Identification of Novel Sites of O-N-Acetylgalactosamine Modification of Serum Response Factor Using Quadrupole Time-of-flight Mass Spectrometry*

Robert J. Chalkley‡ and A. L. Burlingame

The addition of a single N-acetylgalactosamine moiety O-linked to serine and threonine residues of nuclear and cytoplasmic proteins is a widespread post-translational modification. The conventional method for detecting and locating sites of modification is through a multistep radioactivity-based approach. We have recently shown that sites of O-GlcNAc modification can be determined using quadrupole time-of-flight tandem mass spectrometry (Chalkley, R. J., and Burlingame, A. L. (2001) Identification of GlcNAcylation sites of peptides and α-crystallin using Q-TOF mass spectrometry. J. Am. Soc. Mass Spectrom. 12, 1106–1113). In this work utilization of this new approach has revealed previously undetected sites of O-GlcNAc modification of the transcription factor serum response factor. Molecular & Cellular Proteomics 2: 182–190, 2003.

Since O-GlcNAc modification was first detected (1), a wide variety of proteins have been shown to be post-translationally modified with a single N-acetylgalactosamine residue (2, 3). The modification is transient, and the enzymes responsible for its addition (4, 5) and removal (6, 7) have been identified. Although not fully understood, the importance of O-GlcNAcylation is becoming increasingly apparent (8). All known GlcNAc-modified proteins are also potential phosphoproteins. Indeed in many instances the sites of phosphorylation and GlcNAcylation are localized to the same or neighboring residues (9–11). Hence in many instances there appears to be direct interplay between the two post-translational modifications (3, 12).

The major stumbling block in the understanding of O-GlcNAcylation is the difficulty of detection and subsequent determination of sites of modification. The traditional method for characterization of sites of GlcNAcylation is through the use of radioactivity (1, 13). Modified proteins are identified by radio-labeling GlcNAc residues with [3H]galactose in vitro using a galactosyltransferase. Modified proteins are enzymatically digested, and peptides are separated by multiple rounds of HPLC.2 Radiolabeled peptides are located by screening fractions for radioactivity. When a fraction containing only the radiolabeled peptide is isolated the peptide can be identified by Edman sequencing by monitoring for the cycle during which the radioactivity is released from which the modification site can be assigned.

Mass spectrometry has previously been used to indirectly identify sites of GlcNAc modification of synthetic peptides using base-catalyzed β-elimination of GlcNAc (14, 15). This approach has recently been adapted to provide an affinity isolation strategy that has allowed for the identification of novel sites of O-GlcNAc modification of Synapsin I and nuclear pore proteins (16). However, the use of this strategy requires care to prevent misassignment between GlcNAc modification and phosphorylation as both modifications are eliminated chemically by the same mechanism. Removal of all phosphorylation with phosphatases prior to derivatization has been used to avoid this problem (16). However, this prevents the characterization of both modifications simultaneously.

We have shown that the new generation of quadrupole time-of-flight mass spectrometry (Q-TOF) tandem electrospray mass spectrometers can be used to directly identify sites of modification of synthetic peptides and that this approach was also able to identify a previously known site of O-GlcNAc modification on αA-crystallin (17). Samples are digested and then analyzed by LC-MS-MS, and GlcNAc-modified peptides are identified on the basis of the facile production of the GlcNAc oxonium ion. The masses of modified peptides are then used to create an “inclusion list” of peaks in a second LC-MS-MS analysis of the same sample. In this second run, lower collision energies are used to increase the probability of observing glycosylated fragment ions, and MS-MS spectra of each precursor ion are acquired over a longer time period. We

1 The abbreviations used are: HPLC, high performance liquid chromatography; LC, liquid chromatography; CID, collision-induced dissociation; MS, mass spectrometry; MS-MS, tandem MS; ESI, electrospray ionization; GlcNAc, N-acetylgalactosamine; MALDI, matrix-assisted laser desorption/ionization; SRF, serum response factor.

From the Department of Biochemistry and Molecular Biology, University College, London, Gower Street, London WC1E 6BT, United Kingdom and the Ludwig Institute for Cancer Research, 91 Riding House Street, London W1 7BS, United Kingdom

Received, March 19, 2003
Published, MCP Papers in Press, April 8, 2003, DOI 10.1074/mcp.M300027-MCP200

182 Molecular & Cellular Proteomics 2.3
© 2003 by The American Society for Biochemistry and Molecular Biology, Inc.
This paper is available on line at http://www.mcponline.org
now demonstrate the power of this strategy for studies of another GlcNAc-modified protein, serum response factor (SRF).

