the HY dyad, suggesting that this reactivity is broadly applicable to other electrophilic acylating agents.

We reasoned that acetylation of the His-Tyr dyad proceeds through initial acetylation of the histidine imidazole group, followed by intramolecular transfer of the acetyl group to the tyrosine hydroxyl group (Figure 4a). This mechanism is consistent with the known reaction of activated esters with imidazole groups (though the acylated adducts undergo rapid hydrolysis) and with the lack of reactivity with phenols (as is also shown by lack of reaction with other tyrosine-containing peptides in the array). To test this mechanism, we performed reactions of acetic anhydride with four additional peptides: Ac-GYHGC, Ac-GY^{OMe}HGC (where the tyrosine hydroxyl group is methylated to remove the nucleophilic character), Ac-GHGC, and Ac-GYGC. We again allowed each peptide to react with acetic anhydride (100 mM) for times ranging from 0 to 30 min, in 2 or 4 min intervals (Figure 4b).

Our results show that the parent peptide having the YH dyad, Ac-GYHGC, was nearly quantitatively acetylated after 2 min and remained acetylated through the 30 min reaction. The peptides Ac-GY^{OMe}HGC and Ac-GHGC, which do not include the nucleophilic tyrosine residue, were acetylated in approximately 50% yield after 2 min, but the extent of acetylation decreased to approximately 20% after 30 min. The peptide containing no histidine, Ac-GYGC, did not react with the anhydride. These results show that the histidine residue is required for acetylation, but acetylation on histidine is not stable and undergoes hydrolysis to give the native peptide. However, a neighboring tyrosine group competes with hydrolysis through a trans-acetylation reaction. We find no previous reports of an acetyltyrosine post-translational modification or this His-to-Tyr acylation mechanism, though an early report by Moore and co-workers observed efficient acetylation of the angiotensin II peptide (DRVYIHPF) with acetic anhydride, and suggested that the histidine serves as a base to promote direct acetylation and deacetylation of the tyrosine.[24] Finally, we note that, while acetyl transfer from the imidazole side chain of histidine to the hydroxyl group on tyrosine has not been reported in polypeptides, there is a significant precedent for imidazole serving as a catalyst in acylation reactions. These examples include the acylation of cellulose through a similar acyl transfer mechanism[25] as well as for mercaptans, amines, azides, and phenols.[26,27] Similarly, imidazole can catalyze the hydrolysis of esters.[27,28]

We found that the stability of acetylated tyrosine is affected by proximity of histidine residues and salts. We measured the hydrolysis of acetyltyrosine on two peptides, Ac-GY^{ac}GC and Ac-GY^{ac}HGC, in both water and PBS buffer pH 7.4 over 96 h (Figure S4). We found that acetyltyrosine in a peptide lacking the proximal histidine residue was stable over the 96 h in both water and PBS buffer, with little hydrolysis. This stability is

