Available online at www.sciencedirect.com
SCIENCE DIRECT®
International Journal of Mass Spectrometry 231 (2004) 179–187
International Journal of Mass Spectrometry
www.elsevier.com/locate/ijms

Gas phase derivatization in peptide analysis I: the utility of trimethyl borate in identifying phosphorylation sites

Scott Gronert*, Renee Huang, Kathy H. Li

Department of Chemistry and Biochemistry, San Francisco State University, 1600 Holloway Avenue, San Francisco, CA 94132, USA

Received 23 June 2003; accepted 11 October 2003

Dedicated to Professor Jean Claude Tabet on the occasion of his 60th birthday

Abstract

Phosphopeptides can readily lose H_3PO_4 during collision-activated dissociation (CAD) and this process can complicate efforts to use mass spectrometry to identify phosphorylation sites in peptides. As a part of a larger project aimed at surveying the gas phase reactivity of peptide ions, protonated phosphopeptides were derivatized in the gas phase by trimethyl borate (TMB) and subjected to CAD. During the reaction and subsequent CAD, three CH_3OH molecules are lost to give a peptide derivative that contains boron $[M + B - 2H]^+$. This ion undergoes backbone cleavages when subjected to CAD. In simple phosphopeptides, AGGsG, GsGGV, AKsF, and AKtF (lower case indicates phosphorylated residue), the boron derivatization enhances the yields of phosphate-containing sequence ions. In accord with previous work, it appears that the boron can bind to the phosphate as well as nearby nucleophilic sites and inhibit phosphate cleavage. The method was also applied to hexapeptides containing two potential phosphorylation sites (i.e., two serines), VSGAsA and LsGASA. In the former, boron derivatization leads to phosphate-containing sequence ions with greater relative intensities and provides ample information for identifying the phosphorylation site. The effect is less dramatic with LsGASA presumably because there is a greater amount of non-selective derivatization (i.e., away from the phosphate). Overall, the results suggest that boron derivatization is a potentially promising new tool for identifying phosphorylation sites in peptides where more than one site is available.

Keywords: Gas phase derivatization; Peptide analysis; Trimethyl borate; Phosphopeptides; CAD

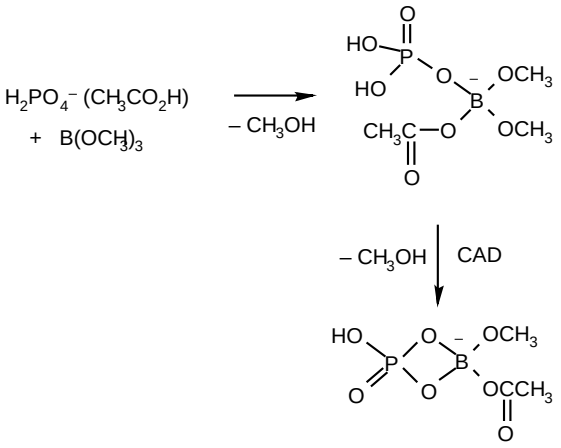
1. Introduction

Protein phosphorylation plays an important role in many biological processes [1–4] and there has been much interest in the development of mass spectrometric methods for identifying phosphorylation sites [5–24]. Because phosphorylation is a post-translational modification, the site cannot be inferred by analysis of the DNA coding of the protein. Instead, the protein itself must be probed. Often the protein can be cleaved by proteolytic enzymes to give peptides whose mass shifts (+80 Da) can be used to identify the phosphorylation sites. This assumes that each peptide contains only one residue that could potentially be phosphorylated. In cases where the phosphorylated peptide contains more than one of the typically phosphorylated residues (i.e., serine, threonine, or tyrosine), the situation is more complicated and the peptide must be probed by collision-activated dissociation (CAD) experiments. Unfortunately, the CAD analysis of peptides containing phosphorylated serines and threonines has been hampered by the fact that the phosphate group is relatively labile and is easily lost during the process in the form of H_3PO_4 [25,26]. Moreover, the product of phosphate loss can be indistinguishable from the product of H_2O loss from a non-phosphorylated serine or threonine residue [12,20] (alkene formation is usually assumed, but O'Hair and co-workers have shown that H_3PO_4 loss in small peptides prefers an internal cyclization pathway rather than an elimination pathway [27]). In any case, H_3PO_4 loss can greatly complicate the determination of peptide phosphorylation sites. More recently, Zubarev [19] and McLafferty [18] have shown that electron-capture dissociation (ECD) is less susceptible to H_3PO_4 loss than CAD.

To address the problem of H_3PO_4 loss during CAD, we have sought gas phase reactions that could be used to sta-

* Corresponding author. Tel.: +1-415-338-7709; fax: +1-415-338-2384.
E-mail address: sgronert@sfsu.edu (S. Gronert).

1387-3806/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijms.2003.10.006



Scheme 1.

bilize the phosphate linkage in peptides containing a phosphorylated serine or threonine, and therefore produce more phosphate-containing sequence ions. This would simplify the analysis of CAD spectra and lead to a more convenient determination of the phosphorylation site in peptides with more than one potential site. In 2001, we reported that trivalent boron species, such as borane and trimethyl borate (TMB), react readily in the gas phase with amino acids [28] and phosphate-containing species such as phosphopeptides [29]. In collaboration with the O'Hair group (University of Melbourne), we also showed that the boron reagents could cross-link multiple nucleophilic groups even in weakly bound complexes [30]. For example, when TMB is allowed to react with the hydrogen-bonded complex of dihydrogen phosphate and acetic acid, addition with loss of CH_3OH is observed. Even during the first CAD cycle, CH_3OH loss is observed rather than expulsion of either of the partners in the initially weakly bound complex (Scheme 1). In addition, CAD of the reaction product of TMB with deprotonated phosphoserine led to only a modest yield of dephosphorylation products, $H_2PO_4^-$ and PO_3^-. More recently, Leary and co-workers have reported several instances where boron derivatization provides useful modifications to the fragmentation behavior of phosphorylated saccharides [31,32].

The results from our lab suggested that boron reagents might be able to stabilize phosphate linkages in phosphorylated peptides by cross-linking the phosphate to other, nearby nucleophilic groups on the peptide. As a result, a greater number of phosphate-containing sequence ions could be realized during CAD analysis, and therefore it would be more straightforward to determine the site of phosphorylation in peptide where the sequence alone did not identify the phosphorylation site. The approach is somewhat akin to methods based on derivatizing the phosphopeptide in solution before mass spectrometric analysis [14–16].