SRF is a ubiquitous transcription factor that binds to the serum response element, a regulatory sequence found upstream of the genes of many proteins transiently expressed upon growth factor stimulation (18). SRF is a member of the MADS box family of transcription factors, which are named after the four original members of the family: MCM1, AG, DEFA, and SRF. Members of this family contain a conserved domain of 56 amino acids. The N-terminal part of this domain defines the DNA binding specificity of the protein, while the C-terminal part effects protein dimerization. Dimers of SRF form ternary complexes on the serum response element along with accessory ternary complex factors such as Elk-1, Sap-1, and Sap-2. SRF can be activated by serum, lysophosphatidic acid, aluminum fluoride, and G proteins via the GTPase RhoA (19, 20) and triggers the expression of key proteins involved in cell cycle progression, differentiation, and development (21–23).

SRF is a 50-kDa protein for which a number of phosphorylation sites have been identified. SRF is phosphorylated on serine 103 in response to stimulation by mitogen-activated protein kinase-activated protein kinase 1 (p90RSK) (24) in response to stress by mitogen-activated protein kinase-activated protein kinase-2 (25) and by calcium/calmodulin-dependent kinases II and IV (26). It is also phosphorylated at serine 83 by casein kinase II (27) and serine 435 by DNA-activated protein kinase (28).

SRF is also O-GlcNAc-modified (29), and Reason et al. (30) reported previously the characterization of a number of GlcNAc modification sites. Using recombinant SRF overexpressed in baculovirus (31), they identified serine 283 as a major site of GlcNAc modification. Serine 316 was also found to be glycosylated, and one of either the serine 307 or serine 309 sites appeared to be modified at very low stoichiometry. The sites were identified by first carrying out sequential enzymatic digestion of the protein using a combination of casein kinase II (27) and serine 435 by DNA-activated protein kinase-activated protein kinase-activated protein kinase (28). SRF is a ubiquitous transcription factor that binds to the serum response element, a regulatory sequence found upstream of the genes of many proteins transiently expressed upon growth factor stimulation (18). SRF is a member of the MADS box family of transcription factors, which are named after the four original members of the family: MCM1, AG, DEFA, and SRF. Members of this family contain a conserved domain of 56 amino acids. The N-terminal part of this domain defines the DNA binding specificity of the protein, while the C-terminal part effects protein dimerization. Dimers of SRF form ternary complexes on the serum response element along with accessory ternary complex factors such as Elk-1, Sap-1, and Sap-2. SRF can be activated by serum, lysophosphatidic acid, aluminum fluoride, and G proteins via the GTPase RhoA (19, 20) and triggers the expression of key proteins involved in cell cycle progression, differentiation, and development (21–23).

SRF is also O-GlcNAc-modified (29), and Reason et al. (30) reported previously the characterization of a number of GlcNAc modification sites. Using recombinant SRF overexpressed in baculovirus (31), they identified serine 283 as a major site of GlcNAc modification. Serine 316 was also found to be glycosylated, and one of either the serine 307 or serine 309 sites appeared to be modified at very low stoichiometry. The sites were identified by first carrying out sequential enzymatic digestion of the protein using a combination of enzymes. Peptides were then separated by HPLC, and fractions were collected. These were analyzed by fast atom bombardment mass spectrometry to detect fractions containing peptides that differed in mass by 203 Da, which could correspond to unmodified and GlcNAc-modified versions of the same peptide. Selected fractions were subjected to in vitro [3H]galactose labeling to tag the GlcNAc residue, and sites of modification were finally determined by Edman degradation analysis by monitoring for the release of the radioactive modified amino acid. This previous work was carried out starting with 5–10 nmol of highly purified protein. In the present study, only 5–10 pmol of gel-purified protein was required to confirm some of the GlcNAc modification sites previously reported (30) and reveal the presence of two novel GlcNAc modification sites and yet another phosphorylation site.

**EXPERIMENTAL PROCEDURES**

**Materials**—Water and acetonitrile (HPLC grade) were from Rathburn (Walkerton, Scotland, UK); formic acid and trifluoroacetic acid were from Romil (Cambridge, UK). Sequencing-grade modified trypsin (Promega, Southampton, UK), sequencing-grade chymotrypsin (Roche Applied Science), and proline-specific endopeptidase (Pro-C) (AMS Biotechnology, Abingdon, UK) were used for protein digestions.

**Serum Response Factor**—Human SRF was a kind gift from Dr. Rob Nicolas (Cancer Research UK, London, UK). The protein was overexpressed in baculovirus and then purified (31). This was in fact the same sample that was previously used to identify O-GlcNAc modification sites using a combination of mass spectrometry, radioactivity, and Edman sequencing (30). The sample was supplied in 20 mM Hepes, pH 7.9, 300 mM KCl, 0.2 mM EDTA, 0.2 mM EGTA, 0.1% Nonidet P-40, 10% glycerol, 1 mM dithiothreitol. The presence of detergents and salts made the sample unsuitable for direct analysis by mass spectrometry. Hence protein was purified by one-dimensional SDS-PAGE on a 10% polyacrylamide gel and visualized using Coomassie Brilliant Blue staining. A single band was observed as published previously (31).

