JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/JACS

Article

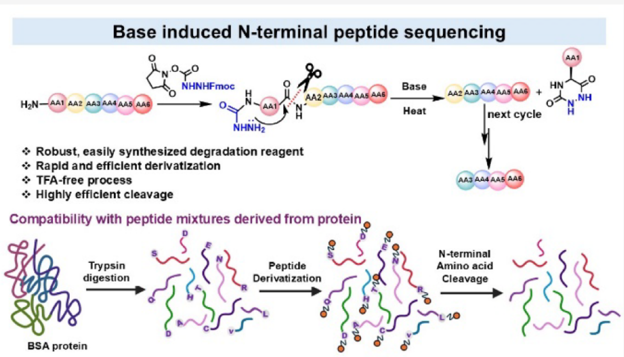
After 75 Years, an Alternative to Edman Degradation: A Mechanistic and Efficiency Study of a Base-Induced Method for N-Terminal Peptide Sequencing

Harnimarta Deol, Ava Raeisbahrami, Phuoc H.T. Ngo, Jagannath Swaminathan, Ophelia Papoulas, Edward M. Marcotte,* and Eric V. Anslyn*

Cite This: J. Am. Chem. Soc. 2025, 147, 13973−13982

Read Online

ABSTRACT: The sequencing of peptides via N-terminal amino acid removal is a classic reaction termed Edman degradation. This method involves repeated treatment of the N-terminal amino group of a peptide with phenyl isothiocyanate (PITC), followed by treatment with trifluoroacetic acid. Spurred by the need for an alternative non-acid-based chemistry for next-generation protein sequencing technologies, we developed a base-induced N-terminal degradation method. Several N-terminal derivatization reagents carrying supernucleophiles were tested. After rounds of iterative designs, compound **DR3**, with a *N*-hydroxysuccinimide as a leaving group and hydrazinecarboxamide as the supernucleophile, demonstrated the highest yield for the peptide derivatization step and the most efficient elimination of the N-terminal amino acid in just 1% of a hydroxide salt. The method successfully removed all 20 amino acids at the N-terminus in high yield. The technique demonstrates compatibility with oligonucleic acids, which differs from Edman degradation due to their inherent sensitivity to acidic environments. To demonstrate the practical application of our approach, we sequenced amino acids sequentially from a peptide, effectively determining the sequence of an unknown peptide. Notably, our methodology was successfully applied to mixtures of peptides derived from protein samples, where a significant fraction of the peptides derivatized with **DR3** underwent elimination of their N-terminal amino acid upon addition of base. Overall, although our method does not outperform Edman degradation in efficiency, it serves as a valuable alternative in cases where base-induced cleavage is advantageous, particularly for preserving acid-sensitive functionalities.



1. INTRODUCTION

Peptide sequencing is an essential step for protein primary structure determination, leading to a better understanding of their function,[1,2] aiding in developing new drugs,[3,4] and advancing biotechnology[5,6] and synthetic biology.[7,8] The demand for innovative sequencing methods is on the rise, such as nanopore methods,[9,10] fluorosequencing,[11] and dye-labeled *N*- terminal amino acid recognizers.[12]

However, the only well-established method for sequencing proteins and peptides through sequential amino acid degradation was pioneered by Pehr Edman in 1950, representing a groundbreaking chemical method[13,14] that has not been altered significantly over the last 75 years. The method uses phenyl isothiocyanate (PITC) to functionalize the N-terminus of a peptide under mild basic conditions, forming a thiourea. Subsequently, this adduct is treated with anhydrous TFA at 50 °C to cleave off a cyclic phenylthiohydantoin, revealing a new N-terminus on the peptide (Figure 1). Edman degradation has long been the gold standard in proteomics for peptide sequencing via chain-end degradation.[15,16]

New protein sequencing technologies, inspired by the success of DNA sequencing, have emerged over the last five years. The primary aim of these technologies is to identify and quantify protein counts in complex mixtures at or near single-molecule sensitivity.[11,17−20] These advancements include optical fluorescence methods, translation of peptide information into DNA,[21,22] and nanopore-derived techniques.[23−25]

We have developed a single-molecule sequencing technique called fluorosequencing, where peptides are selectively labeled on amino acids with fluorophores, placed in a TIRF microscope, and sequenced using Edman degradation while tracking the fluorescence change.[11] We observed that TFA degrades many

Received: February 25, 2025
Revised: April 3, 2025
Accepted: April 4, 2025
Published: April 14, 2025

https://doi.org/10.1021/jacs.5c03385
J. Am. Chem. Soc. 2025, 147, 13973−13982

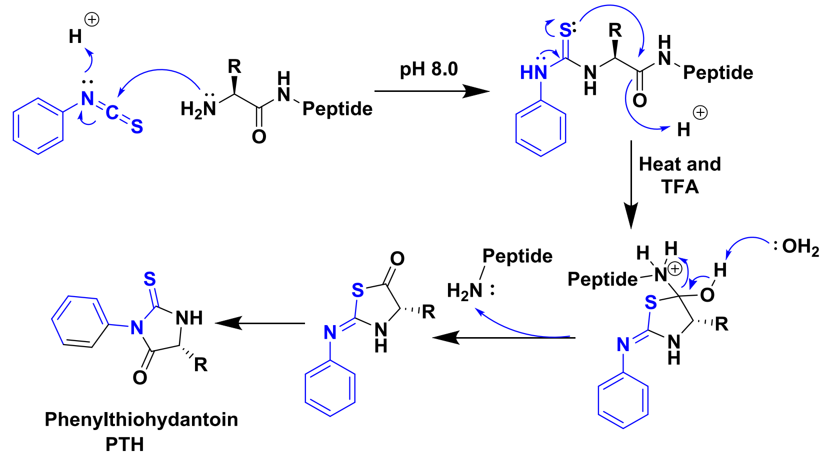


Figure 1. Steps in the Edman degradation.

