Efficient and directed peptide bond formation in the gas phase via ion/ion reactions

William M. McGee and Scott A. McLuckey

Department of Chemistry, Purdue University, West Lafayette, IN 47907

Edited by Peter B. Armentrout, University of Utah, Salt Lake City, UT, and accepted by the Editorial Board December 18, 2013 (received for review September 24, 2013)

Amide linkages are among the most important chemical bonds in living systems, constituting the connections between amino acids in peptides and proteins. We demonstrate the controlled formation of amide bonds between amino acids or peptides in the gas phase using ion/ion reactions in a mass spectrometer. Individual amino acids or peptides can be prepared as reagents by (i) incorporating gas phase–labile protecting groups to silence otherwise reactive functional groups, such as the N terminus; (ii) converting the carboxyl groups to the active ester of N-hydroxysuccinimide; and (iii) incorporating a charge site. Protonation renders basic sites (nucleophiles) unreactive toward the N-hydroxysuccinimide ester reagents, resulting in sites with the greatest gas phase basicities being, in large part, unreactive. The N-terminal amines of most naturally occurring amino acids have lower gas phase basicities than the side chains of the basic amino acids (i.e., those of histidine, lysine, or arginine). Therefore, reagents may be directed to the N terminus of an existing “anchor” peptide to form an amide bond by protonating the anchor peptide’s basic residues, while leaving the N-terminal amine unprotonated and therefore reactive. Reaction efficiencies of greater than 30% have been observed. We propose this method as a step toward the controlled synthesis of peptides in the gas phase.

Significance

The generation of peptide bonds is of longstanding interest both from the standpoints of the origin of life and for synthesis. We describe here a general approach for forming peptide linkages in the gas phase via ion/ion reactions. Examples of the addition of a single amino acid to the N terminus of a peptide and the addition of a polypeptide to the N terminus of a peptide are demonstrated. This work constitutes a unique means for generating peptide linkages that are fast (<1 s), efficient (tens of percent), and flexible. Possible application areas include the generation of peptide libraries for the generation of tandem MS data and for the soft landing of peptide products for microarrays.

Author contributions: W.M.M. and S.A.M. designed research; W.M.M. performed research; and W.M.M. and S.A.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. P.B.A. is a guest editor invited by the Editorial Board.

1To whom correspondence should be addressed. E-mail: mcluckey@purdue.edu.

This article contains supporting information online at www.pnas.orglookup/suppl/doi:10.1073/pnas.1317914111/-/DCSupplemental.
contains a sufficiently delocalized, or from being a competitive channel. Furthermore, if the reagent reagents, the electron affinities are high enough to prevent ET competing pathway to consider. In the case of sulfoNHS ester between ET and electrostatic complex formation is the main agent does not have a low electron affinity, then competition between ET begins to compete with PT (36). However, if the reagent has a low enough electron affinity to transfer an electron, then the reactivity be-

between PT and electrostatic complex formation is the main competing pathway to consider. In the case of sulfoNHS ester reagents, the electron affinities are high enough to prevent ET from being a competitive channel. Furthermore, if the reagent contains a sufficiently delocalized, or “sticky,” charge-bearing site, then electrostatic complex formation tends to be the dominant pathway over PT. In the case of a multiply protonated anchor peptide and an anionic sulfoNHS ester reagent, the strong attraction between the oppositely charged ions results in the formation of a long-lived electrostatic complex. Upon activation of the electrostatic complex via energetic collisions, a nucleophilic attack by the unprotonated N-terminal amine of the anchor peptide on the carbonyl carbon of the sulfoNHS ester, resulting in (8) the formation of an amide bond and loss of a neutral sulfoNHS molecule, forming an extended and protected peptide. Subsequent activation of the new and protected peptide (C) results in loss of the PG with simultaneous transfer of a hydrogen from the PG to regenerate an amine, forming the new and unprotected peptide.

