Peptide profiling by capillary separation techniques coupled to mass spectrometry

Gaspari, Marco

Chapter 6

Protein transacetylase of buffalo liver: characterization and mass spectrometry of the acetylated protein product

Discovery of modified peptides in protein tryptic digests by automated handling of capillary LC-MS data

Kohli E, Gaspari M, Raj HG, Parmar VS, Sharma SK, van der Greef J, Kumari R, Gupta G, Seema, Khurana P, Tyagi YK, Watterson AC, Olsen CE "*Acetoxy drug: protein transacetylase of buffalo liver-characterization and mass spectrometry of the acetylated protein product"* **Biochim Biophys Acta- Proteins and Proteomics** 2004 ;1698(1):55-66.

Gaspari M, Kohli E, Raj HG, Verheij E, van der Greef J"*Discovery of modified peptides in protein tryptic digests by automated handling of capillary LC-MS data"* **Proceedings of the 19th LC-MS Symposium** (Montreux 2002).

Protein transacetylase of buffalo liver: characterization and mass spectrometry of the acetylated protein product

The purification and characterization of the buffalo liver microsomal transacetylase (TAase) catalyzing the transfer of acetyl groups from a model acetoxy drug: 7,8-diacetoxy-4-methylcoumarin (DAMC) to GST3–3 has been described here. The enzyme was routinely assayed using DAMC and cytosolic GST as the substrates and was partially purified from microsomes of the buffalo liver. The action of TAase and DAMC on liver cytosolic GST resulted in the formation of monoacetoxymonohydroxy-4-methylcoumarin (MAMHC) and 7,8-dihydroxy-4-methylcoumarin (DHMC), although the former was the major metabolite. GST3–3 was used as a model protein substrate for the action of TAase using DAMC as the acetyl donor. Seven potential acetylation points were discovered by MALDI-TOF MS analysis of tryptically digested GST 3-3 after incubation with Taase and an acetyl donor. All acetyaltions were confirmed by capillary LC/MS/MS analysis. The structure of acetylated GST revealed that the lysines that underwent acetylation were peripheral in positions.

Introduction

The current knowledge on biological protein acetylation is largely confined to acetyl CoA-dependent acetylation of proteins catalyzed by specific acetyltransferases. Histones, the tumor-specific protein, p53, acyl carrier protein are some examples of proteins that are acetylated by the action of acetyl CoA-dependent acetyl transferases[1-3]. Proteins are also targeted by the acetylating drugs such as aspirin that can acetylate a protein such as cyclooxygenase[4], which is an example of non-enzymatic, independent of acetyl CoA protein acetylation. Puzzlingly, nothing was known about the enzymatic acetylation of protein independent of acetyl CoA. We have discovered a membrane bound enzyme[5,6] catalyzing the transfer of acety groups from acetoxy drugs, such as acetoxycoumarins and other acetyl polyphenols to glutathione S-transferase (a receptor protein). A preliminary report in this regard evidenced the acetylation of GST3–3 by 7,8-diacetoxy-4-methylcoumarin (DAMC) catalyzed by acetoxy drug: protein transacetylase (Taase)[7]. We have in this communication described the purification and characterization of buffalo liver TAase and details of GST3–3 acetylation by the application of matrix assisted laser desorption/ionization–time of flight–mass spectrometry (MALDI-TOF MS) and liquid chromatography coupled to tandem mass spectrometry LC/MS/MS.

Materials and methods

Chemicals

DEAE Sepharose, Sephacryl HR-200 and gel filtration calibration kits were obtained from Amersham Pharmacia Biotech, USA. Acrylamide, bisacrylamide, sodium dodecyl sulfate (SDS) and ammonium persulfate were the products of Bio-Rad Laboratories (India). DAMC was synthesized in our laboratory[5]. Sequencing grade trypsin was purchased from Boehringer Mannheim (Germany), purified rat GST3–3 was kindly provided by Dr. Jan Bogaards, TNO Nutrition and Food Research, Zeist (The Netherlands). Reduced glutathione (GSH), purified GST (Cat. No. 8386) and 1-chloro-2,4-dinitrobenzene (CDNB) were purchased from Sigma Chemical Company, St. Louis, MO (USA). All other laboratory reagents were of high grade and obtained from the local suppliers.

Purification of TAase from buffalo liver

Buffalo liver was obtained from the municipal slaughterhouse, Idgah, Delhi. Homogenates of the liver (20%, w/v) were prepared in 10 mM phosphate buffer containing 0.25 M sucrose and 1.4 mM h-mercaptoethanol at pH 7.0. Microsomes and cytosolic fraction were prepared as described earlier[5] and stored at -20º C till further use.

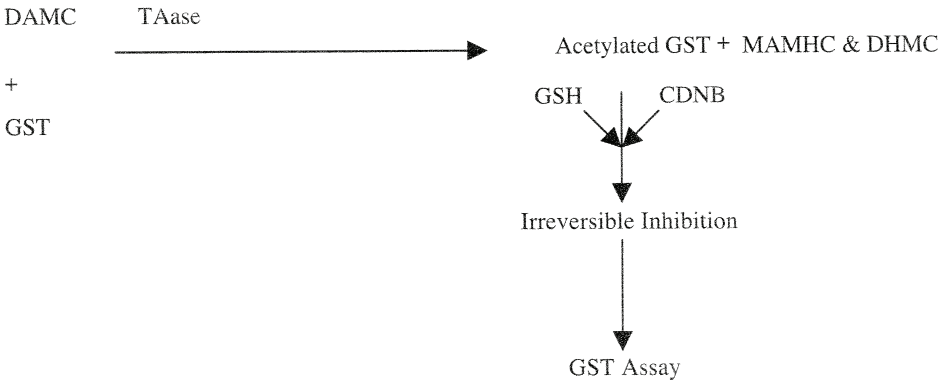
Assay of Taase

TAase was assayed[5] using DAMC and cytosolic GST as the substrates as per the details given in our earlier report. The assay mixture consisted of 0.25 M phosphate buffer (pH 6.5), liver microsomes (25 μg protein), DAMC (100 μM) added in 50 μl DMSO, purified GST (10–15 μg protein) and water to make up 0.5 ml. The contents of the tube, scaled up as per requirement were preincubated at 37º C for various periods. The aliquots were removed periodically into a spectrophotometer cuvette containing CDNB and GSH to make up their concentration (1 mM) in a total volume of 1 ml and the progress of the GST activity was followed at 340 nm using a Cary spectrophotometer (Cary Bio100). In control samples, DAMC was replaced by DMSO. The unit of TAase was expressed in terms of percentage inhibition of GST under the conditions of the assay and ensured that the reaction was linear with respect to enzyme concentration and incubation time.

Table 1 Purification of TAase from buffalo liver.

Preparation stage	Total protein (mg)	Total units	Specific activity (units/mg)	Yield (%)	Fold purification
Homogenate	400	30,000	75	100	—
Solubilized supernatant	104	17,000	163	56.66	2.17
Ion exchange chromatography	5.444	3600	667	21.17	9.0
gel filtration chromatography	0.3	400	1340	11.11	18.00

Purification of TAase from buffalo liver. Buffalo liver was solubilized by 1 M phosphate buffer (pH 7.4) for 30 min, centrifuged at 105,000 x g for 1.25 h; ammonium sulfate was added to the clear supernatant to reach 45 – 75% salt concentration, which had maximum TAase activity, dialyzed and chromatographic procedures were conducted. The unit of TAase was expressed in terms of percentage inhibition of GST under the conditions of the assay.



MAMHC: Monoacetoxy monohydroxy-4-methylcoumarin

DHMC : 7,8-Dihydroxy-4-methylcoumarin

Figure 1 TAase-catalyzed reaction.