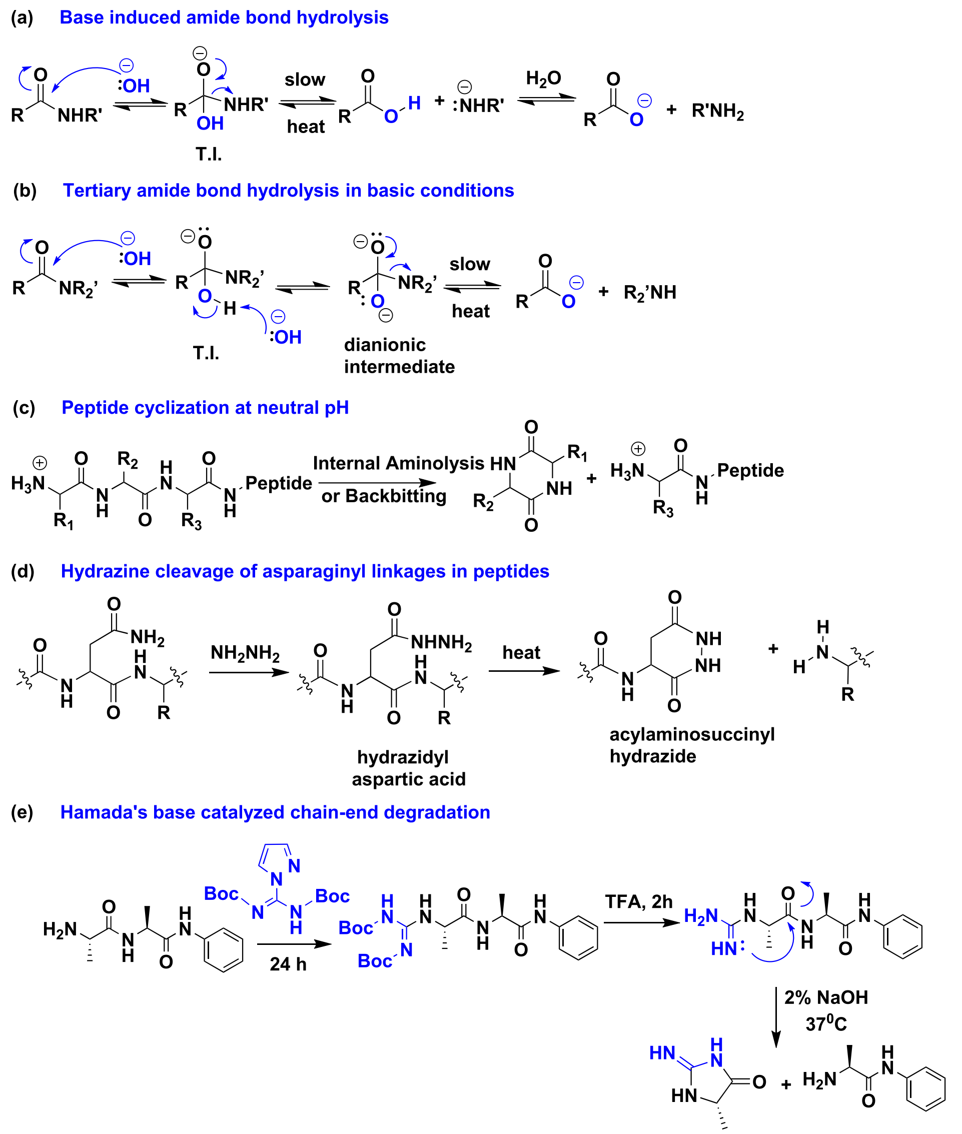


Figure 2. (a) Hydrolysis of amide bonds under basic conditions. (b) Tertiary amide bond hydrolysis. (c) Cyclic transamidation of peptides at a neutral pH. (d) Asparaginyl linkages used for peptide cleavage with hydrazine. (e) Hamada's base catalyzed chain-end degradation.

fluorophores, such as BODIPY- and cyanine-based structures, whereas many of these are stable in base. This issue is not unique to fluorosequencing. Single-molecule protein sequencing methods that use N-terminal recognition units, such as binders or antibodies for imaging, also rely on Edman degradation.[21] Thus, overcoming the Edman degradation acid-requirement

https://doi.org/10.1021/jacs.5c03385
J. Am. Chem. Soc. 2025, 147, 13973−13982

constraint could also facilitate these approaches to sequencing proteins, providing chemists with an alternative approach to chain-end degradative sequencing.

Additionally, maintaining DNA integrity is important during the barcoding of a peptide sequence. Pawlosky uses Edman degradation in DNA barcoding to repeatedly capture and encode newly exposed amino acid positions and identities.[21] During this work, the investigators noted that TFA depurinates DNA in their barcode-cyclic sequencing (BCS) machinery.

Analogously, another study reporting DNA barcoding with Edman chemistry comes from Soh.[22] These authors also noted that conventional Edman conditions cause DNA damage and attempted to replace TFA with $BF_3.Et_2O$. However, depurination persisted, leading them to analyze the depurination products and adopt chemically modified oligonucleotides resistant to degradation.

Overcoming Edman degradation constraints could enable new approaches to sequencing proteins and provide chemists with an interesting opportunity to apply the field's advancements in addressing a major application.

To achieve our objective of an alternative non-acid-catalyzed peptide-end degradative reaction, we took inspiration from three prior themes of work.

1.1. Base-Assisted Amide Bond Hydrolysis. The literature on nonenzymatic methods for amide bond hydrolysis indicates that the mechanism involves nucleophilic attack by hydroxide on the carbonyl carbon, forming a tetrahedral intermediate (Figure 2a). Then, elimination of the leaving group occurs as an amide anion that is protonated instantly or concerted with departure.[26−28] An alternative mechanism has been observed for tertiary amides: these amides undergo cleavage at ambient temperature, as illustrated in Figure 2b. The mechanism again involves a hydroxide ion attacking the carbonyl carbon, but subsequently, deprotonation of the added hydroxide ion forms a dianionic intermediate.[29,30] This intermediate facilitates the elimination of the amide anion leaving group, likely occurring simultaneously with protonation. It is unclear whether the mechanism found for tertiary amides also occurs for other amides or if there is a mix of mechanisms. Either way, the possibility of deprotonation of the tetrahedral intermediate to generate a dianion was a critical feature that we incorporated into our approach to base-induced N-terminal degradation, as described below.

1.2. Internal Cyclization and Aminolysis. As further background to inform our approach, we examined various cyclizations that cleave peptides. For example, at high temperature (130 °C) and neutral pH, a pathway called internal aminolysis or "backbiting" occurs. Here, intramolecular cyclization between the N-terminal amino acid and the adjacent amide results in the formation of a diketopiperazine and the release of the truncated peptide (Figure 2c).[31] Further, when proteins and peptides are heated at 100 °C for 10 h, in the presence of anhydrous hydrazine, cleavage of all peptide bonds occurs.[32−34] However, at lower temperatures (20 °C), hydrazine selectively cleaves the asparaginyl linkages in peptides (Figure 2d).[35] The asparagine amino acid side chain undergoes hydrazinolysis to form *β*-hydrazidyl aspartic acid, wherein the terminal NH_2 group of hydrazine cyclizes on the neighboring amide. This intramolecular cyclization forms a six-membered ring product, accompanied by the liberation of the amino acid or peptide fragment. Similarly, it has been found that hydroxylamine can be used as a nucleophile to cleave asparaginyl-glycyl peptide bonds via an intramolecular cyclization.[36] These intramolecular cyclizations using hydrazine and hydroxylamine are features that we also incorporated into our approach.

1.3. Hamada's Base Catalyzed Chain-End Degradation Method. In 2016, Hamada introduced a chain-end degradation method using *N*,*N*'-bis(tert-butoxycarbonyl)-1*H*-pyrazole-1-carboxamidine for peptide functionalization and 2% NaOH for cleaving the terminal amino acid (Figure 2e).[37] However, a direct extension of their method was not found to be feasible for our objectives because the method still requires TFA for removing the Boc group during the functionalization step. Furthermore, this approach also requires protecting each amino acid side chain during peptide functionalization, and when coupled with the extended 24 h reaction time, significantly limits its practicality and applicability in proteomics. Nevertheless, we took inspiration from their use of the guanidino group as a nucleophile, which targets the carbonyl group, leading to the cyclization of the N-terminal amino acid.

These nonenzymatic degradation methods typically require high temperatures (130 °C) and extended reaction times (often several hours or more) and often result in random peptide degradation. These limitations underscore the need for developing a milder, more controlled, and efficient base-induced sequencing method. Thus, with the lessons from the examples in Figure 2, our primary objective was to develop a non-acid/base-induced degradation method that operates at lower temperatures, requires shorter reaction times, and provides precise N-terminal peptide cleavage.

As reported below, to achieve our objective, we iteratively improved the performance of degradation reagents (**DRs**) by varying leaving groups and nucleophiles until we successfully generated a method to remove each of the 20 N-terminal amino acids in high yield. Using a series of methyl substitutions and deuterium scrambling experiments, we probed the mechanisms of degradation.

Our method effectively enables the stepwise cleavage of amino acids from peptides, providing a valuable tool for peptide sequencing. We also confirmed that oligonucleotides were stable under the sequencing conditions, and we successfully applied the method to peptide mixtures derived from the protein bovine serum albumin. This represents the first entirely base-induced N-terminal degradation method, thereby complementing the nearly 75-year-old Edman degradation method.