**Results and Discussion**

**Peptide Extension: A General Mechanism.** Ion/ion reactions in a mass spectrometer are fast and highly exothermic as a result of, among other things, a strong Coulombic attraction (31–35). Competing pathways in these reactions mainly include proton transfer (PT), electron transfer (ET), and electrostatic complex formation, all of which are heavily dependent on the nature of the two reactants. If the anionic reagent has a low enough electron affinity to transfer an electron, then the reaction between ET begins to compete with PT (36). However, if the reagent does not have a low electron affinity, then competition between PT and electrostatic complex formation is the main competing pathway to consider. In the case of sulfoNHS ester reagents, the electron affinities are high enough to prevent ET from being a competitive channel. Furthermore, if the reagent contains a sufficiently delocalized, or “sticky,” charge-bearing site, then electrostatic complex formation tends to be the dominant pathway over PT. In the case of a multiply protonated anchor peptide and an anionic sulfoNHS ester reagent, the strong attraction between the oppositely charged ions results in the formation of a long-lived electrostatic complex. Upon activation of the electrostatic complex via energetic collisions, a nucleophilic attack by the unprotonated N-terminal amine of the anchor peptide on the carbonyl carbon of the sulfoNHS ester takes place within the complex (Fig. 1A), resulting in the formation of an amide bond at the N terminus of the anchor peptide with concomitant loss of the neutral sulfoNHS molecule (Fig. 1F). At this point, the anchor peptide has been extended; however, the protecting groups, labeled PG in Fig. 1, remain on reactive functionalities. A second activation step results in the loss of the protecting group(s), regenerating the initial functional group(s) (Fig. 1C). The protecting group tert-butylxycarbonyl (boc) is used to protect amines, such as the N-terminal amine, the ε-amine of lysine, and the secondary amine of the imidazole ring of histidine. Several similar protecting groups have been shown to be labile in the gas phase such that, upon activation, these protecting groups are lost and the corresponding functional group is regenerated. A list of protecting groups investigated and the matching functional groups is provided in Table 1. Data demonstrating gas phase deprotection of several functional groups are provided in SI Text along with a table that summarizes the amino acids and di- and tripeptides that have been added to anchor peptides and the reagents used for their coupling.

**Addition of a Tripeptide Reagent to an Anchor Peptide.** Peptide extension is demonstrated here by adding the tripeptide GGG, where G is glycine, to the anchor peptide PKAAAKA, where P is proline, K is lysine, and A is alanine. Unless described as being N-terminally protected with a boc group or C-terminally activated with a sulfoNHS ester, the peptides are otherwise unmodified. We have shown previously that primary amines, such as the ε-amine of lysine and the N-terminal amines, are highly reactive toward sulfoNHS esters via ion/ion reactions (16–18). This is a demonstration of covalent modification of a secondary amine using sulfoNHS esters via gas phase ion/ion reactions. Here, proline was chosen to show that, despite the greater basicity of the secondary amine of proline compared with the primary amines of the remaining naturally occurring amino acids, protonation at the basic residue, lysine in this case, was still more favorable. Consequently, this results in covalent modification occurring at the N-terminal
The gas phase ion/ion reaction between [boc–GGG–sulfoNHS – H]− and [PKAAAKA + 2H]+ initially forms a long-lived electrostatic complex [PKAAAKA + (boc–GGG–sulfoNHS) + H]+ (Fig. 2A). Collision-induced dissociation (CID) of the electrostatic complex (Fig. 2B) can access two competing fragmentation pathways (31–37). One pathway leads to a net PT, as reflected in the detachment of the neutral boc–GGG–sulfoNHS ester, indicating no covalent reaction. The second pathway involves covalent modification, as reflected by the signature loss of a neutral sulfoNHS molecule (16–18). Isolation and monoisotopic activation of the covalently modified peptide [boc–GGGPKAAAKA + H]+ (Fig. 2C) results mainly in the loss of the boc protecting group, generating the unprotected peptide GGGPKAAAKA. Additional fragmentation is observed by the presence of the b8 and b9 ions (Fig. 2C), which are produced from subsequent fragmentation of the unprotected peptide GGGPKAAAKA.

