

Artifactual isoform profile modification following treatment of human plasma or serum with protease inhibitor, monitored by 2-dimensional electrophoresis and mass spectrometry

Mark D. Schuchard¹, Richard J. Mehigh¹, Steven L. Cockrill¹, George T. Lipscomb¹, Jonathan D. Stephan¹, Justin Wildsmith¹, Rafael Valdes-Camin¹, William K. Kappel¹, Alex J. Rai², and Graham B.I. Scott¹

BioTechniques 39:239-247 (August 2005)

The inclusion of protease inhibitors in serum or plasma samples has been found to significantly impact the isoform profile of selected plasma proteins as seen on 2-dimensional electrophoresis (2-DE) gels. With the addition of a protease inhibitor cocktail, several human plasma protein trains [depleted of albumin and immunoglobulin G (IgG)] exhibited higher isoelectric point (pI) isoforms. This shift was especially apparent for apolipoprotein A1 (apo A1), a relatively high abundance protein. The six protease inhibitor components of the cocktail were individually investigated with albumin and IgG depleted human plasma, and it was shown that the observed effects were caused by 4-(2-aminoethyl) benzenesulfonyl fluoride (AEBSF), a serine protease inhibitor that covalently modifies proteins and/or peptides. Several serine- and/or tyrosine-containing peptides of apo A1 were modified with a concomitant mass increase of 183 Da, which is consistent with the mass increase expected following reaction with AEBSF. These modifications were observed with increasing propensity in the higher pI spots. An increase in both the number and proportion of modified peptides with increasing pI was also observed. A model is proposed for the random or stochastic coupling of AEBSF-derived moieties to serine and/or tyrosine residues throughout apo A1 and potentially other plasma proteins.

INTRODUCTION

The utilization of 2-dimensional (2-D) gel electrophoresis followed by mass spectrometry represents a classical paradigm in proteomics-based investigations with far-reaching applications, including the identification of potential disease biomarkers. This is the subject of a great deal of academic, clinical, and pharmaceutical proteomic research, particularly for those involved in the development of targeted protein therapeutic strategies. For example, one of the goals of the Human Proteome Organization—Plasma Proteome Project (HUPO-PPP) is the discovery and subsequent validation of protein biomarkers for early diagnosis of risk, onset, and progression of many important diseases. It is widely recognized that sensitivity to subtle changes in protein profiles (such as isoform distributions) is crucial to the

successful analysis and identification of potential biomarkers. Heterogeneity inherent to protein isoforms may be the result of polymorphisms or variability in posttranslational modifications. Numerous examples of such phenomena have been demonstrated over the past two decades using many different technologies, including 1-D and 2-D gel electrophoresis, mass spectrometry, chromatographic fractionation, immunoblotting, or a combination thereof (1–4).

The pilot phase of the HUPO-PPP was instituted to evaluate various serum/plasma preparation methods and technology platforms (5). Such proteomic exploration often requires the identification and characterization of proteins or peptides found at low levels in biological samples. These protein surveillance activities necessitate some mode of sample depletion or purification to reduce interference from

highly abundant species. Unfortunately the inherent protease complement of serum and plasma raises concerns over sample stability, especially when incorporating such processing steps as depletion, separation, or purification.

Protease inhibitors are commonly used in proteomic analysis to preserve the protein population from endogenous and exogenous proteolytic cleavage. There are a multitude of protease inhibitors, which act upon the various protease classes, including serine proteases (e.g., trypsin, chymotrypsin, plasmin, urokinase, and kallikrein), cysteine proteases (e.g., calpain, papain, cathepsin B, and cathepsin L), aminopeptidases (e.g., leucine and alanyl aminopeptidase), acid proteases (e.g., pepsin, rennin, cathepsin D, chymosin, and protease B), and elastases. Protease inhibitors are commonly combined into cocktails for the purpose of concurrent inhibition of several protease classes

¹Sigma-Aldrich Biotechnology, St. Louis, MO and ²The Johns Hopkins University School of Medicine, Baltimore, MD, USA

during protein sample preparation and purification activities. Moreover, some modified peptides and amino acids (e.g., leupeptin, bestatin, and pepstatin A) or native peptides (e.g., aprotinin) serve as competitive reversible inhibitors, which bind to the active site of proteases but are not cleaved. Other reagents (e.g., sulfonyl fluoride derivatives) are competitive irreversible inhibitors, which covalently attach to critical amino acids in the active sites of proteases.

This study significantly expands upon the outcome of recent investigations into the effects of inclusion or omission of protease inhibitors on depleted plasma samples via a “routine” proteomic approach to protein charac-

terization as part of the HUPO-PPP pilot phase study (6). Analysis of these samples by 2-dimensional electrophoresis (2-DE) showed that several visible isoform trains were shifted to a higher isoelectric point (pI) motif following treatment with a protease inhibitor cocktail. One component of this protease inhibitor cocktail, namely 4-(2-aminoethyl) benzenesulfonyl fluoride (AEBSF), was shown to mediate this effect. This particular inhibitor functions through the covalent modification of serine residues in the active site of serine proteases, resulting in the formation of sulfonate esters (7). AEBSF has also been shown to derivatize other proteins through similar modification of serine residues

(8,9). Similarly, the AEBSF reagent reportedly also modifies tyrosine residues and, to a lesser extent, lysine, histidine, and the protein/peptide N terminus, (see www.abrf.org/index.cfm/dm.details?DMID=235&AvgMass=all&Margin=0). Such modifications result in molecular weight increases of 183 Da, as determined by electrospray ionization mass spectrometry (ESI-MS). Here we describe a previously unreported artifactual modification of protein isoform distributions on 2-DE gels as manifested by the appearance of higher pI species. This potentially deleterious effect is attributed to the increasing appearance of modified serine and/or tyrosine residues in protein isoforms and is further characterized via mass spectrometry. A model is also presented to rationalize the appearance of both modified and unmodified peptides in mass spectra acquired from a single isoform.

MATERIALS AND METHODS

Plasma Samples

Plasma samples were utilized as part of the HUPO-PPP pilot phase study (5), consisting of pooled samples from multiple donors with either citrate, heparin, or EDTA included in the collection tube as the anticoagulant (sample HUPO-PPP nos. BDAA01-Cit, BDAA01-Hep, or BDAA01-EDTA, respectively). Samples were stored at -80°C until use.