**In-gel Digestion**—Excised gel bands were further destained in 25 mM NH4HCO3, 50% acetonitrile, reduced using dithiothreitol, alkylated using iodoacetamide, vacuum-centrifuged to dryness, and then allowed to swell in 5 μl of 25 mM NH4HCO3 containing 4 ng/μl modified trypsin for 5 min. Gel pieces were then overlaid with a further 25 μl of 25 mM NH4HCO3 and digested overnight at 37 °C. Peptides were extracted using two aliquots of 20 μl of 50% acetonitrile, 5% trifluoroacetic acid, vacuum-centrifuged to dryness, and resuspended in 10 μl of water prior to analysis.

**On-line LC-ESI-MS**—HPLC was carried out using the Ultimate/Famos/Switchos suite of instruments (LC Packings, Amsterdam, The Netherlands). Samples were loaded onto a guard column (300-μm inner diameter × 5-cm C18 PepMap) in the injection loop and washed using 0.1% formic acid at 40 μl/min for 2 min using the Switchos pump. Peptides were then separated on a nanoflow column (75-μm inner diameter × 15-cm C18 PepMap) at a flow rate of 200 nl/min using a gradient from 5 to 40% buffer B (80% acetonitrile, 0.1% formic acid) over a period of 32 min. Four 2-s MS-MS scans were performed on each precursor ion. For ions of m/z 400–900 two scans each at 28 and 32 V were acquired, whereas ions observed between m/z 900 and 2000 were fragmented with collision energies of 32 and 35 V.

For subsequent CID-MS-MS analysis of observed GlcNAc-modified peaks, using the automatically acquired CID-MS-MS spectra as a reference, collision energies were adjusted to optimize for the observation of GlcNAc-modified fragment ions. This resulted in reducing the collision energies to 26 and 28 V for the glycosylated peptides spanning residues 303–323 and 396–417. MS-MS was acquired on these precursor ions for the whole duration of their elution (30 s).

**HPLC Fractionation of Chymotryptic Digest**—HPLC was carried out using an ABI 1400 pump and 785A detector (Applied Biosystems). The digest was loaded onto a guard column (300-μm inner diameter × 5-cm C18 PepMap) in the injection loop and washed using 20 μl of 0.1% formic acid. Peptides were then separated on a microcolumn (300-μm inner diameter × 15-cm C18) at a flow rate of 5 μl/min using a gradient from 5 to 40% buffer B (80% acetonitrile) over a period of 32 min. Fractions were collected when peaks were observed on the UV chromatogram. Fractions were vacuum-centrifuged to dryness and resuspended in 4 μl of water for MALDI-MS screening.

**Pro-C Subdigestion**—5 μl (2 millions) of Pro-C in 25 mM NH4HCO3 was added to the HPLC fraction containing the GlcNAc-modified peptide. Digestion was carried out at 37 °C for 1 h.
MALDI-MS—0.5 l of each HPLC fraction was mixed on target with 0.5 l of 2,5-dihydroxybenzoic acid (saturated solution in water). Sample was allowed to dry at room temperature before introduction into the mass spectrometer. Spectra were acquired on the Reflex III (Bruker Daltonics, Bremen, Germany), surveying a range of m/z 600–4000. Spectra were externally calibrated from a nearby spot using a peptide calibration mixture (calibration mixture 2 from Sequazyme kit, Applied Biosystems).

RESULTS

A one-dimensional SDS-PAGE band of 5 pmol of serum response factor was digested with trypsin and analyzed by LC-MS with data-dependent selection of precursor ions for fragmentation analysis.

Characterization of Phosphorylated Ser101–Arg135—A triply protonated peak, [M + 3H]3+/m/z 1092.37 was observed in the LC-MS run and was automatically selected for CID fragmentation. The resulting spectrum is shown in Fig. 1, and the corresponding fragment ion identities are listed in Table I. This peptide component is a phosphorylated version of the peptide spanning residues 101–135 of SRF. This particular sequence is not a predicted tryptic fragment as although residue 135 is a lysine residue, residue 136 is a proline, and trypsin is not predicted to cleave after lysines that are followed by prolines (20). The peptide contains a known phosphorylation site at residue 103 (24–26). None of the “y” ions are observed in a phosphorylated state, but all “b” ions observed are either phosphorylated or have lost H3PO4 to produce an ion 18 Da smaller than the equivalent unmodified fragment ion. The smallest phosphorylated b ion is the b6 ion at m/z 757.42. This defines the phosphorylation to be on either serine 101 or serine 103. b1 and b2 ions are not observed, so there is no information in this spectrum that permits the differentiation between the two possible sites. However serine 103 is more likely as trypsin is unlikely to cleave after arginine 100 if serine 101 is phosphorylated due to steric hindrance of the bulky phosphate group.