Comparison Between Gas Phase–Prepared and Solution-Prepared Peptides. Tandem mass spectrometry (MS/MS) can be performed on the peptide GGGPKAAAKA, prepared via ion/ion reactions to determine the sequence. To validate the identity of the newly formed peptide, its fragmentation spectrum was compared with that of a peptide of the same sequence prepared by NEOBioLab using SPPS. Fig. 3 compares the dissociation behavior of [GGGPKAAAKA + H]+ prepared in the mass spectrometer (Fig. 3A) to the common SPPS method (Fig. 3B). The two spectra are very nearly identical; however, the presence of y-ions labeled with a solid diamond, in the gas phase–prepared peptide spectrum (Fig. 3A), indicates evidence for reaction of a lysine side chain as a minor reaction channel for the peptide reactant ions. This would occur if a proton was bound at the N-terminal amine, leaving the ε-amine of one of the lysine residues unprotonated and therefore reactive toward the reagent. Although it is a much less favorable pathway, linkages to the side chains of lysine residues can be observed. This side reaction has not been observed when histidine or arginine residues are used as the basic proton-bearing sites in the anchor peptide. The precursor ion used to produce the spectrum of Fig. 3A could contain a mixture of the desired peptide, GGGPKAAAKA, and the anchor peptide with a modification at either lysine. The precursors of both the peak locations on the mass scale and their relative abundances produced from fragmentation of the gas phase–prepared peptide (Fig. 3A) to those produced from fragmentation of the SPPS-prepared peptide (Fig. 3B) suggest that the two precursor ions are nearly identical. In this regard, it is likely that the mixture of isomers of Fig. 3A is comprised, in large part, of the desired peptide, GGGPKAAAKA, with any other isomeric products present at under 10%. In general, to achieve separation of isomers that may be produced from reaction at two or more sites, one could envisage the use of ion mobility MS as a purification method for the desired product.

The single asterisk (*) indicates that primary or secondary amines can be protected with boc. The double asterisk (**) indicates that rather than a covalently bound protecting group, the guanidinium–sulfonate interaction is electrostatic.

<table>
<thead>
<tr>
<th>Functional group</th>
<th>Functional structure</th>
<th>Protecting group</th>
<th>Protected structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amine*</td>
<td>H₂N—</td>
<td>N-boc</td>
<td><img src="image" alt="Amine Structure" /></td>
</tr>
<tr>
<td>Sulfhydryl</td>
<td>HS—</td>
<td>Acetamidomethyl (5-acm)</td>
<td><img src="image" alt="Sulfhydryl Structure" /></td>
</tr>
<tr>
<td>Hydroxy</td>
<td>HO—</td>
<td>Tert-butyl ester (O-tBu)</td>
<td><img src="image" alt="Hydroxy Structure" /></td>
</tr>
<tr>
<td>Carboxylic Acid</td>
<td>O—</td>
<td>Tert-butyl ester (O-tBu)</td>
<td><img src="image" alt="Carboxylic Acid Structure" /></td>
</tr>
<tr>
<td>Guanidine**</td>
<td>H₂N—C—N—H</td>
<td>1,3-benzenedisulfonate (BDSA)</td>
<td><img src="image" alt="Guanidine Structure" /></td>
</tr>
</tbody>
</table>

The single asterisk (*) indicates that primary or secondary amines can be protected with boc. The double asterisk (**) indicates that rather than a covalently bound protecting group, the guanidinium–sulfonate interaction is electrostatic.
Reaction Efficiency. The overall reaction efficiency, defined as the percentage of initial peptide reactant ions that are converted to product, is dependent on several factors, such as the extent to which the peptide reactant ion is allowed to react with the reagent (with sufficient reagent anion numbers and reaction time), 100% depletion of the peptide reactant is possible, the extent to which the ion/ion complex is dissociated (under optimal activation conditions, 100% efficiency is possible), the competition between PT and covalent reaction, and the deprotection efficiency. Fig. 4 illustrates the overall efficiency for the addition of diglycine to the anchor peptide using conditions tuned to give relatively high step-wise efficiencies. In this case, an overall efficiency of 31% was measured. This value is based on comparing the abundance of the precursor anchor peptide, [AKAAAKA + 2H]+, to the extended and deprotected peptide ion, [GGAKAAAKA + H]+, generated using the gas phase process and is based on the sum of 20 scans for each spectrum. The various steps leading from the reactants to the final product are listed in Fig. 4, *Inset* and are analogous to those described in Fig. 2.

