Available online at www.sciencedirect.com
SCIENCE DIRECT
Carbohydrate Research 341 (2006) 198–209
Carbohydrate RESEARCH

Asymmetric glycosylation of soybean seed coat peroxidase

James S. S. Gray* and Rex Montgomery

Department of Biochemistry, Roy J. and Lucille A. College of Medicine, University of Iowa, Iowa City, IA 52242, USA

Received 3 October 2005; received in revised form 9 November 2005; accepted 17 November 2005
Available online 9 December 2005

Abstract—Reanalysis of the tryptic digests of soybean seed coat peroxidase (SBP) and its carboxyamidated peptide derivatives in the light of more complete sequence data has thrown light on the diglycosylated tryptic peptides, TP13 (Leu[183–205]Arg) and TP15 (Cys[208–231]Arg). Matrix assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) analyses indicate that although all potential sites carry some glycan substituents, not all sites are fully occupied. Tryptic glycopeptide TP13, carrying two N-glycosylation consensus sequons (Asn185 and Asn197), occurs mainly (85–90%) as the diglycosylated species, the remainder (10–15%) being monoglycosylated. In contrast, tryptic peptide TP15, also with two N-glycosylation sites (Asn211 and Asn216), is primarily monoglycosylated (~90%), with the remainder (10%) being diglycosylated. No non-glycosylated TP13 or TP15 was observed. Some artifacts are noted in the reactions of N-terminal cysteine residues and aspartate/asparagines residues in glycopeptide TP15. Mapping the glycans onto the crystal structure of SBP shows that these are asymmetrically distributed on the molecule, occurring primarily on the substrate-channel face of the enzyme. In contrast, the glycans of HRP, isozyme c, are more uniformly distributed over the enzyme surface.

Keywords: Soybean seed coat peroxidase; N-Glycosylation; Glycan distribution; Modeling

1. Introduction

Soybean peroxidase (SBP) is one of the most abundant proteins present in the soybean seed coat, where it may constitute up to 10% of the total protein,[1] and is sequestered mainly into the hourglass cells of the subepidermis of the seed coat.[2] Soybean cultivars can be classified into high-peroxidase and low-peroxidase phenotypes,[3] the former due to the presence of a dominant gene, designated *Ep*.[4] Soybean cultivars that are homozygous for the recessive gene, *epep*, express peroxidase at only 1% of that of the high-peroxidase phenotypes.[2]

SBP displays remarkable thermal and pH stability.[5] Its stability at low pH is particularly important for the oxidation of compounds such as veratryl alcohol, a lignin-type substrate, although the activity on veratryl alcohol is much lower than that observed for its oxidation by the lignin peroxidase produced by *Phanerochaete chrysosporium*.[6] SBP is a glycoprotein with 18% (w/w) carbohydrate.[7] There is considerable heterogeneity in the glycan composition of SBP, the major glycans belonging to the α-D-Man*p*-(1→6)-(α-D-Man*p*-(1→3)-)-(β-D-Xyl*p*-(1→2)-)-β-D-Man*p*-(1→4)-β-D-Glc*p*NAc-(α-L-Fuc*p*-(1→3)-)-β-D-Glc*p*NAc (hereafter abbreviated N_2M_3FX) family of glycans, with a variable number of terminal Man*p*, Fuc*p*, and Xyl*p* residues (Man*p* 2–4, Fuc*p* 0 or 1, Xyl*p* 0 or 1).[7] One of the N-glycosylation sites is unusual in being substituted primarily with only high-mannose type glycans, N_2M_{5-9}, with $M = 7$ being the major species present.[7,8]

An early attempt[8] to map the glycans onto the sequence of SBP was only partially successful due to errors in the partial sequence published by Huangpu and Graham.[9] The gene was later completely sequenced

Abbreviations: ACTH, adrenocorticotropic hormone; MALDI-TOF-MS, matrix assisted laser desorption/ionization time-of-flight mass spectrometry; HRPc, horseradish peroxidase isozyme c; SBP, soybean peroxidase; RP-HPLC, reverse phase high-performance/pressure liquid chromatography; TP13 (MG), TP15 (MG), monoglycosylated tryptic peptide 13 and 15, respectively; TP13 (DG), TP15 (DG), diglycosylated tryptic peptide 13 and 15, respectively; RZ, Reinheitzahl, ratio of A403/A280; SPME, solid phase microextraction.

* Corresponding author. Tel.: +1 319 335 7908; fax: +1 319 335 9570; e-mail: james-gray@uiowa.edu

0008-6215/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2005.11.016

by Gijzen[2] and the protein was sequenced by Welinder and Larsen.[10] Welinder and Larsen[10] also investigated the occupancy of the putative N-glycosylation sites, and confirmed the occupation of Asn56, Asn130, Asn144, Asn185, Asn197, and Asn211. The occupancy of the seventh N-glycosylation sequon, Asn[216–218]Thr, could not be confirmed, although there was previous evidence for its occupation.[8]

The glycosylation pattern of SBP is reinvestigated, particularly with respect to the partially occupied glycosylation sites on Asn185, Asn197, Asn211, and Asn216, and is reported here. Results from the investigation of TP15 by MALDI-TOF-MS suggest that modifications of the peptide, together with its variable glycan composition, are responsible for the heterogeneity of this peptide.

2. Results

The incomplete sequence of SBP[9] used in the previous study on the glycosylation of this enzyme,[8] and the complete sequence published by Gijzen,[2] are presented in Figure 1. A comparison of the two sequences reveals that the incomplete sequence, in addition to the missing N-terminal portion of the protein, contains errors in amino acid sequence, most notably those covering the sequence of tryptic peptide (hereafter abbreviated TP) 13 viz. Leu191 in place of Asn191, Ile192 instead of Pro191, and His193 rather than Asp193; additionally, Pro194 and Thr195 are absent in this sequence (all numbers are based on the complete sequence published by Gijzen). Consequently, the mass calculated for TP13, and reported as the unknown in the previous report from this laboratory,[8] was incorrect ($[M+H]^+$ (average) = 2482.8 Da rather than 2643.9 Da). Differences are also observed at position 249 (full sequence numbers), where a Leu residue is incorrectly replaced by an Arg residue, and in amino acids 260–266 (Fig. 1). The predicted masses of the tryptic glycopeptides from both the partial[9] and complete[2] SBP gene sequences and their major glycoforms, are presented in Table 1, which also serves to integrate the terminology used in the previous publications[7,8] with the nomenclature in this report.

```
  1 QLTPTFYRet cpnlfpivfg vifdasftdp rIGASLMRlh
  1 ---------- ---------- ---------- --------lh

 41 fhdcfvqgcd gsvllnntdt ieseqdalpn insirGLDVV
 41 fhdcfvqgcd gsvllnntdt ieseqdalpn insirGLDVV

 81 NDIKtavens cpdtvscadi laiaaeiasv lgggpgwpvp
 81 NDIKtavens cpdtvscadi laiaaeiasv agrrsgwpvp

121 lgrRdsltan rTLANQNLPA PFFNLTQLKa sfavqglntl
121 lgrRdsltan rTLANQNLPA PFFNLTQLKa sfavqglntl

161 dlvtlsgght fgrARcstfi nrLYNFSNTG NPDPTLNTTY
161 dlvtlsgght sgrARcstfi nrLYNFSNTG LIH--LDTTY

201 LEVLRarCPQ NATGDNLTNL DLSTPDQFDN Ryysnllqln
201 LEVLRarCPQ NATGDNLTNL DLSTPDQFDN Ryysnllqln


241 gllqsdqelf stpgadtipi vnsfssnqnt ffsnfrVSMI
241 gllqsdqerF STPGADTIPL SIASA-NQNT FFSNFRVSMI

281 Kmgnigvltg degeirLQCN FVNG
281 Kmgnigvltg degeirLQCN FVNG
```

Figure 1. Amino acid sequence for SBP. The complete sequence, as reported by Gijzen[2] is depicted in the upper row, and partial sequence, as reported by Huangpu and Graham,[9] is depicted in the lower row. Alternating upper and lower case letters are used to depict the tryptic peptides. N-Glycosylation sequons are highlighted.