Characterization of Phosphorylated Ala213–Arg235—Two additional triply protonated components were observed in the LC-MS analysis: [M + 3H]3+/m/z 861.77 and [M + 3H]3+/m/z 888.39. These correspond to residues 213–235 of SRF containing an acrylamide-modified cysteine residue and a phosphorylated version of this peptide. The crucial region in the MS-MS spectra that permits determination of the site of phosphorylation is shown magnified in Fig. 2. Both the non-phosphorylated and phosphorylated CID spectra contain a doubly protonated (charged) y11 ion at m/z 635.35. However, whereas the y12 ion in the spectrum of the non-phosphorylated peptide appears at m/z 678.88, in the phosphorylated spectrum a doubly charged peak at m/z 718.84 is observed instead. Also a doubly charged peak at m/z 669.89 is observed, which corresponds to the loss of H3PO4 from the phosphorylated y12 ion. Hence interpretation of this fragmentation spectrum establishes serine 224 as a new site of phosphorylation of SRF.

A number of GlcNAc-modified peptides were also observed in this tryptic digest, and some of these were selected for CID fragmentation. However, none of the resulting automatically acquired MS-MS spectra contained sufficient data to deter-

![Fig. 1. CID spectrum of [M + 3H]3+/m/z 1092.37 from a tryptic digest of SRF. This peptide is residues 249–270 of SRF, and the identities of peaks observed are given in Table I.](image)

| Peak | Match | Peak | Match | Peak | Match |
|------|-------|------|-------|------|-------|
| 147.14 | y1 | 795.55 | GGPEASAAAT-H2O | 1092.76 | y12 |
| 218.19 | y2 | 811.54 | VGGPEASAAA | 1109.25 | y26 |
| 275.23 | y3 | 813.55 | GGPEASAAAT | 1139.61 | b10-H2O |
| 362.30 | y4 | 829.55 | b8-H3PO4 | 1149.77 | y13 |
| 399.28 | b4-H3PO4 | 852.50 | b7-H2O | 1157.65 | b10 |
| 436.83 | y10 | 870.50 | b7 | 1174.74 | y27 |
| 461.38 | y5 | 872.66 | y10 | 1203.33 | y28 |
| 530.36 | b5-H3PO4 | 912.66 | y11-NH3 | 1206.80 | y14 |
| 552.36 | GGPEASAA-H2O | 927.59 | b8 | 1215.74 | b12-H3PO4 |
| 589.49 | y7 | 929.67 | y11 | 1238.76 | b11-H2O |
| 641.39 | b6-H3PO4-H2O | 951.65 | y22 | 1256.73 | b11 |
| 654.52 | y15 | 960.62 | b9-H3PO4 | 1259.77 | y29 |
| 659.41 | b6-H3PO4 | 972.70 | y23-H2O | 1272.76 | b13-H3PO4 |
| 676.51 | y8 | 981.72 | y23 | 1307.91 | y15 |
| 690.03 | y16 | 1001.19 | y24-H2O | 1313.76 | b12 |
| 712.48 | GGPEASAAA | 1010.71 | y24 | 1362.91 | b29-H2O |
| 725.52 | y17 | 1050.78 | y25-H2O | 1370.92 | b13 |
| 739.41 | b6-H2O | 1059.73 | y25 | 1378.86 | y16 |
| 757.42 | b6 | 1086.39 | MEKGMVVGPEASAA | 1449.98 | y17 |
| 772.53 | b7-H3PO4 | 1090.66 | GGPEASAAATGGY |
mine the identification of a particular site of modification (data not shown). This situation appeared to be due to the presence of a basic C-terminal residue formed by tryptic digestion that directed fragmentation toward C-terminally derived fragment ions, whereas overlapping N-terminal and C-terminal fragment ion series are generally required to identify sites of GlcNAc modification (17). Therefore, a second digest of 5 pmol of SRF was performed using chymotrypsin.

Characterization of Glycosylated Leu303–Leu323—A GlcNAc-modified peptide was observed at [M + 2H]^{2+} / m/z 1059.01 in this digest corresponding to a glycosylated peptide of residues 303–323. A CID-MS spectrum of this glycopeptide is shown in Fig. 3, and the relative peak identities are listed in Table II. The presence of four glycosylated y ions in this spectrum establishes that serines 307, 309, and 311 are not modified. A glycosylated b_{13} (b_{13G}) ion is also present in this spectrum at m/z 1369.80. The combination of b_{13G} and y_{13G} ions defines the site of glycosylation among the middle four residues, PSAV. Of these, only serine 313 can bear the modification. Thus interpretation of the ions present in this CID spectrum establishes a new site of GlcNAc modification on serine 313.

Characterization of Glycosylated Thr396–Tyr417—Another GlcNAc-modified peptide was observed at [M + 2H]^{2+} / m/z 1059.01. This peptide corresponds to residues 303–323 of SRF, and peaks observed are identified in Table III. The ions at m/z 1369.80 (b_{13G}) and m/z 1392.76 (y_{13G}) determine the modified residue to be serine 313.