Conclusion
This process has applied common concepts of conjugation chemistry to gas phase ion/ion reactions performed in the mass spectrometer to create a method for rapid directed and efficient formation of amide bonds in peptides. A variety of amino acids and small peptides have been used to form amide bonds to the primary or secondary (proline) N-terminal amines of anchor peptides. Thus far, there has been no restriction to the amino acid(s) added to the anchor peptide or to the sequence of the anchor peptide, provided that certain criteria are met. The criteria for reaction are (i) the anchor peptide must bear n protons such that its absolute charge, [z], is greater than that of the reagent ion, [z–1]; (ii) the number of basic sites, including the N-terminal amine, must be n+1 for n protons; (iii) the reagent ion must have the nucleophilic functional groups protected with a gas phase–labile protecting group (boc in this case); (iv) the carboxyl group must be activated with a reactive ester (NHS); and (v) the reagent must contain a charge-bearing site ( sulfonate). This approach can be modified to create branched peptides by protecting (or protonating) nucleophilic sites on the anchor peptide such as the N terminus or by introducing sites with higher gas phase basicity, such as arginine residues, to favor protonation at those sites, keeping the less basic amines of lysine unprotonated and therefore free to react.

Materials and Methods
Boc-gly-gly–OH and boc-gly–gly–gly–OH were purchased from Bachem Americas, Inc. SulfoNHS and 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC) were purchased from PierceNet. Methanol (MeOH) and N,N-dimethylformamide (DMF) were purchased from Macron Chemicals. The peptides AKAAAKA, PKAAAKA, and GGGPKAAAKA were synthesized by NeoBioSci.

**SulfoNHS Ester Reagent Preparation.** Boc-protected amino acid and small peptides were dissolved in mixtures of MeOH and DMF. SulfoNHS esters were prepared by activating the carboxyl derivatives of the boc-protected amino acids or peptides using an equimolar amount of EDC in a mixture of DMF and H2O. To this, equimolar amounts of sulfoNHS in H2O were added, which resulted in the formation of an ester between the carboxyl group and the N-hydroxy of the succinimide, leaving the boc-protected sulfoNHS ester acid or peptide reagent. Final concentrations of sulfoNHS ester reagents were prepared to between 25 and 50 mM.

**MS.** All MS experiments were performed on QqQ hybrid triple quadrupole/linear ion trap (LIT) mass spectrometers (QTRAP 2000 and QTRAP 4000, AB Sciex), modified with home-built, alternately pulsed nanoelectrospray ionization sources (38). Analyte and reagent ions were sequentially mass-selected in the Q1 mass filter, and then subsequently transferred to the Q2 reaction cell for mutual storage or reactions (39) for periods between 500 ms and 1,000 ms in length. Mutual storage reactions were performed by applying alternating current waveforms to the IQ2 and IQ3 lenses, while ions of both polarities were held in the q2 reaction cell. The long-lived electrostatic complex reaction product ions were then transferred to the Q3 LIT, where they were mass-selected and collisionally activated using ion trap CID by applying a low amplitude dipolar waveform for periods between 100 and 2,000 ms. For all subsequent fragmentation steps, additional quadrupolar radio frequency and direct current voltage isolations followed by ion-trap CID were performed. The ions were then mass-analyzed using mass-selective axial ejection (40).

Activation of the electrostatic complex between the reagent ion and the anchor peptide required the longest activation period with the lowest amplitude. By using a low-amplitude waveform to slowly heat the electrostatic complex, the covalent modification pathway often becomes more favorable than the PT pathway. Although this behavior is highly dependent on the nature of the ions involved in the electrostatic complex, all of the reagents investigated herein have reacted in a manner consistent with this observation.

**Fig. 3.** Comparison of the CID spectra of \( M = [\text{GGGPKAAAKA} + H]^+ \) prepared (A) in the gas phase and (B) via SPPS. The lightning bolt represents the peak that is activated and water loss is labeled as a degree sign (°). The solid diamond represents fragments corresponding to a modification on a lysine.