The glycopeptides derived from TP4, TP9, diglycosylated (abbreviated as DG) TP13, and monoglycosylated (abbreviated as MG) TP15 are observed as strong signals in the MALDI-TOF mass spectrum of an unfractionated trypsin digest of carboxyamidated apo-SBP (Fig. 2) and can be used to identify the corresponding N-glycosylation sites on the enzyme. However, glycopeptides arising from TP8, TP13 (MG), and TP15 (DG) are not easily and routinely observed in these spectra. Glycosylated TP8 ($[M+H]^+$ (average) 1947.9 Da), containing the smallest peptide ($[M+H]^+$ (average), 776.8 Da), is not generally observed in the MALDI-TOF-MS spectra of unfractionated trypsin digests of SBP, as was found previously.[8] The use of a CHCA matrix ('cluster buster'), which suppresses adduct and cluster formation,[11,12] did not improve detection of glycosylated TP8, although the appearance of other, low mass ions was improved in the spectrum (not shown). However, glycosylated TP8 was seen in an unfractionated trypsin digest of SBP analyzed with 2,5-dihydroxybenzoate as a matrix (data not shown).

The detection of TP15 (DG) and TP13 (MG), both of low abundance, is complicated by the presence of the major signals due to TP13 (DG) and TP15 (MG), respectively, and some separation of TP13 and TP15 is necessary for their observation. This can be accomplished by solid phase microextraction (SPME) on an OMIX 100 μL C_{18} tip (Fig. 3), where TP15 (MG and DG) elute in the 5–10% and 10–15% ACN fractions. Although the 15–20% and 20–25% ACN fractions contain TP15 (MG) as well as TP13 (MG and DG), its intensity is sufficiently reduced so that it no longer overwhelms the signal from TP13 (MG) (Fig. 3). The signal due to TP13 (MG) is further improved by the shift of a cluster of signals, $[M+H]^+$ = 3500–4000 Da, to the 30–40% ACN fraction (Fig. 3).

Fractionation of a tryptic digest of carboxyamidated SBP by RP HPLC on a C_8 column produces a complex chromatogram in which all of the predicted glycopeptides, including TP8, can be detected by MALDI-TOF-MS analysis (Fig. 4). These analyses also show that TP13 and TP15 occur in both the di- and monoglycosylated forms, with the former present as the major

Table 1. Predicted masses of SBP and carboxyamidated-SBP tryptic glycopeptides

Tryptic peptide (full SBPc sequence)[a]	Glycosylated Asn residue	Tryptic peptide (partial SBPc sequence)[b]	Glycan pool[c]	Peptide $[M+H]^+$		Glycopeptide $[M+H]^+$	
				Average	Monoisotopic	Average	Monoisotopic
TP4 Leu[39–75]Arg	56	T1′	GP5	4115.49[d]	4112.907	5656.88	5653.435
			GP6	(4231.61)	(4229.594)	(5773.00)	(5769.494)
TP8 Asp[125–131]Arg	130	T7′	GP1	776.82	776.390	1947.90	1946.807
TP9 Thr[132–149]Lys	144	T8′	GP4	2031.36	2030.107	3202.43	3200.524
TP13 Leu[183–205]Arg (monoglycosylated)	185/197	Not identified	Not identified	2643.91	2642.3096	3814.98	3812.727
TP13 Leu[183–205]Arg (diglycosylated)	185/197	Not identified	Not identified	2643.91	2642.3096	4986.06	4983.144
TP15 Cys[208–231]Arg (monoglycosylated)	211/216	T14′	GP3	2650.79[a]	2649.185	3822.87	3820.610
				(2707.84)	(2706.206)	(3878.92)	(3876.623)
TP15 Cys[208–231]Arg (diglycosylated)	211/216	T14′	GP3	2650.79[a]	2649.185	4993.95	4991.028
				(2707.84)	(2706.206)	(5050.00)	(5047.041)

[a] The amino acid numbering is according to the sequence reported by Gijzen.[2]
[b] T1′, T7′, etc.: nomenclature of the tryptic peptides based on the sequence reported by Huangpu and Graham.[9]
[c] GP1, GP2, etc. identity of the parent glycopeptide pools used for studies of glycan structure.[7] This terminology is based on the elution position of the glycopeptide in the HPLC profile and not on its amino acid sequence. The first four columns provide equivalence between the identity of the glycopeptides, based on amino acid sequence, used in this study, and those from previous studies.
[d] *m/z* for reduced disulfide (upper) and carboxyamidated (lower) SBP peptide. Monoisotopic masses are give to three decimal places and average masses to two decimal places.

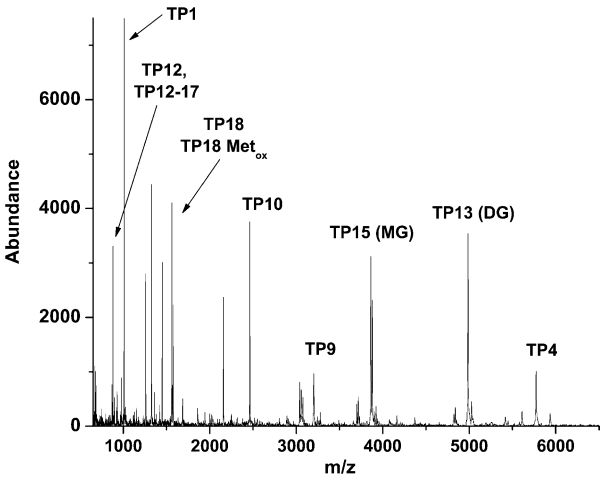

Figure 2. MALDI-TOF-MS spectrum of an unfractionated trypsin digest of SBP. The spectrum was acquired in the linear mode on a PerSeptive BioSystems DE STR MALDI-TOF mass spectrometer. Analysis conditions are presented in the text.

glycopeptide form of TP13, and the latter as the major form of TP15 (Fig. 4). It is also clear that there is considerable heterogeneity in TP15, because the diglycosylated form elutes from the C_8 RP-HPLC column in at least two, and the monoglycosylated form in at least three, peaks (Fig. 4). Importantly, the masses of the diglycosylated forms of TP15 in both RP-HPLC peaks are indistinguishable ($[M+H]^+$ (average) 5090.0 Da), being within the accuracy, with external calibration, of the PerSeptive Biosystems Voyager DE™ STR BioSpectrometry™ Workstation operated in the linear mode (~0.05%, or 2.5 Da at *m/z* 5000). Similarly, the masses for TP15 (MG) in the three RP-HPLC peaks are all similar ($[M+H]^+$ (average) 3880.0 Da). Integration of the RP-HPLC chromatogram affords a rough estimate of the relative proportions of mono- and diglycosylated TP13 and TP15. The monoglycosylated TP13 accounts for <10% of the total, the remainder being diglycosylated TP13. The reciprocal is true for TP15, where the monoglycosylated species represents about 85–90% and the diglycosylated species about 10–15% of the total.