1.4. Reagent Design Criteria. The DRs we designed are urethane-based with two key components: a leaving group for selective N-terminal derivatization and a good nucleophile to facilitate cyclization (Figure 3). The leaving group is required to

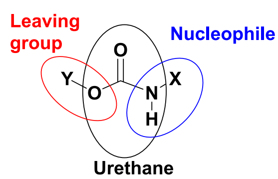


Figure 3. Design of degradation reagents (**DRs**).

selectively react with the N-terminus, such that all the reactive side chains do not necessarily need to be protected. Alternatively, if it is also reactive at the *ε*-amine of lysine, no cleavage would occur (as with PITC in the Edman degradation). Regarding the nucleophile, we used the class of hydroxylamines or hydrazines as supernucleophiles due to the alpha effect,[38] which, as described above, gives either random or asparaginyl-

https://doi.org/10.1021/jacs.5c03385
J. Am. Chem. Soc. 2025, 147, 13973−13982

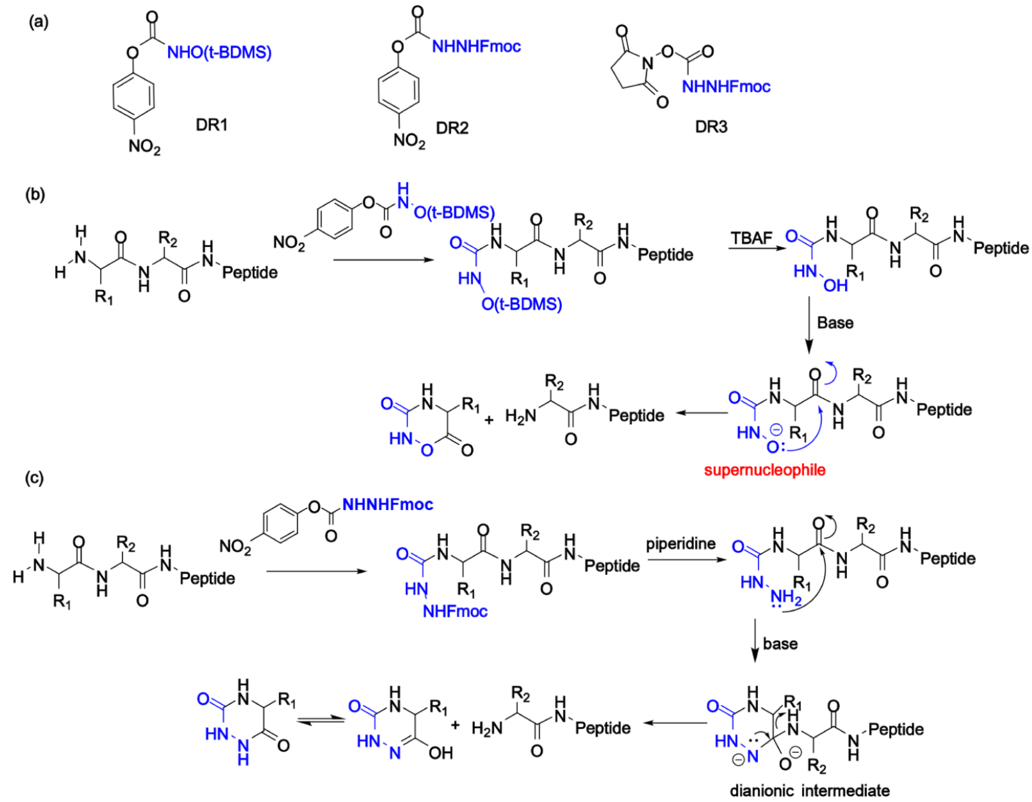


Figure 4. (a) Design refinement, **DR1** to **DR2** to **DR3**. (b) Proposed mechanism for base-induced peptide sequencing using **DR1**. (c) Proposed mechanism for base-induced peptide sequencing using **DR2**.

specific peptide cleavage. Subsequent to derivatization, nucleophilic cyclization of the supernucleophile would form a six-membered ring, as seen in the diketopiperazines or asparaginyl-amino acid literature precedents (Figure 2c,d), favorable by Baldwin's rules,[39] consequently releasing the peptide fragment revealing a new N-terminus.

This strategy aims to overcome the limitations of traditional base-induced and supernucleophilic degradation methods by controlling the site of derivatization and facilitating neighboring amide degradation.

2. RESULTS AND DISCUSSION

Using these design criteria, we began our search for base-induced N-terminal cleavage by comparing two degradation reagents, **DR1** and **DR2**, both featuring a 4-nitrophenyl leaving group, and 1-hydroxy urea or hydrazinecarboxamide as the nucleophiles, respectively (Figure 4a). To protect the nucleophilic end of the reagent from reacting with others of the same molecule, the −OH and $-NH_2$ groups in **DR1** and **DR2** were protected with *tert*-butyldimethylsilyl(*t*-BDMS) and fluorenylmethoxycarbonyl (Fmoc), respectively. We envisioned that **DR1** or **DR2** would react at the N-terminus of a peptide by nucleophilic attack and the loss of 4-nitrophenoxide. In the case of **DR1**, deprotection by tetrabutyl ammonium fluoride (TBAF) would generate a 1-hydroxy urea (an oxime analogue). Because the pK_a of the −OH group of a hydroxylamine is lower than that of water,[40] deprotonation in the presence of hydroxide would generate an aminoalkoxide anion. This anion could cyclize at the adjacent amide of the peptide (Figure 4b), analogous to the cleavage discussed above for asparaginyl-glycine. This reaction generates a six-membered ring and eliminates the N-terminal amino acid. Similarly, in the case of **DR2**, derivatization and subsequent deprotection would generate a hydrazinecarboxamide with a terminal $-NH_2$ nucleophile to attack the adjacent amide (Figure 4c), just as in the precedent of Figure 2d. However, unlike **DR1**, this nucleophile can add without deprotonation. Importantly, in the presence of a base, we hypothesized it could lead to the formation of a dianionic tetrahedral intermediate (Figure 4c), analogous to the mechanism depicted in Figure 2b. If so, this intermediate would facilitate cleavage of the amide bond, resulting in degradation of the peptide.

DR1 was synthesized by the reaction of 4-nitrophenyl chloroformate and O-(*tert*-butyldimethylsilyl) hydroxylamine (Figures S64−S66), consisting of a 4-nitrophenyl group as a leaving group and 1-hydroxy urea as a nucleophile, which is protected by *t*-BDMS. We chose YGFWVY (**1**) as a model peptide, possessing a tyrosine at the N-terminus, thereby allowing us to track the cyclized product cleaved from the peptide in an LC trace. For this derivatization reaction, **DR1** was added to a solution of peptide **1** in DMF, and the reaction was stirred for 2 h at ambient temperature. Subsequently, TBAF was introduced, and the mixture was stirred for an additional 2 h for deprotection (Figure 5). The reaction progress was monitored using LC-MS, and the derivatized peptide **2** was purified via preparative reversed-phase high-performance liquid chromatography (RP-HPLC). For the sequencing step, initial attempts involved tris buffer with a pH of 9.5 and temperatures ranging from 37 to 60 °C for 5 h and 0.4 mM of the derivatized peptide. This resulted in conversion to the truncated peptide **3** of only ∼10%, with more than 90% of the derivatized peptide **2** remaining unreacted. Further, to improve the yield, we employed Cs_2CO_3 and DBU as bases, but this did not improve conversion yields, suggesting the need for a stronger base to

https://doi.org/10.1021/jacs.5c03385
J. Am. Chem. Soc. 2025, 147, 13973−13982

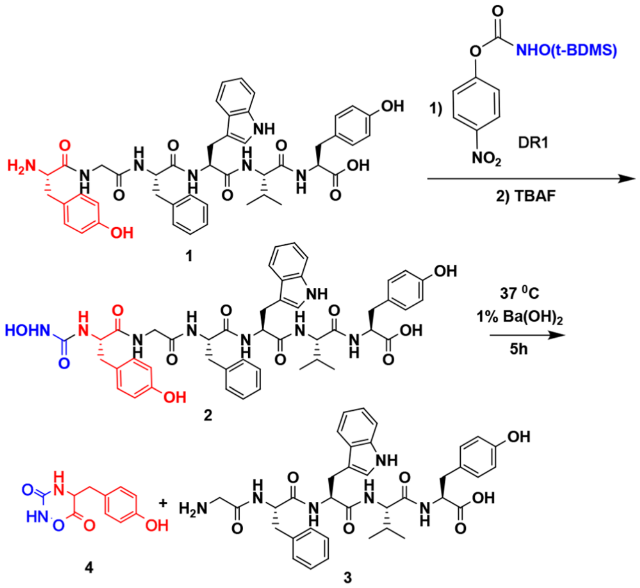


Figure 5. Base-induced sequencing of the N-terminal amino acid in peptide **1** using **DR1**.

cyclize onto the N-terminal amino acid. Subsequently, 1% NaOH at 37 °C was employed for the sequencing step, achieving 60% conversion of the derivatized peptide into truncated peptide **3**. Further, upon using 1% $Ba(OH)_2$, the conversion improved to 70% (Figure 5). Additionally, we identified a peak corresponding to the cyclized form of the amino acid (tyrosine) cleaved from the peptide (**4**) (Figures S1 and S2).