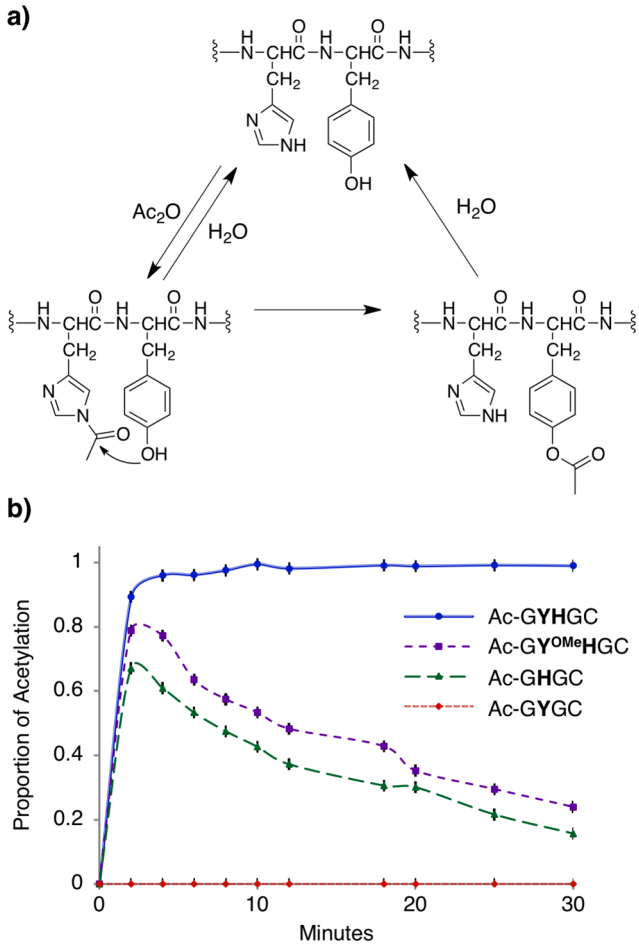


Figure 4. (a) Mechanism of acetylation of the histidine-tyrosine dyad. The histidine residue is first acetylated, followed by intramolecular transfer of the acetyl group to the phenol of the neighboring tyrosine, to give a stable adduct. (b) Kinetic profile for acetylation of four additional peptides after reaction with acetic anhydride (100 mM in water) for 30 min. The peptide containing both histidine and tyrosine, Ac-GYHGC (blue), was nearly quantitatively acetylated after 2 min and remained acetylated. The two peptides, Ac-GY^{OMe}HGC (purple, containing histidine and an O-methylated tyrosine) and Ac-GHGC (green, containing only histidine), were acetylated but showed significant hydrolysis over 30 min. The peptide containing only tyrosine, Ac-GYGC (red), remained unacetylated.

consistent with microscopic reversibility for the forward reaction and also with prior work showing that acetylated phenols have half-lives of several days.[24,29] However, the Ac-GY^{ac}HGC peptide was found to hydrolyze more rapidly, with half-lives of 157.5 h in water and 30.4 h in PBS, similar to the reported values by Moore and co-workers.[24] This lowered stability in buffered solutions could be one of the reasons why this modification has not been observed in proteomic studies, as it may not be stable throughout the sample preparation process.

We next asked whether the acylation of the HY dyad could be further accelerated (or inhibited) by varying the amino acids surrounding this motif. We synthesized a second peptide library having the sequence Ac-GXYHZGRC, where X and Z are variable positions and again represent all 19 amino acids (except for cysteine). We prepared an array comprising these 361 peptides as described above and treated the array with acetic anhydride (100 mM in water) for 1 h (Figure 5a). The