Protease Inhibitor Cocktail

Protease inhibition was accomplished using a 100 \times cocktail stock solution (Sigma-Aldrich, St. Louis, MO, USA) consisting of 104 mM AEBSF, 80 μM aprotinin, 2 mM leupeptin, 4 mM bestatin, 1.5 mM pepstatin A, and 1.4 mM trans-epoxy-succinyl-L-leucylamido-(4-guanidino) butane (E 64) in dimethylsulfoxide (DMSO). Samples were treated immediately upon thawing with either water (control), inhibitor cocktail, or individual protease inhibitor components (solubilized in DMSO) at a 1% (v/v) level (i.e., the final concentration

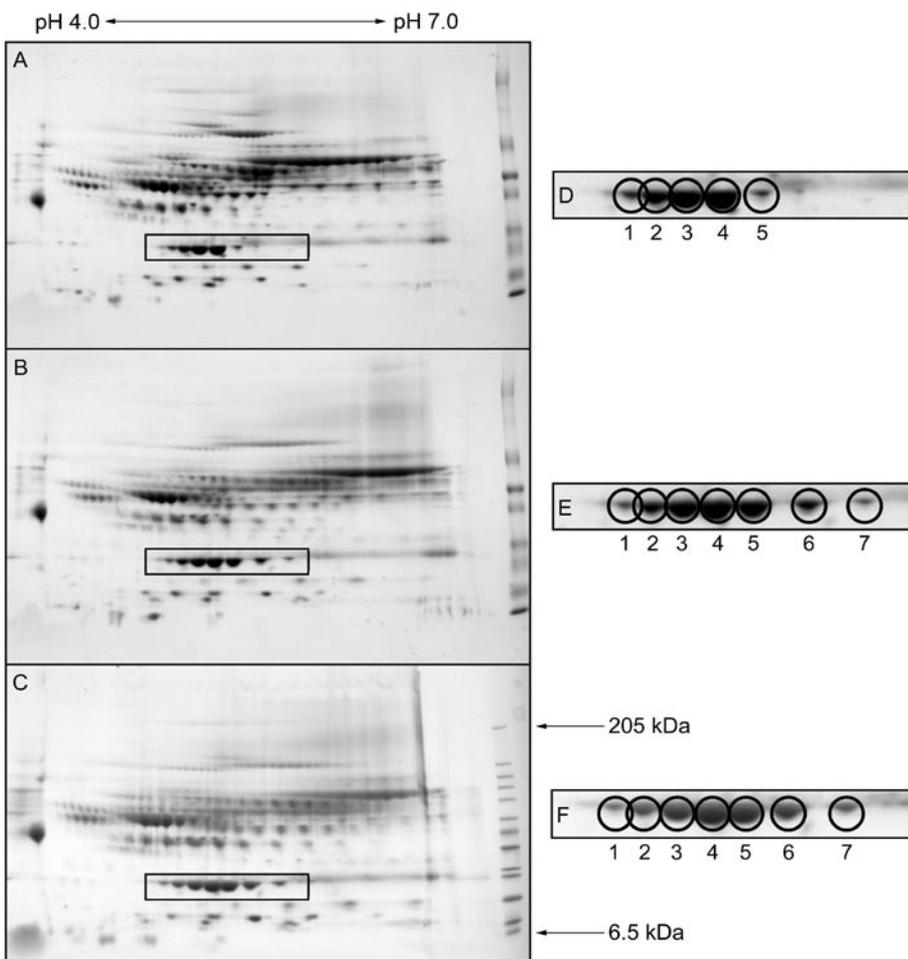


Figure 1. Impact of a protease inhibitor cocktail on the 2-DE pattern of human plasma. Three samples (50 μL) of Human Proteome Organization—Plasma Proteome Project (HUPO-PPP) citrated plasma (BDAA01-Cit) were subjected to the addition of either (A) water, (B) protease inhibitor cocktail, or (C) 4-(2-aminoethyl) benzenesulfonyl fluoride (AEBSF), and then depleted of albumin and immunoglobulin G (IgG). The depleted samples (100–200 μg) were separated on 2-dimensional electrophoresis (2-DE) gels as described in the Methods and Methods section. Panels D, E, and F are magnifications of the apolipoprotein A1 (apo A1) train from gels in panels A, B, and C, respectively.

in the plasma samples was 1/100th of the values in the stock solution).

Albumin/Immunoglobulin G Depletion

Aliquots (50 μ L) of each plasma sample were depleted of albumin and immunoglobulin G (IgG) with ProteoPrep™ Immunoaffinity Albumin and IgG Depletion kit (Sigma-Aldrich) and equilibrated with buffer containing 1% (v/v) of either water, protease inhibitor cocktail, or individual cocktail components, as detailed above. Total protein loads of 100–200 μ g for 2-DE were diluted to 130 μ L with water and added to 120 μ g of powdered protein extraction reagent type 4 (Sigma-Aldrich) to produce a final volume of 200 μ L in 7 M urea, 2 M thiourea, 1% (w/v) C7BzO, and 40 mM Trizma® Base.

Protein Quantitation and Reduction/Alkylation

Protein quantitation of each native and depleted sample was determined via Bradford assay (Sigma-Aldrich) with a 1 mg/mL bovine serum albumin (BSA) solution (Sigma-Aldrich) used as a standard. Each depleted plasma sample was reduced and alkylated with 5 mM tributylphosphine and 15 mM iodoacetamide (Sigma-Aldrich) for 30 min at room temperature.

First Dimension Isoelectric Focusing

Bromophenol blue (Sigma-Aldrich) was added as a tracking dye to each sample at a final concentration of 0.001% (w/v). The 11-cm, pH 4.0–7.0 immobilized pH gradient (IPG) strips (GE Healthcare, Piscataway, NJ, USA) were passively rehydrated with the protein samples for 4 h at room temperature, until the entire sample absorbed into the strip. The strips were then focused at 6000 V for a total of 85,000 Vh at 20°C using a PROTEAN® IEF cell (Bio-Rad Laboratories, Hercules, CA, USA).

Second Dimension SDS-PAGE

The IPG strips were equilibrated with sodium dodecyl sulfate (SDS)

using IPG equilibration buffer (Sigma-Aldrich) for 15 min. Equilibrated strips were subsequently loaded onto 4%–20% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gels (Bio-Rad Laboratories) and electrophoresed at 170 V for 80 min in Tris-glycine-SDS running buffer (Sigma-Aldrich). ColorBurst™ markers or SigmaMarker™ Wide Range (both from Sigma-Aldrich) were added to the marker lane on the extreme right of the gels. Gels were Coomassie® blue-stained via incubation with EZBlue™ gel staining reagent (Sigma-Aldrich) for 1 h and subsequently destained with water.