To investigate the reaction mechanism, we designed a modified reagent, **DR1′**, in which the hydroxyl group was blocked with a methyl group (Supplementary Figures 67 and 68). Peptide **1** was derivatized with **DR1′** and subjected to a sequencing reaction using 1% $Ba(OH)_2$ under the optimized conditions. As anticipated, the reaction did not proceed, supporting our hypothesis that the reaction requires nucleophilic attack of the oximate anion at the carbonyl carbon (Figures S3 and S4). This also indicated that deprotonation of the α-NH group in **DR1′** does not lead to cyclization and formation of a 5-membered ring, a topic we return to below.

To further improve the yield and better understand the nature of the side product in the reaction, we explored other amino acids at the N-terminus and discovered after derivatization and base treatment that we consistently observed a side product that accounted for the majority of the remaining mass, with a 15-unit mass difference from the derivatized peptide. This prompted us to characterize the side product. In order to do so, we synthesized the dipeptide WY, derivatized it with **DR1**, and then treated it with 1% $Ba(OH)_2$. Subsequently, we isolated the side product and characterized it using HRMS (Figures S5−S7). The results revealed the formation of a carbamate anion at the N-terminus, likely due to hydrolysis of the urethane under basic conditions with the departure of hydroxylamine. Given this side product, we shifted our focus to investigate another reagent, DR2, which possesses a hydrazine nucleophile.

DR2 was synthesized through a reaction between 4-nitrophenyl chloroformate and 9-fluorenylmethyl carbazate (Figures S69 and S70). In this case, the NH_2 group is protected with fluorenylmethoxycarbonyl (Fmoc).

Again, we conducted our first studies of **DR2** using YGFWVY (**1**) in DMF in the presence of DIPEA as the base at room temperature, followed by piperidine addition for Fmoc removal. The reaction took 18 h, which is considerably longer compared to **DR1**, but our primary interest lay in evaluating the conversion rate during the base-induced cleavage step. Initially, we conducted the sequencing reaction in an aqueous environment devoid of any base, which resulted in a 10% degradation of the peptide after 4 h (Figure S8). Subsequently, we increased the pH to 9 and monitored the sequencing yield. At this point, 30% cleavage had occurred, while 70% of the derivatized peptide remained unreacted. However, by employing a 1% $Ba(OH)_2$ solution at 37 °C, we observed a 93% elimination of degraded peptide **3** after 5 h (Figures S9 and S10). Additionally, we identified a peak corresponding to the cyclized form of the amino acid (tyrosine) cleaved from the peptide (**6**). The cyclized product **6** was isolated using HPLC and characterized by HRMS (Figure S11). From these findings, we concluded that the nucleophile was reactive enough for cyclization and that the initial adduct does not lead to any significant side products in the presence of the base. However, the prolonged derivatization time was a limitation, which prompted us to consider changing the leaving group of **DR2** to enhance the reactivity of the derivatization step. Thus, we replaced 4-nitrophenolate with *N*-hydroxysuccinimide, generating **DR3**.

DR3 was synthesized by allowing *N*,*N′*-disuccinimidyl carbonate to react with 9-fluorenylmethyl carbazate (Figures S71−S73). To assess **DR3** reactivity in the derivatization step, we performed the same procedure with the YGFWVY peptide **1**, consisting of DMF with DIPEA as the base at room temperature, giving a 92% isolated yield of the derivatized peptide in just 1 h. After confirmation of the derivatization reaction by LC-MS, piperidine was added to deprotect the Fmoc group (Figure 6).

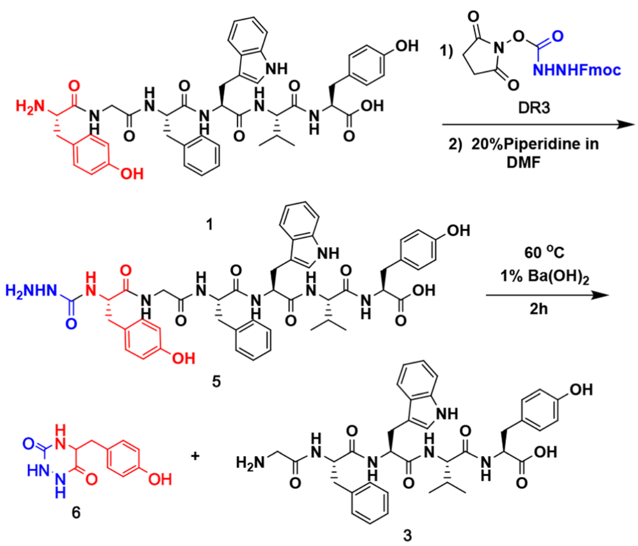


Figure 6. Base-induced sequencing of the N-terminal amino acid in peptide **1** using **DR3**.

The desired derivatized peptide was isolated from the reaction mixture by simply adding ether for precipitation, which eliminates the need for HPLC purification and enhances the efficiency of the approach. Although **DR2** and **DR3** are two distinct reagents, the derivatized peptides exhibit identical molecular structures and formulas; therefore, consistent results were obtained during the sequencing reaction conducted in the presence of 1% $Ba(OH)_2$ at 37 °C for 5 h. From these

https://doi.org/10.1021/jacs.5c03385
J. Am. Chem. Soc. 2025, 147, 13973−13982

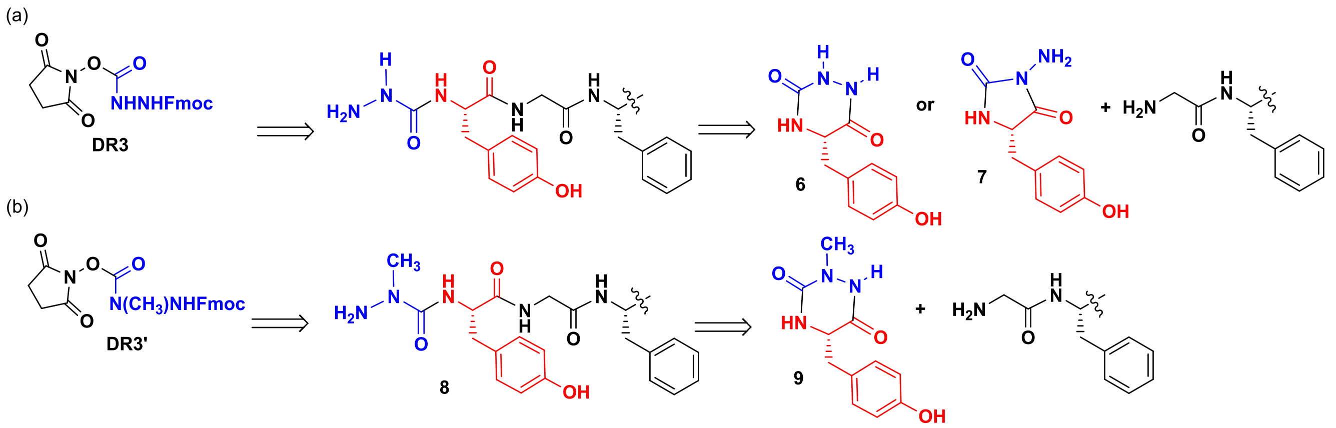
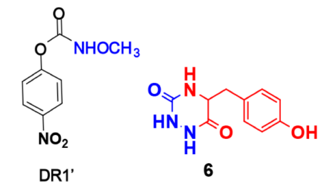


Figure 7. (a) Two possible heterocyclic products with **DR3**. (b) Only one possible heterocyclic product with **DR3′**.

experiments, we concluded that **DR3** is an optimal reagent for analyzing all 20 amino acids at the N-terminus.

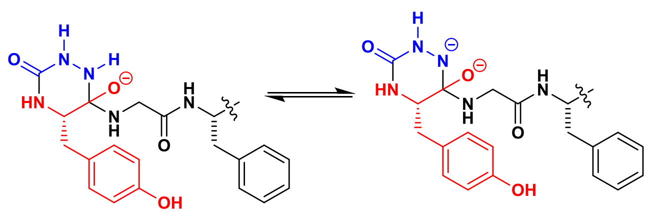


Next, to increase the efficiency of the sequencing step, we aimed to reduce the sequencing time by increasing the temperature. We performed the sequencing step on the derivatized peptide synthesized by using **DR3** at an elevated temperature of 60 °C (Figure 6). This adjustment reduced the time to achieve >92% conversion to degraded peptide **3** to only 2 h (Figure S12). We observed two or more peaks in the LC trace of the sequencing step with identical mass, all corresponding to degraded peptides. These peaks are due to epimerization[41,42] caused by the higher temperature and the alkaline pH of the solution during sequencing of reaction. When monitoring the efficiency of the repeated base-induced chain-end cleavage (see below), we find no dropoff of efficiency of the process, indicating that the diastereomers generated have little effect on the derivatization or cyclization steps, likely due to their distance from the centers of reactivity.