DOI: 10.1021/acs.biochem.9b00022
Biochemistry 2019, 58, 1810−1817

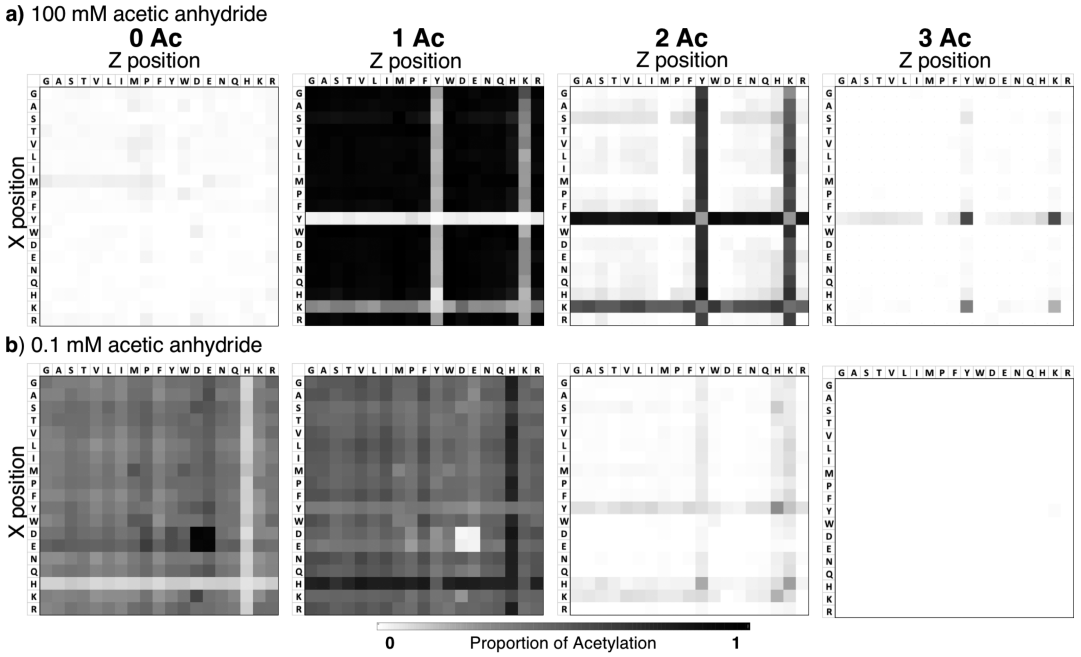


Figure 5. A peptide array based on the sequence Ac-GXYHZGRC treated with acetic anhydride in water at either 100 mM (a) or 0.1 mM (b) and analyzed by SAMDI mass spectrometry. Separate heat maps show the amount of each peptide that did not react (0 Ac) and that underwent acetylation one (1 Ac), two (2 Ac), or three (3 Ac) times.

average data values and standard deviations for each peptide are shown in Figure S5. As expected, we found that all peptides underwent efficient acetylation. However, peptides with two or three tyrosine groups were found to be di- or triacetylated, respectively. Further, peptides containing lysine in either one or two variable positions were also acetylated more than once, but to a lesser extent. This result shows that a single histidine residue can catalyze the acetylation of multiple tyrosine residues in the same peptide, as would be expected for its role as a catalyst.

We then repeated this experiment with a lower concentration of acetic anhydride (0.1 mM) to decrease the yield of acetylation and better resolve the relative reactivities of the peptides with the anhydride (Figure 5b). The average data values and standard deviations for each peptide are shown in Figure S6. We found that peptides containing acidic residues, either aspartic acid or glutamic acid, in both variable positions were not acetylated. Interestingly, a single acidic residue did not significantly affect the reaction rate. A recent report by Denu and colleagues measured the kinetics of nonenzymatic lysine acetylation using acetylphosphate and acetyl-CoA on 90 lysine sites within eight purified proteins.[30] They also found that lysines with a glutamic or aspartic acid in the +1 and −1 positions were acetylated at lower rates than were lysines without proximal acidic residues and suggested this tempered reactivity was due to salt bridge formation. Our similar observation with the acetylation of the HY dyad could similarly be due to a salt bridge or to a higher pK_a value for the protonated imidazole due to an electrostatic interaction.

The array also revealed that peptides containing a histidine in one or both of the variable positions underwent acetylation more rapidly, possibly by increasing the amount of acylated histidine intermediate. Similar reactivity was also found with propionic anhydride (Figure S7a) and m-dPEG-NHS ester (Figure S7b); however, the NHS ester was significantly less reactive across the whole array. These data again show that acetylation of the histidine-tyrosine dyad by acetic anhydride is a reasonably specific reaction that is most favorable in the presence of a neighboring histidine residue and absence of nearby acidic residues.