Western Blot Analysis

Gels for Western blot immunodetection were incubated in Bjerrum and Schafer-Nielsen transfer buffer [1× BSN; 48 mM Tris, 39 mM glycine, 0.00375% (w/v) SDS, 20% methanol; Sigma-Aldrich] for 20 min. The gel proteins were transferred semidry to a polyvinylidene difluoride (PVDF) membrane (Immobilon™ P, 0.45 μ m; Sigma-Aldrich) at 12 V for 40 min. Following transfer, the blots were washed with water and allowed to air-dry prior to immunodetection. The blots were wet briefly with 100% methanol, soaked in water for 15 min, stained with Ponceau S (Sigma-Aldrich) for 5 min, and finally washed with water to confirm protein transfer. The Ponceau S staining was then completely stripped with phosphate-buffered saline (PBS)-Tween® 20 (Sigma-Aldrich) for 30 min, and the membrane was blocked with PBS-5% milk (Sigma-Aldrich) for 30 min. Primary antibody incubation was carried out with immunoglobulin Y (IgY) anti-apolipoprotein A1 (anti-apo A1; Genway, San Diego, CA, USA) in 15 mL of PBS-5% milk (1:2000 dilution) for 90 min at room temperature. The blots were washed twice for 5 min each with PBS-Tween 20 and then incubated with anti-IgY horseradish peroxidase (HRP) conjugate secondary antibody (Genway) in 24 mL of PBS-5% milk (1:16000 dilution) for 60 min at room temperature. The blots were finally washed three times for 5 min each with PBS-Tween 20 and developed with 3,3',5,5'-tetramethyl-

benzidine (TMB) substrate (Sigma-Aldrich) for 1 min.

Gel Analysis and In-Gel Digestion

Stained gels were imaged and processed using the Phoretix 2D Expression™ imaging software (Nonlinear Dynamics, Durham, NC, USA). Spots of interest were manually cut from the gel, and the proteins were tryptically digested overnight at 37°C using the Trypsin Profile IGD kit (Sigma-Aldrich). Extracted peptides were dried at 30°C using a vacuum centrifuge (Eppendorf, Westbury, NY, USA).

Protein Identification by Mass Spectrometry

Tryptic peptides extracted from each spot were submitted for identification by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) peptide mass fingerprinting and sequence query by post-source decay (PSD) analysis. Dried samples were reconstituted in 0.1% trifluoroacetic acid (TFA; typically 10–20 μ L; Sigma-Aldrich) and mixed 1:1 with matrix solution (α -cyano-4-hydroxycinnamic acid, 5 or 10 mg/mL in 70% acetonitrile, 0.03% TFA; Sigma-Aldrich). Aliquots (1 μ L) were spotted onto a stainless steel MALDI target and allowed to dry at room temperature before introduction into the mass spectrometer.

Mass spectra were acquired in positive ion reflectron mode after close-external calibration using bradykinin 1–7 and insulin-oxidized B chain as standards (Sigma-Aldrich). Spectra were summed over approximately 100 shots, using an Axima CFRplus™ instrument (Shimadzu-Biotech, Manchester, UK). PSD mass spectra were acquired by the summation of 200–500 shots, with an appropriate ion gate setting for mass-to-charge ratio (m/z) discrimination. Fragmentation was augmented by collision-induced dissociation (CID) using air as the collision gas.

Protein identification was performed by searching the NCBI Inr database using the MASCOT search algorithm,

with identification confidence reported as probability-based MOWSE scores. For consistency, peaks were automatically parsed from mass spectrum files for peptide mass fingerprint analysis using the MASCOT Wizard software application. Database searching, scoring, and software applications were obtained from www.matrixscience.com. Software settings for peak picking included a signal-to-noise threshold of 10 and a minimum peak $m/z = 500$. Database search parameters were restricted to *Homo sapiens* taxonomy, and enzyme selection was trypsin, with up to one missed cleavage permitted. Carbamidomethylation of cysteines was selected as a fixed modification. Protein mass was unrestricted, and peptide mass accuracy tolerance was set at ± 250 parts per million (ppm).

RESULTS

Thawed plasma samples were subjected to the addition (1% by volume) of either water (Figure 1, A and D), protease inhibitor cocktail (Figure 1, B and E), or AEBSF (Figure 1, C and F). Following albumin and IgG depletion, these samples (100–200 μ g; citrated) were separated on 2-DE gels (Figure 1). Obvious changes in the isoform profiles of several proteins were observed, as exemplified by apo A1, in which the intensity of the two highest pI isoforms were substantially increased with the inclusion of the protease inhibitor cocktail. Figure 1, panels D, E, and F, magnify the apo A1 seven isoform train from the non-inhibited sample (Figure 1, A and D) and the inhibited samples (Figure 1, B, C, E, and F), respectively. The non-inhibited sample (Figure 1, A and D) shows significantly fewer high pI isoforms as compared with samples with inhibitors present (Figure 1, B, C, E, and F). The relative intensities of the protein spots on the gels from Figure 1, A and B, were compared using the 2-DE gel analysis software. The intensity of the low pI isoform spots decrease with the addition of the protease inhibitors, and there is a concomitant increase of the high pI isoforms (data not shown). Each of the six components of the protease inhibitor cocktail

Table 1. Identification of Apo A1 Isoforms by MALDI-TOF Mass Spectrometry

Isoform/Spot No.	NCBI nr Database Protein ID	Accession No.	MOWSE Score	Sequence Coverage (%)
1	proapolipoprotein	AAA51747	118	61
2	proapolipoprotein	AAA51747	148	65
3	proapolipoprotein	AAA51747	157	65
4	proapolipoprotein	AAA51747	196	76
5	proapolipoprotein	AAA51747	184	69
6	proapolipoprotein	AAA51747	177	67
7	proapolipoprotein	AAA51747	124	59

This table details the criteria for identifying apolipoprotein A1 (apo A1) from the isoform train (i.e., spots 1–7) as shown in Figure 1E. Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectra were searched against the NCBI nr database using the MASCOT search algorithm. Only proteins identified as the top hit are reported, along with the corresponding Accession no. Probability-based MOWSE scores for each protein are also reported, and sequence coverage derived from identified peptides is indicated.