Given that either the −NH− or $-NH_2$ group of the derivatized peptides could cyclize on the adjacent amide, we sought to investigate whether a 5-membered or 6-membered heterocyclic product was formed, respectively (**6** or **7**, Figure 7a). To this end, we synthesized **DR3′**, a methylated reagent derivative of **DR3**, in which the NH group was blocked, reducing its nucleophilicity and eliminating the ability for deprotonation before or after the formation of the tetrahedral intermediate (Figures S74−S76). Thus, we carried out the derivatization reaction of the model peptide with **DR3′**. However, we observed a low yield of derivatized peptide **8**, likely due to steric hindrance. Irrespective of this, our mechanistic question concerned the cleavage reaction; hence, we moved forward by purifying **8** using RP-HPLC.

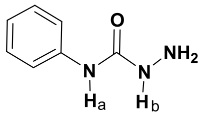
Afterward, we conducted the sequencing reaction in the presence of 1% $Ba(OH)_2$ for 4 h. The formation of the degraded peptide **3**, along with the generation of heterocycle **9**, was observed in good yield (Figure 7b) (Figures S13 and S14), supporting our hypothesis of six-membered ring formation and that the terminal $-NH_2$ of **DR3** performs the nucleophilic attack. Given that the terminal $-NH_2$ of the hydrazine carboxamide group is the nucleophile, it was reasonable to ask if a base was even required in the degradation step. After all, this $-NH_2$ is not likely to be protonated to any significant extent at neutral pH, given that its pK_a of its conjugate acid is likely similar to that of a hydrazide, i.e., around 5.[43] Thus, it could cyclize without the assistance of NaOH or $Ba(OH)_2$. However, when we performed the sequencing reaction in an aqueous environment without base, it led to only 10% degradation of the peptide after 4 h, whereas with base in 2 h, conversion was above 90%. Additionally, there was still a lingering question of whether the tetrahedral intermediate, once formed, undergoes deprotonation as has been reported in amide hydrolysis reactions (Figure 4c).

To explore this question and to determine if the terminal $-NH_2$ group can be deprotonated and thereby generate a better nucleophile, we recorded the ^1H NMR spectra in acetonitrile (where no deuterium exchange can occur) of 4-phenylsemicarbazide (**PSC**) as a model of the derivatized peptides. Then, to induce base-catalyzed deuterium exchange, we added NaOD (40 wt % in D_2O) to generate a 2% solution. The signals of NH_a and NH_b as well as the terminal $-NH_2$ group are due to deuteration over a few minutes, indicating their facile deprotonation. To further confirm that this deuterium exchange depends upon base concentration, we increased it to 10% NaOD. Indeed, the intensity of the signals diminished faster (Figure S15). This deuterium scrambling experiment revealed that terminal NH_2 can be deprotonated by NaOH. However, as discussed above, the cleavage reaction of the **DR3**-derivatized peptide can proceed without the addition of base, indicating that direct nucleophilic attack by the terminal $-NH_2$ group can initiate degradation, albeit at a far lower level of efficiency. This coupled with the fact that hydrazine can cleave peptides at asparagine linkages at neutral pH (Figure 2d) leads us to believe that one role of the base is to deprotonate the tetrahedral intermediate (eq 1), thereby facilitating leaving group departure.



In further support, we note that when using **DR1**, the cleavage reaction also proceeds at lower efficiency, potentially because a dianionic tetrahedral intermediate is not possible (albeit this is complicated by the fact that a side product also lowers the efficiency). Thus, while the base is needed to improve efficiency, we are not entirely confident of the role it plays. We find that it can deprotonate the terminal $-NH_2$ group to create a significantly better nucleophile, yet this is not necessary in our studies or in literature examples. Additionally, our results show that deprotonating the tetrahedral intermediate is also not necessarily required. But, when using **DR3**, both deprotonating the terminal $-NH_2$ and the tetrahedral intermediate are possible, and this derivative gives us the highest efficiency cleavage.

To explore the broad applicability of using **DR3** to remove N-terminal amino acids, we tested all 20 amino acids. For example, tryptophan is at the N-terminus (Table 1, entry 1).



PSC

Table 1. Sequencing of Peptides Having Different N-Terminal Amino Acids

Entry	Peptides with different N-terminal amino acids	Time for sequencing step	% Conversion of degraded peptide
1	WYFWVY	4 h	82
2	HYFWVY	2 h	92
3	PYFWVY	4 h	87
4	FYFWVY	4 h	92
5	GYFWVY	4 h	85
6	AYFWVY	2 h	91
7	VYFWVY	4 h	93
8	LYFWVY	4 h	87
9	IYFWVY	4 h	85
10	SYFWVY	2 h	92
11	TYFWVY	2 h	90
12	RYFWVY	2 h	93
13	MYFWVY	4 h	87
14	DYFWVY	2 h	91
15	NYFWVY	4 h	70
16	KYFWVY	2 h	92
17	C*YFWVY	4 h	60
18	EYFWVY	2 h	90
19	QYFWVY	2 h	93

% conversion calculated using LC-MS traces, with toluic acid used as internal standard. * cysteine (C) was modified with iodoacetamide.

We observed a peak in the LC-MS spectra corresponding to the cyclized amino acid adduct, which was subsequently isolated and characterized by HRMS (Figures S16−S18). With lysine at the N-terminus, we observed disubstitution in the derivatization step.

Nevertheless, it successfully underwent cleavage under the optimized conditions (Table 1, entry 16) (Figures S47 and S48). However, the derivatization step for cysteine had a relatively poor conversion due to the reactivity of the thiol group with **DR3**. To address this, we capped the thiol with iodoacetamide, a commonly practiced reaction in protein sciences. We then performed sequencing and observed only 60% conversion rate but with an uncharacterized side product. Unfortunately, we were unable to fully identify the side product at this stage (Table 1, entry 17, Figures S49 and S50). Further, previous literature has reported that both glutamic acid and glutamine convert into pyroglutamic acid under basic conditions.[44] In our case, during the derivatization step, we observed more than 50% formation of pyroglutamic acid when we have glutamic acid at the N-terminus, while lowering the equivalents of DIPEA and decreasing the temperature from 25 to 0 °C improved the yield of the derivatized peptide to 60%. Consequently, we purified this derivatized peptide and subjected it to the sequencing conditions, resulting in successful elimination of the N-terminal Glu and generation of the degraded peptide (Table 1, entry 18, Figures S51 and S52). However, in the case of glutamine, only ∼5% pyroglutamic acid formation was observed during the derivatization reaction using modified conditions, while the desired derivatized peptide gave 93% degradation (Table 1, entry 19). These exact same issues are encountered with cysteine[45] and glutamic acid[46] at the N-terminus when using Edman degradation because of the basic conditions in the functionalization step. For other N-terminal amino acids, there was varying reactivity, but we did not observe any significant side products. For consistency, we capped the sequencing reaction at 4 h, although, in most cases, the N-terminal amino acid was cleaved off within 2 h. However, for tryptophan, glycine, leucine, isoleucine, and asparagine, we observed unreacted **DR3**-derivatized peptides even after 4 h, which contributed to conversions of less than 90%. Overall, all amino acids at the N-terminus worked well, resulting in good conversions for elimination of the degraded peptide (Table 1 and Figures S16−S54). Based on these results, we conclude that **DR3** effectively derivatizes each of the 20 amino acids when placed on the N-terminus, as well as successfully removes all 20 amino acids with a high yield in base (Figure 8). Furthermore, the detection of cyclic amino acids cleaved from peptides in LC-MS spectra, supported by mechanistic studies, substantiated the elements required from our original design criteria.