The masses of the glycopeptides in all of these MALDI-TOF-MS analyses, that is, those of the unfractionated (Fig. 2), OMIX C_{18} fractionated (Fig. 3), and C_8 RP-HPLC fractionated (Fig. 4) tryptic digests, are consistent with previous results[7,8] showing that the major glycan occupying Asn130, Asn144, Asn185, Asn197, Asn211, and Asn216 is N_2M_3FX and that at Asn56, is N_2M_7. Associated with each of these major signals are a variable number of satellite signals due to the different glycoforms of the glycopeptides arising from variations in their Man*p* (Δ162 Da), Fuc*p* (Δ146 Da), and Xyl*p* (Δ132 Da) content. The distribution of the glycoforms in each RP-HPLC peak is not uniform and can differ

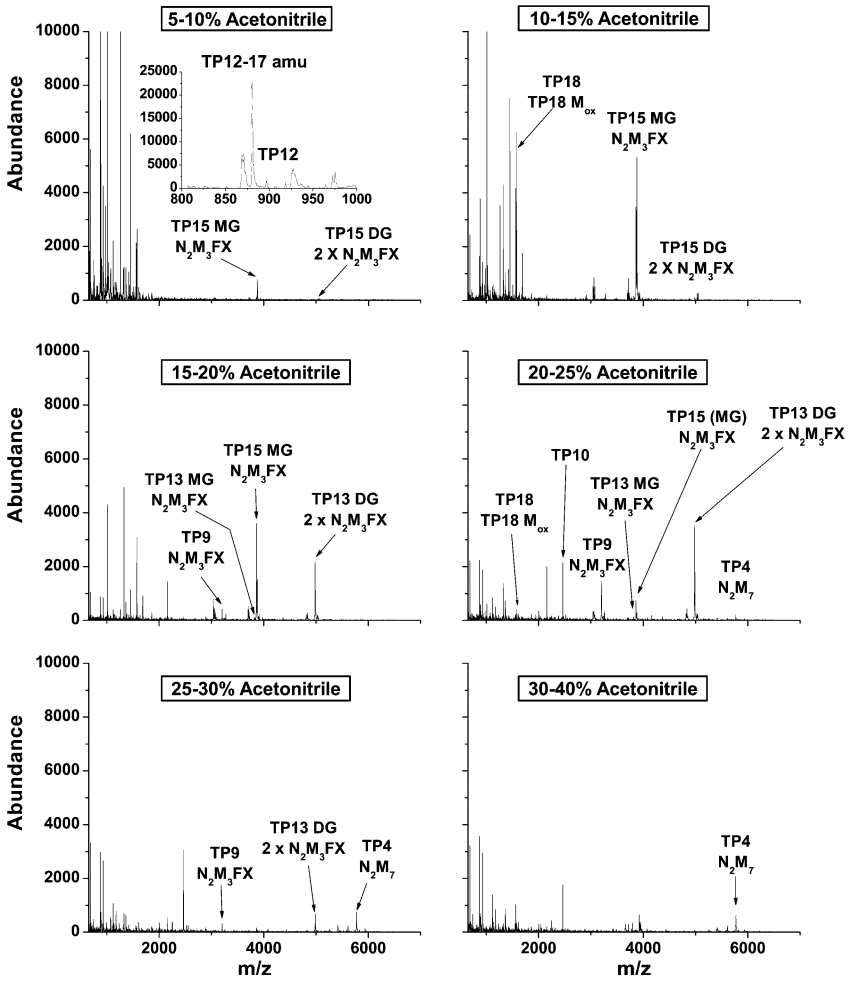

Figure 3. Fractionation of a trypsin digest of carboxyamidated SBP with an OMIX C_{18} tip. The trypsin digest was bound to the tip in 2% CH_3CN/1.0% trifluoroacetic acid, which after washing with 2% CH_3CN/0.1% trifluoroacetic acid (2×20 μL), was sequentially eluted with 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 50%, and 80% CH_3CN/0.1% trifluoroacetic acid (2×20 μL). Fractions were analyzed by MALDI-TOF-MS (linear mode) using the conditions described in the Methods section. All the fractions were analyzed; only those containing glycopeptides are presented here.

between fractions on the leading and the tailing edges of the peaks.

Prominent $[M+H]^+$-17 Da peaks are observed for both TP15 (MG) and TP15 (DG), but not for TP13 (MG) or TP13 (DG). This mass difference of 17 Da is also observed in the satellite signals arising from the different glycoforms of TP15, suggesting that peptide modification, rather than glycan modification, is responsible for this signal. One of the consequences of this variation is, as previously observed,[8,13] that the HPLC peaks containing the glycopeptides are broadened.

The isotope pattern for fraction 74 (RP-HPLC fraction collected between 73 and 74 min) is depicted in Figure 5A. This reflector mode MALDI-TOF mass spectrum, acquired with a resolution ($\Delta m/m$, where Δm = width of the isotope peak at half-height) of about 12,500, was calibrated with an internal standard composed of ACTH clips 1–17 ($[M+H]^+$ 2093.087), 18–39 ($[M+H]^+$ 2465.199), and 7–38 ($[M+H]^+$ 3657.929). It is clear that the predicted monoisotopic peak for monoglycosylated TP15 ($[M+H]^+$ 3876.623 Da) is absent, and that the second isotope peak ($[M+H]^+$ 3877.650) is the base peak. A comparison of the experimental (Fig. 5A) and the simulated (Fig. 5B) isotope patterns clearly shows that the experimental envelope is too broad to arise from a single species with a monoisotopic mass of 3876.623 Da, and is better described by the presence of at least three species. This is borne out by the examining of the deisotoped spectrum (Fig. 5C), showing that the isotope envelope can be explained by the presence of four species, the major species (58.6% of the total) having a $[M+H]^+$ of 3880.567 Da. Similar analyses for the other RP-HPLC peaks containing monoglycosylated TP15 are presented in Table 2, where the predicted monoisotopic species are absent in all of the fractions. The results in Table 2 also indicate that differences in the chromatographic behavior of the different TP15 glycoforms can be explained by peptide

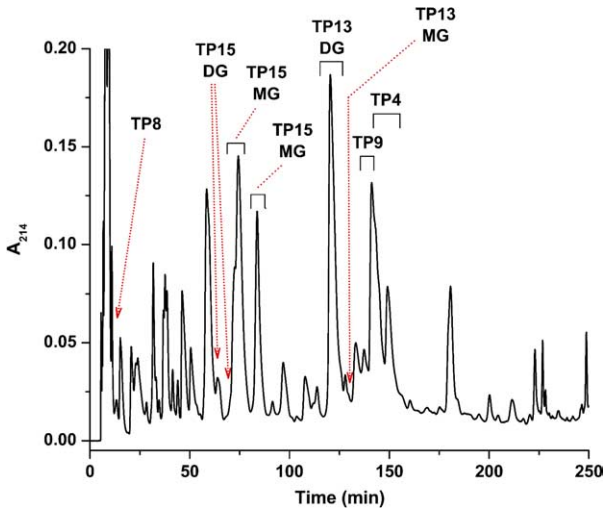

Figure 4. Purification of glycopeptides from a trypsin digest of carboxyamidated SBP on a Beckman Ultrasphere C_8 column (4.6 × 250 mm). Elution conditions are presented in the Methods section. Fractions (1 min) were collected and analyzed by linear MALDI-TOF-MS (conditions in the Methods section). Peaks containing glycopeptides, labeled with the parent tryptic peptide. TP13 (MG), TP15 (MG), monoglycosylated tryptic peptides 13 and 15, respectively; TP13 (DG), TP15 (DG), diglycosylated tryptic peptides 13 and 15, respectively.