**Fig. 4.** Determining the overall efficiency of the addition of diglycine (GG) to the anchor peptide based on the ratio of the final product ion, [GGAKAAAKA + H]+, shown in red on the right, to the precursor anchor peptide ion, [AKAAAKA + 2H]+, shown in blue on the left. The process by which GG was added is shown in the top right of the figure.
Following the formation of an amide bond between the protected reagent ion and the anchor peptide, subsequent activation steps may vary in the amplitude and time required to initiate gas phase deprotection of functional groups. In general, shorter activation periods with higher amplitudes may be used to initiate gas phase deprotection; however, the higher the amplitude used, the more likely it becomes that cleavages occurring along the peptide backbone will be observed in competition with deprotection events. Therefore, extended activation periods with low amplitudes may also be used for deprotection to maximize the formation of the unprotected peptide.

ACKNOWLEDGMENTS. Dr. Ryan T. Hilger, Dr. Boone M. Prentice, and Dr. John R. Stutzman are acknowledged for helpful discussions. This work was supported by National Institutes of Health Grant GM 45372.

Supporting Information

McGee and McLuckey 10.1073/pnas.1317914111

SI Text

Additional Materials and Methods. The following is a list of materials used: tert-butyloxycarbonyl (boc)–ala–OH, boc–phe–OH, boc–pro–OH, boc–lys(boc)–OH, boc–cys(acm)–OH, and 1,3-benzenedisulfonic acid (BDSA) were purchased from Sigma Aldrich. Boc–aspart(boc) ester (OrBu)–OH was purchased from Anaspec, Inc. Boc–arg–OH was purchased from Bachem Americas, Inc. Boc–ser(OrBu) –OH, boc–his(boc)–OH, and boc–asn–OH were purchased from P3biosystems. N-hydroxysuccinimide (NHS), N-hydroxysulfosuccinimide (sulfoNHS), and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) were purchased from PierceNet. Methanol and N,N-dimethylformamide (DMF) were purchased from Macron Chemicals. The peptide AKAAAKA was synthesized by NeoBioSci.

NHS Reagent Preparation. The NHS ester reagents were prepared in a similar fashion to the sulfoNHS esters described in the main text. The reactivity of NHS esters has been shown to be essentially identical to that of sulfoNHS esters (1). The NHS reagent boc–arg(BDSA)–NHS was prepared by activating the carboxyl group of boc–arg–OH with EDC in a mixture of DMF and H₂O followed by esterification using NHS in DMF. Next, this solution was mixed with a 5 mM solution of aqueous BDSA, producing a final reagent concentration of about 25 mM. This was followed shortly by ionization via nanoelectrospray ionization (nESI). A list of reagents prepared to date is given in Table 1.

Activation of Protecting Groups. In this report, all ions are activated using an alternating current (AC) auxiliary waveform with a frequency corresponding to the secular frequency of the mass-to-charge ratio (m/z) of the desired ions. A schematic of the instrument design is given in Fig. S1. The presence of reactive groups on amino acids and peptides results in the enormous variability found in peptides and proteins. In some cases, these reactivities are desirable; however, in other cases this is problematic. The preparation of gas phase reagent ions using amino acids or peptides as the central species requires protection of certain functional groups to allow the activation of others. The use of activated carboxyl groups for coupling reactions requires that the reactivity of nucleophilic functional groups be addressed. Without protection, some of these sites could undergo undesirable side reactions with themselves or another reagent. For example, with the reagent phe-sulfoNHS, the α-amino of phenylalanine could initiate a nucleophilic attack on either its own carbonyl carbon or that of a neighboring phenylalanine. The result of which is a highly diminished reagent concentration. In every case, the amine protecting group boc is used on the N-terminal amine and any side chain amines, preventing a nucleophilic reaction on the carboxyl group once it is activated. This is demonstrated in Fig. S2 and is discussed in more detail below.

There are a variety of other functional groups found in amino acids, some of which need to be protected in order for the preparation of these reagents to be successful. Sulphydrys and hydroxyls are functional groups that should also be protected for reagent preparation. Additionally, the acidic functionalities of aspartic and glutamic acids should be protected to ensure the appropriate carboxyl group is activated during reagent preparation. Table S1 provides a list of gas phase–labile protecting groups and the corresponding functional groups they are protecting.

