COMPOUNDS AND METHODS TO STUDY *O*-GLCNAC MODIFICATION OF PROTEINS

By

HONGCHAO ZHENG, B.Sc., M.Sc.

A Thesis

Submitted to the School of Graduate Studies in Partial

Fulfillment of the Requirements

For the Degree

Master of Science

McMaster University

Master of Science (2007) McMaster University
Hamilton, Ontario

(Chemistry)

TITLE:

Compounds and Methods to Study *O*-GlcNAc Modification of Proteins

AUTHOR: Hongchao Zheng, M.Sc. (Tsinghua University, P.R.China)

SUPERVISOR: Dr. Graham A. McGibbon

NUMBER OF PAGES: xv, 103

ABSTRACT

A variety of important cell functions rely on *O*-GlcNAcylation of proteins, a kind of post-translational glycosylation that modifies nuclear and cytoplasmic proteins on serine or threonine residues by the addition of the single sugar moiety β-*N*-acetylglucosamine (*O*-GlcNAc). Generally, two enzymes can catalyze the formation and cleavage of the *O*-GlcNAc *O*-glycosidic linkage. The *O*-GlcNAc transferase (OGT) catalyzes the transfer of *N*-acetylglucosamine from uridine diphosphate *N*-acetylglucosamine to the hydroxyl group of specific Thr and Ser residues. On the other hand, hydrolysis of the sugar moiety from proteins is achieved via the bi-functional nuclear cytoplasmic *O*-GlcNAcase and acetyltransferase (NCOAT). Since *O*-GlcNAcylation plays very important roles in many cell processes such as transcription, translation and protein-protein interaction, my research mainly focused on preparing compounds and developing methods to study the *O*-GlcNAcylation process and discover the structural information of OGT and NCOAT.

In this thesis, the important enzyme substrates (YSDSPSTST for OGT and YSDSP(*O*-GlcNAc-Ser)TST for NCOAT) were prepared by solid phase synthesis and purified by preparative reverse phase HPLC. Their structure and sequence were confirmed by ESI-MS and tandem MS (MS/MS). The building block *S*-GlcNAc-Ac_3-Ser-COOBn for preparing *S*-GlcNAcylated peptides was prepared by a multistep procedure. Two potent photoaffinity labeled probes for mapping the active site of sOGT were designed and synthesized, and their IC_{50} values for sOGT were measured by CapLC and a beta-elimination method. Also, many analytical methods were developed for studying the *O*-glcNAcylation, including discovering the *O*-Glycosylation site of peptides by MS/MS and beta-elimination methods, sequencing peptides by MS/MS, analyzing the protein digests by CapLC-MALDI MS and studying the ion mobility of peptides and glycopeptides using ion mobility spectrometry.

ACKNOWLEDGEMENTS

I would like to extend my sincere appreciation to my supervisor Dr. McGibbon, for his support and help in finishing my M.Sc. research project. With his offer, I have had this opportunity to come to McMaster University to enjoy an advanced education. In addition, I truly believe Dr. McGibbon's direction and profound suggestions will help me in my future career. I greatly appreciate my committee members, Dr. Valliant, Dr. Harrison and Dr. Green, for their useful suggestions and help throughout my graduate studies. I would also like to thank another group member, Mr. Xi Wang, for his contribution on purifying and expressing NCOAT and sOGT proteins, as well as Ms. Leslie from Dr. Valliant's group for her assistance in solid phase syntheses of my enzyme substrates, Ms. Yuzhu Guo from Dr. Siu's group at York University for her help in performing IMS analyses of my samples, Ms. Lihua Liu for her help in the MS analyses of my samples, Dr. Li Gan and Dr. Yang Chen for their helpful discussions about the synthetic routes for my target compounds and Dr. Jiaxi Wang for his useful suggestions for MS sample preparations. The Mass Spectrometry Lab staff (Ms. Gina) and the NMR Lab staff (Dr. Hughes *et al.*) are appreciated for their assistance.