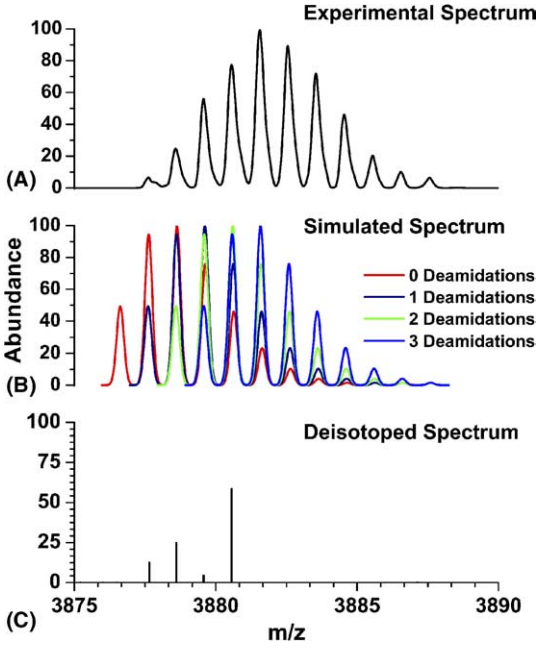

Figure 5. Comparison of the isotope pattern of monoglycosylated TP15 in RP-HPLC fraction number 74 acquired by reflector MALDI-TOF-MS (A), with the simulated spectrum (B). The deisotoped spectrum is depicted in (C). Reflector MALDI-TOF-MS conditions are presented in the Methods section.

modification. The absence of the monoisotopic peak is however variable, and is observed as a minor peak in some preparations of monoglycosylated TP15 (data not shown).

As a control, the isotope pattern of ACTH clip 7–38 ($[M+H]^+$ (monoisotopic), 3657.929 Da), one of the internal calibration standards, in each fraction reported in Table 2 was analyzed and compared to the theoretical isotope distribution (data not shown). The isotope pattern in every fraction was consistent with the presence of a single peptide species. Moreover, the area of each isotope peak was close to that predicted by the spectrum simulated with the same mass resolution.

The glycans were modeled onto the crystal structure of SBP (deposited with the Protein Data Bank, PDB, as 1FHF[14]) and HRPc (PDB 1W4W[15]), using the program GlyProt (http://www.glycosciences.de/glyprot/).[16] In SBP, N_2M_3FX was attached to Asn130, Asn144, Asn185, Asn197, Asn211, and Asn216, and N_2M_9 to Asn56. The major glycan, N_2M_3FX was modeled onto all eight N-linked glycosylation sites of HRPc shown from previous work to be glycosylated[13,17] viz. Asn13, Asn57, Asn158, Asn186, Asn198, Asn214, Asn255, and Asn268. The files were imported into and manipulated with PyMOL v. 0.98.[18]

The glycosylation of SBP and HRPc, viewed along the axis from the proximal end of the molecules, is presented in Figure 6. It is clear that the glycans are distributed primarily on the substrate-channel face of SBP (Fig. 6, left pane) whereas they are distributed more evenly in HRP (Fig. 6, right pane).

3. Discussion

The SBP glycosylation study reported here extends previous studies from this laboratory,[7,8] confirms the study by Welinder and Larson,[10] and addresses the partial occupation of four of the N-glycosylation sites present on the enzyme, viz. Asn185, Asn197, Asn211, and Asn216.

Some of the problems associated with MALDI-TOF-MS, including signal suppression and crowded/overlapping signals from peptides with similar masses, are obviated by fractionation of the tryptic digest of SBP with a C_{18} OMIX tip. This SPME procedure is rapid, inexpensive, and as demonstrated here, allows the identification of some of the minor glycosylated peaks, in this case, monoglycosylated TP13 and diglycosylated TP15. This technique is therefore useful for qualitative work, where identification of minor species is desired, but does not lend itself readily to quantitative studies.

Fractionation of tryptic digests of SBP by RP-HPLC on a C_8 column not only allowed the separation of the major glycopeptides, but also the separation of glycopeptides clearly related to each other by mass, but different because of glycan composition (e.g., the separation of TP13 (MG and DG) and TP15 (MG and DG)) or

Table 2. Analysis of the deamidation of monoglycosylated glycopeptide TP15 present in C_8 RP-HPLC fractionation of a trypsin digest of carboxyamidated SBP[a]

Fraction	Calculated $[M+H]^+$ (monoisotopic)						
	3876.623 (**0**)	3877.607 (**1**)	3878.591 (**2**)	3879.575 (**3**)	3880.559 (**4**)	3881.543 (**5**)	3882.527 (**6**)
Intensity (area) of isotope peak (as % of total)							
73	0.0	8.4	29.7	43.5	0.0	18.4	0.0
74	0.0	12.4	24.8	4.3	58.6	0.0	0.0
75	0.0	24.7	6.2	40.6	28.6	0.0	0.0
76	0.0	19.5	12.8	39.0	20.5	8.2	0.0
84	0.0	13.1	27.0	27.2	29.7	3.0	0.0
85	0.0	24.6	15.0	18.0	42.4	0.0	0.0
86	0.0	0.0	79.2	0.0	0.0	20.8	0.0

Number of deamidations are in bold and given within parentheses.

[a] Selected fractions, supplemented with an internal standard consisting of about 0.5 pmol of ACTH clips 1–17 ($[M+H]^+$ 2093.087), 18–39 ($[M+H]^+$ 2465.199), and 7–38 ($[M+H]^+$ 3657.929) were spotted with CHCA as matrix by the dried drop method and analyzed by MALDI-TOF-MS in the reflector mode. Conditions are the same as presented in the legend to Figure 5.

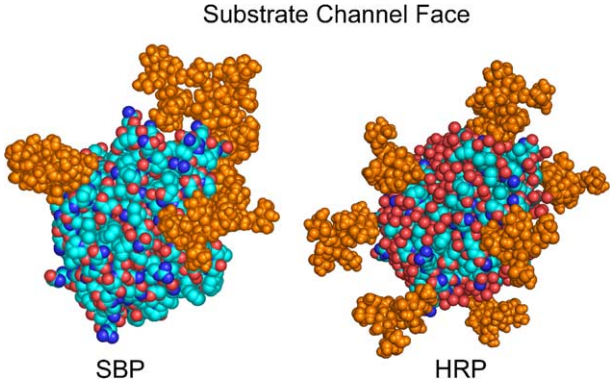

Figure 6. The distribution of glycans (orange spheres) on the surface of SBP (left) and HRP (right), viewed from the proximal end of the molecules.

peptide modification (the appearance of TP15 (DG) in two fractions, and TP15 (MG) in three fractions). Similar chromatographic profiles are observed when tryptic digests from a number of different SBP preparations are analyzed.

3.1. Peptide modification

Tryptic glycopeptide TP15 is interesting, not only because of its existence in both the mono- and diglycosylated forms, but because peptide modification appears to be responsible for its chromatographic behavior.

Ten of the 24 amino acids in the tryptic peptide TP15 (presented below in one letter code) are either Asn (4 residues), Gln (2 residues), or Asp (4 residues).

CPQNATGDNLTNLDLSTPDQFDNR

TP15

Two type of elimination reactions can be invoked to explain the 17 Da loss from monoglycosylated TP15; the formation of aspartyl succinimide, with the loss of NH_3, during the deamidation of Asn,[19–21] and the formation of a cyclic thiazine ((*R*)-5-oxoperhydro-1,4-thiazine-3-carboxylic acid) from the N-terminal carboxyamidated Cys.[22] Both reactions are depicted in Figure 7. The formation of aspartyl succinimide from aspartyl residues occurs, but at a rate 30-fold slower than for asparaginyl residues.[20,23,24] The one exception is the enhanced formation of aspartyl succinimide in the sequence, -Asp-Asn-Ile-Thr-, when the Asn residue is glycosylated.[25] A similar motif, with Leu replacing Ile, occurs in TP15. Glutamine residues also undergo deamidation, via a glutarimide intermediate, but at a much lower rate.[26]

The products produced from the hydrolysis of the aspartyl succinimide are α-Asp and its isomer, β-Asp (isoaspartate) in the ratio of 1:3.[20] Racemization of the α-C of the cyclic imide occurs at a low rate, resulting in the formation of D-Asp and D-isoAsp after hydrolysis.[20]

3.1.1. Isotope pattern analysis and deamidation. The loss of the amide and its replacement with a hydroxyl group in Asn and Gln residues increases the nominal mass by 1 Da. The mass of monoglycosylated TP15 can therefore increase by 5 Da if the Asn residues not involved in *N*-glycan formation and the Gln residues are deamidated.