It is desirable to use protecting groups that are stable in the solution phase under conditions required to prepare the reagent but are labile in the gas phase upon activation. Several protecting groups for multiple functional groups have been identified and used as gas phase–labile protecting groups. That is, upon activation in the gas phase, the protecting groups undergo a rearrangement that leads to the loss of the protecting group and the complete regeneration of the original functional group that was being protected. Competition between the initial covalent bond formation and deprotection of functional groups has never been observed; therefore, the highly favorable amide bond formation will occur before rearrangement and loss of protecting groups.

The main reactive group to be protected is the amine functionality. This group is found on essentially every amino acid, peptide, and protein. In both the condensed and gas phases, this functionality is highly nucleophilic and can initiate a nucleophilic attack on electropositive carbonyl carbon atoms of activated carboxyl groups. The protecting group boc is covalently bound to the nitrogen of a primary or secondary amine. In solution, boc is removed by addition of acids; in the gas phase, however, this is labile in the positive polarity with protonated peptides. The pervasiveness of the amine functional group in amino acids and peptides requires a constant presence of the gas phase–labile boc protecting groups among these reagents. Activation of the boc-protected amines may result in two fragmentation paths, both leading to the loss of the boc protecting group and the regeneration of the unprotected amine, illustrated in Scheme S1. The first fragmentation path involves the total loss of boc via the loss of isobutylen and carbon dioxide, leading directly to the deprotection of the amine. The second and less frequently observed fragmentation path involves the initial loss of isobutylene to form carbamic acid. Activation of the highly unstable carbamic acid results in loss of carbon dioxide and the regeneration of the unprotected amine, as illustrated in Scheme S1. Fig. S2 illustrates the gas phase deprotection of boc. The less reactive sulfhydryl group of cysteine may need to be protected under certain conditions, especially when it is desirable to avoid disulfide bond formation in the condensed phase. The protecting group used for sulphydrys that has been found to be gas phase–labile is the acetamidomethyl (acm) group. Upon activation of the acm-containing peptide, the protecting group is lost and the sulfhydryl is regenerated. Fig. S3 illustrates the gas phase deprotection of both boc and acm. The isolated protected dipeptide [boc–C(acm)K + H]⁺ (Fig. S3.4) is subjected to collision-induced dissociation (CID) to produce a boc loss, an acm loss, and a loss of both (Fig. S3B). CID of the boc loss produces the unprotected dipeptide [CK + H]⁺ (Fig. S3C) following acm loss.

The next common nucleophilic functional group found in amino acids and peptides is the hydroxyl, whether from serine, threonine, or the more acidic phenol of tyrosine. In the condensed phase and depending on the solvent, this functional group may be nucleophilic enough to react with activated carboxyl groups. A gas phase–labile protecting group is therefore used to address this reactivity. The protecting group used is a tBu covalently bound to the oxygen of the hydroxyl. In a similar fashion to boc, the tert-butyloxycarbonyl group undergoes a rearrangement upon activation and is lost while the hydroxyl group is simultaneously regenerated. Fig. S4 illustrates the gas phase deprotection of both boc and tBu. The isolated protected dipeptide [boc–S(OrBu)K + H]⁺ (Fig. S4.4) is subjected to CID to produce a boc loss, a tBu loss, and the loss of both (Fig. S4B). CID of the tBu loss produces both the unprotected dipeptide [SK + H]⁺ and a large loss of CO₂ (Fig. S4C), with the latter lost as the dissociation of carbamic acid from the boc group (vide supra). CID of the boc loss produces...
a loss of \(tBu\) to produce the unprotected dipeptide \([SK + H]^+\) (Fig. S4D).