Finally, I would also like to thank my parents for their support and encouragement throughout my undergraduate and graduate years of study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES AND SCHEMES

ABBREVIATIONS

Asp – Aspartic acid

BzCl – Benzoyl chloride

CapLC – capillary HPLC

DIPEA – *N*, *N*-Diisopropylethylamine

DMF – *N*,*N*-dimethyl formamide

DMSO – Dimethylsulfoxide

ESI – Electronic spray ionization

Fmoc – Fluorenyl methoxy-carbonyl

GlcNAc – *N*-acetylglucosamine

HBTU–2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium Hexafluorophosphate

HPLC – High pereformance liquid chromatography

HRMS – High Resolution Mass Spectrum

IMS – Ion mobility spectrometry

MALDI – Matrix-assisted laser desorption/ionization

MS – mass spectrometry

MW – Molecular Weight

NaOMe – Sodium Methoxide

NCOAT – nuclear cytoplasmic O-GlcNAcase and acetyltransferase

NMR – Nuclear Magnetic Resonance

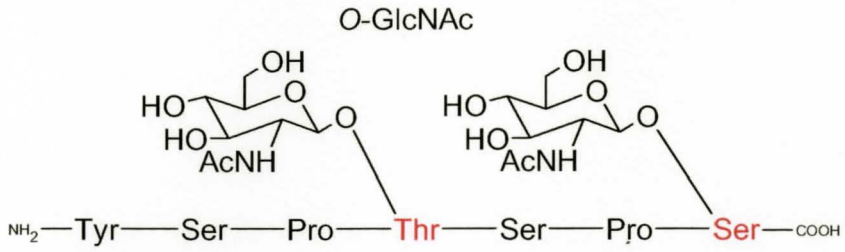


Figure 1.1 *O*-GlcNAc modification model

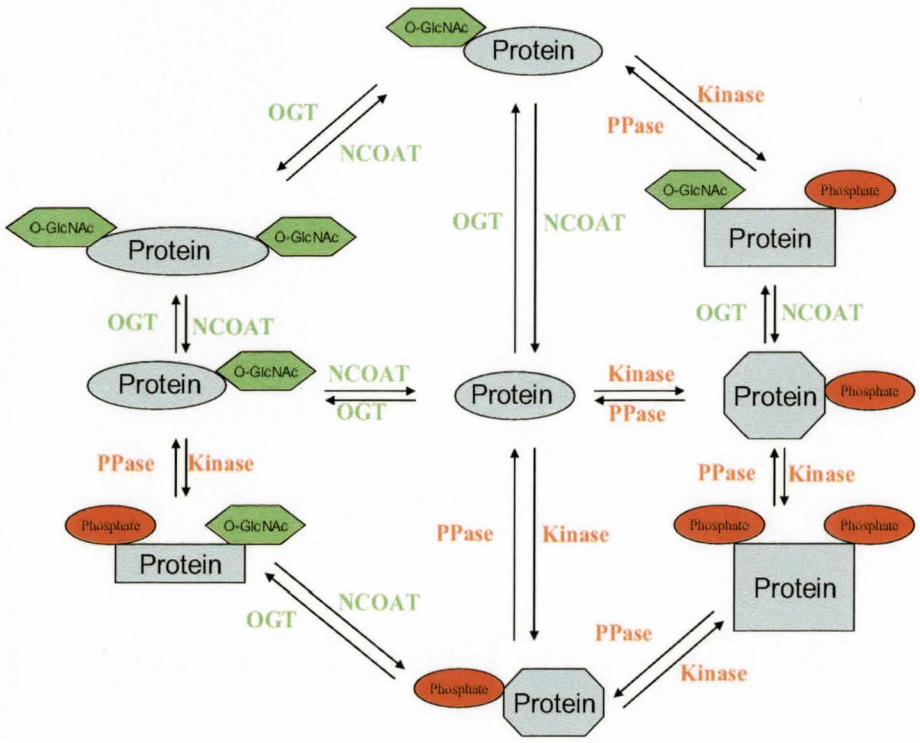


Figure 1.2 The interplay between *O*-GlcNAc and *O*-phosphate (adapted from reference 13)