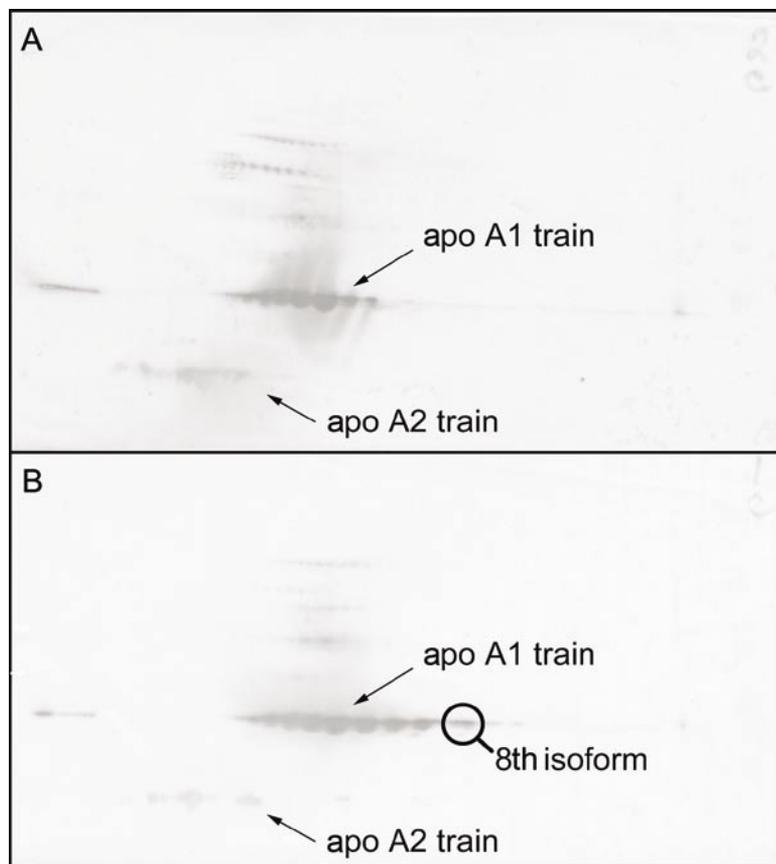


Figure 2. Identification of apolipoprotein A1 (apo A1) isoforms by Western blot analysis. Two samples (50 μ L) of Human Proteome Organization—Plasma Proteome Project (HUPO-PPP) citrated plasma (BDAA01-Cit) were subjected to the addition of either (A) water or (B) protease inhibitor cocktail and then depleted of albumin and immunoglobulin G (IgG). The depleted samples (100 μ g) were separated on 2-dimensional electrophoresis (2-DE) gels and then transferred to polyvinylidene difluoride (PVDF) membranes as described in the Methods and Methods section. Both blots were subjected to analysis by Western blot analysis using an anti-apo A1 primary antibody, anti-immunoglobulin Y (IgY) horseradish peroxidase (HRP) secondary antibody, and then colorimetrically detected using 3,3',5,5'-tetramethylbenzidine (TMB) substrate.

was individually and systematically added to the depleted plasma samples, and this phenomenon of isoform perturbation was found to be due solely to AEBSF (data not shown for other inhibitor components). Specifically, the intensity of the high pI forms in the apo A1 protein train increased when AEBSF was added in isolation from the other constituents of the protease inhibitor cocktail (Figure 1, C and F). The results obtained with citrated plasma were also similarly observed using serum and plasma treated with EDTA or heparin (data not shown).

The circles in Figure 1E depict the seven isoform spots identified as apo A1 by MALDI-TOF MS peptide mass fingerprinting (Table 1). The seven spots were cut from a Coomassie-stained 2-DE gel of an inhibitor cocktail-treated plasma sample (BDAA01-Cit) and were in-gel digested as described in

the Methods and Methods section. All seven spots were identified with high confidence (probability-based MOWSE score ≥ 118 for peptide mass fingerprinting) and high sequence coverage ($\geq 59\%$).

The apo A1 spot identifications were also confirmed by Western blot analysis (Figure 2). Isoforms 1–4 are seen in high abundance both without (Figure 2A) or with (Figure 2B) the protease inhibitor cocktail. The relative amounts of isoforms 5–7 are significantly higher with the inhibitor cocktail, and isoform 7 is essentially not seen in the absence of the cocktail. Significantly, an eighth isoform (highest pI), detected via Western blot analysis, was not observed on a Coomassie-stained gel. Other proteins were detected via interrogation with the anti-apo A1 antibody, such as the lower molecular weight train highlighted in Figure 2,

which have molecular masses and pIs consistent with apolipoprotein A2 (apo A2). Higher molecular weight trains were observed in Figure 2 but were not identified and, consequently, may form the basis for future investigation.

The characterization of the apo A1 isoform train was accomplished by MALDI-TOF MS evaluation of peptides derived from each of the seven spots. Four of the nine apo A1 tryptic peptides identified were found to have a modified partner peptide with a molecular weight increase of 183 Da. The four partner peptides were observed with higher intensity in the mass spectra of high pI isoforms (spots 5–7). It was found that the relative abundance of the partner peptides observed in the mass spectra was correlated with increasing isoform number and hence higher pI of the protein, as shown in Figure 3. Interestingly, no evidence was found for peptides with multiple modifications, regardless of the number of potential derivatization sites. Sequence analysis by MALDI-TOF PSD was executed on the T₂₇ peptide ($[M+H]^+ = 1301$ Da) and its partner peptide ($[M+H]^+ = 1484$ Da). The sequence of the T₂₇ peptide was confirmed through the appearance of the complete y-ion series, as shown in Figure 4 (lower panel). The partner peptide at +183 Da exhibited an equivalent fragmentation pattern up to the serine residue, thus verifying sequence homology for the partnered peptides. The y-ion series for the partner peptide then breaks at the serine or tyrosine residue, indicating that the higher molecular weight partner is apparently an altered version of the native peptide, modified at either the serine or adjacent tyrosine residue. Furthermore, peaks corresponding to the predicted higher mass y-ions of the modified peptide were observed, as indicated in Figure 4 (upper panel).