To test this hypothesis, we initially have focused on model systems where the peptide is small and contains a single phosphorylated residue. As a derivatization reagent, we have used TMB. By comparison with the CAD spectra of the underivatized peptide, we can assess the efficacy of the borate in generating a greater number and relative intensity of phosphate-containing sequence ions. In addition, we have included in the study peptides with two serines, one that is phosphorylated and one that is not. This provides a direct test of the method in a system where the phosphorylation site would not be evident from the amino acid sequence.

2. Methods

All experiments were completed in a modified Finnigan-LCQ quadrupole ion trap with an electrospray ionization (ESI) source. The instrument is equipped with an external gas manifold that allows reagents to be mixed into the helium buffer gas of the ion trap [33–35]. Liquid reagents are introduced into a measured, fast flow of helium by a syringe pump. Although no rates were measured in the present study, the apparatus is suitable for kinetic studies [35]. We developed this approach a number of years ago and it provides a convenient method for studying gas phase ion molecule reactions at near thermal energies [33].

The peptides were obtained from United Biochemical Technologies (Seattle, WA, USA) and used without further purification. Typical conditions were used in the ESI source to generate the peptide ions. For the derivatization experiments, the external manifold was used to spike the helium buffer gas with the derivatization reagent. The peptide ion was isolated with a notched waveform and then held in the trap for a preset time delay to allow for reactions with the derivatization reagent. By varying the time delay and the derivatization reagent's partial pressure, the extent of reaction is conveniently controlled. The appropriate product ion was then isolated by applying a notched waveform and was subjected to CAD. This process was repeated until sequence ions were generated. Because the derivatization reagent is in the trap throughout the experiment, secondary reaction product ions are formed, but this problem can be minimized by careful choice of the time delays (typically 2–3 s for reactions and 30 ms for subsequent CAD activation) and reagent pressure. CAD spectra in the absence of the derivatization reagent were generated in the normal way.

TMB as well as the ESI solvents were obtained from commercial sources and used without further purification.

3. Results

3.1. AGGsG

This peptide contains only simple amino acids with no side chain functional groups except for a single phosphorylated serine. Fig. 1a shows the CAD spectrum of protonated AGGsG (lower case 's' indicating phosphorylated serine). The parent ion is at $m/z = 428$ and the spectrum is dominated by a peak at $m/z = 330$, which corresponds to the loss

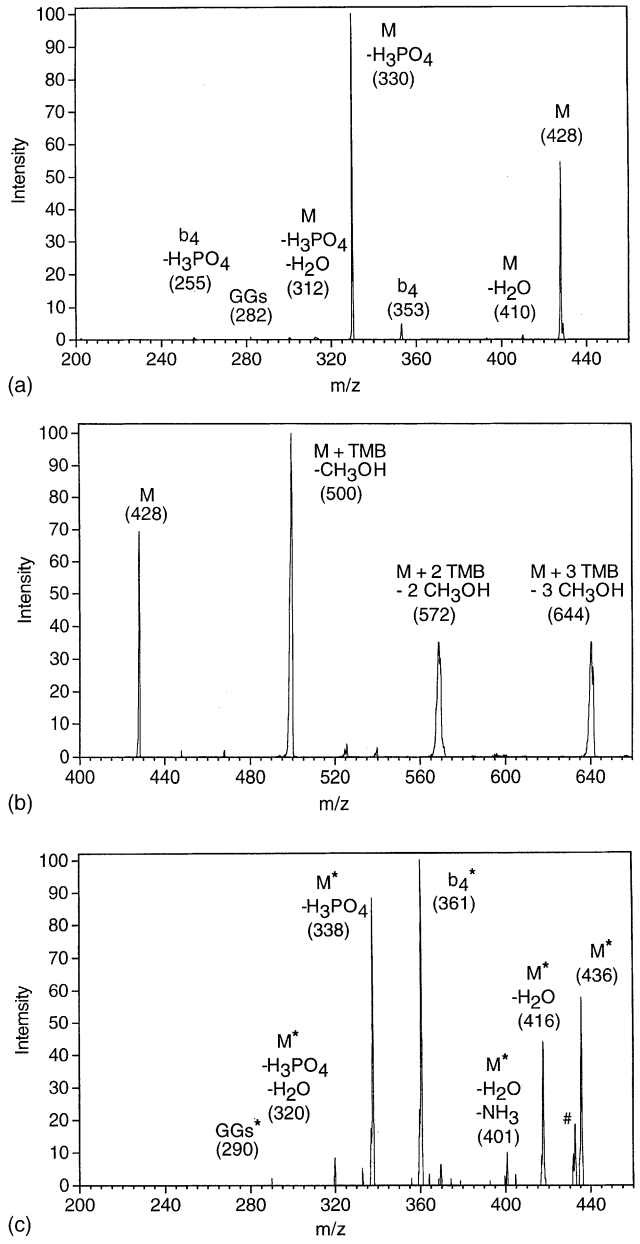
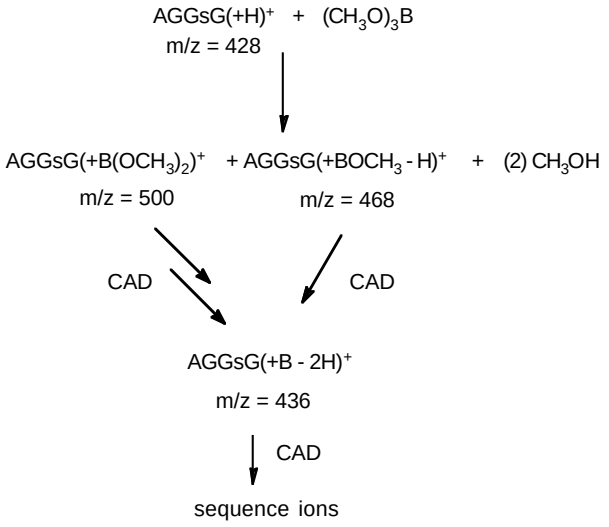


Fig. 1. (a) CAD spectrum of protonated AGGsG. Parent ion at $m/z = 428$. (b) Spectrum from reaction of protonated AGGsG with TMB. Parent ion at $m/z = 428$. Small mass shifts seen with multiple addition products are a mass scanning artifact that occasionally occurs with adducts in the ion trap. (c) CAD spectrum of $AGGsG(+B - 2H)^+$ ion ($m/z = 436$). The signal labeled with a symbol (#) is a secondary reaction product, $b_4 + TMB - 2CH_3OH$. Sequence ions are labeled, but some ions resulting from the loss of small molecules (i.e., H_2O) are not.

of H_3PO_4. The only significant sequence ion is at $m/z = 353$ (b_4). Very small signals occur at $m/z = 410$ (loss of H_2O), $m/z = 312$ (loss of H_3PO_4 and H_2O), $m/z = 282$ (GGs), and $m/z = 255$ ($b_4 - H_3PO_4$).