The resolution of the PerSeptive Biosystems Voyager DE™ STR BioSpectrometry™ Workstation mass spectrometer is sufficient to resolve the isotope peak of the monoglycosylated TP15 present in RP-HPLC fraction 74. Surprisingly, the predicted monoisotopic mass for monoglycosylated TP15 ($[M+H]^+ = 3876.623$ Da) in fraction 74, is absent from the spectrum. Inspection of the breadth of the isotope envelope shows that this is too wide to be accounted for by a single species with a monoisotopic mass, $[M+H]^+$, of 3877.650 Da. This is borne out by comparing the experimental isotope

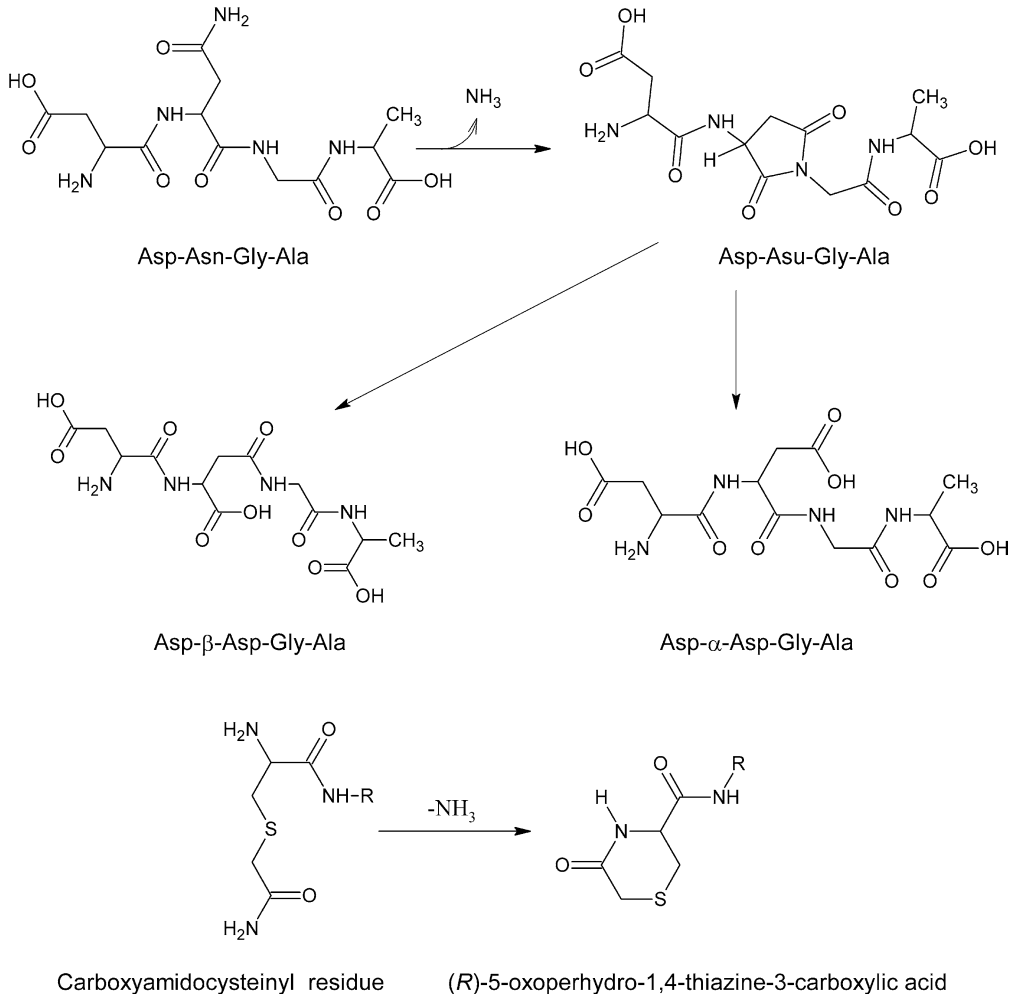

Figure 7. Reaction sequences leading to aspartyl succinimyl formation from Asn and (*R*)-5-oxoperhydro-1,4-thiazine-3-carboxylic acid formation from N-terminal carboxyamidated Cys.

pattern for TP15 (MG) from fraction 74 (Fig. 5A) with that simulated for this glycopeptide with up to three deamidations (Fig. 5B). The deisotoped spectrum (Fig. 5C) reveals that four glycopeptide species are present in this fraction. It is clear from the data presented in Table 2 that the experimental isotope patterns can incorporate up to five deamidations. Moreover, the results in Table 2 also show that each fraction contains a different pattern of modification, for example, about 80% of the glycopeptides in fraction 86 have masses consistent with 2 deamidations. Similarly, other fractions contain a variety of different species, and it appears as if this variability contributes to the chromatographic behavior of these glycopeptides.

How well do the experimental and theoretical isotope patterns agree? Can the isotope pattern be used to estimate the number of discrete species in a sample? This was answered, in part, by comparing the experimental and theoretical isotope patterns for ACTH clips 7–38 present, as an internal standard, in each of the fractions analyzed in Table 2, where good correlation was found between the experimental and theoretical intensities of the isotope peaks. In addition, a single deisotoped peak with the expected monoisotopic mass (±10 ppm), was obtained for each of the analyses. Consequently, it can be argued that the results presented in Table 2 represent a reasonable estimate of the number and proportion of TP15 species present in each fraction.

An important question is the source of this modification. Does it arise from modification of the SBP before extraction and purification, during SBP purification and storage, or is it a consequence of trypsin digestion and storage of the tryptic digest? Although the answer to this question requires long-term experiments to be performed, it can be partially answered by considering some of the factors that influence deamidation of Asn and Gln residues.

The susceptibility of Asn residues to deamidation is influenced by amino acid sequence and by tertiary structures of the protein.[27–32] Bulky groups on the α-car-

boxyl-side of Asn reduce the rate of imide formation significantly, whereas small groups, such as Gly enhance its formation.[20,27,28,30,33] In addition to sequence, secondary and tertiary structures also play a role in regulating the rate of Asn deamidation.[29–32,34,35] A computer program to calculate the susceptibility of Asn residues to deamidation was written by Robinson[28,29] and is available on the World Wide Web (http://www.deamidation.org). The program uses an extensive database containing data on the influence of neighboring residues on the rates of Asn deamidation, together with the three dimensional structure of the protein, to calculate the deamidation coefficient, C_D (defined as equal to one hundredth the half-life of the amide, in days), for each Asn residue in the protein. A number of additional parameters are used to calculate C_D, including a parameter representing the effect of 3D structure on deamidation, and a parameter incorporating other factors, such as H-bonding, that reduce the rate of deamidation.