Like amines, carboxylic acids are found in nearly all unmodified amino acids, peptides, and proteins. The side chains of aspartic and glutamic acids contain carboxylic acids that must be protected to specifically activate the C-terminal carboxyl group. Furthermore, in the gas phase, activation of peptides and proteins containing these acidic residues tends to show favorable cleavages on the C-terminal amide bond under certain conditions. It is therefore necessary to protect these functional groups. In the same manner that hydroxyls are protected, a \(tBu\) is used. Upon activation of the \(tBu\)-containing peptide in the gas phase, the tert-butyl group undergoes a rearrangement and is lost while the carboxylic acid is simultaneously regenerated. Fig. S5 illustrates the gas phase deprotection of both \(tBu\) and \(tBu\) in a nearly identical manner as Fig. S4. The isolated protected dipeptide \([boc-D(OrBu)K + H]^+\) (Fig. S5A) is subjected to CID to produce a \(tBu\) loss, a \(boc\) loss, and the loss of both (Fig. S5B). CID of the \(tBu\) loss produces both the unprotected dipeptide \([DK + H]^+\) and a small loss of \(CO_2\) (Fig. S5C), with the latter lost as the dissociation of carbamic acid from the \(boc\) group (vide supra). CID of the \(boc\) loss produces a loss of \(tBu\) to produce the unprotected dipeptide \([DK + H]^+\) (Fig. S5D).

Most functional groups can be protected using covalently bound protecting groups, as has been discussed for amines, hydroxyls, sulfhydryls, and carboxylic acids. However, the highly basic guanidine functional group of arginine introduces a unique challenge. We are unaware of a protecting group on the guanidine functionality that, upon activation, undergoes a rearrangement to lose the protecting group while simultaneously regenerating the unprotected group. Although a recently published study investigating the fragmentation behavior of gas phase–modified arginine has introduced the possibility of a functional group that can be lost to regenerate unmodified arginine (2), this same process is not accessible in the reagent preparation phase. An electrostatically bound protecting group was therefore explored for guanidinium functionalities. It is well known that the interaction between a protonated guanidinium (guanidinium) of arginine and a deprotonated sulfonate is exceedingly stable (3, 4), both in solution and in the gas phase. Therefore, BDSA was used to generate a strongly bound noncovalent interaction with the guanidinium of an arginine NHS ester in solution such that two sulfonates are in the presence of a single guanidinium, resulting in an overall negative charge. This noncovalent interaction is capable of surviving ionization using negative nESI to produce [(boc–R–NHS) + BDSA – H]. This reagent behaves similarly to sulfoNHS reagents with covalently bound protecting groups in that the first process upon activation involves covalent modification with subsequent activation steps resulting in the loss of the noncovalently bound protecting group. The main consequence to this method is that the guanidinium functionality will be neutral, rather than protonated, and is therefore highly fragile. The deprotection of a BDSA-protected guanidinium is illustrated in Fig. S6. The reagent anion [(boc–R(BDSA)–NHS – H)]反应 with the anchor peptide \([AKAAAKA + 2H]^+\) to form an electrostatic complex [(boc–R(BDSA)–NHS) + AKAAAKA + H]^+ (Fig. S6E). CID of the electrostatic complex results in both proton transfer and covalent modification, observed as the loss of NHS, to produce [boc–R(BDSA)AKAAAKA + H]^+ (Fig. S6B). CID of the NHS loss produces mainly the loss of BDSA, with a much smaller loss of \(boc\). The product of the BDSA loss is the boc-protected peptide \([boc–RAKAAAKA + H]^+\) (Fig. S6C). CID of the boc-protected peptide induces the deprotection of boc, resulting in the formation of the unmodified peptide \([RAKAAAKA + H]^+\) (Fig. S6D).

Fig. S2. Gas phase deprotection of the boc-protected N-terminal amine of alanine. Activation of (A) the isolated dipeptide [boc–AK + H]^+ results in (B) deprotection of boc to regenerate the amine functionality resulting in the formation of the unprotected dipeptide [AK + H]^+.