The observation of a large number of fragment ions, the only glycosylated fragment ions present were y_{18G}–y_{20G}. There are also a large number of internal fragment ions whose formation was favored by the presence of multiple proline residues in the sequence, and three of these internal ions are observed in a glycosylated state. These glycosylated ions discount the four most N-terminal residues as sites of modification and restrict the GlcNAc modification site to one of threonine 401, threonine 402, or serine 411.

To determine which of these residues bear the modification, a second chymotryptic digest of SRF was carried out, this time using 10 pmol of gel-purified protein. This digest was separated by HPLC, and fractions were collected. Each fraction was then screened by MALDI-MS to locate the fraction that contained the GlcNAc-modified peptide corresponding to residues 396–417. The mass spectrum of fraction 9 is shown in Fig. 5. This fraction contains several peptides, but unfortunately the ions at m/z 2350.24, m/z 2366.20, m/z 2382.20, and m/z 2398.13 correspond to the non-glycosylated peptide 396–417 itself as well as analogs containing one, two, and three oxidized methionine residues, respectively. However, weak ions are observed at m/z 2553.45, m/z 2569.38, and m/z 2585.42, which correspond to the GlcNAc-modified analogs of this peptide itself and analogs bearing one and two oxidized methionine residues.

This fraction was then subdigested with Pro-C, which
cleaves peptides C-terminal to proline residues and analyzed by LC-MS. If this enzyme had worked efficiently, it should have digested the GlcNAc-modified peptide present to produce the sequence TTVGGHMMYP. Fig. 6A shows the extracted ion chromatograms of m/z corresponding to doubly charged peaks for this peptide in non-glycosylated and GlcNAc-modified states containing one and two oxidized methionine residues. A peak is observed in the chromatogram of the non-glycosylated, singly oxidized peptide, and a weaker peak is observed for the doubly oxidized version of this peptide. However, there are no peaks for GlcNAc-modified versions of this peptide since the peak after 28.5 min in the chromatogram of m/z relates to a singly charged peak at m/z (data not shown). Fig. 6B confirms that the two peaks in the ion chromatograms of the unmodified species relate to doubly charged peptides of the correct mass values. Hence apparently this peptide is present only in a GlcNAc-modified state.

Together these results suggest that the Pro-C enzyme normally cleaves this peptide after proline 400 to produce the peptide spanning residues 401–410. However, the presence of the GlcNAc moiety prevents cleavage at this site, leading to the N-terminally extended peptide spanning residues 396–410. Thus, it appears that the presence of a GlcNAc moiety provides sufficient steric hindrance to prevent the enzyme from attacking this site. This strongly suggests the site of modification is the immediately neighboring threonine 401. Indeed Reason et al. (30) proposed that a GlcNAc on a neighboring residue would prevent enzymatic cleavage. This previously unknown site is the first modification of threonine identified thus far in SRF; all other known sites are on serine residues.
In this work a well characterized phosphorylation site on serine 103 (24–26) was detected and confirmed, and a previously unobserved site of phosphorylation was identified at serine 224. This is immediately contiguous to a region determined to be important for DNA binding (32). Thus, phosphorylation at this site may regulate the ability of the protein to bind to the serum response element.

A GlcNAc modification site in the peptide spanning residues 303–323 was identified as serine 313. This represents a previously undetected site of modification because it was not observed in the work of Reason et al. (30). One of the glycopeptides Reason et al. focused on spans residues 313–324 and was reported to be modified on serine 316. The probable reason they did not observe GlcNAcylation of serine 313 may relate to the approach they used to detect the presence of GlcNAc-modified peptides. They used a sequential digestion approach using cyanogen bromide followed by trypsin and then a proline-specific endopeptidase (Pro-C). The resulting peptide mixture was then separated by HPLC, and fractions were collected and screened for pairs of peaks that differed by 203 Da in mass, which is characteristic of an unmodified and GlcNAc-modified peptide. However, if residue 313 was GlcNAc-modified, this would most likely prevent Pro-C from cleaving after proline 312 due to steric effects from the proximal sugar residue (further evidence for this effect is presented above in the analysis of the GlcNAc-modified peptide spanning residues 396–417). Hence this site will be cleaved only when serine 313 is unmodified, but if there is a bulky sugar moiety attached it will be inaccessible, so a glycosylated peptide of residues 306–324 would be the product of this combined enzymatic digestion. Thus, there would not be a pair of peptides differing by 203 Da in mass, so this particular modified peptide would not have been detected by their screening strategy.