When protonated AGGsG is allowed to react with TMB, the major product is addition with the loss of CH_3OH (Fig. 1b). In addition, signals are seen for multiple additions of TMB. As in our earlier work [30], CAD of the



Scheme 2.

addition products mainly leads to the sequential loss of all the boron ligands as CH_3OH molecules (Scheme 2) rather than peptide cleavages. Applying CAD conditions to the reaction product ion, $AGGsG(+B(OCH_3)_2)^+$, eventually leads to $AGGsG(+B - 2H)^+$. Therefore, the derivatization with TMB effectively adds 8 Da to the ion as a result of incorporating boron (isotope 11) and losing three hydrogens from the peptide to form the three CH_3OH molecules that are expelled. As shown previously, these hydrogens are likely to come from relatively acidic sites in the peptide [28]. The "+8" ions will be the key players in this study because CAD of them leads to sequence ions. All of the protonated peptides in the study follow the general reaction/CAD pattern outlined in Scheme 2.

Fig. 1c shows the CAD spectrum of the derivatized peptide ion, $AGGsG(+B-2H)^+$. The effect of the TMB derivatization is dramatic and the dominant ion now is a b_4 species containing the boron ($m/z = 361$). Throughout the text, boron-derivatized ions will be indicated by an asterisk (e.g., b_4^*). Loss of H_2O ($m/z = 418$) and H_3PO_4 ($m/z = 338$) are also large peaks in the spectrum. Smaller signals are seen at $m/z = 401$ (loss of H_2O and NH_3) and $m/z = 320$ (loss of H_3PO_4 and H_2O). Fig. 1a and c have no peaks in common so it is clear that the boron generally is being retained during the fragmentation process and is greatly affecting the pathways. More importantly, it is inhibiting the loss of phosphate during CAD.

3.2. GsGGV

This peptide is much like the previous one except that the phosphorylated serine is near the N-terminus rather than the C-terminus. The CAD spectrum of protonated GsGGV ($m/z = 456$) is very simple and the only strong peak is the result of H_3PO_4 loss (Fig. 2a). No significant phosphate-containing sequence ions are seen in the spec-

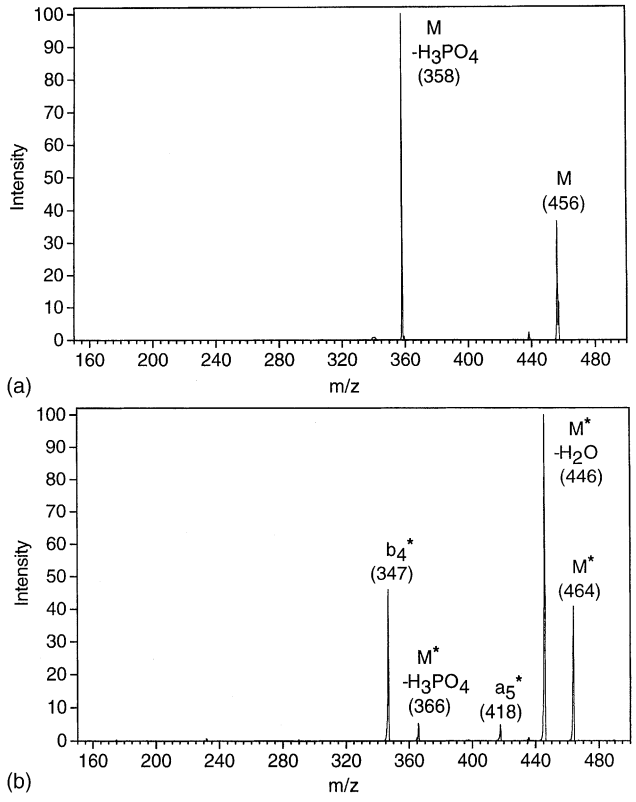


Fig. 2. (a) CAD spectrum of protonated GsGGV. Parent ion at $m/z = 456$. (b) CAD spectrum of GsGGV$(+B-2H)^+$ ion ($m/z = 464$). Sequence ions are labeled, but some ions resulting from the loss of small molecules (i.e., H_2O) are not.

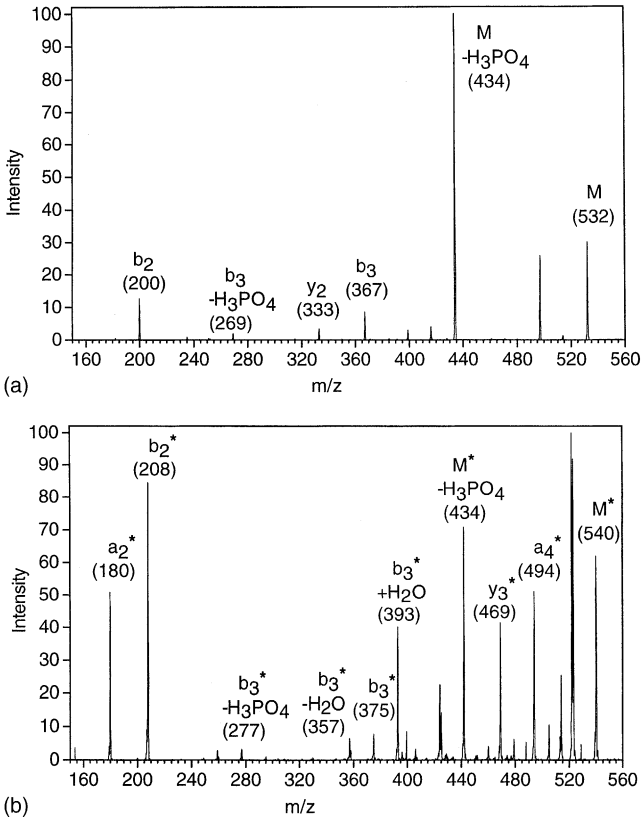


Fig. 3. (a) CAD spectrum of protonated AKsF. Parent ion at $m/z = 532$. (b) CAD spectrum of AKsF$(+B-2H)^+$ ion ($m/z = 540$). Sequence ions are labeled, but some ions resulting from the loss of small molecules (i.e., H_2O) are not.

trum! When treated with TMB, addition with the loss of two CH_3OH molecules dominates, and this product yields GsGGV$(+B-2H)^+$ during CAD. When subjected to CAD conditions, the boron-derivatized peptide ($m/z = 464$) also produces a relatively simple spectrum (Fig. 2b), but H_3PO_4 loss is almost completely inhibited. Instead the dominant products are at $m/z = 446$ (loss of H_2O) and $m/z = 347$ (b_4^*). A small signal is seen for an a_5^* ion at $m/z = 418$ (i.e., loss of H_2O and CO). The b_4^* ion is nearly 10 times more intense in this spectrum than the ion corresponding to H_3PO_4 loss. Clearly, boron-derivatization can have a dramatic impact on the stability of the phosphate linkage in a phosphopeptide.