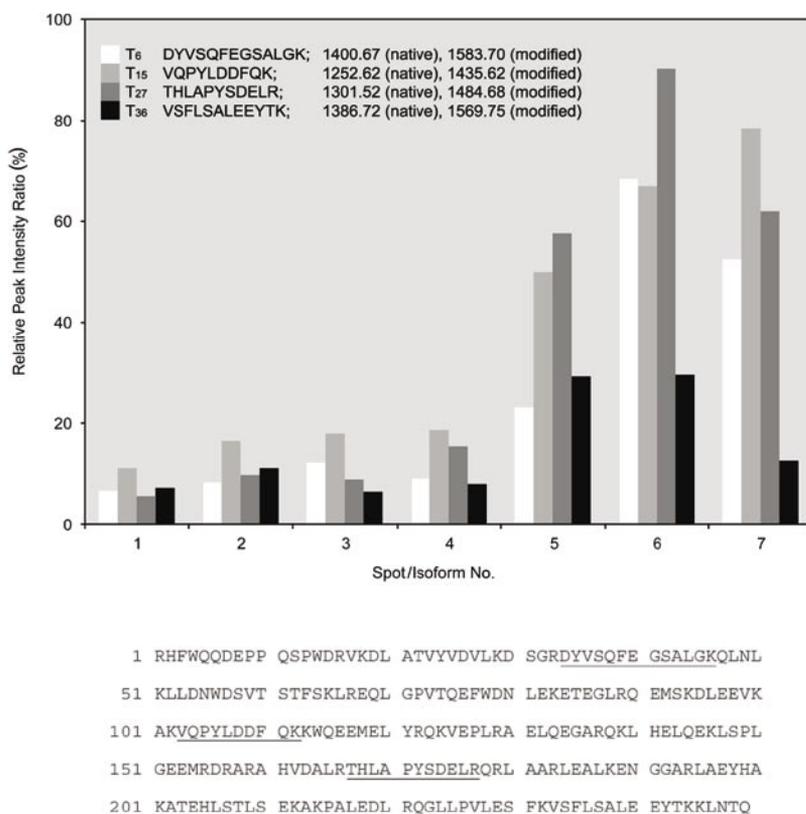


Figure 3. Comparative tryptic peptides for each isoform. Relative ratios of modified-to-native for four tryptic peptides were observed in matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectra from protein spots 1–7. Modified partner peptides appear at $\Delta m/z = +183$ Da, corresponding to sulfonate ester formation via the reaction with 4-(2-aminoethyl) benzenesulfonyl fluoride (AEBSF). Note the presence of serine and/or tyrosine residues in the highlighted peptide sequences, (either of which hold potential for modification by AEBSF). Moreover, the appearance of modified peptides in higher isoelectric point (pI) spots is consistent with the incorporation of additional charged species following reaction with AEBSF. The locations of observed peptides in the amino acid sequence of apolipoprotein A1 (apo A1) are underscored. *m/z*, mass-to-charge ratio.

DISCUSSION

In evaluating the use of a protease inhibitor cocktail for the analysis of a reference plasma sample using 2-DE, changes were observed in the isoform profiles of several proteins [including apo A1, apo A2, haptoglobin

(α and β subunits), transthyretin, and fibrinogen]. Studying apo A1 as a model protein, a systematic investigation of the various components of this cocktail identified AEBSF as being responsible for this observation. AEBSF (also called Pefabloc[®]) is a serine protease inhibitor that affects the activity of enzymes such as trypsin, chymotrypsin, plasmin, kallikrein, thrombin (3), and other inhibited proteases including subtilisin-like endoproteases and tissue- and urokinase-type plasminogen activators (10). Another protein inactivated by AEBSF at 0.1 mM is platelet-activating factor degrading acetylhydrolase (PAF-AH) (11). Inhibition of these proteases takes place by covalent modification of serine residues in the active site. The modification of each serine residue results in the formation of sulfonate esters and produces an increase in molecular weight of 183 Da, as determined by ESI-MS (8,9). The AEBSF reagent has also been reported to modify tyrosine residues and, to a lesser extent, lysine, histidine, and the protein/peptide N terminus. Note that this observed isoform perturbation phenomena might be gradually mitigated if progressively lower concentrations of AEBSF were added to the serum or plasma sample; however, formulating such a titration curve was beyond the scope of the current study.

The covalent modification of amino acid residues by AEBSF incorporates a positively charged amine group, which shifts modified proteins to higher pI forms. In the current work, a higher frequency of AEBSF-modified peptides was found in isoforms of increasing pI (Figure 3). It is proposed that high pI isoforms have a greater number of positively charged modifications than low pI isoforms and that sequential modification of the available susceptible residues shifts the protein to the next highest pI spot.

From Figure 5, it is apparent that each modified spot has a mixture of unmodified ($m/z = 1301$) and modified ($m/z = 1484$) peptides. The question then arises as to why a single isoform might have a population of both modified and unmodified specific residues, if the modification creates a shift to a different isoform. Apo A1 has

a complement of 22 serine and tyrosine residues, and if all are equally accessible and reactive, each has the potential to be modified by AEBSF. Provided that each apo A1 molecule in a specific isoform spot carries an equivalent overall charge imparted by the same net

number of modifications, there is no requirement that each discrete apo A1 molecule necessarily carries precisely the same modifications at the same amino acid residues. Figure 6 illustrates this model and rationalizes the mixed population of modified and unmodified