After examining each of the amino acids at the N-terminus, we were interested in exploring the application of the method for sequencing a series of amino acids sequentially. We performed this test in the same manner as the original report by Edman, isolating the products from each cleavage and resubjecting the truncated peptides to the derivatization and cleavage conditions. To ensure an adequate quantity of material for subsequent steps, we scaled up the derivatization step using peptide **1** (YGFWVY, 0.02 mmol), successfully obtaining a 90% yield of the desired product. In the sequencing step, the reaction was initially performed on a 0.4 mM scale using 100 μL of 1% $Ba(OH)_2$. However, in the present case, we began with 0.01 mmol of the **DR3**-derivatized peptide. As the reaction scale increased, additional solvent was required to fully dissolve this peptide. Consequently, the volume of 1% $Ba(OH)_2$ was increased to 500 μL to maintain the solution's pH. After cleavage, we neutralized the reaction mixture with 1% formic acid and purified the truncated peptide via RP-HPLC to eliminate salts and other impurities before proceeding to the next cycle of derivatization and cleavage (Figures S55- S60). We demonstrated the sequential elimination of three amino acids, with potential for further extension as required (Figure 9).

Aiming to reduce the amount of base used in the above experiment, we devised a strategy of parallel reactions while maintaining the same reaction scale (Page no S11). This ability

https://doi.org/10.1021/jacs.5c03385
J. Am. Chem. Soc. 2025, 147, 13973−13982

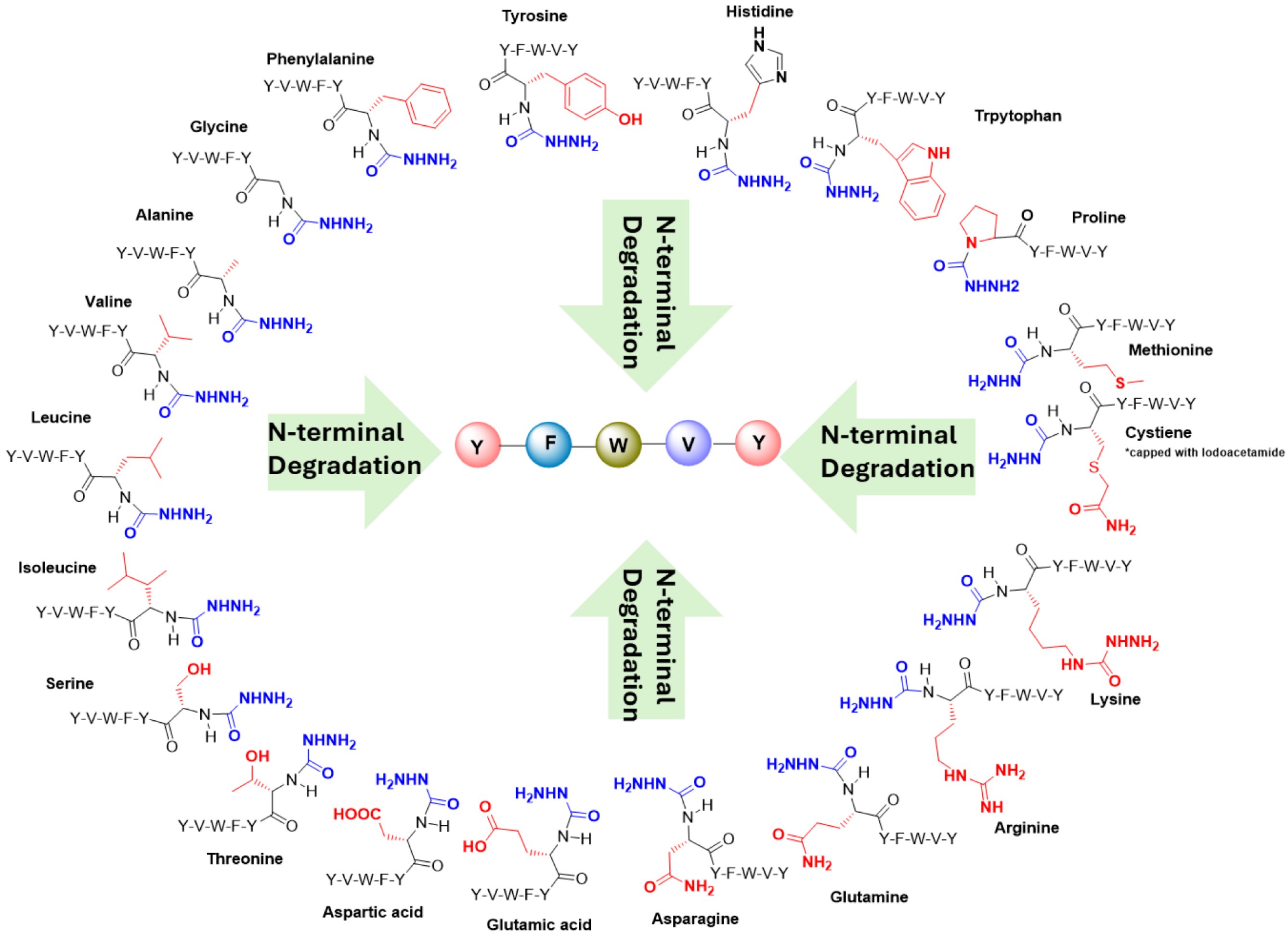


Figure 8. Schematic of base-induced peptide sequencing using **DR3** for peptides with all 20 different amino acids at their N-termini.

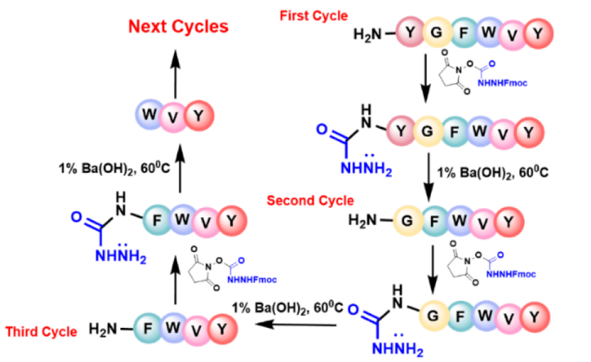


Figure 9. Sequential removal of three amino acids of peptide **1** in solution using **DR3**.

enables the successful identification of amino acid sequences in unknown peptides.

To evaluate the versatility and robustness of our method across a broad spectrum of peptide sequences, we also tested it on a complex peptide mixture with diverse N-terminal residues. Specifically, we selected a predigested bovine serum albumin (BSA) sample, which consists of more than 50 unique peptides, providing an ideal model for assessing the method's applicability in proteomics workflows. We employed **DR3** for peptide derivatization and analyzed the products by tandem mass spectrometry (MS/MS). The results were highly promising, with an average of 84% of the peptides successfully derivatized with **DR3**, and individual reaction efficiencies ranging from 42% to as high as 98%, depending on the N-terminal amino acid (Figure S61). When we subsequently treated the derivatized peptide mixture with 1% $Ba(OH)_2$ and heated it for 4 h (Figure 10), we found that 52% of the derivatized peptides were N-terminally cleaved (Supplementary Section 2.11 details the analysis process utilized for the data presented on the complex mixture of BSA peptides and provides a link to the raw MS data set). We suspect that the lower observation of N-terminal cleavage in this complex mixture is primarily due to low peptide recovery during the desalting and chromatography processes required for mass spectrometry analysis. This is supported by the fact that we did not observe any remaining derivatized peptides and that there were significantly fewer overall ionizable peptide signals.