Finally, we examined the distance dependence for the transfer of the acetyl group from the histidine to tyrosine. We first synthesized four peptides having the sequence Ac-GY(G)$_n$HRGC, where n is 1, 2, 3, or 4 and therefore increases the distance between the histidine and tyrosine residues. We treated those four peptides as well as Ac-GYHRGC with acetic anhydride (100 mM) for 25 min (Figure 6). Each of the peptides reached a maximum yield at about 12 min, likely because the acetic anhydride undergoes hydrolysis during the reaction. We observed a slight decrease in the extent of acetylation for the peptides having one and two glycine residue

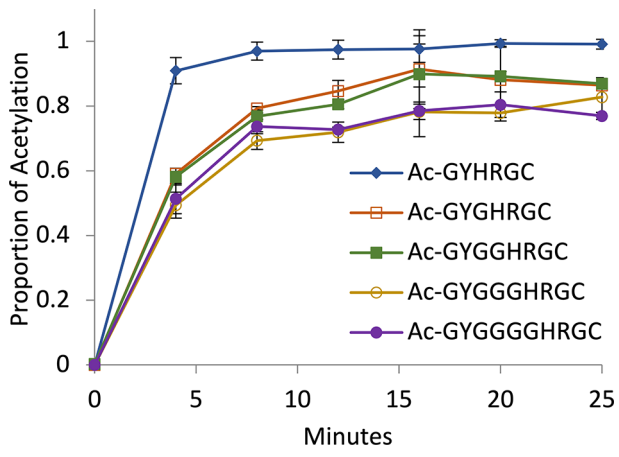


Figure 6. Five peptides, where the distance between the tyrosine and histidine residues varied, each treated with acetic anhydride (100 mM in water) for 25 min. Analysis of the extent of acetylation at several time points shows the relative rates for acetylation.

DOI: 10.1021/acs.biochem.9b00022
Biochemistry 2019, 58, 1810−1817

spacers and a greater decrease for the peptides having three and four glycine residue spacers. However, the acetylation did not drop off significantly, even with a peptide having four glycine residues, indicating that the acetyl transfer can occur over longer distances (at least for peptides having a conformationally unconstrained oligoglycine linker).

Lastly, we address the possibility that the acetyl transfer from histidine to tyrosine could occur intermolecularly. We coimmobilized equal amounts of two peptides, Ac-GHGC and Ac-GYGC to monolayers with varying densities of maleimide, so that the mean distance between neighboring peptides could be varied. We expected that at high densities of peptide, the former could undergo acetylation and transfer the acetyl group to the latter, but that transfer would not be possible for low densities of peptides. The alkanethiolate chains in a monolayer have an approximate density of 4.5/nm^2.[31,32] Therefore, the density of immobilized peptides for monolayers having maleimide groups presented at a percentage of the total alkanethiolates are 0.045/nm^2 at 1%, 0.23/nm^2 at 5%, 0.45/nm^2 at 10%, 0.68/nm^2 at 15%, 0.90/nm^2 at 20%, and 1.13/nm^2 at 25%. We treated the monolayers with acetic anhydride (again, at 100 mM in water) for 1 h and analyzed the reactions using SAMDI-MS (Figure 7). Indeed, we found

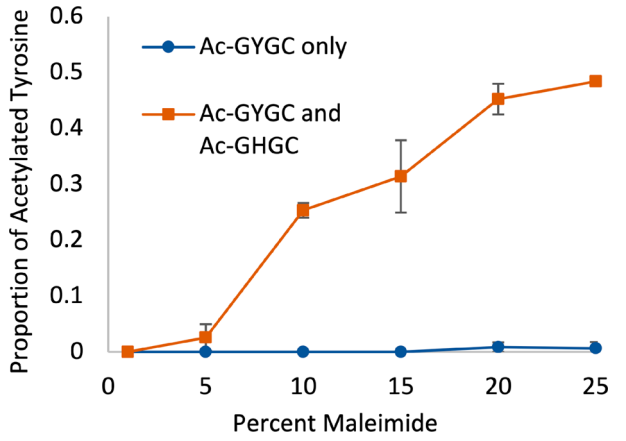


Figure 7. Monolayers presenting a mixture of GYGC and GHGC peptides (orange) or only the GYGC peptide (blue) for densities of peptide ranging from 1 to 25% treated with acetic anhydride in water (100 mM) for 1 h. The yield for acetylation of the Ac-GYGC peptide was determined by SAMDI mass spectrometry.

that the tyrosine-containing peptide showed very little acetylation when the peptides were present at a density of 1 or 5%, but significant levels of acetylation for densities above 10%. Monolayers presenting only the tyrosine-containing peptide, by contrast, showed no acetylation at any of the densities. This example demonstrates that histidine-mediated acetylation of tyrosine can occur in a distance-dependent intermolecular pathway and by extension suggests that, in proteins, this transfer could occur not only when histidine and tyrosine are close in primary sequence but also when proximal in space.