The crystal structure for SBP (Brookhaven Protein Database, PDB entry 1FHF) was submitted to this site, generating the following results for the Asn residues present in TP15 (C_D, the deamidation coefficient, is presented in parentheses): Gln-Asn211-Ala (9.417), Asp-Asn216-Leu (10.574), Thr-Asn219-Leu (165.556), and Asp-Asn230-Arg (1238.720). Asn 211 and Asn 216, part of the N-glycosylation sequons, are more susceptible to deamidation by at least one order of magnitude than are Asn219 and Asn 230. Thus, glycosylation of either of these residues may be important in their stability. From data presented above and by Welinder and Larson,[10] the glycan occupancy of Asn216 is low. It thus appears that Asn216 may be deamidated in situ, that is, on the protein, but that the deamidation of Asn219 and Asn230 occurs during trypsin digestion and peptide purification and storage. Thus, deamidation of Asn216 may occur during storage of SBP in the hourglass cells of seed coat, during protein purification, and during storage of stock solutions of enzyme. The other deamidation reactions suggested by the complex isotope pattern are probably artifacts generated during trypsin digestion, and peptide purification and storage. Nonetheless, the overall conclusions regarding the occupation of the N-glycosylation consensus sequons are valid provided that all of the peaks associated with each peptide are identified and summed for their contribution to glycosylation site occupation.

3.1.2. Thiazine formation. The $[M+H]^+$-17 Da peak observed in the spectrum of carboxyamidated TP15 is considered to arise from the formation of the cyclic thiazine for two reasons. Firstly, the aspartyl succinimide intermediate is relatively unstable and is rapidly degraded in solution and so is not expected to persist in the sample. Secondly, TP12, which is also N-terminal carboxyamidated, is observed mainly as the $[M+H]^+$-17 and not as the $[M+H]^+$ species (881.0 and 898.0 Da (average mass), respectively; Fig. 3, first panel inset). Because the formation of the cyclic thiazine requires an N-terminal carboxyamidated residue for formation, its occurrence in these samples must be regarded as an artifact generated during trypsin digestion and storage of the tryptic digests and HPLC fractions derived from carboxyamidated SBP.

3.2. Glycosylation of SBP

The two N-glycosylation sequons on TP15 are close to one another, being separated by two amino acids. It is known that translation and glycosylation are concomitant, with glycosylation occurring when the nascent peptide chain is about 30 residues from the ribosome.[36] It can be speculated that, because Asn211 is exposed to the oligosaccharyl transferase before Asn216, and because of the close proximity of the two N-glycosylation sequons, that glycosylation at Asn211 would interfere with that at Asn216 by factors such as steric hindrance. That Asn211 is preferentially glycosylated is supported by the studies of Welinder and Larson,[10] where no glycosylation of Asn216 was observed. That Asn216 can be glycosylated is clear from the data presented here and in a previous study,[8] where TP15 (DG) is observed, albeit as a minor species. This apparent contradiction can be resolved by the observation that TP15 (MG) and TP15 (DG) are separated by RP-HPLC. Consequently, it is possible that the HPLC fractions selected by Welinder and Larson[10] for sequencing did not contain diglycosylated TP15.

In addition to steric hindrance, other factors influence the efficiency of glycosylation at a particular Asn residue. Recently, Petrescu et al.[37] performed a statistical analysis of the factors influencing N-glycosylation of a large number of glycoproteins contained in the Brookhaven Protein Data Bank (RCSB PDB, http://www.rcsb.org). The most significant conclusions from this study were: Thr, rather than Ser, is more often the third member of occupied N-glycosylation sites (about 7:3 ratio), that regions of tertiary structural change favor glycosylation, that aromatic amino acids often precede occupied glycosylation sites, and that acidic residues immediately preceding the Asn residue in the N-glycosylation consensus sequon, decrease site occupancy. They also found that surface topology of the protein in the vicinity of the glycan was important, with most occupied sites being either exposed on convex surfaces, or on flat regions of the protein.

Both of the N-glycosylation consensus sequons of TP15 terminate in Thr, and so this is unlikely to influence glycosylation at either site. However, glycan occupancy at Asn216 can be decreased due to the preceding Asp residue. Because the negative effect of the preceding Asp residue on the oligosaccharyl transferase

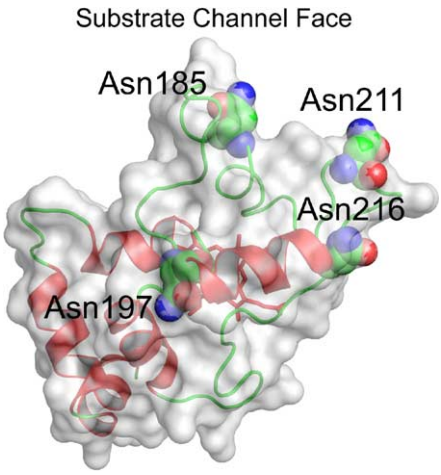

Figure 8. The topography of the SBP protein/solvent contact surface in the vicinity of Asn185, Asn197, Asn211, and Asn216. The view is from the proximal surface and similar to that depicted in Figure 6. The Asn residues that are part of the putative N-glycosylation consensus sequons, are shown as their van der Waals radii. The orientation of the amide N of Asn216 is away from the surface in contrast to those of Asn residues 185, 197, and 211.

activity is independent of steric hindrance, influencing either its affinity or its catalytic activity, the two effects will be additive. The overall effect would be to narrow the time window in which oligosaccharide transfer to Asn216 can occur. Subsequent folding of the peptide chain may further reduce the ability of the oligosaccharyl transferase to transfer an oligosaccharide to Asn216.

Examination of the surface of SBP in the vicinity of Asn185, Asn197, Asn211, and Asn216, shows that the three former Asn residues are either on convex surfaces, or on a ridge, whereas the amide N of Asn216 appears to be buried and therefore less accessible (Fig. 8).

3.3. Asymmetric distribution of the glycans on SBP

The distribution of the glycans on the surface of SBP is asymmetrical, with most of them occupying the substrate-channel face of the enzyme (Fig. 6). In contrast, the glycans are distributed more evenly over the surface of HRP, both as a consequence of there being two more *N*-glycans attached to HRP, and also because these are distributed over a greater proportion of the sequence.[13] The structures depicted in Figure 6 represent single conformers of the glycans; in practice, the glycans are present as an ensemble of conformers, and sweep out a much greater surface area than depicted here. The asymmetric distributions of glycans on the SBP surface, may explain many of the atypical behaviors of the enzyme, including its tendency to bind to glass surfaces, its quantitative distribution into an acetone/buffer cosolvent from aqueous ammonium sulfate solution,[38] and its stability when bound to the nanotubes.[39]

The asymmetric glycosylation of SBP has some potential uses. For example, derivatizing the glycans, after generation of reactive groups by mild periodate oxidation, would allow linkers to be attached that could be used to insert SBP into membranes with a specific orientation. Alternatively, appropriate derivatization of the glycans could be used to orient SBP at the interface between two immiscible solvents, allowing reactions to occur in the aqueous phase and extraction of products into the organic phase. Similarly, use of derivatized glycans could be used for constructions of biosensors, where the short diffusion path afforded by SBP bound in the same orientation, could result in greater sensitivity and higher response times.

3.4. Conclusion

In conclusion, the results reported here show that, although all the sites of SBP can be glycosylated, not all sites are equally populated. N-Glycosylation appears essentially complete at Asn56, Asn130, and Asn144, as the corresponding, non-glycosylated peptides are not observed in any of the analyses reported here. About 10% of the glycosylation sites on Asn185 and N197, based on the RP-HPLC results, are unoccupied; in contrast, about 35–40% of Asn211 and Asn216 are not occupied, that is, 85–90% of TP15 is detected as the monoglycosylated species, and 10–15% of the glycopeptide as the diglycosylated species. For reasons presented above, it is believed that Asn211 is the primary glycosylation site, and that Asn216 is the minor glycosylation site. Analysis of the results also lead to the conclusion that much of the heterogeneity observed during the MALDI-TOF-MS analysis of the tryptic peptides, particularly with respect to deamidation and cyclic thiazine formation, are experimental artifacts, and also conclude that glycosylation site occupancy can be estimated correctly if these are taken into account.