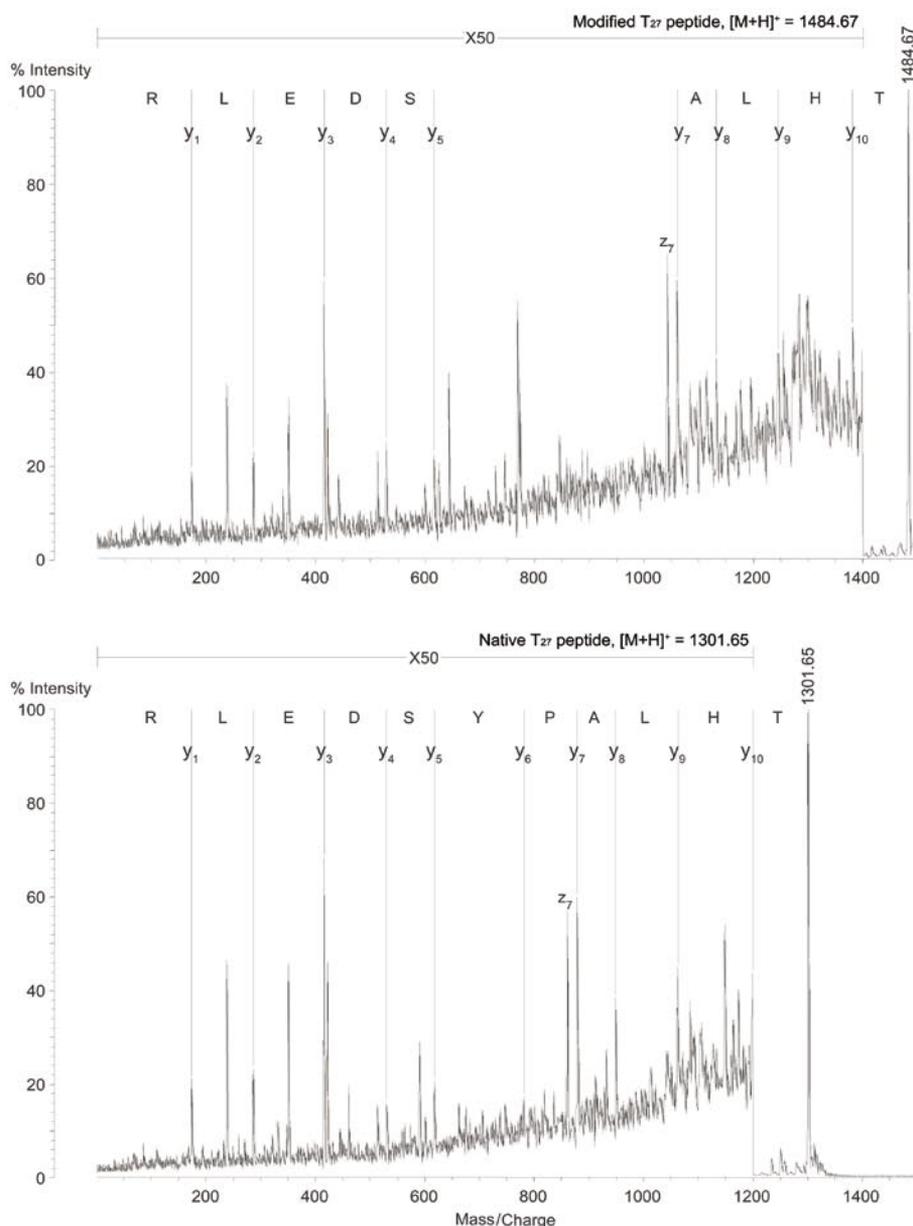


Figure 4. Post-source decay (PSD) matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectra of native T₂₇ (lower) and modified T₂₇ (upper) peptides. PSD mass spectra of the two T₂₇ partner peptides (1301 and 1484 Da) show amino acid sequence homology for each peptide up to the serine residue. A complete y-ion series is observed for the native peptide, confirming the sequence as predicted from the peptide mass fingerprint. The y₅ ion is observed for both modified and native T₂₇ peptides, however an incongruity is observed for the modified peptide in the region toward the N terminus. This disparity can be explained by either 4-(2-aminoethyl) benzenesulfonyl fluoride (AEBSF) adduct loss from the serine residue during fragmentation or, alternatively, by modification of the adjacent tyrosine residue. Higher mass y-ions match the predicted series for the AEBSF-modified peptide, derivatized at the serine or tyrosine residue. *m/z*, mass-to-charge ratio.

peptides observed in each isoform spot in accordance with the data presented in Figure 3. Moreover, the data provided in Figure 3 indicates that increasing numbers of modified peptides are found in higher pI spots, which also supports this model. Despite a variable number of serine or tyrosine residues in the tryptic peptides highlighted, modification of only a single serine or tyrosine was observed for any given peptide. It seems reasonable to suggest that covalent modification may be limited as a result of steric hindrance or

by kinetic limitations. Preliminary data in ongoing research has demonstrated that, at equivalent concentrations of sulfonyl fluoride derivatives, modification of amino acid residues appears to be subject to kinetic constraints.

In general, this type of chemical modification can be considered deleterious to the proteomic evaluation of plasma proteins. Some changes in clinically significant plasma proteins may be due to differential posttranslational modifications (PTMs), which have the potential for changing the

isoform profile of those proteins. Artifactual modification of proteins with AEBSF or other types of charged protease inhibitors that associate with proteins through covalent bonding could potentially mask or even derail such PTM evaluation. Other sulfonyl fluoride serine protease inhibitors, such as phenylmethyl sulfonyl fluoride (PMSF), are uncharged. Amino acid modification with PMSF should not significantly impact the isoform profile observed through 2-DE gel separation, but could result in mass shifts for observed peptides. Regrettably, PMSF does not possess some of the desirable properties of AEBSF (e.g., highly water soluble, stability in solution, low toxicity), but experiments with PMSF need to be carried out in future work.

This work has shown that high pI variants of protein isoform trains may in some cases be the direct result of sample handling procedures. Using apo A1 as a model protein, we have demonstrated that protein isoforms are covalently modified at serine or tyrosine residues by AEBSF. The evidence for this assertion includes the well-characterized nature of AEBSF to react with such residues (7), the finding that at least four serine- or tyrosine-containing peptides show paired peaks shifted by 183 Da, and finally, the PSD data showing y-ion sequence termination at the serine or tyrosine residue for the modified peptide. A model has been proposed whereby AEBSF adds in a stochastic manner to available residues in a given protein (Figure 6). For each high pI protein isoform, this model necessarily produces a mixed population of modified and unmodified peptides following proteolytic digestion, resulting in the observed presence of both native and modified serine- or tyrosine-containing peptides in a single mass spectrum.

The researcher must therefore balance the benefits and drawbacks of using protease inhibitors, depending upon the analysis methods and targets being evaluated. Inclusion of protease inhibitors of noncovalent mechanism should be considered, especially when using an approach