To assess the compatibility of chemistry with DNA, such as that attached to barcode peptides, we investigated the stability of single-stranded DNA (ssDNA) in our sequencing protocol. A 0.1 mM solution of a 15-mer ssDNA was subjected to 1% $Ba(OH)_2$ at 60 °C for 1−4 h. Analysis via denaturing TBE-urea gel electrophoresis showed no significant degradation products (Supplementary Figure 62). Additionally, MALDI-MS demonstrated that the 15-mer ssDNA substrate experienced only minor depurination, primarily at the 5′ adenosine base, with no substantial degradation observed (Supplementary Figure 63). These findings confirm the robustness of our sequencing

https://doi.org/10.1021/jacs.5c03385
J. Am. Chem. Soc. 2025, 147, 13973−13982

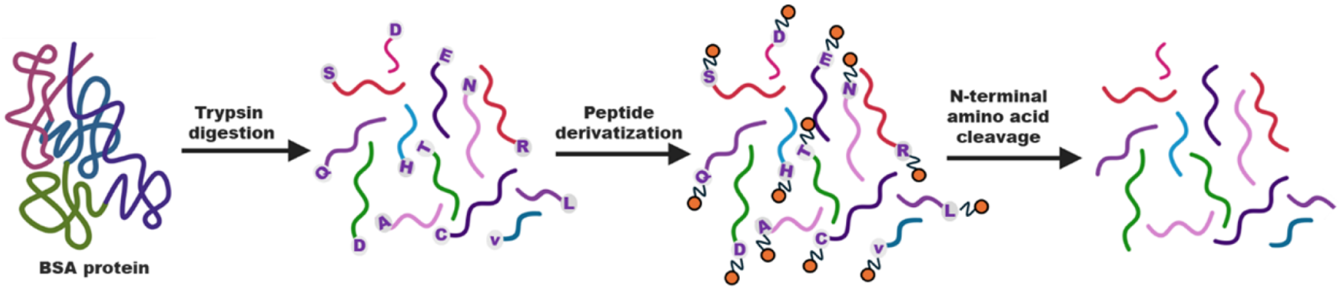


Figure 10. Schematic representation of derivatization followed by the sequencing process for peptide mixtures derived from digested BSA peptides.

method with DNA substrates, underscoring its potential for applications involving conjugated peptide-DNA in barcoding.

3. CONCLUSION

The derivatization of N-terminal amino acids of a peptide using **DR3** followed by the addition of a base provides a complementary method to the Edman degradation. We investigated the reaction mechanism using methyl substitution and isotope scrambling. Further, the six-membered ring cyclized adduct of the amino acids cleaved from the peptide was characterized. We also achieved the sequencing of three amino acids, one by one, with minimal loss in yield demonstrating the effective identification of the sequence of an unknown peptide. Additionally, we explored the stability of oligonucleotides under the sequencing conditions. Our methodology was successfully extended to peptide mixtures derived from protein samples, where all peptides were derivatized with **DR3** and underwent parallel elimination of their N-terminal amino acids upon base addition. Currently, we are focusing on base-induced N-terminal degradation on solid supports, such as glass slides or beads, to eliminate the need for purification after each sequencing step, thereby improving the method to rival the highest yields of Edman sequencing. Overall, this study highlights how mechanistic physical organic chemistry can be used to devise new methods that can be applied to proteomics.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.5c03385.

Detailed information on the materials and methods, experimental procedures, LC-MS, HRMS, MS-MS, MALDI, and NMR spectra (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Eric V. Anslyn − *Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712, United States;* orcid.org/0000-0002-5137-8797; Email: anslyn@austin.utexas.edu

Edward M. Marcotte − *Department of Molecular Biosciences, The University of Texas at Austin, Austin, Texas 78712, United States;* orcid.org/0000-0001-8808-180X; Email: marcotte@utexas.edu

Authors

Harnimarta Deol − *Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712, United States;* orcid.org/0000-0003-2798-564X

Ava Raeisbahrami − *Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712, United States*

Phuoc H.T. Ngo − *Department of Chemistry, The University of Texas at Austin, Austin, Texas 78712, United States*

Jagannath Swaminathan − *Erisyon Inc, Austin, Texas 78752, United States*

Ophelia Papoulas − *Department of Molecular Biosciences, The University of Texas at Austin, Austin, Texas 78712, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.5c03385

Notes

The authors declare the following competing financial interest(s): EVA and EMM are co-founders and shareholders of Erisyon, Inc. HD and EVA are co-inventors of a patent to this work which is currently pending. JS is co-founder, shareholder, and an employee of Erisyon, Inc.

■ ACKNOWLEDGMENTS

This work was supported by grants from the NIH (1R35GM149308 to E.V.A., R35GM122480 to E.M.M.), a sponsored research agreement from Erisyon, Inc. (UTA 18-000440 to E.V.A. and E.M.M.), the Army Research Office (W911NF-12-1-0390 to E.M.M.), and the Welch Foundation (F-1515 to E.M.M., Welch Regents Chair F-0046 to E.V.A.).

■ REFERENCES

(1) Vogel, C.; Bashton, M.; Kerrison, N. D.; Chothia, C.; Teichmann, S. A. Structure Function, and Evolution of Multidomain Proteins. *Curr. Opin. Struct. Biol.* **2004**, *14*, 208−216.

(2) Dephoure, N.; Gould, K. L.; Gygi, S. P.; Kellogg, D. R. Mapping and Analysis of Phosphorylation Sites: A Quick Guide for Cell Biologists. *Mol. Biol. Cell* **2013**, *24*, 535−542.

(3) Fosgerau, K.; Hoffmann, T. Peptide therapeutics: Current status and future directions. *Drug Discovery Today* **2015**, *20*, 122−128.

(4) Lau, J. L.; Dunn, M. K. Therapeutic peptides: Historical perspectives, current development trends, and future directions. *Bioorg. Med. Chem.* **2018**, *26*, 2700−2707.

(5) Qi, G.-B.; Gao, Y.-J.; Wang, L.; Wang, H. Self-assembled peptide-based nanomaterials for biomedical imaging and therapy. *Adv. Mater.* **2018**, *30* (22), 1703444.

(6) Nguyen, D. N.; Becker, G. W.; Riggin, R. M. Protein mass spectrometry: Applications to analytical biotechnology. *J. Chromatogr. A* **1995**, *705*, 21−45.

(7) Torres, M. D. T.; Cao, J.; Franco, O. L.; Lu, T. K.; de la Fuente-Nunez, C. Synthetic biology and computer-based frameworks for antimicrobial peptide discovery. *ACS Nano* **2021**, *15*, 2143−2164.

(8) Sword, T.; Abbas, G.; Bailey, C. B. Cell-free protein synthesis for nonribosomal peptide synthetic biology. *Front. Nat. Prod.* **2024**, *3*, 1353362.