Solubilization of buffalo liver microsomes

The method described by Dey et al.[8] was adopted. Microsomal pellet was thawed and resuspended by homogenization in 1.0 M potassium phosphate buffer of pH 7.4 (0.5 mg/ml protein). The mixture was stirred on a magnetic stirrer for 30 min at 4º C and then centrifuged at 105,000 g for 1.25 h. The clear supernatant was decanted and used as the source of TAase. The clear supernatant was dialyzed against 10 mM potassium phosphate buffer of pH 7.2. The solubilized enzyme was subjected to ammonium sulfate fractionation in a stepwise manner by the addition of solid ammonium sulfate

at 4º C with continuous stirring to reach the desired saturation of the salt. The mixture was stirred for 40 min and centrifuged at 10,000 g for 30 min. The precipitated protein was kept separately and ammoni ammonium sulfate saturation was further raised. Accordingly, the supernatant was raised to 25%, 45%, 50% and up to 95% ammonium sulfate saturation. The precipitated proteins were separately dissolved in a known volume of the loading buffer containing 1.4 mM h-mercaptoethanol and vigorously dialyzed and TAase activity and protein was assayed in all the fractions. Maximum TAase activity was precipitated between 45% and 75% ammonium sulfate saturation.
The active fraction obtained above was loaded onto a 3 ml DEAE-Sepharose column at 4º C and unbound proteins were eluted with loading buffer (3 ml fractions) and read at 280 nm. The fractions with appreciable absorbency at 280 nm were assayed for TAase activity, if any. The enzyme was eluted with increasing molar concentration of NaCl in the loading buffer. The fractions of 1 ml volume were read at 280 nm followed by the assay of TAase. The active fractions were pooled and diluted four times with loading buffer and subjected to FPLC (Amersham Pharmacia Biotech) using Resource Q column. The NaCl gradient was run for 10 min with elution volume of 15 ml, 1 ml fractions were collected and assayed for TAase activity; the enzyme was optimally eluted with 0.25 M NaCl. The active fractions were pooled and fractionated by gel filtration using preswollen Sephacryl HR-200 which was equilibrated with 10 mM potassium phosphate buffer, pH 7.2 containing 0.02% sodium azide and eluted by collecting 1 ml fractions and their absorption at 280 nm was monitored. The active fractions were pooled and dialyzed against 10 mM phosphate buffer containing 1.4 mM h-mercaptoethanol. SDS PAGE (Figure 2) was performed by following the procedure of Lamelli[9].

Characterization of buffalo liver microsomal TAase

Kinetics of buffalo liver microsomal Taase Effect of varying substrate concentration on liver microsomal TAase. The assay described above with the addition of varying amounts of partially purified enzyme was carried out. Partially purified TAase (25 Ag protein) was mixed separately with the fixed concentration of DAMC (50–400 AM, in 0.05 ml DMSO), 0.25 M phosphate buffer (pH 6.5), purified GST (12 Ag protein) and water to make up the volume of 0.8 ml. The contents of the tube were incubated for 10 min, followed by the addition of GSH and CDNB required for the assay of the cytosolic GST as described earlier. The percentage inhibition of GST (initial rate) was plotted against the concentration of DAMC to compute the kinetics parameters Km

and Vmax. The effect of varying concentration of the second substrate (GST) was quantified by performing the assay using a fixed concentration of DAMC.

Identification of the products of buffalo liver Taase-catalyzed reaction The incubation mixture consisted of Tris–HCl (0.025 M, pH 7.0), partially purified TAase (25 μg protein), DAMC (100 μM) and GST (0.095 units), the volume was adjusted to 1.6 ml with water. The mixture was incubated at 37º C for 10 min. The tubes were later covered with glass marbles and placed in boiling water for 10 min, cooled and 2 ml of ethyl acetate was added. The contents of the tubes were mixed using a vortex shaker in order to extract the metabolites. The tubes were centrifuged at 2000 rpm for 10 min and a clear organic layer was aspirated in another tube. The aqueous mixture was re-extracted twice, the organic layers were pooled and evaporated to dryness by blowing N2 gas. The products were dissolved in 1 ml methanol, and the metabolites were separated by HPLC using Waters 996 Chromatograph fitted with C-18 column. Methanol extract (20 μl) prepared as described earlier was injected and isocratically eluted with methanol/water (60:40 v/v) using a diode array detector. Authentic samples of DAMC and DHMC were also chromatographed to identify them in the sample (Figure 4).

Incubation of GST and TAase for MS characterization

The reaction mixture containing partially purified buffalo liver TAase (50 μg), recombinant GST3–3 isoform (50 μg), and DAMC (200 μM) in 0.25 M phosphate buffer (pH 6.5) was incubated for 30 min at 37º C, and the reaction mixture was subjected to SDS-PAGE in order to separate modified GST3–3 from other contaminating proteins. The reaction mixture to which DMSO was added in place of DAMC served as the control. The gel plugs (modified and unmodified GST3–3) were separately washed and treated with DTT and iodoacetamide to alkylate the cysteine residues present in the proteins.

In-gel protein reduction and alkylation

After staining and destaining of gel, the gel was rinsed with water and the band of interest (a GST3–3 subunit) was excised with a clear scalpel. In order to reduce background gel, the bands were cut close to the edges. The gel bands were processed in a clear-air cabinet to prevent keratin contamination. The excised bands were chopped into little pieces of approximately 1x1 mm and gel particles were transferred to the 0.5 ml microcentrifuge. The gel particles were washed for 4 min with 100 μM

NH_4HCO_3, vortexed and 100 μl acetonitrile added and gel particles were incubated for 10 min until the shrunk gel pieces became opaque and stuck together. The liquid was discarded and washing was repeated. Avolume of 10 mM DTT solution in 100 mM NH_4HCO_3 sufficient to cover the gel pieces (approximately 150 μl) was added and the proteins were reduced at 56º C for 45 min. After cooling at room temperature, the DTT solution was replaced with roughly the same volume of alkylation buffer containing 55 mM iodoacetamide in 100 mM NH_4HCO_3. After 45 min of incubation at ambient temperature in the dark with occasional vortexing, the gel pieces were washed with 150 μl of 100 mM NH_4HCO_3 for 10 min, dehydrated by addition of acetonitrile, swollen by rehydration in 100 mM NH_4HCO_3 and shrunk again by addition of acetonitrile. The liquid phase was removed and gel was completely dried.

In-gel digestion and MALDI-TOFMS analysis

After reduction and alkylation, the gel plugs were incubated overnight with sequence grade trypsin[10]. The peptide mixture obtained after trypsin digestion was diluted four times with the matrix 2,5-dihydroxybenzoic acid (20 mg/ml) in 1:1 acetonitrile/0.1% TFA v/v (1 μl of the mix), and deposited on a MALDI target and allowed to dry in order to perform MALDI-TOF MS analysis using Bruker Biflex III MALDI-TOF mass spectrometer (Bruker Daltonics, Bremen, Germany) equipped with a nitrogen laser emitting pulsed UV light at 337 nm and operated in a reflectrone mode. External linear calibration was achieved using a mixture of angiotensin II, substance P, bombesin, adrenocorticotropic hormone (fragments 18–39) and somatostatin. Spectra were acquired in the positive ion mode and the signal was accumulated and averaged over 100–200 shots. For intact protein measurements, samples were purified by zip tips (Millipore Corporation, Bedford, MA) according to the procedure suggested by the manufacturer. Desalted samples were mixed 1:1 with matrix, consisting of a saturated solution of sinapinic acid in 0.1% trifluoroacetic acid/acetonitrile (2:1, v/v), and deposited on the MALDI target. Acquisition was performed in linear positive mode. Tuning and calibration of the mass spectrometer were established using a standard myoglobin solution.

LC-MS/MS

LC-MS/MS measurements were conducted on an LCQ DECA (Thermoquest, San Jose, CA, USA) using conven-tional ESI source in positive ion mode detection. The spray voltage was set at 3.8 kV and the heated capillary temperature at 300 jC. The eluent flow of 25 μl/min was provided by an Eldex micro-LC (separations). Analysis

was conducted on a 15 cm 800 µm i.d. column packed with 5 µm spherisorb C_{18} reverse phase material (LC Packings, Amsterdam, The Netherlands). The injection volume was 10 µl. Gradient elution was performed by using the following mobile phases: (A) 10 mM ammonium acetate in 0.1% HCOOH (v/v); (B) 10 mM ammonium acetate in 0.15% HCOOH (v/v) and 80% acetonitrile (v/v). Gradient: from 5% to 30% B in 25 min, from 30% to 60% B in 10 min, from 60% to 100% B in 5 min, down to 5% B again in 2 min. Two full-scan (*m/z* range 200–2000) LC-MS runs were initially acquired for modified and non-modified GST, respectively. Potentially acetylated peptides previously detected by MALDI-TOF experiments were found back in LC-MS total ion current (TIC) of the modified GST as singly or doubly protonated molecules, though their signal intensity in full-scan MS was rather weak compared to other peptides present in the digest mixture. A second LC-MS experiment with acquisition windows was designed in order to perform MS/MS on these seven potentially acetylated peptides. In this second LC-MS experiment, the mass spectrometer program was switched to MS/MS mode in correspondence with the retention time of the abovementioned peaks. Precursor ions were selected with a window of 2.5 *m/z*.