4. Experimental

4.1. Materials

Soybean seed coat peroxidase, RZ (Reinheitzahl, ratio of absorbance at 403 and 280 nm) 2.7, was obtained from Bioresearch Products, Inc. (North Liberty, IA). Modified porcine trypsin was from Promega US (Madison, WI). The Sequazyme Mass Standards Kit and the MALDI-MS Calibration Kit, MSCAL-2 were purchased from Applied Biosystems (Foster City, CA) and Sigma (Sigma–Aldrich, St. Louis, MO), respectively. α-Cyano-4-hydroxycinnamic acid (CHCA) (Fluka), 2,5-dihydroxybenzoate, Angiotensin I, adrenocorticotropic hormone (ACTH) clips 1–17, 18–39, and 7–39, dithiothreitol (DTT, SigmaUltra) and iodoaceta-

mide (SigmaUltra) were obtained from Sigma (Sigma–Aldrich, St. Louis, MO). The OMIX 100 μL C_{18} tips were the kind gift of Varian, Inc. (Lake Forest, CA). HPLC grade acetonitrile (Burdick and Jackson), was obtained from Fisher Scientific (Hampton, NH). All other reagents were of analytical quality.

4.2. Purification of SBP

SBP (1 mL of a 10 mg/mL in water) was chromatographed on a Sephadex superfine G-75 (Pharmacia, Uppsala, Sweden) column (1 × 40 cm) equilibrated with water, and eluted with water at a flow rate of 18 mL/h. Fractions (0.5 mL) were collected and their absorbances, after dilution with water if necessary, were determined at 280 and 403 nm on a Perkin Elmer (Norwalk, CN) Lambda 3B UV/Vis Spectrophotometer. Fractions with RZ > 2.9 were pooled, concentrated to about 10 mg/mL using an Amicon Ultra-4 (Millipore, Bedford, MA) PL-30 ultrafilter (MWCO 30,000), and stored at 4 °C.

4.3. Removal of heme

Heme was removed with acidic acetone (2% concd HCl (v/v) in acetone) at −20 °C, as described previously.[7,40]

4.4. Carboxyamidation and trypsin digestion

Urea (8 M) in 0.4 M NH_4HCO_3, pH 7.8–8.0, and 50 mM dithiothreitol (DTT) in this buffer solution, were freshly prepared immediately before use. Holo- or Apo-SBP (1 mg) was dissolved in 200 μL of 8 M urea/0.4 M NH_4HCO_3 and 20 μL of 50 mM DTT, mixed gently, and heated at 60 °C for 60 min, cooled, and allowed to react with freshly prepared 100 mM ICH_2CONH_2 in water (20 μL) in the dark for 30 min at room temperature. The excess ICH_2CONH_2 was quenched by adding 40 μL of 50 mM DTT. The urea concentration was reduced to <1 M with 50 mM NH_4HCO_3 (1960 μL), and Promega frozen modified porcine trypsin (530 μg/mL) was added to a final SBP/trypsin ratio of 40:1. A few microliters of toluene was added to the tube, and digestion was allowed to proceed for 16 h at 37 °C. The extent of digestion was checked by MALDI-TOF-MS.

4.5. Peptide fractionation on an Omix C_{18} 100 μL tip

The Omix tip was used according to the manufacturer's instructions. Briefly, the tip was activated with 50% (v/v) CH_3CN (2 × 100 μL), equilibrated with 2% (v/v) CH_3CN in 0.1% (v/v) CF_3COOH. The tryptic digest of carboxyamidated apo-SBP (~70 μg) was made to 2% CH_3CN/1% CF_3COOH (v/v) and bound to the tip by repeatedly aspirating and dispensing the samples. The tip was washed with 2 × 20 μL 2% (v/v) CH_3CN in 0.1% (v/v) CF_3COOH. The peptides were sequentially eluted from the tip with 2 × 20 μL of 5%, 10%, 15%, 20%, 25%, 30%, 40%, 50%, and 80% (v/v) CH_3CN in 0.1% CF_3COOH and analyzed by MALDI-TOF-MS.

4.6. High-performance liquid chromatography (HPLC)

Fractionation of the tryptic digests of SBP by Reverse Phase HPLC was performed on a Beckman (Altex) Ultrasphere Octyl column (4.6 × 250 mm) and a Dionex (Dionex Corporation, Sunnyvale, CA) DX300 HPLC equipped with a VDM-II variable wavelength detector. The tryptic digest of SBP (up to 800 μg) in 80 μL 2% CH_3CN/0.1% CF_3COOH (v/v) was filtered through a 0.2 μm filter (Millex-GV4, Millipore, Bedford, MA), and 60–80 μL was loaded onto the column and eluted at 0.5 mL/min with a binary gradient generated as follows (A = 0.1% (v/v) CF_3COOH; B = 100% CH_3CN/0.085% CF_3COOH (v/v)): 0–10 min, 5% B; 10–25, linear gradient from 5% to 18% B; 25–210 min, linear gradient 18–30% B; 210–250 min, linear gradient 30–50% B. The column was washed with 60% B for 30 min and equilibrated for 30 min with 5% B. The absorbance at 214 nm was continuously viewed and collected using Peaknet version 5.2 (Dionex Corporation, Sunnyvale, CA). Fractions (0.5 mL) were collected in polypropylene microfuge tubes, and analyzed by MALDI-TOF-MS. The volume of the fractions was reduced to ~50 μL on a SpeedVac (Savant) and stored at −20 °C.

4.7. MALDI-TOF-MS

MALDI-TOF-MS spectra were acquired on a PerSeptive Biosystems Voyager DE™ STR BioSpectrometry™ Workstation (Applied Biosystems, Foster City, CA) fitted with a N_2 laser operating at 337 nm and pulsed (3 ns) at 3 Hz. Data were recorded with a Tektronix 784A oscilloscope (Tektronix, Inc., Gaithersburg, MD), downloaded and processed using software supplied with the instrument (Biospectrometry Workstation Software version 5.1 and Data Explorer version 4.1). The vacuum of the source and the reflector regions was $<10^{-7}$ Torr. Linear spectra were obtained under the following conditions: accelerating voltage, 25,000 V; guide wire, 6.25 V (0.025%); grid voltage, 23,625 V (94.5%); low mass gate, 700 amu. Extraction delay was set at 150 ns for masses below 2000 amu, and at 260 ns for masses >3000 amu. Between 50 and 250 individual shots were averaged for spectra accumulated in the linear mode. Reflector spectra were acquired under the following conditions: accelerating voltage, 20,000 V; guide wire, 0.02 V (0.001%); grid voltage, 14,300 V (71.5%); low mass gate, 700 amu. Seven spectra of 150 shots from different points on the sample spot were acquired for each sample. The spectra were saved individually, together with the accumulated spectrum. The MALDI-TOF-MS was

calibrated using either the Sequazyme Mass Standards Kit (Applied Biosystems, Foster City, CA) or the ProteoMass MALDI-MS Calibration Kit, MSCAL-2 (Sigma, St. Louis, MO).