(9) Singh, A. Nanopores for sequencing proteins. *Nat. Methods* **2023**, *20* (12), 1870.
(10) Zhang, Y.; Yi, Y.; Li, Z.; Zhou, K.; Liu, L.; Wu, H.-C. Peptide Sequencing Based on Host−Guest Interaction-Assisted Nanopore Sensing. *Nat. Methods* **2024**, *21*, 102−109.
(11) Swaminathan, J.; Boulgakov, A. A.; Hernandez, E. T.; Bardo, A. M.; Bachman, J. L.; Marotta, J.; Johnson, A. M.; Anslyn, E. V.; Marcotte, E. M. Highly Parallel Single-Molecule Identification of Proteins in Zeptomole-Scale Mixtures. *Nat. Biotechnol.* **2018**, *36*, 1076−1082.
(12) Reed, B. D.; Meyer, M. J.; Abramzon, V.; Ad, O.; Adcock, P.; Ahmad, F. R.; Alppay, G.; Ball, J. A.; Beach, J.; Belhachemi, D.; Bellofiore, A.; Bellos, M.; Beltrán, J. F.; Betts, A.; Bhuiya, M. W.; Blacklock, K.; Boer, R.; Boisvert, D.; Brault, N. D.; Buxbaum, A. Real-Time Dynamic Single-Molecule Protein Sequencing on an Integrated Semiconductor Device. *Science* **2022**, *378*, 186−192.
(13) Edman, P.; Högfeldt, E.; Sillén, L. G.; Kinell, P.-O. Method for Determination of the Amino Acid Sequence in Peptides. *Acta Chem. Scand.* **1950**, *4*, 283−293.
(14) Edman, P.; Begg, G. A Protein Sequenator. *Eur. J. Biochem.* **1967**, *1* (1), 80−91.
(15) Miyashita, M.; Presley, J. M.; Buchholz, B. A.; Lam, K. S.; Lee, Y. M.; Vogel, J. S.; Hammock, B. D. Attomole Level Protein Sequencing by Edman Degradation Coupled with Accelerator Mass Spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 4403−4408.
(16) Henzel, W. J.; Billeci, T. M.; Stults, J. T.; Wong, S. C.; Grimley, C.; Watanabe, C. Identifying Proteins from Two-Dimensional Gels by Molecular Mass Searching of Peptide Fragments in Protein Sequence Databases. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 5011−5015.
(17) Restrepo-Pérez, L.; Joo, C.; Dekker, C. Paving the Way to Single-Molecule Protein Sequencing. *Nat. Nanotechnol.* **2018**, *13*, 786−796.
(18) van Ginkel, E.; Filius, M.; Szczepaniak, M.; Tulinski, P.; Meyer, A. S.; Joo, C. Single-Molecule Peptide Fingerprinting. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (13), 3338−3343.
(19) Filius, M.; van Wee, R.; de Lannoy, C.; Westerlaken, I.; Li, Z.; Kim, S. H.; de Agrela Pinto, C.; Wu, Y.; Boons, G.-J.; Pabst, M.; de Ridder, D.; Joo, C. Full-Length Single-Molecule Protein Fingerprinting. *Nat. Nanotechnol.* **2024**, *19*, 652−659.
(20) Alfaro, J. A.; Bohländer, P.; Dai, M.; Filius, M.; Howard, C. J.; van Kooten, X. F.; Ohayon, S.; Pomorski, A.; Schmid, S.; Aksimentiev, A.; Anslyn, E. V.; Bedran, G.; Cao, C.; Chinappi, M.; Coyaud, E.; Dekker, C.; Dittmar, G.; Drachman, N.; Eelkema, R.; Goodlett, D.; Hentz, S.; Kalathiya, U.; Kelleher, N. L.; Kelly, R. T.; Kelman, Z.; Kim, S. H.; Kuster, B.; Rodriguez-Larrea, D.; Lindsay, S.; Maglia, G.; Marcotte, E. M.; Marino, J. P.; Masselon, C.; Mayer, M.; Samaras, P.; Sarthak, K.; Sepiashvili, L.; Stein, D.; Wanunu, M.; Wilhelm, M.; Yin, P.; Meller, A.; Joo, C. The Emerging Landscape of Single-Molecule Protein Sequencing Technologies. *Nat. Methods* **2021**, *18* (6), 604−617.
(21) Hong, J. M.; Gibbons, M.; Bashir, A.; Wu, D.; Shao, S.; Cutts, Z.; Chavarha, M.; Chen, Y.; Schiff, L.; Foster, M.; et al. ProtSeq: Toward High-Throughput, Single-Molecule Protein Sequencing via Amino Acid Conversion into DNA Barcodes. *iScience* **2022**, *25* (1), 103586.
(22) Zheng, L.; Sun, Y.; Eisenstein, M.; Soh, H. T. Peptide Sequencing via Reverse Translation of Peptides into DNA. *bioRxiv* **2024**, *05*, 31.
(23) Brinkerhoff, H.; Kang, A. S. W.; Liu, J.; Aksimentiev, A.; Dekker, C. Multiple Rereads of Single Proteins at Single-Amino Acid Resolution Using Nanopores. *Science* **2021**, *374*, 1509−1513.
(24) Hu, Z.-L.; Huo, M.-Z.; Ying, Y.-L.; Long, Y.-T. Biological Nanopore Approach for Single-Molecule Protein Sequencing. *Angew. Chem., Int.* **2021**, *60*, 14738−14749.
(25) Howorka, S.; Siwy, Z. S. Reading Amino Acids in a Nanopore. *Nat. Biotechnol.* **2020**, *38*, 159−160.
(26) Smith, R. M.; Hansen, D. E. The pH-Rate Profile for the Hydrolysis of a Peptide Bond. *J. Am. Chem. Soc.* **1998**, *120*, 8910−8913.
(27) East, A. L. L. On the Hydrolysis Mechanisms of Amides and Peptides. *Int. J. Chem. Kinet.* **2018**, *50*, 705−709.
(28) Ian, Y.; Engel, M. H.; Macko, S. A.; Carpenter, S.; Deming, J. W. Kinetics of Peptide Hydrolysis and Amino Acid Decomposition at High Temperature. *Geochim. Cosmochim. Acta.* **1993**, *57*, 3281−3293.
(29) Gassman, P. G.; Hodgson, P. K. G.; Balchunis, R. J. Base-Promoted Hydrolysis of Amides at Ambient Temperatures. *J. Am. Chem. Soc.* **1976**, *98*, 1275−1276.
(30) Drake, D.; Schowen, R. L.; Jayaraman, H. Amide hydrolysis. VI. Diffusion-controlled proton transfer and heavy-atom reorganization in the general acid, specific base catalyzed hydrolysis of amides. *J. Am. Chem. Soc.* **1973**, *95* (2), 454−458.
(31) Steinberg, S. M.; Bada, J. L. Peptide Decomposition in the Neutral pH Region via the Formation of Diketopiperazines. *J. Org. Chem.* **1983**, *48* (13), 2295−2298.
(32) Akabori, S.; Ohno, K.; Ikenaka, T.; Okada, Y.; Hanafusa, H.; Haruna, I.; Tsugita, A.; Sugae, K.; Matsushima, T. Hydrazinolysis of Peptides and Proteins. II. Fundamental Studies on the Determination of the Carboxyl-Ends of Proteins. *Bull. Chem. Soc. Jpn.* **1956**, *29*, 507−518.
(33) Bradburry, J. H. An Improvement of the Hydrazine Method for Determination of C-Terminal Amino Acids. *Nature* **1956**, *178* (4539), 912−913.
(34) Morishita, M.; Sowa, T.; Sakiyama, F.; Narita, K. The Non-Enzymatic Cleavage of Peptide Bonds. I. Deacylation of r-Oxoacyl Amino Acids and Peptide by Hydrazines. *Bull. Chem. Soc. Jpn.* **1967**, *40*, 632−639.
(35) Miyatake, N.; Satake, K.; Kamo, M.; Tsugita, A. Specific Chemical Cleavage of Asparaginyl and Glycyl-Glycine Bonds in Peptides and Proteins by Anhydrous Hydrazine Vapor. *J. Biochem.* **1994**, *115*, 208−212.
(36) Bornstein, P.; Balian, G. Cleavage at Asn-Gly Bonds with Hydroxylamine. *Methods Enzymol.* **1977**, *47*, 132−145.
(37) Hamada, Y. A Novel N-Terminal Degradation Reaction of Peptides via N-Amidination. *Bioorg. Med. Chem. Lett.* **2016**, *26*, 1690−1695.
(38) Fina, N. J.; Edwards, J. O. The Alpha Effect: A Review. *Int. J. Chem. Kinet.* **1973**, *5*, 1−26.
(39) Baldwin, J. E. Rules for Ring Closure. *J. Chem. Soc., Chem. Commun.* **1976**, *20*, 734−736.
(40) Makarycheva-Mikhailova, A. V.; Stanbury, D. M.; McKee, M. L. Kinetic Isotope Effect and Extended pH-Dependent Studies of the Oxidation of Hydroxylamine by Hexachloroiridate(IV): ET and PCET. *J. Phys. Chem. B* **2007**, *111*, 6942−6948.
(41) El-Faham, A.; Albericio, F. Peptide Coupling Reagents: More than a Letter Soup. *Chem. Rev.* **2011**, *111*, 6557−6602.
(42) Duengo, S.; Muhajir, M. I.; Hidayat, A. T.; Musa, W. J. A.; Maharani, R. Epimerization in Peptide Synthesis. *Molecules* **2023**, *28*, 8017.
(43) Raddatz, S.; Mueller-Ibeler, J.; Kluge, J.; Wäß, L.; Burdinski, G.; Havens, J. R.; Onofrey, T. J.; Wang, D.; Schweitzer, M. Hydrazide Oligonucleotides: New Chemical Modification for Chip Array Attachment and Conjugation. *Nucleic Acids Res.* **2002**, *30*, 4793−4802.
(44) Liu, Z.; Valente, J.; Lin, S.; Chennamsetty, N.; Qiu, D.; Bolgar, M. Cyclization of N-Terminal Glutamic Acid to Pyroglutamic Acid Impacts Monoclonal Antibody Charge Heterogeneity Despite Its Appearance as a Neutral Transformation. *J. Pharm. Sci.* **2019**, *108*, 3194−3200.
(45) Yan, J. X.; Kett, W. C.; Herbert, B. R.; Gooley, A. A.; Packer, N. H.; Williams, K. L. Identification and Quantitation of Cysteine in Proteins Separated by Gel Electrophoresis. *J. Chromatogr. A* **1998**, *813*, 187−200.
(46) Walker, J. M. *Methods in Molecular Biology: proteins*; Humana Press: New York, 1984, Vol. *1*.