Results

Acetoxy drug: TAase was purified from buffalo liver (Table 1). The enzyme was found to have a molecular weight of around 65 kDa by SDS-PAGE (Figure 2).

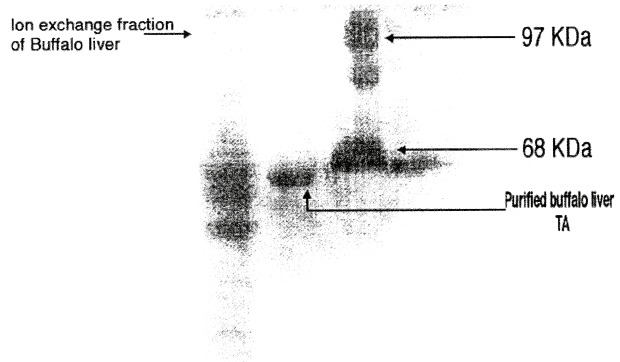


Figure 2 SDS-PAGE of buffalo liver TAase.

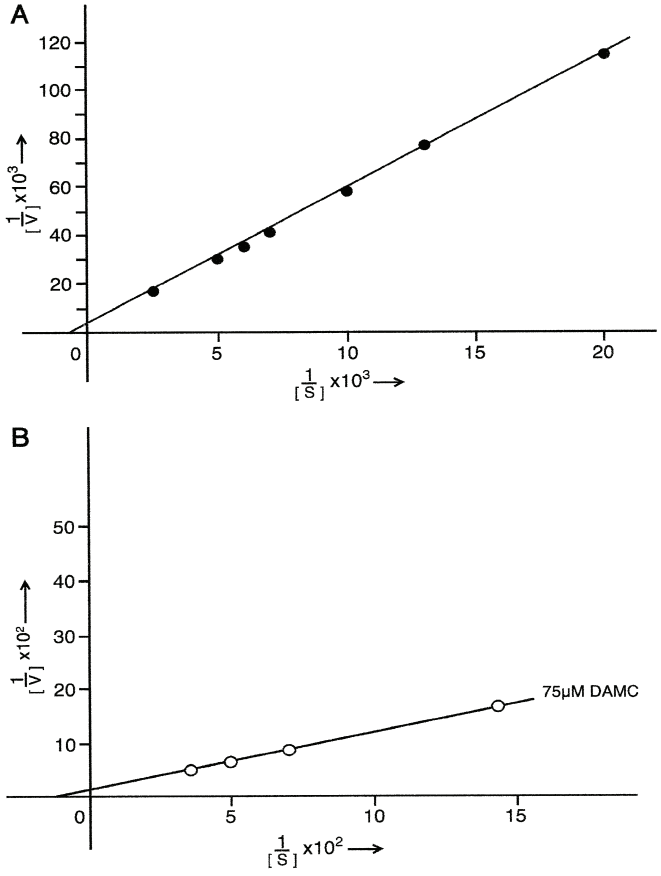


Figure 3 (a) Kinetics of buffalo TAase: effect of the concentration of DAMC (50– 400 μM). The concentration of purified GST was 2 μg. (b) Kinetics of buffalo TAase: effect of varying concentrations of purified protein (5– 30 μg). Concentration of DAMC was 75 μM.

The kinetics of partially purified enzyme is highlighted in Figure 3ab. The enzyme exhibited hyperbolic kinetics, yielding Km (1667 μM) and Vmax (192 units), when the concentration of one of the substrates, i.e. DAMC was varied keeping the concentration of GST constant. Similar kinetics patterns were also apparent when the concentration of GST was varied, keeping the concentration of DAMC constant. The GST preparation used in the kinetics analysis consisted of a mixture of isoforms; hence, the molar concentration of GST was not expressed. The aim of the kinetic analysis was to demonstrate that GST acted as the substrate for the purified TAase, which is clear from the presented data in Figure 3b.

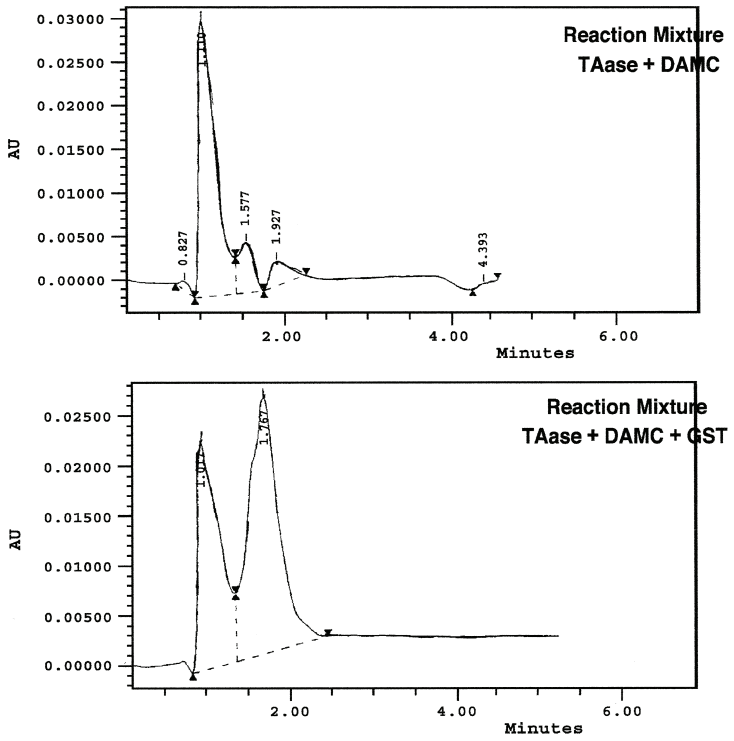


Figure 4 Product profile of buffalo TAase-catalyzed reaction. Purified TAase (25 µg protein) was incubated with Tris –HCl (0.025 M), pH 7.0, DAMC (100 µM) and purified GST (25 µg protein) for 10 min at 37º C. The products were extracted and separated by HPLC as described in Materials and Methods.

The products of buffalo liver TAase-catalyzed reaction were analyzed by HPLC as described in Figure 4. Buffalo liver TAase when incubated with DAMC and GST was found to result in the conversion of DAMC to DHMC, significantly enhanced formation of the monohydroxy product (peak II) having the retention time in between DAMC and DHMC was noticed. Hence, the product represented by peak II was tentatively termed monoacetoxy monohydroxycoumarin (MAMHC) with concomitant decrease in the concentration of DAMC (Table 2). These results indicate that in the presence of receptor protein GST, DAMC is acted upon by the TAase resulting in the enhanced formation of the hydroxycoumarin, MAMHC.

Table 2 Product profile of buffalo liver TAase-catalyzed reaction

Reaction mixture	Product formed (%)		
	DHMC	MAMHC	DAMC
(1) Enz + DAMC	77.30	7.60	13.17
(2) Enz + DAMC + GST	39.95	61.05	–

Demonstration of the acetylation of intact GST3–3 by DAMC catalyzed by TAase

In these studies, recombinant GST3–3 was used as the substrate for TAase. Accordingly, GST3–3 was incubated with purified buffalo liver TAase followed by purification of GST3–3 by zip tip and MALDI-TOF MS on intact GST was performed. Unmodified GST is detected around 25.8 kDa (theoretical molecular weight 25,783 Da) (Figure 5a). The second peak at around 26.0 kDa and the shoulder at 26.2 kDa are likely to be adducts with the sinapinic acid matrix (MW= 224 Da). Modified GST appears as a broad, unre- solved peak composed of several species with a different degree of acetylation (Figure 5b). The maximum is at 26.0 kDa, accounting for multiple acetylation. No matrix adducts are observed due to overlapping of peaks and insufficient resolution.