Fig. S3. Gas phase deprotection of the acm-protected sulfhydryl and boc of the N-terminal amine of cysteine. Activation of (A) the isolated dipeptide [boc–C (acm)K + H]^+ results in (B) deprotection of both protecting groups to regenerate the sulfhydryl and amine functionalities, with the loss of boc being more labile. (C) Activation of the boc loss results in deprotection of the sulfhydryl, resulting in the formation of the unprotected dipeptide [CK + H]^+.
Fig. S4. Gas phase deprotection of the tert-butyl–protected hydroxyl and boc of the N-terminal amine of serine. Activation of (A) the isolated dipeptide [boc-S((OtBu)K + H)]\(^{+}\) results in (B) deprotection of both protecting groups to regenerate the hydroxyl and amine functionalities. (C) Activation of the tert-butyl loss induces two separate fragmentation pathways. One pathway results in deprotection of the amine observed as a boc loss, resulting in the formation of the unprotected dipeptide [SK + H]\(^{+}\). The second pathway results in a loss of 44 Da, corresponding to the mass of CO\(_2\). This likely arises from the loss of the tert-butyl group of boc, resulting in the formation of a highly unstable carbamic acid. Subsequent activation of this peak results in the loss of the remaining tert-butyl moiety. (D) Activation of the boc loss results in the deprotection of the hydroxyl, observed as a tert-butyl loss, resulting in the formation of the unprotected dipeptide [SK + H]\(^{+}\).

Fig. S5. Gas phase deprotection of the tert-butyl–protected carboxylic acid and boc-protected N-terminal amine of aspartic acid. Activation of (A) the isolated dipeptide [boc-D(O-tBu)K + H]\(^{+}\) results in (B) deprotection of both protecting groups to regenerate the carboxylic acid and amine functionalities. (C) Activation of the tert-butyl loss results in the deprotection of the amine observed as a boc loss, resulting in the formation of the unprotected dipeptide [DK + H]\(^{+}\). Likewise, (D) activation of the boc loss results in deprotection of the carboxylic acid observed as a tert-butyl loss, which also results in the formation of the unprotected dipeptide [DK + H]\(^{+}\).
Fig. S6. Gas phase addition and deprotection of the boc- and BDSA-protected reagent boc–R(BDSA)NHS to the anchor peptide [AKAAAKA + 2H]^{2+}. Activation of (A) the isolated electrostatic complex [(boc–R(BDSA)NHS) + AKAAAKA + H]^+ results in (B) both covalent bond formation, observed as a NHS loss, and proton transfer. (C) Activation of the sulfoNHS loss results mainly in loss of BDSA, resulting in the formation of the protected peptide [boc–RAKAAKA + H]^+. (D) Activation of the boc-protected peptide results mainly in the loss of boc, resulting in the formation of the deprotected peptide [RAKAAKA + H]^+.

Scheme S1. Possible gas phase deprotection pathways of N-boc in the presence of O-tBu. Upon activation, the full loss of boc (left) lost as isobutylene and carbon dioxide is commonly observed. In the presence of tBu-protected oxygen atoms, such as from hydroxyl or carboxyl groups, loss of isobutylene may occur in a single event to form carbamic acid. Subsequent activation results in loss of carbon dioxide to deprotect the amine.
Table S1. Reagent amino acids or peptides added to anchor peptides and the reagents used for successful coupling

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Activated precursor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine</td>
<td>Boc-Ala–sulfoNHS</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>Boc-Phe–sulfoNHS</td>
</tr>
<tr>
<td>Lysine</td>
<td>Boc-Lys(Boc)–sulfoNHS</td>
</tr>
<tr>
<td>Histidine</td>
<td>Boc-His(Boc)–sulfoNHS</td>
</tr>
<tr>
<td>Cysteine</td>
<td>Boc-Cys(Acm)–sulfoNHS</td>
</tr>
<tr>
<td>Proline</td>
<td>Boc-ProsulfoNHS</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>Boc-Asp(tBu)–sulfoNHS</td>
</tr>
<tr>
<td>Asparagine</td>
<td>Boc-Asn–sulfoNHS</td>
</tr>
<tr>
<td>Arginine</td>
<td>Boc-Arg(BDSA)–NHS</td>
</tr>
<tr>
<td>Diglycine</td>
<td>Boc-Gly–Gly–sulfoNHS</td>
</tr>
<tr>
<td>Triglycine</td>
<td>Boc-Gly–Gly–Gly–sulfoNHS</td>
</tr>
</tbody>
</table>