CONCLUSIONS

This work illustrates how peptide arrays and SAMDI-MS can be used to discover sequence-specific reactions of peptides. We found that acetylation of peptides by acetic anhydride occurs efficiently on tyrosine residues that have a neighboring histidine. It is surprising that this reaction has not been previously reported, even though acylation of proteins has been thoroughly studied and the mechanistic basis of this reaction is well-established. We believe that this approach will be important for the discovery of additional chemistries for site-specific protein modification and conjugation, as well as for optimization of those that have already been reported.[9,10,33−38] Current approaches often do not assess the global sequence specificity and also are limited to identifying anticipated reactivities. Our method uses nonbiased peptide arrays comprising of two variable positions with mass spectrometry to characterize all possible sequences and identify unforeseen reactivities, allowing for a complete understanding of the specificity of the reaction. This example demonstrates the benefits of using peptide arrays and SAMDI-MS to discover peptide-modifying reactions and to understand their sequence specificity.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.9b00022.

Heat maps, MALDI-MS spectra, and pseudo-first-order rate plot of the hydrolysis of *O*-acetyltyrosine (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: milan.mrksich@northwestern.edu.

ORCID

Lindsey C. Szymczak: 0000-0002-0667-4730

Milan Mrksich: 0000-0002-4964-796X

Author Contributions

All authors designed the experiments and wrote the manuscript. L.C.S. performed all experiments.

Funding

This work was supported by the Office of Naval Research, Multidisciplinary University Research Initiative, under award no. W911NF1810200 and the Department of Defense, Defense Advanced Research Projects Agency, under award no. W911NF-18-2-0030.

Notes

The authors declare no competing financial interest.

ABBREVIATIONS

CoA, coenzyme A; MS, mass spectrometry; SAMDI-MS, self-assembled monolayers for matrix-assisted laser desorption/ionization mass spectrometry; MALDI, matrix-assisted laser desorption/ionization; SAM, self-assembled monolayer; HY, histidine-tyrosine; YH, tyrosine-histidine; NMM, *N*-methyl morpholine; DMF, dimethylformamide; TFA, trifluoroacetic acid; AUC, area under the curve

REFERENCES

(1) Speers, A. E., and Cravatt, B. F. (2004) Profiling Enzyme Activities In Vivo Using Click Chemistry Methods. *Chem. Biol. 11*, 535−546.

(2) Xue, L., Karpenko, I. A., Hiblot, J., and Johnsson, K. (2015) Imaging and manipulating proteins in live cells through covalent labeling. *Nat. Chem. Biol. 11*, 917.

(3) Griffin, B. A., Adams, S. R., and Tsien, R. Y. (1998) Specific Covalent Labeling of Recombinant Protein Molecules Inside Live Cells. *Science 281*, 269.