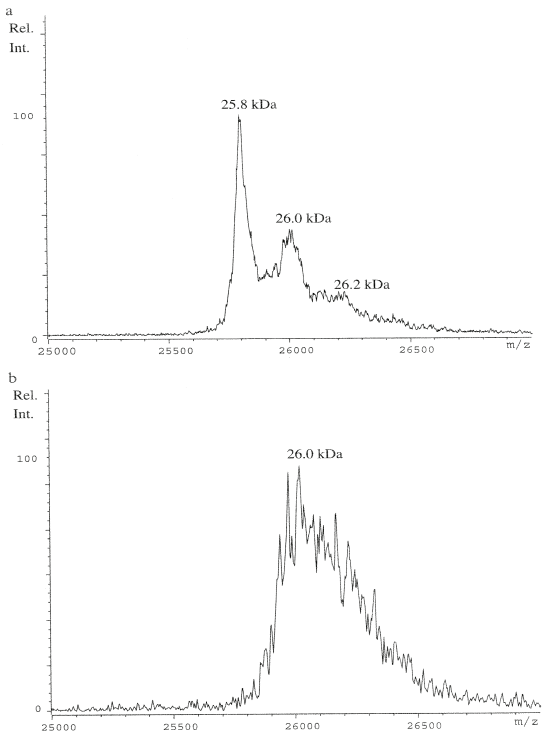


Figure 5 (a) MALDI-TOFMS spectrum of intact control GST3– 3. (b) MALDI-TOFMS spectra of intact modified GST3– 3.

Identification of TAase-catalyzed acetylation of amino acid residues of GST3–3 by DAMC

The GST3–3 incubated with DAMC and purified Taase was separated on SDS-PAGE. Tryptic peptides were extracted from the gel pieces and analyzed by MALDI-TOF MS. The peptides were identified using the Mascot Search Engine www.matrixscience.com). The peptide map so obtained matched to GTM1-RAT: glutathione S-transferase YB1, taxonomy Rattus norvegicus, nominal mass of protein is 25,937 Da.

MALDI-TOFMS peptide maps represented an excellent starting point in order to address possible covalent modifications in the amino acid sequence of GST3–3 incubated with DAMC and purified TAase. These modifications could be individuated by the appearance on new peaks in the modified GST3–3 tryptic map, not actually present in the control tryptic map. Overlying MALDI peptide maps of modified and control GST recognized seven new peaks corresponding to potentially acetylated peptides, recognized in the modified maps. The mass value of each of them, listed in Table 3, was 42 Da higher than the theoretical mass of native GST peptides, strongly suggesting acetylation (Figure 6a–g).

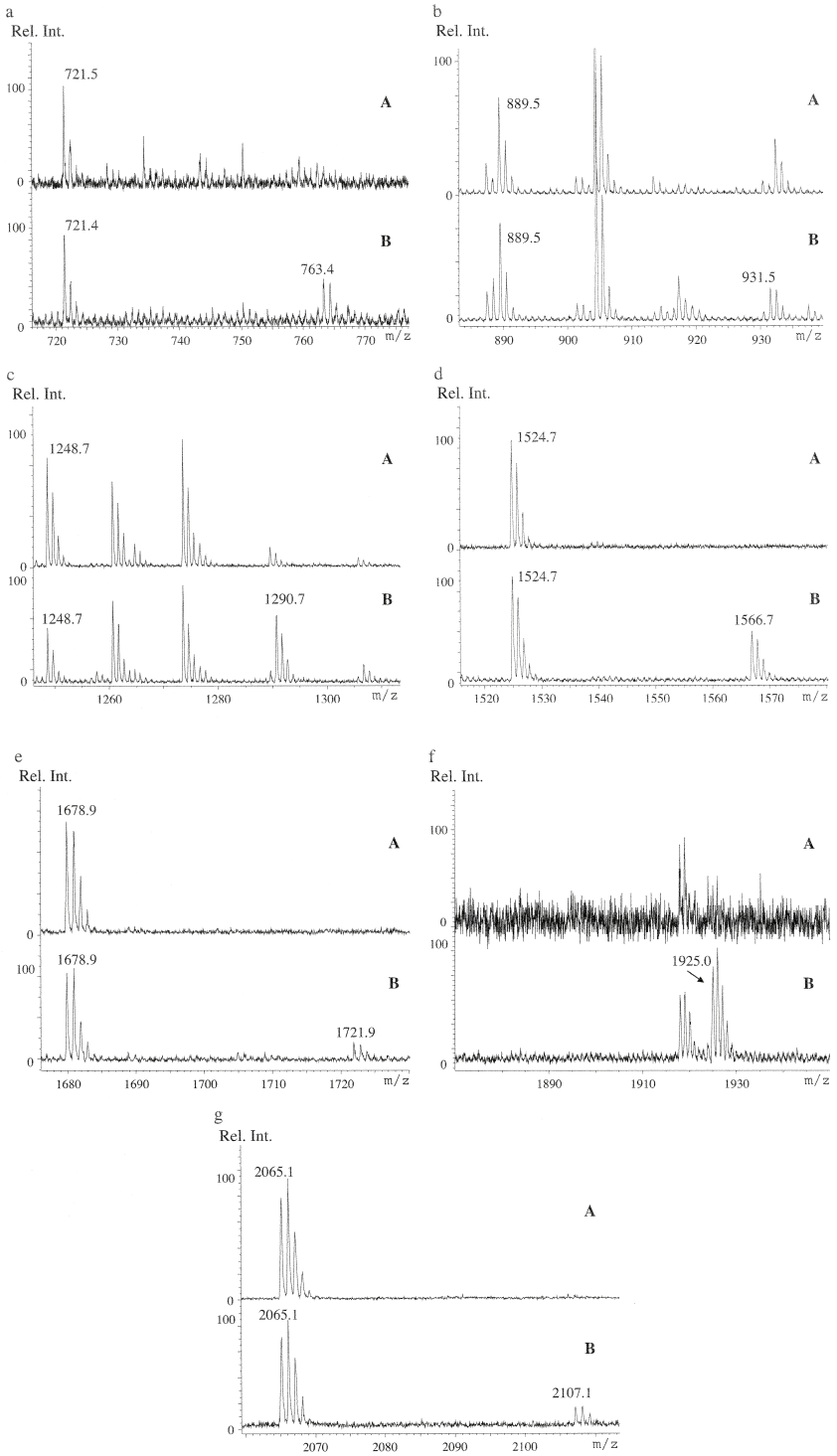


Figure 6 (a– g) Relevant sections of MALDI-TOFMS tryptic maps for (A) control and (B) modified GST3– 3. New peaks appearing in the modified map are labeled together with their putative corresponding nonmodified tryptic peptide at +42 *m/z*. Peptide all detected as $[M+H]^+$ ions, are labeled with their corresponding *m/z* values.

Table 3 Assignment of the peptide masses observed in the MALDI-TOFMS spectra of in-gel tryptic digest extracts of control and modified GST3 – 3 (masses and delta masses are in Da).

Control GST3 – 3	Modified GST3 –3	Predicted mass for GST3 – 3 digest	Delta mass	Comment
1247.7	1247.7	1247.7	0.0	T1–10
–	1289.7	1247.7	42.0	T1–10 possibly acetylated
792.5	792.5	792.5	0.0	T11 – 17
1588.8	1588.8	1588.8	0.0	T18 – 30
1744.9	1744.9	1744.9	0.0	T18 – 31
2158.0	2158.0	2157.9	0.1	T32 – 49
903.5	903.4	903.4	0.0	T43 – 49
2064.1	2064.1	2064.1	0.0	T50 – 67
–	2106.1	2064.1	42.0	T50 – 67 possibly acetylated
1789.0	1789.0	1788.9	0.1	T52 – 67
1917.1	1917.0	1917.0	0.0	T52 – 68
1160.7	1160.7	1160.6	0.1	T68 – 77
1032.6	1032.5	1032.5	0.0	T69 – 77
1523.7	1523.7	1523.7	0.0	T82 – 93
–	1565.7	1523.7	42.0	T82 – 93 possibly acetylated
1671.9	1671.9	1671.8	0.1	T94 – 107
1402.7	1402.7	1402.7	0.0	T96 – 107
1800.9	1800.9	1800.8	0.1	T108 – 121
888.5	888.5	888.5	0.0	T122 – 128
–	930.5	888.5	42.0	T122 – 128 possibly acetylated
955.5	955.5	955.5	0.0	T136 – 143
975.5	975.5	975.5	0.0	T144 – 151
2588.4	2588.4	2588.3	0.1	T152 – 172
1076.6	1076.6	1076.5	0.1	T173 – 181
1678.9	1678.9	1678.9	0.0	T173 – 186
–	1720.9	1678.9	42.0	T173 – 186 possibly acetylated
620.4	620.4	620.3	0.1	T182 – 186
720.5	720.4	720.4	0.0	T187 – 192
–	762.4	720.4	42.0	T187 – 192 possibly acetylated
–	1924.1	1882.0	42.1	T202 – 217 possibly acetylated
1055.6	1055.6	1055.6	0.0	T202 – 210
845.4	845.4	845.4	0.0	T211 – 217