It is interesting that this site appears to be heavily modified, but we observed no evidence of the presence of any modification at serine 313. This suggests that the stoichiometry of modification at serine 313 is considerably higher than any possibly present at serine 316. In fact, none of the previously identified GlcNAc modification sites in the peptide spanning residues 303–324 were detected in these studies, although these experiments use exactly the same protein preparation of SRF (31). This
may show different selectivity for these two very different approaches to identifying GlcNAc modification sites in a large protein and advocates the use of multiple GlcNAc detection methods to maximize information about a post-translational glycosylation state of the protein. By using radioactivity to locate GlcNAc-modified peptides, sites modified at lower stoichiometry might be detected and identified, although much more total protein was required. The mass spectrometry-based approach used in this work, although extremely sensitive, struggles to find sites that are modified at very low stoichiometry as it is a prerequisite of this approach that GlcNAc-modified peptides must have been detected and selected for MS-MS in the initial “screening” LC-MS run. To find these, selective enrichment of modified peptides could be used to augment the dynamic range accessible and simplify the sample (8, 33, 34).

A further site of GlcNAc modification was detected during this study. Fragmentation data from a GlcNAc-modified peptide observed in a chymotryptic digest (Fig. 4) narrowed this site of modification to either threonine 401, threonine 402, or serine 411, and then data from a Pro-C subdigest of this peptide provided strong evidence that threonine 401 is indeed the modified residue. It may be possible to confirm unambig-

**Fig. 6.** A, extracted ion chromatograms of the doubly charged versions of the peptide TTVGGHMMYP from an LC-MS analysis of a Pro-C digest of HPLC fraction 9 from a chymotryptic digest of SRF. A peak is observed after 23.7 min in the chromatogram for the non-glycosylated singly methionine-oxidized peptide (m/z 555.3), and a weak peak is observed after 21.2 min in the chromatogram of the non-glycosylated doubly methionine-oxidized peptide (m/z 563.3), but there was no evidence for GlcNAc-modified versions of this peptide at either oxidation state. (The peak after 28.5 min in the chromatogram of m/z 656.8 related to a singly charged peak at m/z 657.37.) B, ESI-MS spectra showing the peak in the extracted ion chromatogram of m/z 555.3 corresponds to [M + 2H]^2+ m/z 555.23, the predicted mass for the singly methionine-oxidized non-glycosylated peptide TTVGGHMMYP, and the peak in the extracted ion chromatogram of m/z 563.3 is formed by [M + 2H]^2+ m/z 563.27, the doubly methionine-oxidized version of the same peptide.

**Fig. 7.** A, extracted ion chromatograms of doubly charged versions of the peptide TSSVPTTVGGHMMYP from an LC-MS run of a Pro-C digest of HPLC fraction 9 from a chymotryptic digest of SRF. A peak is observed after 24.7 min in the chromatogram for the GlcNAc-modified singly methionine-oxidized peptide (m/z 892.4), and a peak is also observed after 23.2 min in the chromatogram of the GlcNAc-modified doubly methionine-oxidized peptide (m/z 563.3). However, there is no evidence for non-glycosylated versions of this peptide at either oxidation state. (The peak after 24.3 min in the chromatogram of m/z 892.4 relates to a singly charged peak at m/z 892.39.) B, ESI-MS spectra showing the peak in the extracted ion chromatogram of m/z 900.4 is formed by [M + 2H]^2+ m/z 900.39, the doubly methionine-oxidized version of the same peptide.
Novel GlcNAc Modification Sites of SRF

This work has successfully identified novel sites of O-GlcNAc modification of a protein without the use of radioactivity or derivatization of the modification site. Our results demonstrate that sufficient sensitivity can be achieved that it may no longer be necessary to resort to recombinantly overexpressing a protein to be able to characterize sites of O-GlcNAc modification. This study also represents the first reported instance when both phosphorylation and GlcNAcylation sites have been characterized from a single analytical strategy. This approach will hopefully stimulate increased interest in the study of protein GlcNAcylation that will improve our understanding of the function and activity of this protein.

CONCLUSIONS

This work has successfully identified novel sites of O-GlcNAc modification of a protein without the use of radioactive or derivatization of the modification site. Our results demonstrate that sufficient sensitivity can be achieved that it may no longer be necessary to resort to recombinantly overexpressing a protein to be able to characterize sites of O-GlcNAc modification. This study also represents the first reported instance when both phosphorylation and GlcNAcylation sites have been characterized from a single analytical strategy. This approach will hopefully stimulate increased interest in the study of protein GlcNAcylation that will improve our understanding of the function and activity of this protein.

Acknowledgment—We thank Rob Nicolas (Cancer Research UK, London, UK) for the generous gift of the recombinant SRF used for this study.

* This work was financially supported by the Ludwig Institute for Cancer Research, the Medical Research Council, and National Center for Research Resources Grant RR01614. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed: Dept. of Pharmaceutical Chemistry, University of California San Francisco, 521 Parnassus Ave., Rm. C18, San Francisco, CA 94143-0446. Tel.: 415-476-4893; Fax: 415-502-1655; E-mail: robertc@itsa.ucsf.edu.