3.3. AKsF

In the previous peptides, all of the amino acids, with the exception of the phosphorylated serine, had alkyl side chains. In this peptide, a much more reactive residue, lysine, has been incorporated to test the effect of added nucleophilic groups (i.e., the amino side chain of the lysine) on the ability of the boron to inhibit phosphate cleavage. Fig. 3a shows the CAD spectrum of protonated AKsF. The dominant ion in the spectrum is $m/z = 434$, which corresponds to the loss of H_3PO_4. The other major fragment is $m/z = 497$, which corresponds to a loss of H_2O and NH_3. Smaller signals are seen at $m/z = 514$ (loss of H_2O), $m/z = 416$ (loss of H_3PO_4 and H_2O), and $m/z = 399$ (loss of H_3PO_4, H_2O, and NH_3). The only significant sequence ions appear at $m/z = 367$ (b_3), $m/z = 333$ (y_2), $m/z = 269$ ($b_3-H_3PO_4$), and $m/z = 200$ (b_2). Of these, only b_3 and y_2 correspond to phosphate-containing ions and their intensities are only about 5–10% relative to the loss of H_3PO_4 ($m/z = 434$).

When treated with TMB, addition occurs and up to three CH_3OH molecules are lost (Eq. (1)). We have noticed that when the peptide contains many nucleophilic groups, it is possible to lose all the boron ligands during the initial reaction. As before, ions that have not lost all three CH_3OH molecules will undergo CH_3OH expulsion during CAD to eventually give the "+8" ion.

$$\underset{m/z=532}{\mathrm{AKsF}(+H)^+} + (CH_3O)_3B \rightarrow \underset{m/z=540}{\mathrm{AKsF}(+B-2H)^+} + 3CH_3OH \quad (1)$$

When the ion at $m/z = 540$ is isolated and subjected to CAD, the spectrum shown in Fig. 3b results. The spectrum is much richer than with the underivatized peptide and there are many peaks with similar intensities. The dominant peaks are $m/z = 522$ and 523 (loss of H_2O and NH_3, respec-

tively), but $m/z = 442$ (loss of H_3PO_4) is also strong. The most interesting products in the spectrum are the sequence ions. At $m/z = 393$, there is a relatively strong signal for a ($b_3^* + H_2O$) ion that retains the phosphate and the boron derivatization (this type of C-terminal cleavage ion is often seen with metal cationized peptides) [36,37]. Large peaks also appear for a_4^* ($m/z = 494$) and y_3^* ($m/z = 469$) ions. Signals for derivatized b_3^* and $b_3^* - H_2O$ ions are seen at $m/z = 375$ and 357. At $m/z = 277$, a weak signal is seen for a $b_3^* + H_3PO_4$ ion. Finally, strong signals are also seen for derivatized b_2^* and a_2^* ions ($m/z = 208$ and 180, respectively). The latter two ions are interesting because they retain the boron despite losing the phosphate with the serine residue at position 3. Clearly, the reaction of the borate with a peptide is complicated and addition may occur at more than one site. Here, it is possible that the boron is interacting with the lysine at position 2 and directing the cleavage at that residue to give the b_2^* and a_2^* ions. In any case, the enhanced yield of phosphate-containing sequence ions again suggests that the borate generally helps to stabilize the phosphate linkage.

Spectra were also taken in the negative ion mode for this peptide. Under these circumstances, loss of H_3PO_4 was the exclusive fragmentation pathway from the underivatized peptide. TMB readily derivatized the deprotonated peptide, but during CAD, H_3PO_4 loss still dominates the spectrum although the boron is retained with the peptide (spectra not shown). The only significant difference in the fragmentation pattern is that H_2O loss is also seen in the spectrum as a minor ion. From these and other results, we have concluded that the lability of the phosphate group in deprotonated peptides will not allow the boron derivatization to significantly enhance phosphate retention.

3.4. AKtF

In this peptide, the only change is that the phosphate is borne by a threonine rather than a serine residue. The CAD spectrum of the protonated peptide is shown in Fig. 4a. The dominant fragment ion occurs at $m/z = 448$, which is the result of H_3PO_4 loss. Other prominent ions in the spectrum correspond to the loss of H_2O and NH_3 ($m/z = 511$), the loss of H_3PO_4 and H_2O ($m/z = 430$), a b_3 ion ($m/z = 381$), a y_2 ion ($m/z = 347$), a $b_3 - H_3PO_4$ ion ($m/z = 283$), a $y_2 - H_3PO_4$ ion ($m/z = 249$), and a b_2 ion ($m/z = 200$). Of the sequence ions, only the b_3 and y_2 ions retain the H_3PO_4. Fig. 4b shows the CAD spectrum of the boron-derivatized peptide. At $m/z = 554$ is the parent ion resulting from the addition of TMB followed by the loss of three CH_3OH molecules. The dominant peaks in the spectrum correspond to loss of NH_3 ($m/z = 537$), H_2O ($m/z = 536$) and H_3PO_4 ($m/z = 456$). There are several significant sequence ions in the spectrum, including a_4^* ($m/z = 508$), y_3^* ($m/z = 483$), $b_3^* + H_2O$ ($m/z = 407$), b_3^* ($m/z = 389$), $b_3^* - H_2O$ ($m/z = 371$), $b_3^* - H_3PO_4$ ($m/z = 291$), b_2^* ($m/z = 208$), and a_2^* ($m/z = 180$). Overall, the two derivatized peptides, AKsF and AKtF, give similar fragmentation patterns with almost equivalent information content. Moreover, it is clear that the derivatization can lead to additional phosphate-containing sequence ions from the peptides.