Confirmation of the acetylation of the GST tryptic peptides by LC-MS/MS

Acetylation at several points in the GST sequence was suggested by MALDI, but final confirmation was possible by using LC-MS/MS. Tryptic peptides originated from the modified GST in-gel digestion were separated by micro- RPLC and peptides of interest were fragmented in an ion trap mass spectrometer in order to get structural information and to pinpoint amino acid modifications. Good quality MS/MS spectra were obtained for seven peptides for which putative acetylation had been previously assigned by MALDI-TOF MS. By examining the fragmentation patterns and by comparing the experimental and predicted values for MS/MS daughter ions, the identity of the seven acetylated GST tryptic peptides was confirmed. The exact

location of acetylation was identified in all seven peptides: the N-terminal proline and six lysines: Lys_{51}, Lys_{82}, Lys_{123}, Lys_{181}, Lys_{191} and Lys_{210}.

Table 4 Summary of MS/MS results for confirmation of acetylation in peptides of interest.

Peptide	*m/z*	Peptide mass	Peptide sequence	Comments	N-terminal fragments (y ions)	C-terminal fragments (b and a ions)
T187 –192	763.4 (charge + 1)	762.4	FEGLK(Ac)K	Lys191 acetylation is confirmed by y2 and b5 ions	[M + H]+ (763.3); y4 (487.3); y2 (317.2)	b5 (617.2); b5j (599.2); a5 (589.2); b4 (447.1)
T122 –128	466.4 (charge + 2)	930.8	QK(Ac)PEFLK	Lys123 acetylation is confirmed by b2 and y6 ions	[M + 2H-H2O]2+ (457.5); y6 (803.6); y5 (633.3); y3 (407.2)	b5j (654.2); b2 (299.2); b2j (282.1)
T1–10	1290.6 (charge + 1)	1289.6	P(Ac)MILGYWNVR	N-terminal acetylation is confirmed by b3 and y9 ions	[M + H]+ (1290.6); [M + H-42]+ (1248.3); y9 (1151.4); y8 (1020.5); y7 (907.4); y6 (794.4)	b9 (1116.3); b7 (903.3); b6 (717.1); b5 (554.1); b4 (497.1); b3 (384.0)
T82 – 93	784.2 (charge + 2)	1566.4	K(Ac)HHLCGETEE ER	y10 and b2 ion point the acetylation either on Lys82 or His83. The acetylated amino acid is likely to be Lys82	y10 (1259.4); y9 (1122.3); y8 (1009.3); y7 (849.3); y6 (792.4); y5 (663.3); y3 (433.1); y2 (304.1)	b11 (1392.3); b10 (1263.3); b9 (1134.4); b7 (904.4); b6 (774.9); b5 (718.3); b4 (558.2); b3 (445.2); b2 (308.2)
T173 –186	861.7 (charge + 2)	1721.4	CLDAFPNL K(Ac)DFLAR	b9 and y6 confirm Lys181 acetylation	y12 (1448.6); y11 (1333.5); y10 (1262.4); y9 (1115.5); y8 (1018.4); y7 (904.4); y6 (791.3); y5 (621.3); y4 (506.2)	b13 (1547.3); b10 (1216.3); b9 (1102.2); b8 (931.2); b6 (704.2); b5 (607.1); b4 (460.0); b3 (389.1); b2 (274.0)
T202 –217	963.4 (charge + 2)	1924.8	YLSTPIFSK(Ac)LA QWSNK	y8 and b9 confirm Lys210 acetylation	y14 (1648.5); y13 (1561.4); y12 (1460.5); y11 (1363.4); y10 (1250.4); y9 (1103.4); y8 (1016.4); y7 (846.3); y6 (733.2); y5 (662.2); y4 (534.2); y3 (348.1)	b15 (1778.4); b14 (1664.6); b13 (1577.5); b12 (1391.4); b11 (1263.3); b10 (1192.2); b9 (1079.3); b7 (822.1); b6 (675.2); b4 (465.0); b3 (363.9); b2 (276.9)
T50 – 67	1054.2 (charge + 2)	2106.4	FK(Ac)LGLDFPNLP YLIDGSR	b2 ion points the acetylation either on Phe50 or on Lys51. The acetylated amino acid is likely to be Lys51	y15 (1676.5); y14 (1619.5); y13 (1506.5); y12 (1391.5); y11 (1244.5); y10 (1147.4); y9 (1033.4); y8 (920.4); y7 (823.5); y6 (660.4); y5 (547.3); y4 (434.2); y3 (319.1)	b15 (1788.5); b14 (1673.6); b13 (1560.4); b12 (1448.1); b10 (1187.4); b9 (1074.4); b8 (959.7); b7 (863.3); b6 (716.3); b5 (601.2); b4 (488.1); b3 (431.1); b2 (318.0)

In the last two columns, m/z values for the corresponding ions are reported in parentheses. A circle above a ‘‘b’’ ion indicates loss of water.

A complete list of matching ‘‘a’’, ‘‘b’’ and ‘‘y’’ ions[11] for modified peptides is provided in Table 4. The results have convincingly demonstrated enzymatic acetylation of substrate GST3–3. The acetylation of GST3–3 strictly necessitated the presence of DAMC and TAase. Due to the qualitative character of data acquired, no indication on the extent of acetylation could be obtained. Nevertheless, covalent modification was certainly not quantitative, as it can be observed in the MALDI-TOFMS map of modified GST3–3 (trace B, Figure 6a–g), where peak pairs of both acetylated and non-acetylated forms separated by 42 Da are visible in six spectra out of the seven. Only for acetylated peptide $T_{202-217}$ whose $[M+H]^+$ signal is displayed in

Figure 6f, trace B at m/z 1925.0, no corresponding non-acetylated form is visible in either of the maps (control and modified GST). This is because Lys_{210} was quantitatively cleaved by trypsin if not acetylated, originating peptides T202 – 210 (mass 1054.6 Da) and $T_{211-217}$ (mass 845.4 Da). Both peptides are observable both in control GST and modified GST peak lists (see Table 3), indicating that Lys_{210} was also not quantitatively acetylated. Figure 7a and b were obtained by using the software Molecular Operating Environment (MOE) developed by Chemical Computing Group Inc., Montreal, Canada. These figures clearly show the backbone and sites of acetylation of GST3–3 subunit.

Discussion

The enzymatic acetylation of protein without the involvement of acetyl-CoA is a grey area, although nonenzymatic protein modification[12] of varied kinds, including acetylation is familiar[4]. Our earlier investigations[5,6,13,14] conclusively highlighted for the first time the presence of an unique enzyme in liver and other tissues catalyzing the transfer of acetyl groups from a model acetoxy drug DAMC and other acetyl polyphenols to receptor proteins such as GST, cytochrome P-450 linked mixed function oxidases and NADPH cytochrome c reductase resulting in the modulation of their catalytic activities. The enzyme was termed acetoxy drug (DAMC): protein transacetylase (TAase). In order to understand the nature of the TAase-catalyzed reaction, efforts were made to purify TAase from buffalo liver microsomes. The partially purified TAase exhibited molecular weight of 65 kDa according to SDS-PAGE. Similar results were also obtained from size exclusion gel chromatography. The initial rate of Taasecatalyzed reaction was found to vary with the concentration of the substrates DAMC/cytosol. These results confirm that cytosolic GST is an effective substrate for the buffalo liver TAase. TAase-catalyzed transfer of acetyl group from DAMC to the receptor protein resulted in the formation of hydroxycoumarins (Figures 1 and 4). We have earlier reported the presence of a membrane-bound deacetylase in liver microsomes catalyzing the deacetylation of DAMC yielding hydroxycoumarin[15]. It is conceivable that Taase preparation has the contaminating presence of deacetylase, which acts on DAMC forming dihydroxycoumarin (DHMC). But the addition of GST to the TAase reaction mixture resulted in the drastic enhancement (60%) in the formation of MAMHC. These results convey the following conclusions: (a) the action of deacetylase on DAMC yields both MAMHC and DHMC, although the latter was in larger amounts[5]; (b) the nature of TAase action on DAMC is distinct from the deacetylase in that acetyl groups are preferentially transferred from a particular