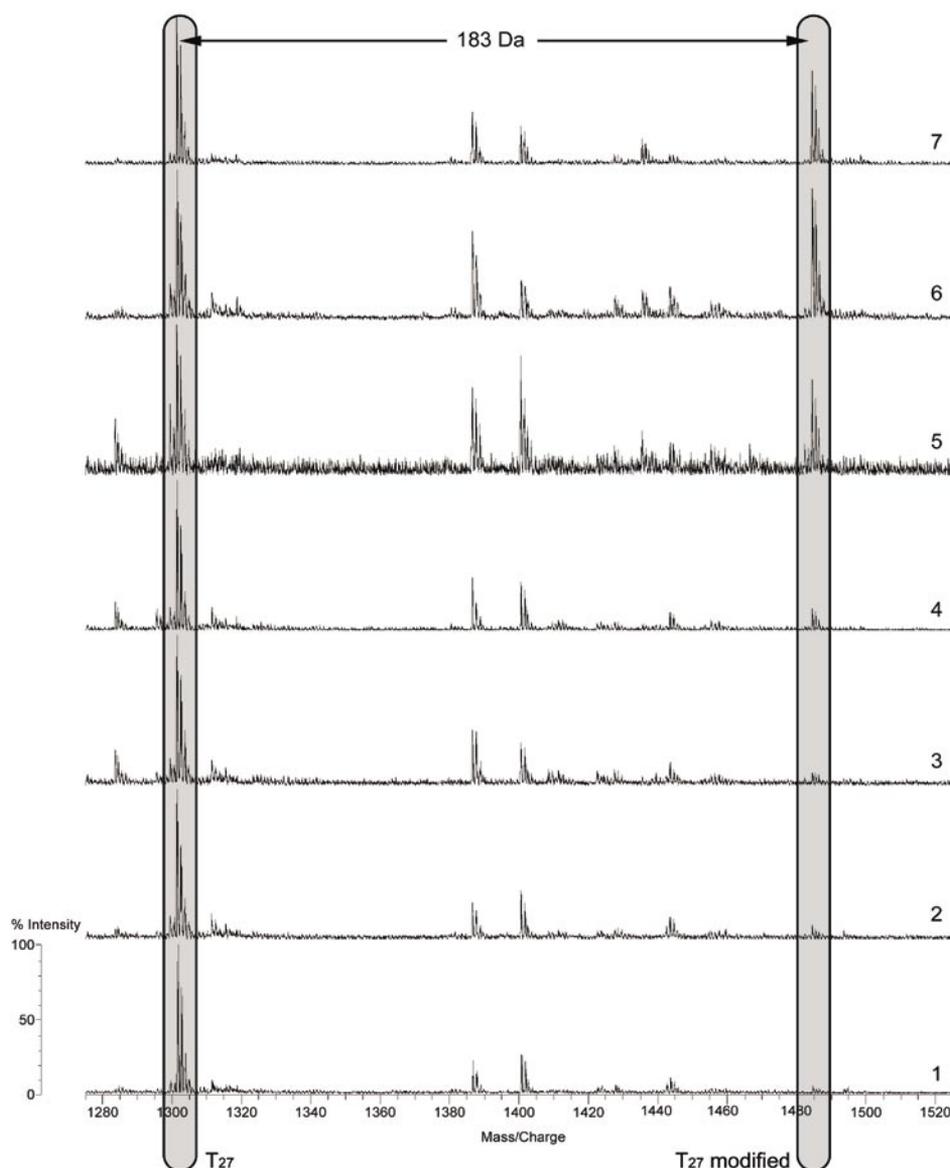


Figure 5. Modification of the apolipoprotein A1 (apo A1) T_{27} peptide by 183 Da. Positive ion reflectron matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectra are shown in the mass-to-charge ratio (m/z) range of 1275–1525 for apo A1 isoforms 1–7 [low to high isoelectric point (pI), respectively]. All spectra are normalized to 100% of the unmodified T_{27} peak intensity.