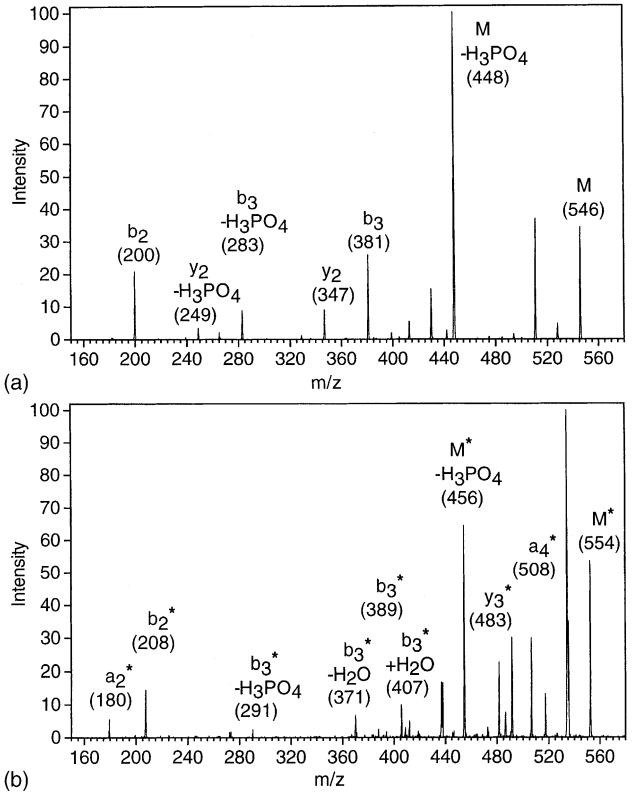


Fig. 4. (a) CAD spectrum of protonated AKtF. Parent ion at $m/z = 546$. (b) CAD spectrum of AKtF$(+B - 2H)^+$ ion ($m/z = 554$). Sequence ions are labeled, but some ions resulting from the loss of small molecules (i.e., H_2O) are not.

3.5. VSGAsA

VSGAsA provides a more challenging peptide because it contains two potential sites for phosphorylation, the two serines. In this case, the serine near the C-terminus is phosphorylated. The CAD spectrum of the protonated peptide is shown in Fig. 5a. Loss of H_3PO_4 dominates this spectrum. The significant sequence ions in the spectrum are b_5 ($m/z = 482$), $b_5 - H_3PO_4$ ($m/z = 384$), y_4 ($m/z = 385$), b_4 ($m/z = 315$), and $y_4 - H_3PO_4$ ($m/z = 287$). Of these, the b_5 and y_4 ions retain the phosphate, but are weak signals compared to H_3PO_4 loss. Their intensities are only 3 and 7%, respectively, relative to the peak corresponding to H_3PO_4 loss. The critical peak is the y_4 ion because it retains the phosphate, but does not include the other serine (residue 2). Consequently this ion could be used as definitive proof that the phosphate is located on the serine at residue 5.

The spectrum resulting from the fragmentation of the derivatized peptide ($+TMB - 3CH_3OH$) is shown in Fig. 5b.

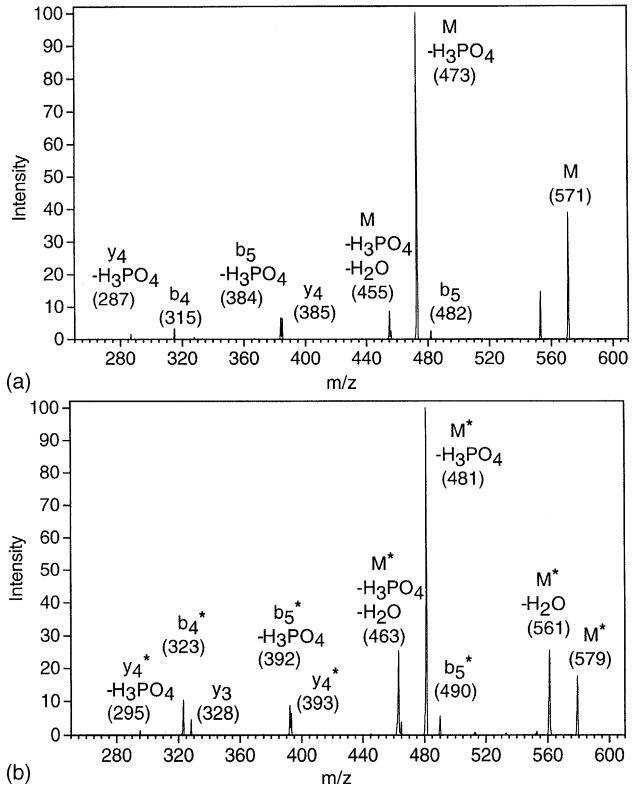


Fig. 5. (a) CAD spectrum of protonated VSGAsA. Parent ion at $m/z = 571$. (b) CAD spectrum of VSGAsA$(+B - 2H)^+$ ion ($m/z = 579$). Sequence ions are labeled, but some ions resulting from the loss of small molecules (i.e., H_2O) are not.

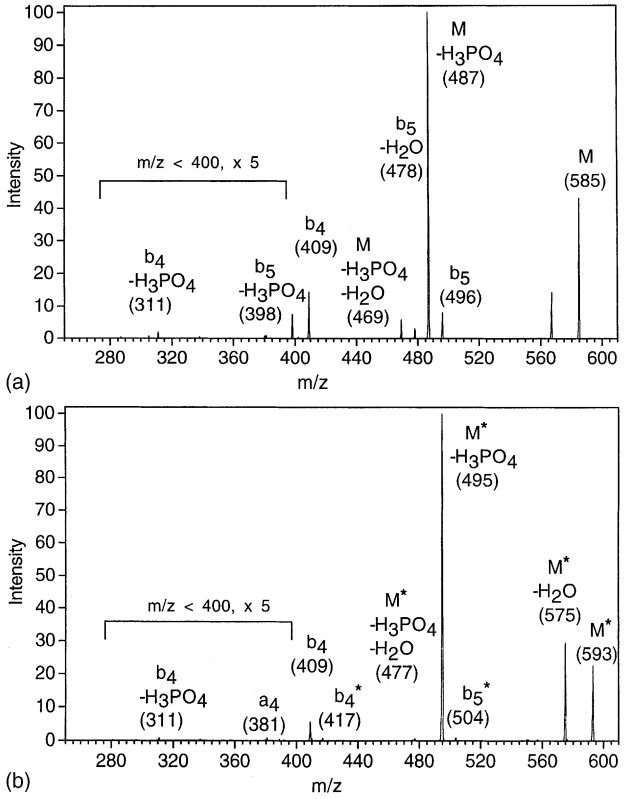


Fig. 6. (a) CAD spectrum of protonated LsGASA. Parent ion at $m/z = 585$. (b) CAD spectrum of LsGASA$(+B - 2H)^+$ ion ($m/z = 593$). Sequence ions are labeled, but some ions resulting from the loss of small molecules (i.e., H_2O) are not.