acetoxy group to the receptor protein leading to the accumulation of MAMHC; and (c) it substantiates the TAase-catalyzed reaction (Figure 1). Also, DAMC-deacetylase was strongly inhibited by the protease inhibitor, phenylmethanesulfonylfluoride (PMSF), while the TAase-catalyzed reaction was unaffected by the inhibitor (unpublished data). The aforementioned results highlighted the distinct nature of Taase compared to the deacetylase. After having characterized the TAase-catalyzed reaction in terms of the nature of the substrates and products of the reaction, efforts were made to identify acetylated protein, which is another product of the reaction by the application of MALDI-TOF MS and LC-MS/MS. To identify the TAase-catalyzed acetylation of GST with DAMC as the acetyl donor, the GST3–3 was used as a model protein substrate. Acetylation of GST in the first place was confirmed by performing MALDI-TOF MS on the tryptic peptide isolated from the gel pieces. The trypsinized modified GST exhibited acetylation of peptides. The modified amino acid residues were determined by LC-MS/MS. In the modified GST, N-terminal proline and six lysines (Lys_{51}, Lys_{82}, Lys_{123}, Lys_{181}, Lys_{191} and Lys_{210}) were found to be acetylated. The flat ribbon structure of acetylated GST showed that the amino acids acetylated are peripheral in position (Figure 7a,b). The fact of the matter is that the acetylation of GST does result in the inhibition of catalytic activity. Although a definitive explanation is difficult from the present data, it is possible that the lysines acetylated distinct from the active site could play a role in causing the inhibition of GST activity. There exists several types of transacetylases such as platelet-activating factor (PAF): sphingosine transacetylase, PAF: lysoplasmalogen transacetylase, N-arylamine acetyltransferase and histone acetyltransferase[1,16,17]. TAase appears to be distinct from such transacetylases in terms of the molecular weights and as well as the substrate specificity. Several enzymes, deacetylases, mixed function oxidases, and carboxyl esterases metabolize acetoxycoumarins. Our studies substantiated the role of TAase in the metabolism of acetoxycoumarins in particular and acetyl polyphenols in general. Moreover, TAase appears promising in the modification of functional proteins by way of acetylation utilizing the appropriate acetoxy drugs as the acetyl donors which can lead to altered physiological/pharmacological effects.

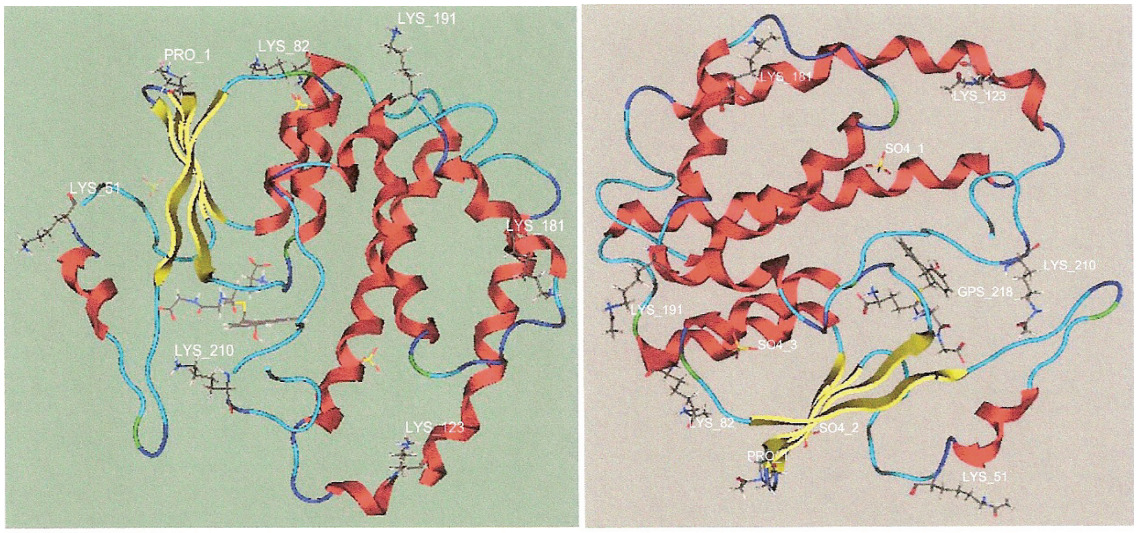


Figure 7 (a) Structure of unmodified subunit of GST3– 3. (b) Structure of modified (acetylated) subunit of GST3– 3.

Acknowledgements

The financial assistance of the Department of Biotechnology (DBT, Govt. of India) and the Danish International Development Agency (DANIDA) are gratefully acknowledged. R.K. was awarded CSIR NET Fellowship of the Govt. of India. Dr. Sonja Jespersen offered helpful discussions.

References

1. C.A. Mizzen, J.E. Brownell, R.G. Cook, C.D. Allis, *Methods Enzymol.* 304, **1999**, 675– 704.
2. W. Gu, R.G. Roeder, *Cell,* 90, **1997**, 595– 606.
3. S.J. Wakil, *J. Biochem.* 45, 28(11), **1981**, 23–30.
4. J.R. Vane, *Nature*, 231, **1971**, 232–234.
5. H.G. Raj, V.S. Parmar, S.C. Jain, E. Kohli, N. Ahmad, S. Goel, Y.K. Tyagi, S.K. Sharma, J. Wengel, C.E. Olsen, *Bioorg. Med. Chem*. 8, **2000**, 1707– 1712.
6. I. Singh, E. Kohli, H.G. Raj, K. Gyanda, S.K. Jain, Y.K. Tyagi, G. Gupta, R. Kumari, A. Kumar, G. Pal, A.K. Prasad, R.C. Rastogi, C.E. Olsen, S.C. Jain, V.S. Parmar, *Bioorg. Med. Chem.* 10, **2002**, 4103– 4111.
7. E. Kohli, G. Gaspari, H.G. Raj, V.S. Parmar, J. Vander Greef, G. Gupta, R. Kumari, A.K. Prasad, S. Goel, G. Pal, Y.K. Tyagi, S.C. Jain, N. Ahmad, A.C. Watterson, C.E. Olsen, *FEBS Lett.* 530, **2002**, 139– 142.
8. A.C. Dey, S. Rahal, R.L. Rimsay, I.R. Senciall, *Anal. Biochem.*110, **1981**, 373.
9. U.K. Lamelli, *Nature*, 227, **1970**, 680–685.
10. A. Shevchenko, M. Wilm, O. Vorm, M. Mann, *Anal. Chem.* 68, **1996**, 850– 858.
11. P. Roepstroff, J. Fohlman, *Biomed. Mass Spectrom.* **1984**, 601.
12. E. Gianazza, *J. Chromatogr.* 705, **1995**, 67–87.
13. H.G. Raj, V.S. Parmar, E. Kohli, Y.K. Tyagi, C.E. Olsen, *FASEB J.* 14, **2000**, A1445.
14. E. Kohli, R. Kumari, G. Gupta, H.G. Raj, V.S. Parmar, N. Ahmad, S.C. Jain, G. Pal, C.E. Olsen, *Drug Metab. Rev.* 33, **2001**, 39.
15. H.G. Raj, V.S. Parmar, S.C. Jain, S. Goel, Y.K. Tyagi, S.K. Sharma,

C.E. Olsen, J. Wengel, *Bioorg. Med. Chem*. 8, **2000**, 2333– 2336.
16. Y. Uemura, T.C. Lee, F. Synder, *J. Biol. Chem.* 266, **1991**, 8268– 8272.
17. W.W. Weber, D.W. Hein, *Pharmacol. Rev.* 37, **1985**, 25– 79.