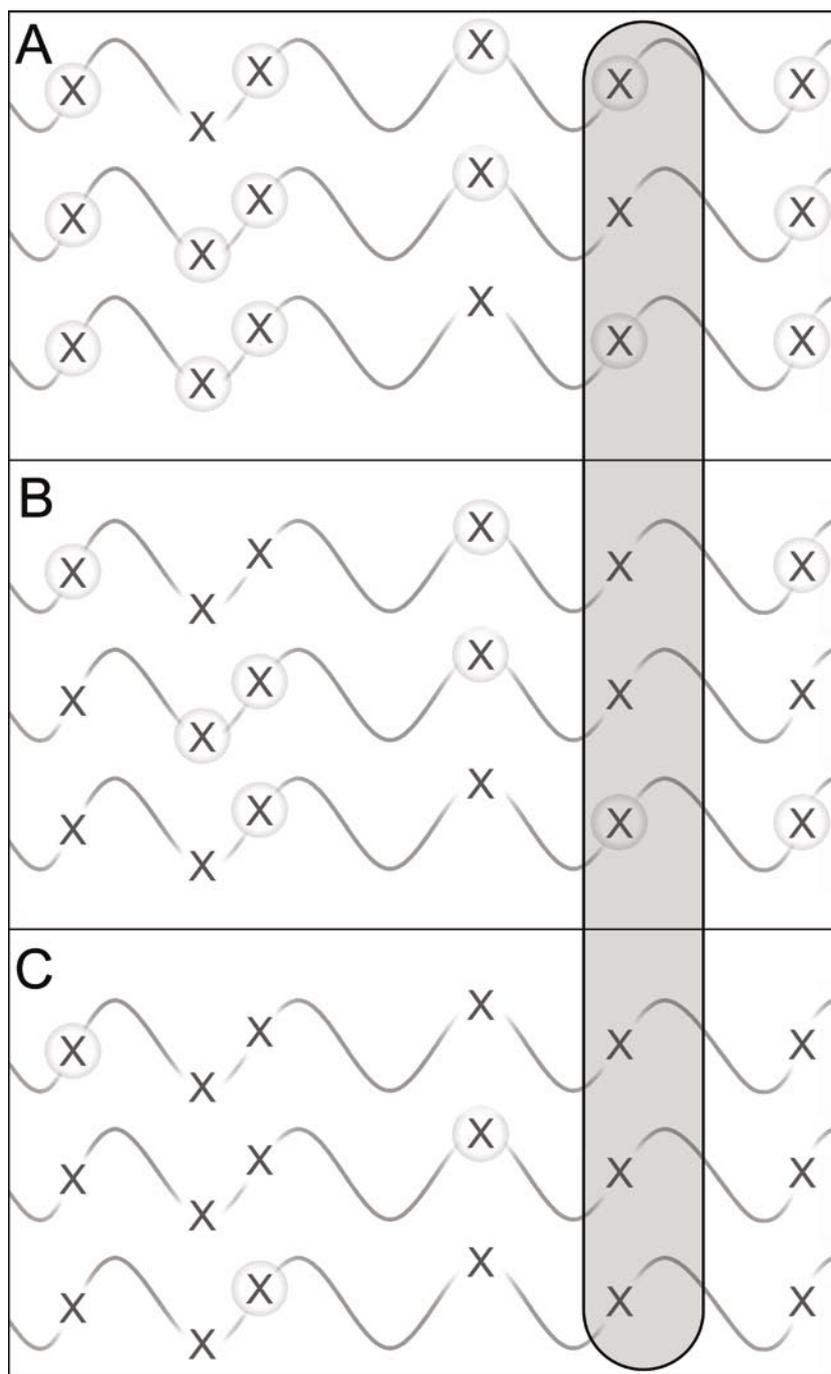


Figure 6. Model of peptide modification within isoforms. This illustration represents three isoforms of a hypothetical peptide chain: (A) high isoelectric point (pI), (B) medium pI, and (C) low pI. The X represents an unmodified serine or tyrosine residue, whereas the circled X represents 4-(2-aminoethyl) benzenesulfonyl fluoride (AEBSF)-modified residues. The three peptide chains in each panel have the same overall number of modified serine and/or tyrosine residues and, therefore, are constituents of the same isoform. The total number of modified (i.e., circled) residues decreases (5, 3, and 1 for panels A, B, and C, respectively), denoting the variation between different isoforms present in the train. The highlighted band corresponds to a hypothetical tryptic peptide incorporating a single serine or tyrosine residue and includes a mixture of both modified and unmodified peptides. Isoforms differ according to the overall charge carried by each protein, and as such, each isoform must be comprised of proteins with the same net charge. For proteins within each isoform, it is postulated that the net number of charge-inducing modifications is constant, but their location on the protein is randomized. For isoforms at higher pIs, a greater statistical proportion of potential derivatization sites (i.e., serine and/or tyrosine residues) carry the modification. This model rationalizes the appearance of both modified and unmodified peptides in a single mass spectrum of protein derived from the same pI spot.

in which the proteins are fractionated prior to proteolytic digestion. Protease inhibitor cocktails are commercially available that contain all reagents analyzed in the current work, except AEBSF. Inclusion of peptide-based protease inhibitors can also be problematic when evaluating the low molecular weight peptide complement (i.e., the so-called peptidome) in plasma samples. This is because the relatively high concentrations of peptide protease inhibitors can mask the native peptides of lower abundance in plasma. While ensuring the integrity of biological samples is critical to meaningful analysis, it is suggested that selection of protease inhibitors be administered with careful consideration of both the analysis and the ultimate objective. Hence, for proteomic analysis using 2-DE, inhibitor cocktails composed of large peptides or proteins should be avoided in favor of those consisting of small molecule species, which will not be retained by the gel matrix.

In conclusion, this study identifies an important issue regarding the utility of AEBSF as a protease inhibitor in the processing and characterization of biological samples for proteomic and/or biomarker discovery endeavors, specifically due to the illustrated modification of isoform trains observed in 2-DE gels. Clearly, as stated above, certain protease inhibitors such as AEBSF must be used judiciously prior to experimentation if one is to avoid the generation of artifactual spots in protein isoform distributions.

ACKNOWLEDGMENTS

The authors wish to acknowledge the efforts of the reviewers for their careful reading of the manuscript and their helpful suggestions for improvement.

COMPETING INTERESTS STATEMENT

M.D.S., R.J.M., S.L.C., G.T.L., J.D.S., J.W., R.V.-C., W.K.K., and G.B.I.S. are employed by Sigma-Aldrich Corporation, a supplier of a number of products used in this study. A.J.R. declares no competing interests.

REFERENCES

1. **Altland, K., T. Roeder, H.M. Jakin, H.G. Zimmer, and V. Neuhoff.** 1982. Demonstration of alpha 1-acid glycoprotein (orosomuroid) by double one-dimensional slab gel electrophoresis: evidence for intra- and interindividual variability of the micro-heterogeneity pattern in health and disease. *Clin. Chem.* 28:1000-1010.
2. **Janckila, A.J., Y.R. Nakasato, D.H. Neustadt, and L.T. Yam.** 2003. Disease-specific expression of tartrate-resistant acid phosphatase isoforms. *J. Bone Miner. Res.* 18:1916-1919.
3. **Johnson, P.J.** 2001. The role of serum alpha-fetoprotein estimation in the diagnosis and management of hepatocellular carcinoma. *Clin. Liver Dis.* 5:145-159.
4. **Thompson, P.A. and C. Ambrosone.** 2000. Molecular epidemiology of genetic polymorphisms in estrogen metabolizing enzymes in human breast cancer. *J. Natl. Cancer Inst. Monogr.* 27:125-134.
5. **Omenn, G.S.** 2004. The Human Proteome Organization Plasma Proteome Project Pilot Phase: reference specimens, technology platform comparisons, and standardized data submissions and analyses. *Proteomics* 4:1235-1240.
6. **Rai, A.J., C.A. Gelfand, B.C. Haywood, D. Warunek, C. Skobe, M.D. Schuchard, R.J. Mehig, S.L. Cockrill, et al.** Human Proteome Organization—Plasma Proteome Project Specimen Collection and Handling Committee Report: towards the standardization of parameters for plasma proteome samples. *Proteomics* (In press).
7. **Mintz, G.R.** 1993. Technical note: an irreversible serine protease inhibitor. *BioPharm.* 6:34-38.
8. **Sweeney, B., K. Proudfoot, A.H. Parton, L.M. King, P. Slocombe, and M.J. Perry.** 1997. Purification of the T-cell receptor ζ -chain: covalent modification by 4-(2-aminoethyl)-benzenesulfonyl fluoride. *Anal. Biochem.* 245:107-109.
9. **Bergen, H.R., 3rd, M.G. Klug, M.E. Bolander, and D.C. Muddiman.** 2004. Informed use of proteolytic inhibitors in biomarker discovery. *Rapid Commun. Mass Spectrom.* 18:1001-1002.
10. **Chu, T.M. and E. Kawinski.** 1998. Plasmin, subtilisin-like endoproteases, tissue plasminogen activator, and urokinase plasminogen activator are involved in activation of latent TGF-beta 1 in human seminal plasma. *Biochem. Biophys. Res. Commun.* 253:128-134.
11. **Dentan, C., A.D. Tselepis, M.J. Chapman, and E. Ninio.** 1996. Pefabloc, 4-[2-aminoethyl] benzenesulfonyl fluoride, is a new, potent nontoxic and irreversible inhibitor of PAF-degrading acetylhydrolase. *Biochim. Biophys. Acta* 1299:353-357.

Received 9 November 2004; accepted 7 March 2005.

Address correspondence to:

Graham B.I. Scott
Sigma-Aldrich Biotechnology
P.O. Box 14508
St. Louis, MO 63178, USA
e-mail: gscott@sial.com

To purchase reprints
of this article, contact
apfeffer@BioTechniques.com