DOI: 10.1021/acs.biochem.9b00022
Biochemistry 2019, 58, 1810−1817

(4) Hochuli, E., Bannwarth, W., Döbeli, H., Gentz, R., and Stüber, D. (1988) Genetic Approach to Facilitate Purification of Recombinant Proteins with a Novel Metal Chelate Adsorbent. *Nat. Biotechnol. 6*, 1321.
(5) de Goeij, B. E. C. G., and Lambert, J. M. (2016) New developments for antibody-drug conjugate-based therapeutic approaches. *Curr. Opin. Immunol. 40*, 14−23.
(6) Dozier, J. K., and Distefano, M. D. (2015) Site-Specific PEGylation of Therapeutic Proteins. *Int. J. Mol. Sci. 16*, 25831−25864.
(7) Mackenzie, K. J., and Francis, M. B. (2013) Recyclable Thermoresponsive Polymer−Cellulase Bioconjugates for Biomass Depolymerization. *J. Am. Chem. Soc. 135*, 293−300.
(8) Witus, L. S., and Francis, M. B. (2011) Using Synthetically Modified Proteins to Make New Materials. *Acc. Chem. Res. 44*, 774−783.
(9) Boutureira, O., and Bernardes, G. J. L. (2015) Advances in Chemical Protein Modification. *Chem. Rev. 115*, 2174−2195.
(10) deGruyter, J. N., Malins, L. R., and Baran, P. S. (2017) Residue-Specific Peptide Modification: A Chemist's Guide. *Biochemistry 56*, 3863−3873.
(11) Gurard-Levin, Z. A., and Mrksich, M. (2008) Combining Self-Assembled Monolayers and Mass Spectrometry for Applications in Biochips. *Annu. Rev. Anal. Chem. 1*, 767−800.
(12) Mrksich, M. (2008) Mass Spectrometry of Self-Assembled Monolayers: A New Tool for Molecular Surface Science. *ACS Nano 2*, 7−18.
(13) Szymczak, L. C., Kuo, H.-Y., and Mrksich, M. (2018) Peptide Arrays: Development and Application. *Anal. Chem. 90*, 266−282.
(14) Castaneda, C. A., Lopez, J. E., Joseph, C. G., Scholle, M. D., Mrksich, M., and Fierke, C. A. (2017) Active Site Metal Identity Alters Histone Deacetylase 8 Substrate Selectivity: A Potential Novel Regulatory Mechanism. *Biochemistry 56*, 5663.
(15) Kightlinger, W., Lin, L., Rosztoczy, M., Li, W., DeLisa, M. P., Mrksich, M., and Jewett, M. C. (2018) Design of glycosylation sites by rapid synthesis and analysis of glycosyltransferases. *Nat. Chem. Biol. 14*, 627−635.
(16) Gurard-Levin, Z. A., Scholle, M. D., Eisenberg, A. H., and Mrksich, M. (2011) High-Throughput Screening of Small Molecule Libraries using SAMDI Mass Spectrometry. *ACS Comb. Sci. 13*, 347−350.
(17) Høeg-Jensen, T., Havsteen Jakobsen, M., Olsen, C. E., and Holm, A. (1991) Formation of peptide thioamides by use of Fmoc amino monothioacids and PyBOP. *Tetrahedron Lett. 32*, 7617−7620.
(18) Houseman, B. T., Gawalt, E. S., and Mrksich, M. (2003) Maleimide-Functionalized Self-Assembled Monolayers for the Preparation of Peptide and Carbohydrate Biochips. *Langmuir 19*, 1522−1531.
(19) Gurard-Levin, Z. A., Kilian, K. A., Kim, J., Bähr, K., and Mrksich, M. (2010) Peptide Arrays Identify Isoform-Selective Substrates for Profiling Endogenous Lysine Deacetylase Activity. *ACS Chem. Biol. 5*, 863−873.
(20) Szymczak, L. C., Huang, C.-F., Berns, E. J., Mrksich, M., and Allen, K. N. (2018) Combining SAMDI Mass Spectrometry and Peptide Arrays to Profile Phosphatase Activities. *Methods Enzymol. 607*, 389−403.