The base peak is for loss of H_3PO_4 ($m/z = 481$) and significant peaks also are seen for H_2O loss ($m/z = 561$) as well as H_3PO_4 combined with H_2O loss ($m/z = 463$). The major derivatized sequence ions are b_5^* ($m/z = 490$), y_4^* ($m/z = 393$), $b_5^* - H_3PO_4$ ($m/z = 392$), b_4^* ($m/z = 323$), and $y_4^* - H_3PO_4$ ($m/z = 295$). Each of these signals is moderately strong and has an intensity that is about 5–10% relative to the base peak (H_3PO_4 loss). As noted above, the key ions are those that retain the phosphate, but have lost the non-phosphorylated serine. For the derivatized peptide, y_4^* fits these criteria. In addition, a y_3 ion ($m/z = 328$) also appears in the spectrum and provides evidence of the phosphorylation site despite having lost the boron derivatization with the neutral fragment.

3.6. LsGASA

This hexapeptide has nearly the same structure as the previous one, but the site of phosphorylation has been shifted from residue 5 to residue 2. In this case, the most valuable sequence ions will be b_n ions containing the phosphorylated serine. The CAD spectrum of the protonated peptide is shown in Fig. 6a. The parent is at $m/z = 585$ and again the major fragment results from loss of H_3PO_4 ($m/z = 487$). A large signal is also seen for loss of H_2O ($m/z = 567$). The most intense sequence ions are b_5 ($m/z = 496$), $b_5 - H_2O$ ($m/z = 478$), b_4 ($m/z = 409$), $b_5 - H_3PO_4$ ($m/z = 398$), and $b_4 - H_3PO_4$ ($m/z = 311$). Only one of these ions definitively identifies the phosphorylation site, b_4, but it is relatively intense in this spectrum.

The spectrum resulting from CAD of the TMB-derivatized peptide is shown in Fig. 6b. Here, the parent ion (LsGASA + TMB $- 3CH_3OH + H^+$) is at $m/z = 593$. Strong signals are seen in the spectrum for the loss of H_2O ($m/z = 575$) and H_3PO_4 ($m/z = 495$). The major sequence ions are b_5^* ($m/z = 504$), b_4^* ($m/z = 417$), b_4 ($m/z = 409$), a_4 ($m/z = 381$), and $b_4 - H_3PO_4$ ($m/z = 311$). The latter two signals are very weak. Like the spectrum of the underivatized peptide, the b_4 ion provides sufficient information to determine the site of phosphorylation. The b_4^* ion also points out the phosphorylation site, but is very weak. Here, derivatization offers little advantage because clear evidence of the phosphorylation site was apparent in the absence of derivatization. Moreover, the phosphorylated sequence ions appear to be relatively weaker in the CAD spectrum of the derivatized peptide. Because the derivatization process significantly reduces the overall ion signal (unwanted ion production and losses from multiple CAD cycles), it has a negative impact if the CAD spectrum does not contain additional information.

4. Discussion

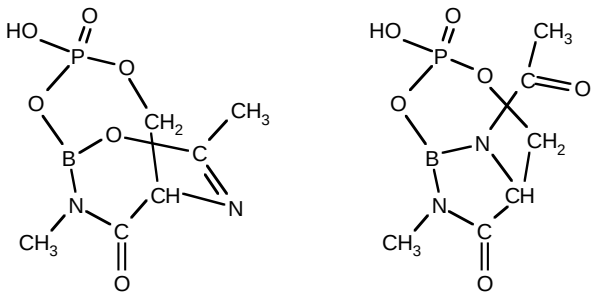
It is clear that the boron derivatization can have an impact on the fragmentation patterns of the phosphorylated peptides, and across the series of six peptides, the boron derivatization has a variety of effects on the CAD spectra. In the simplest of the peptides, AGGsG and GsGGV, the impact is very impressive and greatly enhances the yield of phosphate-containing sequence ions, specifically b_4 ions. Here, the boron derivatization does not alter the fragmentation pathways (i.e., the same sorts of ions generally are seen), but phosphate loss is inhibited. With the lysine-containing species, AKsF and AKtF, phosphate loss is inhibited to a large extent, and richer spectra result from the boron-derivatized ions. In addition, there is evidence of new pathways (i.e., the a_2^* ions) and the retention of boron on the non-phosphorylated portion of the peptide. In the most complicated peptides, VSGAsA and LsGASA, the impact of the boron is more limited. In the former, there is an enhancement in the relative intensities of the phosphate-containing ions (about twofold), but H_3PO_4 loss is the dominant pathway. In the latter, the derivatization has almost no impact on the CAD spectrum of the peptide. To better understand the effect of boron derivatization, three questions need to be addressed: (i) where does the boron add to the peptide chain, (ii) what is its mode of action, and (iii) why is it more effective in some cases rather than others.

4.1. Site of derivatization

From our previous work, we found that boron has a reasonably high affinity for phosphate groups and can have a bidentate linkage with the phosphate [30]. There seems to be a fairly delicate balance between the boron being tri- or tetracoordinate so the peptide must provide three to four nucleophiles to bind to the boron. Finally, the boron linkages are relatively labile so during CAD, extensive structural rearrangements are possible. Even in a small peptide, like AKsF, there are seven potential nucleophilic sites (C-terminus, N-terminus, three amide bonds, lysine side chain, and the phosphate). This leads to numerous, possible bonding schemes. Moreover, there is no reason to believe that they would retain their structural integrity during CAD. As a result, it is not possible to make any definitive statements about the coordination of the boron to the peptide. However, some broad generalizations can be ascertained from the CAD spectra. First, although the boron is retained in the great majority of the fragment ions, the phosphate is lost from 10 to 70% of the time, most often with the boron remaining on the peptide. This indicates that the boron does not always bind tightly to the phosphate. In the collision complex, it might simply encounter other nucleophiles on the peptide and react with them before the phosphate. For example, we have shown in the past that trivalent boron will react with the neutral amine group of an amino acid [28]. If the boron is localized on another part of the peptide, it will not have a significant effect of the stability of the phosphate linkage. The greater effectiveness of the boron in the simpler peptides offers some support for this analysis. As the size of the peptide grows and it incorporates more nucleophilic sites, the probability of non-selective binding to the peptide also grows. Of course, this has implications on the potential utility of the method and we are presently investigating size effects. Overall, it appears that a mixture of selective and non-selective binding is occurring in the reactions of these peptides with TMB.