Discovery of modified peptides in protein tryptic digests by automated handling of capillary LC-MS data

Introduction

Visual inspection of LC-MS data from complex mixtures, like protein tryptic digests, is a tiring and time-consuming procedure which nevertheless is sometimes necessary to fully extract information contained in those data. This fact explains why peptide maps from purified protein digests are preferentially acquired on MALDI-TOF mass spectrometers. MALDI-TOF peptide maps are easy to acquire and allow a much quicker visual comparison of two or more spectra from different samples by spectra overlaying. In fact, researchers faced with the problem of finding small differences between tryptic digests of two closely related proteins or protein isoforms, normally rely on the following approach. A preliminary analysis by MALDI-TOF MS on both samples is performed. MALDI-TOF maps are overlaid and visually inspected. Peaks of interest, exclusively present in one of the maps, are selected for further analysis. If no MS/MS capabilities are available on the MALDI-TOF instrument, targeted LC-MS/MS analysis on the peaks of interest is performed in the search for post-translational modifications (PTMs) or covalent modifications/degradations. Such "classic" approach has been used in the work described in Chapter 6.1, where it was successfully applied to find seven sites of acetylation for the model protein GST, after reaction with a novel transacetylase enzyme. In this work, the same sample set has been subjected to direct capillary LC-MS (capLC-MS) analysis and automated data handling in order to demonstrate the power of newly developed data analysis tools, described in Chapter 5, in the evaluation of complex MS data. Automated data analysis allows for the direct comparison of peptide maps obtained by the highly informative LC-MS analyses directly, without the need of multiple steps and the use of complementary techniques.

Methods

Control (three different incubations) and modified GST (one incubation) were isolated on a 1D PAGE gel in duplicates and digested by trypsin as described in Chapter 6.1. A schematic of the experimental protocol is depicted in Figure 1. Three different incubations were run for the control experiment because they represented (i) GST alone; (ii) GST + liver microsomes (no acetyl donor, DAMC, added); (iii) GST + acetyl donor DAMC (no microsomes added). Tryptic peptides were analysed by

capLC-MS using the same conditions as described in Chapter 6.1, but without the use of targeted or data-dependent MS/MS mode. Full scan mode only was performed. Raw data were converted to cdf format before processing. Data analysis was performed as described in Chapter 5. Briefly, raw data were reduced in complexity by assigning to each integer mass trace (i.e. for each *m/z* value ranging from 300 to 2000) an IQ value 1, the IQ being a normalised value (between 0 and 1) indicating the likelihood that in a mass trace one or more peaks were present. With data reduction, a *n*-dimensional vector (*n*=1700) was generated for each capLC-MS data file acquired. Using principal component analysis (PCA), the vectors corresponding to control and modified GST were plotted on a bidimensional plane.

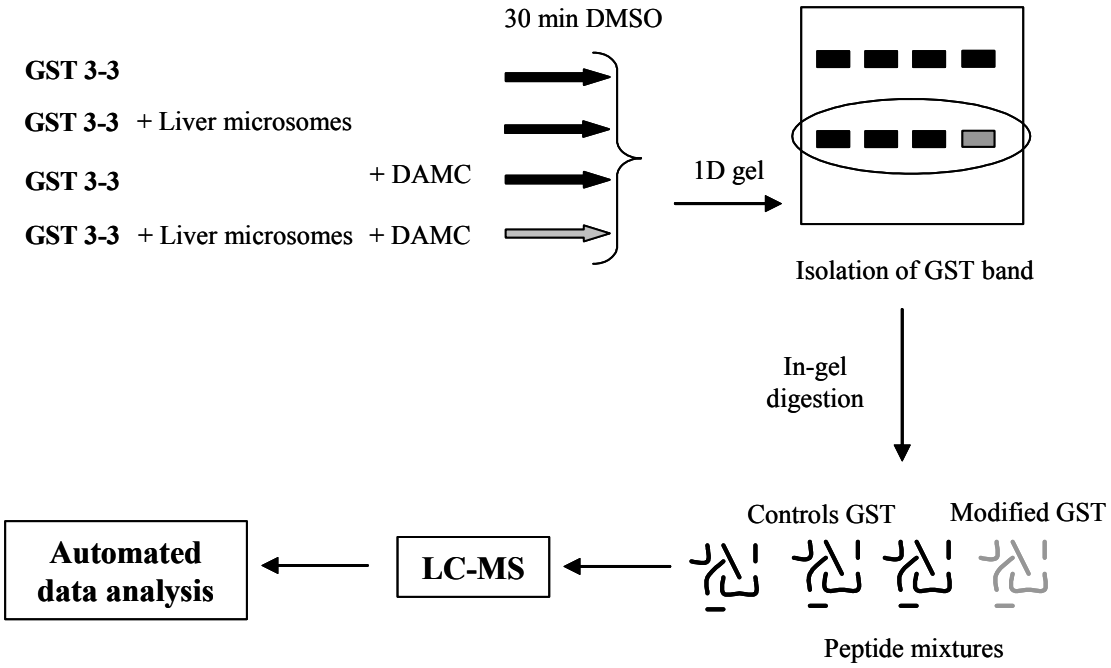


Figure 1 Schematic of the analytical method used. Three different control GST samples were used.

Results and discussion

In this work, automated evaluation of capLC-MS data files in order to find small differences in tryptic peptide mixtures is demonstrated. Raw LC-MS data went through a pre-processing routine (binning and assignment of the IQ value). After data pre-processing, PCA analysis was run in order to discover trends in the data set. The dataset was composed of 16 LC-MS data files, 4 from modified GST (2 LC-MS analyses X 2 gel bans) and 12 from control GST (2 LC-MS analyses X 6 gel bands).

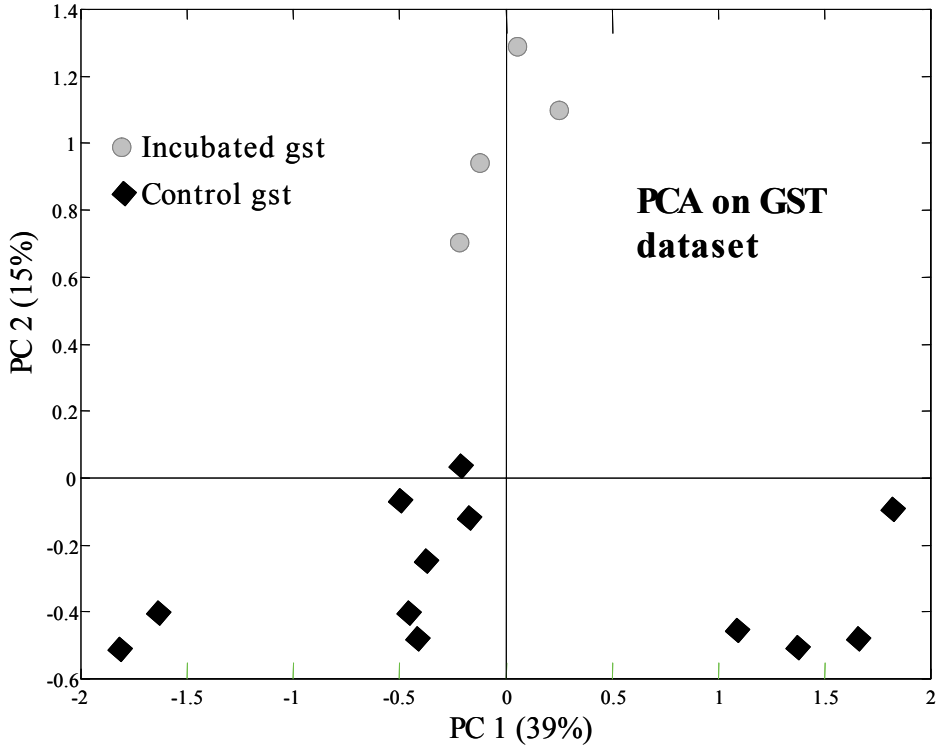


Figure 2 PCA plot of the 16 capLC-MS data files acquired. Dimension PC2 is the one of most interest, because it separates the control group from the modified GST group, containing acetylated peptides as demonstrated in Chapter 6.1.

PCA separation of the 16 data files is shown in Figure 1. Control samples (rhombs) were partially separated in PC1. This dimension, though carrying most of the variance (39%), was not relevant for this study. PC1 mainly accounted for differences among the control group, probably originated by variation in the sample processing (parallel incubation and in-gel digestion). On the contrary, dimension PC2 was very interesting to explore, because it separated the "modified GST" group from the control group, representing 15% of the total variance within the data set. Mass traces possessing a high loadings factor in PC2 were very likely to contain a new peak in the LC-MS data from the modified GST group with respect to the ones obtained from the control samples. Those new peaks were potentially generated by the covalent modifications (acetylations) which were the focus of the biological question behind the study.