(21) Garcia, B. A., Mollah, S., Ueberheide, B. M., Busby, S. A., Muratore, T. L., Shabanowitz, J., and Hunt, D. F. (2007) Chemical derivatization of histones for facilitated analysis by mass spectrometry. *Nat. Protoc. 2*, 933.
(22) Saravanan, P., and Singh, V. K. (1999) An efficient method for acylation reactions. *Tetrahedron Lett. 40*, 2611−2614.
(23) Dmuchovsky, B., Vineyard, B. D., and Zienty, F. B. (1964) The Mechanism of the Base-Catalyzed Addition of Thiols to Maleic Anhydride. *J. Am. Chem. Soc. 86*, 2874−2877.
(24) Moore, G. J. (1985) Kinetics of acetylation-deacetylation of angiotensin II. *Int. J. Pept. Protein Res. 26*, 469−481.
(25) Nawaz, H., Pires, P. A. R., and El Seoud, O. A. (2013) Kinetics and mechanism of imidazole-catalyzed acylation of cellulose in LiCl/N,N-dimethylacetamide. *Carbohydr. Polym. 92*, 997−1005.
(26) Jencks, W. P., and Carriuolo, J. (1959) Imidazole Catalysis: III. General base catalysis and the reactions of acetyl imidazole with thiols and amines. *J. Biol. Chem. 234*, 1280−1285.
(27) Jencks, W. P., and Carriuolo, J. (1960) Reactivity of nucleophilic reagents toward esters. *J. Am. Chem. Soc. 82*, 1778−1786.
(28) Bender, M. L., and Turnquest, B. W. (1957) General Basic Catalysis of Ester Hydrolysis and Its Relationship to Enzymatic Hydrolysis1. *J. Am. Chem. Soc. 79*, 1656−1662.
(29) Kirsch, J. F., and Jencks, W. P. (1964) Nonlinear Structure-Reactivity Correlations. The Imidazole-Catalyzed Hydrolysis of Esters. *J. Am. Chem. Soc. 86*, 837−846.
(30) Baeza, J., Smallegan, M. J., and Denu, J. M. (2015) Site-Specific Reactivity of Nonenzymatic Lysine Acetylation. *ACS Chem. Biol. 10*, 122−128.
(31) Dubois, L. H., and Nuzzo, R. G. (1992) Synthesis, Structure, and Properties of Model Organic. *Annu. Rev. Phys. Chem. 43*, 437−463.
(32) Love, J. C., Estroff, L. A., Kriebel, J. K., Nuzzo, R. G., and Whitesides, G. M. (2005) Self-Assembled Monolayers of Thiolates on Metals as a Form of Nanotechnology. *Chem. Rev. 105*, 1103−1170.
(33) Joshi, N. S., Whitaker, L. R., and Francis, M. B. (2004) A Three-Component Mannich-Type Reaction for Selective Tyrosine Bioconjugation. *J. Am. Chem. Soc. 126*, 15942−15943.
(34) MacDonald, J. I., Munch, H. K., Moore, T., and Francis, M. B. (2015) One-step site-specific modification of native proteins with 2-pyridinecarboxyaldehydes. *Nat. Chem. Biol. 11*, 326.
(35) Zhang, C., Welborn, M., Zhu, T., Yang, N. J., Santos, M. S., Van Voorhis, T., and Pentelute, B. L. (2016) π-Clamp-mediated cysteine conjugation. *Nat. Chem. 8*, 120.
(36) Lee Hong, G., Lautrette, G., Pentelute, B. L., and Buchwald, S. L. (2017) Palladium-Mediated Arylation of Lysine in Unprotected Peptides. *Angew. Chem., Int. Ed. 56*, 3177−3181.
(37) Beatty, K. E., Fisk, J. D., Smart, B. P., Lu, Y. Y., Szychowski, J., Hangauer, M. J., Baskin, J. M., Bertozzi, C. R., and Tirrell, D. A. (2010) Live-Cell Imaging of Cellular Proteins by a Strain-Promoted Azide−Alkyne Cycloaddition. *ChemBioChem 11*, 2092−2095.
(38) Saxon, E., and Bertozzi, C. R. (2000) Cell Surface Engineering by a Modified Staudinger Reaction. *Science 287*, 2007.

DOI: 10.1021/acs.biochem.9b00022
Biochemistry 2019, 58, 1810−1817