4.2. Mode of action

As noted in Section 1, our hope was that the boron would link to the phosphate as well as other nucleophilic sites on the peptide and create a bridge to inhibit phosphate expulsion. To explore the possible bonding schemes, we have turned to molecular modeling on a simple system, *O*-phosphoserine capped on the N-terminus with an acetyl group and on the C-terminus with an *N*-methylamide. We have chosen these end groups to produce a compact species with two peptide bonds. In the modeling, it was assumed that the boron would interact with one or both of the peptide bonds as well as the phosphate. From previous computational work, we have found that the boron can form either a mono- or bidentate linkage with the phosphate so both were considered [30]. In addition, it was assumed that the boron could interact with the nitrogen or oxygen of a deprotonated amide. These options lead to eight distinct bonding schemes. The conformational space of each one was explored in a Monte Carlo search using the MMFF force field in MacSpartan [38]. The geometries of all of the resulting conformers were refined at the PM3 level in MacSpartan and the five best conformers of each bonding scheme were subjected to B3LYP/6-31 + G(d) single point calculations in Gaussian98 [39]. The best representative of each bonding scheme finally was optimized at the B3LYP/6-31 + G(d) level in Gaussian98. Although this approach is not guaranteed to identify the global minimum, it does serve to give an overview of the bonding possibilities and points out the more stable arrangements. The two most stable structures are shown in Fig. 7 and outlined in Scheme 3. In each case, there is monodentate coordination to the phosphate and connections to the deprotonated amide bonds are either through oxygen or nitrogen. This bonding



Scheme 3.

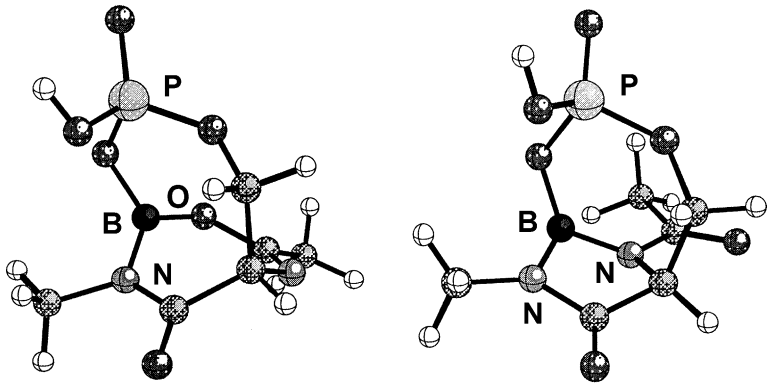


Fig. 7. B3LYP/6-31 + G(d) optimized structures of models for the boron derivatization product of the *N*-methylamide of *N*-acetyl-*O*-phosphoserine. See Scheme 3 for an outline of the bonding scheme.

leads to a bicyclic, cage-like structure. There is not a large variation in the energies of the different bonding schemes, and the structure on the left is about 2 kcal/mol more stable than the one on the right. However, the species with a monodentate boron/phosphate linkage tend to be about 10–20 kcal more stable than the ones with a bidentate linkage. It is easy to imagine how bonding schemes, like those shown in Fig. 7, would aid in phosphate retention. Instead of a single bond cleavage to release the phosphate, two bonds must fracture. Entropically, this is less favorable and it is likely that it also makes phosphate cleavage enthalpically less favorable. The fact that it is rare in the spectra to see the combined loss of the phosphate and the boron also suggests that if the boron binds to the phosphate, it is also making strong linkages to other sites so that it cannot be expelled simply by cleaving the phosphate bond. Of course, these model systems are very simple and do not allow for the possibility of boron bonding to a side chain functional group or a terminal functional group (amine or carboxylic acid), but they do provide a glimpse into the possible mode of action of the boron derivatization.

4.3. Structure/reactivity relationships

Given that only a few peptide sequences have been surveyed to date, it is not possible to make definitive statements about the scope of the method or the effect of side chain functional groups on the fragmentation patterns. The boron derivatization had the greatest effect on the two sequences that did not contain side chain nucleophiles, AGGsG and GsGGV. It was also effective in the compact, lysine-containing peptides, AKsF and AKtF. There are two factors that seem to be responsible for the greater efficacy in these peptides. First, the smaller size limits the number of nucleophilic sites in the peptide and makes it more probable that the phosphate is one of the ligands on the boron. Second, if the nucleophilic side chain is near the phosphate, it could coordinate to it through the boron and lead to stabilization of the phosphate linkage. This type of coordination could occur, but to form a linkage from the boron to the phosphate as well as the lysine's amino group would require a 14-membered ring, which is not favored entropically relative to the formation of smaller rings, such as those in Scheme 3. However, the unusually strong signals for boron-derivatized a_2 and b_2 ions in the spectra suggest that the boron may facilitate cleavage of backbone bonds on the C-terminal side of lysines. In this role, it is likely that the boron does not interact with the phosphate group because it is not lost with the phosphate-bearing residue. Boron reagents are known to act as Lewis acids and can catalyze nucleophilic cleavage of carboxylic acid derivatives, such as amides [40].

The hexapeptides with two serines gave less satisfactory results. In VSGAsA, the boron is somewhat effective in retaining the phosphate on the peptide fragments. However, it has no positive impact on the CAD spectra of LsGASA and some of the fragment ions do not contain the boron. One rationalization is that in this peptide, there are two good nucleophilic and acidic sites near the C-terminus (serine hydroxyl and the C-terminal carboxylic acid) and they compete with the phosphate in the derivatization process. This is supported by the fact that CAD of derivatized LsGASA leads to several a and b ions that do not contain the boron. Assuming much of the derivatization occurs away from the phosphate, it is not surprising that it is less effective in retaining the phosphate in this peptide. As noted above, this result does lead to some questions about the scope and generality of the approach, especially in larger peptides with many nucleophilic sites. We currently are investigating this issue with a greater variety of peptides.

5. Conclusions

TMB reacts readily with small protonated peptides to give addition products that sequentially lose methanol molecules in the reaction and during CAD. After all the boron ligands

have been displaced, backbone cleavages are observed and sequence ions are formed. In most cases, the boron derivatization inhibits phosphate cleavage and leads to greater relative intensities for the phosphate-containing sequence ions. As a result, treatment with TMB can aid in the analysis of phosphorylation sites in peptides. As a test, the boron derivatization methodology was applied to two small peptides, LsGASA and VSGAsA, that contain more than one possible phosphorylation site (i.e., two serines). In VSGAsA, the boron enhances the yield of phosphate-containing sequence ions and allows for a straightforward identification of the serine that is phosphorylated. In contrast, boron-derivatization has a limited impact on LsGASA and provides little additional information. However, identification of the phosphorylation site is also straightforward in this peptide because it produces an intense b_4 ion in the presence and absence of boron. These results suggest that boron derivatization is not wholly selective for phosphates, but could be very useful in some circumstances. Presently, we are testing the generality of the methodology with a wider range of peptides and exploring the utility of other trivalent boron reagents.