REFERENCES

1. Torres, C., and Hart, G. (1984) Topography and polypeptide distribution of terminal N-acetylgalactosamine residues on the surface of intact lymphocytes. J. Biol. Chem. 259, 3308–3317
2. Haltiwanger, R. S., Bussy, S., Grove, K., Li, S., Mason, D., Medina, L., Moloney, D., Philippsberg, G., and Scarottozzi, R. (1997) O-Glycosylation of nuclear and cytoplasmic proteins: regulation analogous to phosphorylation? Biochem. Biophys. Res. Commun. 231, 237–242
3. Hart, G. W. (1997) Dynamic O-linked glycosylation of nuclear and cytoskeletal proteins. Annu. Rev. Biochem. 66, 315–335
4. Haltiwanger, R. S., Blomberg, M. A., and Hart, G. W. (1992) Glycosylation of nuclear and cytoplasmic proteins. Purification and characterization of a uridine diphospho-N-acetylgalactosamine:polypeptide β-N-acetylgalactosaminyltransferase. J. Biol. Chem. 267, 9005–9013
5. Kroppel, L. K., Blomberg, M. A., and Hart, G. W. (1997) Dynamic glycosylation of nuclear and cytosolic proteins. Cloning and characterization of a unique O-GlcNAc transferase with multiple tetrasacchidecabeptide repeats. J. Biol. Chem. 272, 9308–9315
6. Dong, D., and Hart, G. (1994) Purification and characterization of an O-GlcNAc selective N-acetyl-β-linked glycosaminidase from rat spleen cytosol. J. Biol. Chem. 269, 19321–19330
7. Gao, Y., Wells, L., Comer, F. I., Parker, G. J., and Hart, G. W. (2001) Dynamic O-glycosylation of nuclear and cytosolic proteins. Cloning and characterization of a neutral, cytosolic β-N-acetylgalactosaminidase from human brain. J. Biol. Chem. 276, 9838–9845
8. Vosseller, K., Sakabe, K., Wells, L., and Hart, G. W. (2002) Diverse regulation of protein function by O-GlcNAc: a nuclear and cytoplasmic carbohydrate post-translational modification. Curr. Opin. Chem. Biol. 6, 851–857
9. Chou, T., Hart, G., and Dang, C. (1995) c-Myc is glycosylated at threonine 58, a known phosphorylation site and a mutational hot spot in lymphomas. J. Biol. Chem. 270, 18961–18965
10. Cheng, X., Cole, R. N., Zaia, J., and Hart, G. W. (2000) Alternative O-glycosylation/O-phosphorylation of the murine estrogen receptor β. Biochemistry 39, 11609–11620
11. Cole, R. N., and Hart, G. W. (1999) Glycosylation sites flank phosphorylation sites on synapsin I: O-linked N-acetylgalactosamine residues are localized within domains mediating synapsin I interactions. J. Neurochem. 73, 418–428
12. Comer, F. I., and Hart, G. W. (2000) O-Glycosylation of nuclear and cytosolic proteins. Dynamic interplay between O-GlcNAc and O-phosphate. J. Biol. Chem. 275, 29179–29182
13. Roquemore, E. P., Chou, T. Y., and Hart, G. W. (1994) Detection of O-linked N-acetylgalactosamine (O-GlcNAc) on cytoplasmic and nuclear proteins. Methods Enzymol. 230, 443–460
14. Greis, K. D., Hayes, B. K., Comer, F. I., Kirk, M., Barnes, S., Lowary, T. L., and Hart, G. W. (1996) Selective detection and site-analysis of O-GlcNAc-modified glycopeptides by β-elimination and tandem electrospray mass spectrometry. Anal. Chem. 234, 38–49
15. Rademaker, G., Pergantis, S., Blok-Tip, L., Langridge, L., Kleen, A., and Thomas-Oates, J. (1998) Mass spectrometric determination of the sites of O-glycan attachment with low picomolar sensitivity. Anal. Biochem. 257, 149–160
16. Wells, L., Vosseller, K., Cole, R. N., Cronshaw, J. M., Matnis, M. J., and Hart, G. W. (2002) Mapping sites of O-GlcNAc modification using affinity tags for serine and threonine post-translational modifications. Mol. Cell. Proteomics 1, 791–804
17. Chalkley, R. J., and Burlingame, A. L. (2001) Identification of GlcNAcylated sites of peptides and α-crystallin using O-TOF mass spectrometry. J. Am. Soc. Mass Spectrom. 12, 1106–1113
18. Treisman, R. (1990) The SRE: a growth factor responsive transcriptional regulator. Semin. Cancer Biol. 1, 47–58
19. Hill, C. S., Wynne, J., and Treisman, R. (1995) The Rho family GTPases Rap1, Rac1, and GDH2s regulate transcriptional activation by SRF. Cell 81, 1159–1170
20. Thiede, B., Lamer, S., Mattow, J., Siejak, F., Dimmler, C., Rudel, T., and Jungblut, P. R. (2000) Analysis of missed cleavage sites, tryptophan
Novel GlcNAc Modification Sites of SRF

oxidation and N-terminal pyroglutamylation after in-gel tryptic digestion. *Rapid Commun. Mass Spectrom.* **14**, 498–502