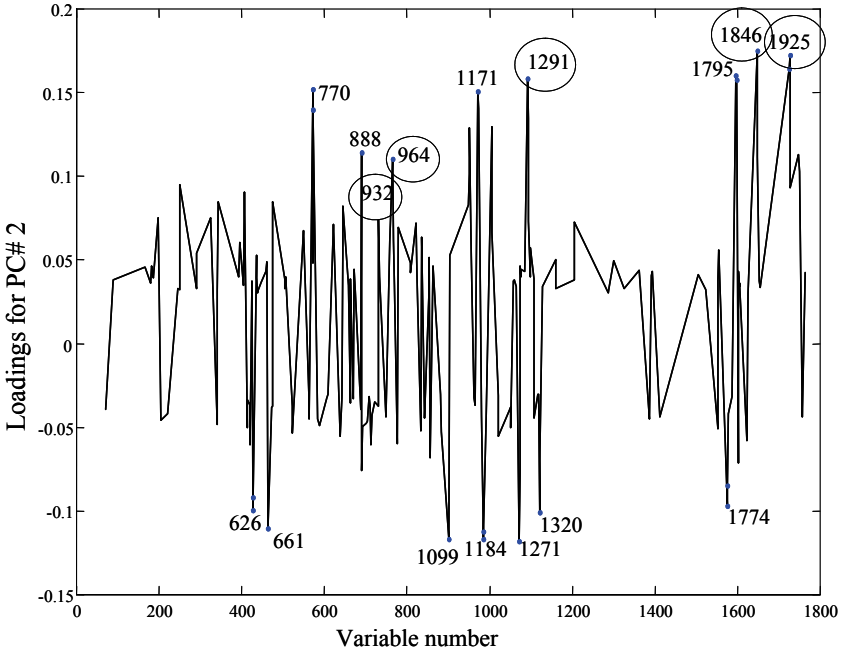


Figure 3 Loadings plot for PC2. Acetylated peptides are circled.

Loadings factors for PC2 separation are reported in Figure 2. Mass traces with a circle represent *m/z* values previously demonstrated to correspond to acetylated peptides (see Chapter 6.1). As shown in Figure 2, the *m/z* value of five of the seven acetylated peptides previously discovered by MALDI-TOF and LC-MS/MS was found back in the highest loadings of PC2.

As an example of the meaning of IQ and high loadings factor in PC2, Figure 4 and 5 illustrate two mass traces with a high PC2 loadings, thus carrying differences between control and modified GST samples. Those mass traces point at two peaks belonging to previously identified acetylated peptides, as described in Chapter 6.1. As it can be seen in Figure 4, mass trace 1925 gets a high IQ value in the modified GST samples because a relevant peak is detected at ~25 minutes; on the other hand, only noise is visible in the control GST mass trace 1925, giving a poor IQ value of 0.39. The difference in IQ value generates a positive PC2 loading of 0.16, one of the highest for this dimension. Similarly, a positive loading factor of 0.07 is registered for the mass trace 932. An extra peak at 15.98 min is detectable in the modified GST LC-MS trace 932. This peak causes the IQ value to raise from 0.38 to 0.44 between control and modified GST samples.

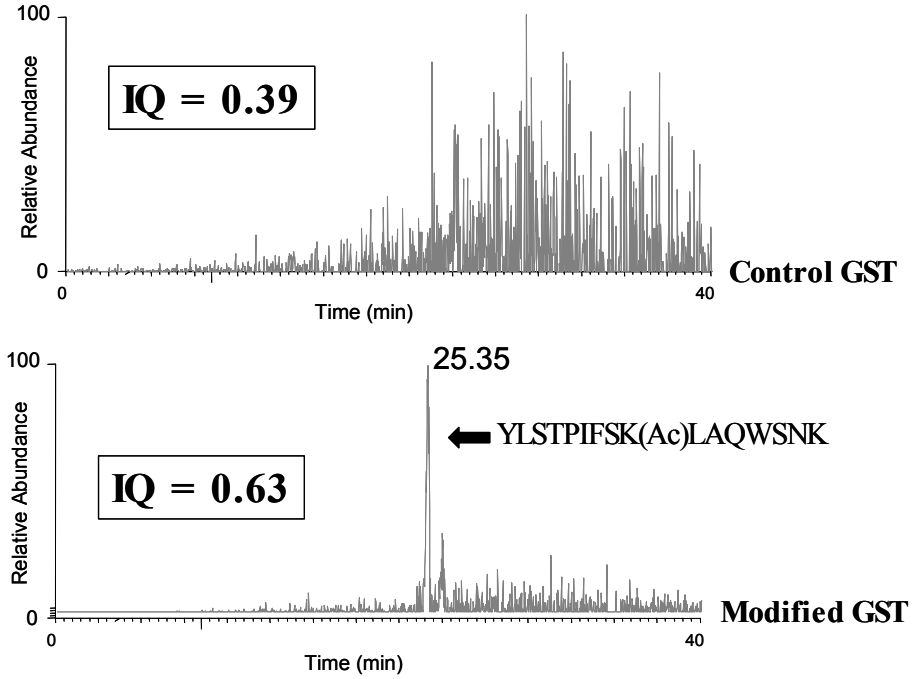


Figure 4 Mass trace 1925 in one control (top panel) and one modified GST (bottom panel) LC-MS data file. Loadings factor of 1925 in PC2 was 0.16.

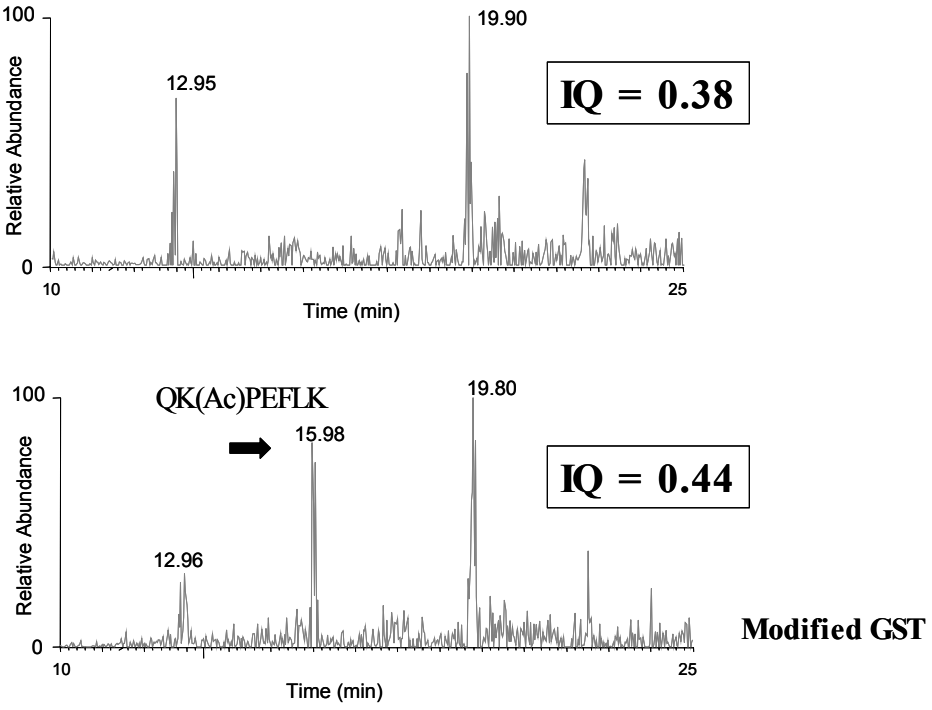


Figure 5 Mass trace 932 in one control (top panel) and one modified GST (bottom panel) LC-MS data file. Loadings factor of 932 in PC2 was 0.07.

Conclusions

This study demonstrates the feasibility of a direct LC-MS approach with automated data handling for comparative analysis of purified protein digests. Peptide maps could be directly generated by LC-MS in automated fashion, and the data screened by a dedicated software in the search for small differences indicative of covalently modified peptides. Peaks of interest could be addressed in a second step by targeted MS/MS. In future applications, real-time, data-dependent MS/MS routines could be implemented during the first sample acquisition in order to build an MS/MS information "repository" that could eventually be used after the data analysis step. The approach was successfully tested on the sample set generated for the study described in Chapter 6.1.