Acknowledgements

Support from the National Institutes of Health (NIH MBRS SCORE 5 SO6 GM52533) is gratefully acknowledged. S.G. thanks Dr. Richard A.J. O'Hair for helpful discussions and providing the inspiration for the project during a sabbatical visit by S.G. to the University of Melbourne.

References

[1] M.J. Hubbard, P. Cohen, Trends Biochem. Sci. 18 (1993) 172.
[2] T. Hunter, Cell 100 (2000) 113.
[3] F. Marks (Ed.), Protein Phosphorylation, VCH, Weinheim, 1996.
[4] T. Pawson, J.D. Scott, Science 278 (1997) 2075.
[5] J.W. Flora, D.C. Muddiman, J. Am. Chem. Soc. 124 (2002) 6546.
[6] R. Ruijtenbeek, C. Versluis, A.J.R. Heck, F.A.M. Redegeld, F.P. Nijkamp, R.M.J. Liskamp, J. Mass Spectrom. 37 (2002) 47.
[7] S.C. Moyer, R.J. Cotter, A.S. Woods, J. Am. Soc. Mass Spectrom. 13 (2002) 274.
[8] K.L. Bennett, A. Stensballe, A.V. Podtelejnikov, M. Moniatte, O.N. Jensen, J. Mass Spectrom. 37 (2002) 179.
[9] G.H. Talbo, D. Suckau, M. Malkoski, E.C. Reynolds, Peptides 22 (2001) 1093.
[10] S. Ogueta, R. Rogado, A. Marina, F. Moreno, J.M. Redondo, J. Vazquez, J. Mass Spectrom. 35 (2000) 556.
[11] T. Kinumi, H. Niwa, H. Matsumoto, Anal. Biochem. 277 (2000) 177.
[12] A. Tholey, J. Reed, W.D. Lehmann, J. Mass Spectrom. 34 (1999) 117.
[13] P. Cao, J.T. Stults, J. Chromatogr. A 853 (1999) 225.
[14] Y. Oda, T. Nagasu, B.T. Chait, Nat. Biotechnol. 19 (2001) 379.
[15] H. Zhou, J.D. Watts, R. Aebersold, Nat. Biotechnol. 19 (2001) 375.
[16] M.B. Goshe, T.P. Conrads, E.A. Panisko, N.H. Angell, T.D. Veenstra, R.D. Smith, Anal. Chem. 73 (2001) 2578.
[17] A. Schlosser, R. Pipkorn, D. Bossemeyer, W.D. Lehmann, Anal. Chem. 73 (2001) 170.
[18] S.D.H. Shi, M.E. Hemling, S.A. Carr, D.M. Horn, I. Lindh, F.W. McLafferty, Anal. Chem. 73 (2001) 19.
[19] A. Stensballe, O.N. Jensen, J.V. Olsen, K.F. Haselmann, R.A. Zubarev, Rapid Commun. Mass Spectrom. 14 (2000) 1793.
[20] J.P. DeGnore, J. Qin, J. Am. Soc. Mass Spectrom. 9 (1998) 1175.
[21] X. Zhang, V.J. Herring, P.R. Romano, J. Szczepanowska, H. Brzeska, A.G. Hinnebusch, Q. J. Anal. Chem. 70 (1998) 2050.
[22] S.A. Carr, M.J. Huddleston, R.S. Annan, Anal. Biochem. 239 (1996) 180.
[23] M.J. Huddleston, R.S. Annan, M.F. Bean, S.A. Carr, J. Am. Soc. Mass Spectrom. 4 (1993) 710.
[24] R.S. Annan, S.A. Carr, Anal. Chem. 68 (1996) 3413.
[25] C. Dass, in: B.S. Larsen, C.N. McEwen (Eds.), Mass Spectrometry of Biological Materials, Dekker, New York, 1998, p. 247.
[26] D.B. Kassel, R.K. Blackborn, B. Antonsson, in: B.S. Larsen, C.N. McEwen (Eds.), Mass Spectrometry of Biological Materials, Dekker, New York, 1998, p. 137.
[27] G.E. Reid, R.J. Simpson, R.A.J. O'Hair, J. Am. Soc. Mass Spectrom. 11 (2000) 1047.
[28] S. Gronert, R. Huang, J. Am. Chem. Soc. 123 (2001) 8606.
[29] S. Gronert, R. Huang, R.A.J. O'Hair, in: 49th American Society for Mass Spectrometry Meeting, Chicago, IL, 2001.
[30] S. Gronert, R.A.J. O'Hair, J. Am. Soc. Mass Spectrom. 13 (2002) 1088.
[31] M.D. Leavell, G.H. Kruppa, J.A. Leary, Int. J. Mass Spectrom. 222 (2003) 135.
[32] M.D. Leavell, G.H. Kruppa, J.A. Leary, Anal. Chem. 74 (2002) 2608.
[33] S. Gronert, J. Am. Soc. Mass Spectrom. 9 (1998) 845.
[34] A.E. Flores, S. Gronert, J. Am. Chem. Soc. 121 (1999) 2627.
[35] S. Gronert, L.M. Pratt, S. Mogali, J. Am. Chem. Soc. 123 (2001) 3081.
[36] R.P. Grese, R.L. Cerny, M.L. Gross, J. Am. Chem. Soc. 111 (1989) 2835.
[37] W.Y. Feng, S. Gronert, K.A. Fletcher, A. Warres, C.B. Lebrilla, Int. J. Mass Spectrom. 222 (2003) 117.
[38] B.J. Deppmeier, A.J. Driessen, W.J. Hehre, J.A. Johnson, P.E. Kluzinger, L. Lou, J. Yu, Wavefunction Inc., Irvine, CA, 1996.
[39] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery, R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P.M.W. Gill, B.G. Johnson, W. Chen, M.W. Wong, J.L. Andres, M. Head-Gordon, E.S. Replogle, J.A. Pople, Revision A.7 ed., Gaussian, Inc., Pittsburgh, PA, 1998.
[40] S.E. Thomas, Organic Synthesis: The Roles of Boron and Silicon, Oxford, New York, 1991.