21. Arsenian, S., Weinhold, B., Oelgeschläger, M., Ruther, U., and Nordheim, A. (1998) Serum response factor is essential for mesoderm formation during mouse embryogenesis. *EMBO J.* **17**, 6289–6299

22. Croissant, J. D., Kim, J. H., Eichele, G., Goering, L., Lough, J., Prywes, R., and Schwartz, R. J. (1996) Avian serum response factor expression restricted primarily to muscle cell lineages is required for *H9251*-actin gene transcription. *Dev. Biol.* **177**, 250–264

23. Gauthier-Rouviere, C., Cavadore, J. C., Blanchard, J. M., Lamb, N. J., and Fernandez, A. (1991) p67SRF is a constitutive nuclear protein implicated in the modulation of genes required throughout the G1 period. *Cell Regul.* **2**, 575–588

24. Rivera, V. M., Miranti, C. K., Misra, R. P., Ginty, D. D., Chen, R. H., Blenis, J., and Greenberg, M. E. (1993) A growth factor-induced kinase phosphorylates the serum response factor at a site that regulates its DNA-binding activity. *Mol. Cell. Biol.* **13**, 6260–6273

25. Heidenreich, O., Neininger, A., Schratt, G., Zinck, R., Cahill, M. A., Engel, K., Kotlyarov, A., Kraft, R., Kostka, S., Gaestel, M., and Nordheim, A. (1999) MAPKAP kinase 2 phosphorylates serum response factor *in vitro* and *in vivo*. *J. Biol. Chem.* **274**, 14434–14443

26. Miranti, C. K., Ginty, D. D., Huang, G., Chatila, T., and Greenberg, M. E. (1995) Calcium activates serum response factor-dependent transcription by a Ras- and Elk-1-independent mechanism that involves a Ca2+/calmodulin-dependent kinase. *Mol. Cell. Biol.* **15**, 3672–3684

27. Manak, J. R., de Bisschop, N., Kris, R. M., and Prywes, R. (1990) Casein kinase II enhances the DNA binding activity of serum response factor. *Genes Dev.* **4**, 955–967

28. Liu, S. H., Ma, J. T., Yueh, A. Y., Lees-Miller, S. P., Anderson, C. W., and Ng, S. Y. (1993) The carboxyl-terminal transactivation domain of human serum response factor contains DNA-activated protein kinase phosphorylation sites. *J. Biol. Chem.* **268**, 21147–21154

29. Schröter, H., Mueller, C. G., Meese, K., and Nordheim, A. (1990) Synergism in ternary complex formation between the dimeric glycoprotein p67SRF, polypeptide p62TCF and the c-fos serum response element. *EMBO J.* **9**, 1123–1130

30. Reason, A. J., Morris, H. R., Panico, M., Marais, R., Treisman, R. H., Haltiwanger, R. S., Hart, G. W., Kelly, W. G., and Dell, A. (1992) Localization of O-GlcNAc modification on the serum response transcription factor. *J. Biol. Chem.* **267**, 16911–16921

31. Marais, R. M., Hsuan, J. J., McGuigan, C., Wynne, J., and Treisman, R. (1992) Casein kinase II phosphorylation increases the rate of serum response factor-binding site exchange. *EMBO J.* **11**, 97–105

32. Norman, C., Runswick, M., Pollock, R., and Treisman, R. (1988) Isolation and properties of cDNA clones encoding SRF, a transcription factor that binds to the c-fos serum response element. *Cell* **55**, 989–1003

33. Hayes, B. K., Greis, K. D., and Hart, G. W. (1995) Specific isolation of O-linked N-acetylgalcosamine glycopeptides from complex mixtures. *Anal. Biochem.* **228**, 115–122

34. Haynes, P. A., and Aebersold, R. (2000) Simultaneous detection and identification of O-GlcNAc-modified glycoproteins using liquid chromatography-tandem mass spectrometry. *Anal. Chem.* **72**, 5402–5410

35. Horne, W. C., Prinz, W. C., and Tang, E. K. (1990) Identification of two cAMP-dependent phosphorylation sites on erythrocyte protein 4.1. *Biochim. Biophys. Acta* **1055**, 87–92

36. Inaba, M., and Maede, Y. (1989) O-N-Acetyl-o-glucosamine moiety on discrete peptide of multiple protein 4.1 isoforms regulated by alternative pathways. *J. Biol. Chem.* **264**, 18149–18155