Samples (1–10 pmol) made up in 10 μL^{-1} CHCA (10 mg/mL in 50% CH_3CN/0.3% CF_3COOH (v/v)), were spotted on gold-plated targets and allowed to air dry. Spots were rinsed with 2 μL 0.1% (v/v) CF_3COOH for 10 s, which was removed and discarded. Samples with predominantly low mass ions were analyzed using 25 mM CHCA in 10 mM ammonium phosphate, pH 4.5 containing 50% CH_3CN/0.1% (v/v) CF_3COOH as matrix.[11,12] MALDI-TOF-MS analyses with 2,5-dihydroxybenzoate as matrix were performed as described previously.[8,13]

The internal standard (IS) contained 10 pmol each of ACTH clips 1–17 ($[M+H]^+$ (monoisotopic), 2093.0867 Da), 18–39 ($[M+H]^+$ (monoisotopic), 2465.1989 Da), and 7–38 ($[M+H]^+$ (monoisotopic), 3657.9294 Da). Matrix/IS (1 part IS + 4 parts matrix) prepared immediately before use, was mixed with sample (1 part sample + $\geqslant$4 parts matrix/IS) and spotted on a gold plate. The final concentration of the IS in the sample ranged from 0.3 pmol/μL to values approaching 1 pmol/μL for samples analyzed with higher matrix/sample ratios.

Data Explorer version 4.1 was used for data analysis. Calibration, using an external standard, is applied automatically during data acquisition. Spectra were calibrated, using an internal standard, after baseline correction and noise reduction. Deisotoping, centroiding, and the simulation of spectra were performed using the functions built into Data Explorer. The input for the simulation of the isotope pattern for monoglycosylated TP15 was the molecular formula calculated for the peptide substituted with one molecule of N_2M_3FX.

Acknowledgments

This material is based upon work supported by the Cooperative State Research, Education, and Extension Service, US Department of Agriculture (USDA), under Agreement No. 2004-34188-15067. Any opinions, findings, conclusions, or recommendations expressed herein are those of the authors and do not necessarily reflect the view of the USDA.

References

1. Gillikin, J. W.; Graham, J. S. *Plant Physiol.* **1991**, *96*, 214–220.
2. Gijzen, M. *Plant J.* **1997**, *12*, 991–998.
3. Buttery, B. R.; Buzzell, R. I. *Crop Sci.* **1968**, *8*, 722–725.
4. Buzzell, R. I.; Buttery, B. R. *Crop Sci.* **1969**, *9*, 387–388.
5. McEldoon, J. P.; Dordick, J. S. *Biotechnol. Prog.* **1996**, *12*, 555–558.
6. McEldoon, J. P.; Pokora, A. R.; Dordick, J. S. *Enzyme Microb. Technol.* **1995**, *17*, 359–365.
7. Gray, J. S. S.; Yang, B. Y.; Hull, S. R.; Venzke, D. P.; Montgomery, R. *Glycobiology* **1996**, *6*, 23–32.
8. Gray, J. S. S.; Montgomery, R. *Glycobiology* **1997**, *7*, 679–685.
9. Huangpu, J.; Graham, M. C.; Graham, J. S. *Plant Physiol.* **1996**, *110*, 714.
10. Welinder, K. G.; Larsen, Y. B. *Biochim. Biophys. Acta* **2004**, *1698*, 121–126.
11. Smirnov, I. P.; Zhu, X.; Taylor, T.; Huang, Y.; Ross, P.; Papayanopoulos, I. A.; Martin, S. A.; Pappin, D. J. *Anal. Chem.* **2004**, *76*, 2958–2965.
12. Zhu, X.; Papayannopoulos, I. A. *J. Biomol. Technol.* **2003**, *14*, 298–307.
13. Gray, J. S. S.; Yang, B. Y.; Montgomery, R. *Carbohydr. Res.* **1998**, *311*, 61–69.
14. Henriksen, A.; Mirza, O.; Indiani, C.; Teilum, K.; Smulevich, G.; Welinder, K. G.; Gajhede, M. *Protein Sci.* **2001**, *10*, 108–115.
15. Carlsson, G. H.; Nicholls, P.; Svistunenko, D.; Berglund, G. I.; Hajdu, J. *Biochemistry* **2005**, *44*, 635–642.
16. Bohne-Lang, A.; von der Lieth, C. W. *Nucleic Acids Res.* **2005**, *33*, W214–W219.
17. Welinder, K. G. *Eur. J. Biochem.* **1979**, *96*, 483–502.
18. DeLano, W. L. *The PyMOL Molecular Graphics System*; DeLano Scientific: San Carlos, CA, USA, 2002.
19. Capasso, S.; Mazzarella, L.; Sica, F.; Zagari, A.; Salvadori, S. *J. Chem. Soc., Perkin Trans. 2* **1993**, 679–682.
20. Geiger, T.; Clarke, S. *J. Biol. Chem.* **1987**, *262*, 785–794.
21. Reissner, K. J.; Aswad, D. W. *Cell. Mol. Life Sci.* **2003**, *60*, 1281–1295.
22. Geoghegan, K. F.; Hoth, L. R.; Tan, D. H.; Borzillerl, K. A.; Withka, J. M.; Boyd, J. G. *J. Proteome Res.* **2002**, *1*, 181–187.
23. Stephenson, R. C.; Clarke, S. *J. Biol. Chem.* **1989**, *264*, 6164–6170.
24. De Boni, S.; Oberthur, C.; Hamburger, M.; Scriba, G. K. E. *J. Chromatogr.* **2004**, *A1022*, 95–102.
25. Hoffmann, R.; Craik, D. J.; Bokonyi, K.; Varga, I.; Otvos, L. *J. Pept. Sci.* **1999**, *5*, 442–456.
26. Capasso, S.; Mazzarella, L.; Sica, F.; Zagari, A. *J. Chem. Soc., Chem. Commun.* **1991**, *23*, 1667–1668.
27. Li, B.; Gorman, E. M.; Moore, K. D.; Williams, T.; Schowen, R. L.; Topp, E. M.; Borchardt, R. T. *J. Pharm. Sci.* **2005**, *94*, 666–675.
28. Robinson, N. E. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5283–5288.
29. Robinson, N. E.; Robinson, A. B. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4367–4372.
30. Robinson, N. E.; Robinson, Z. W.; Robinson, B. R.; Robinson, A. L.; Robinson, J. A.; Robinson, M. L.; Robinson, A. B. *J. Pept. Res.* **2004**, *63*, 426–436.
31. Xie, M.; Aube, J.; Borchardt, R. T.; Morton, M.; Topp, E. M.; Vander Velde, D.; Schowen, R. L. *J. Pept. Res.* **2000**, *56*, 165–171.
32. Xie, M. L.; Schowen, R. L. *J. Pharm. Sci* **1999**, *88*, 8–13.
33. Robinson, N. E.; Robinson, A. B. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 944–949.
34. Athmer, L.; Kindrachuk, J.; Georges, F.; Napper, S. *J. Biol. Chem.* **2002**, *277*, 30502–30507.
35. Robinson, N. E.; Robinson, A. B. *J. Pept. Res.* **2004**, *63*, 437–448.
36. Marth, J. D. *N*-Glycans. In *Essentials of Glycobiology*; Varki, A., Cummings, R., Esko, J., Freeze, H., Hart, G., Marth, J. D., Eds.; Cold Spring Harbor Laboratory Press: New York, 1999; pp 85–100.

37. Petrescu, A. J.; Milac, A. L.; Petrescu, S. M.; Dwek, R. A.; Wormald, M. R. *Glycobiology* **2004**, *14*, 103–114.
38. Liu, J.; Zhang, Y.; Qiu, L.; Ye, L.; Xia, Y.; Su, Z. *Chem. Biochem. Eng. Q.* **2005**, *19*, 199–205.
39. Karajanagi, S. S.; Vertegel, A. A.; Kane, R. S.; Dordick, J. S. *Langmuir* **2004**, *20*, 11594–11599.
40. Chibbar, R. N.; Van Huystee, R. B. *Phytochemistry* **1983**, *22*, 1721–1724.