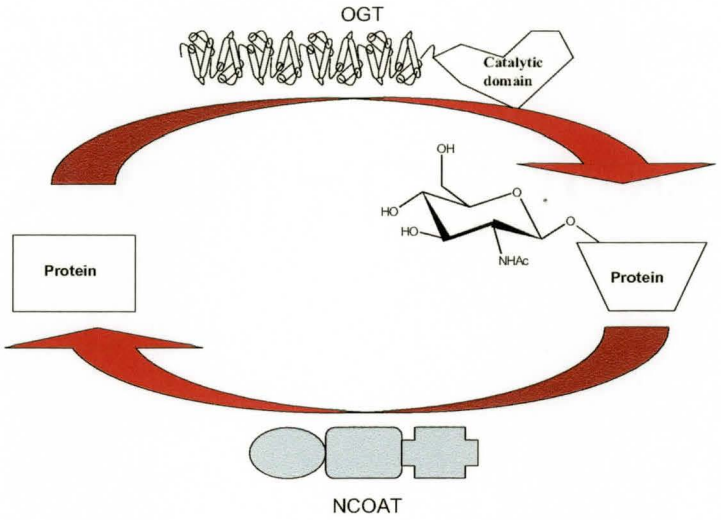


Figure 1.3 A dynamic post-translational modification of proteins

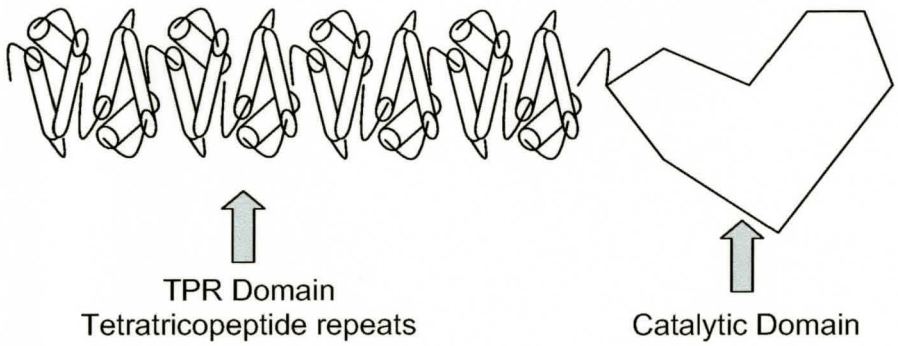


Figure 1.4 Schematic representation of the structure of OGT

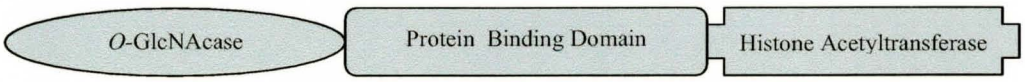


Figure 1.5 Schematic representation of the structure of NCOAT

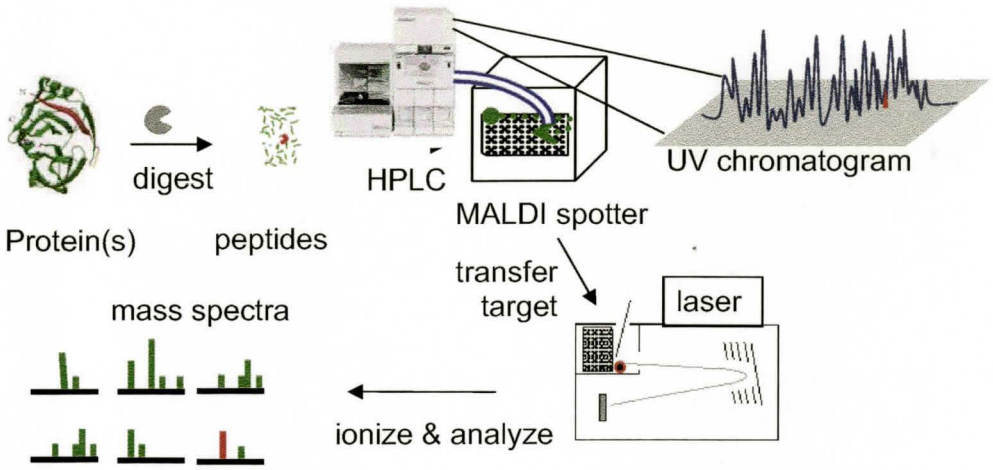


Figure 1.6 CapLC MALDI coupling system

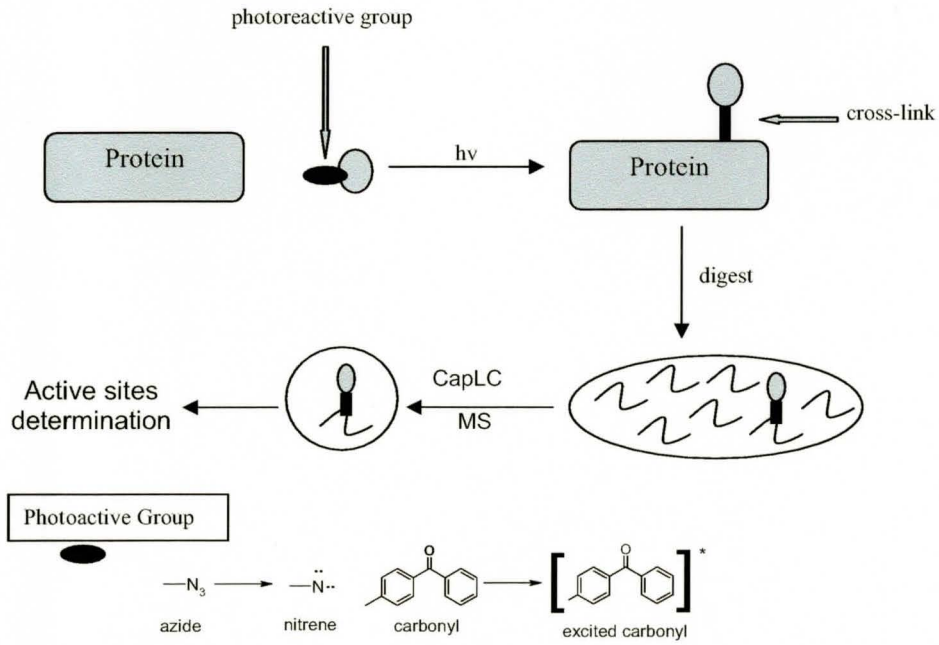


Figure 1.7 Photoaffinity labeling approaches

Chapter 2

Solid Phase Syntheses and HPLC Purification of Enzyme Substrates

2.1 Enzyme Substrate Selection

OGT and NCOAT catalyze the addition and removal of *O*-GlcNAc to or from many nuclear and cytosolic proteins [1-4] and this kind of glycosylation is believed to affect many cellular processes such as translation, transcription, protein-protein interactions and organelle targeting [5]. To study the *O*-GlcNAcylation and the structure of NCOAT and OGT, suitable enzyme substrates should be selected. Some nuclear and cytosolic proteins which can be transiently *O*-GlcNAcylated by OGT seem to be good choices. However, most of these proteins with this type of modification are present in low quantities in the cell, thus complicating purification of large amounts for use as enzyme substrates. Also, the yield of *O*-GlcNAcylation is very low and it is very difficult to detect the target proteins directly by MS, thus making it very hard to study the *O*-GlcNAcylation by MS if proteins are used as enzyme substrates in these experiments. To circumvent these problems, some suitable peptides which experience the *O*-GlcNAcylation can be selected as the enzyme substrates instead of proteins. In 1990, Hart and co-workers reported an optimized substrate YSDSPSTST, which could be *O*-GlcNAcylated on one of its Serine residues

(Figure 2.1) [6]. Thus, YSDSPSTST **1** was chosen as the substrate for OGT and its *O*-GlcNAcylated product YSDSP(*O*-GlcNAc-Ser)TST **2** was selected as the substrate for NCOAT. In this chapter, the solid phase syntheses and HPLC purification of these two enzyme substrates are described.

2.2 Solid Phase Syntheses of Substrates

For syntheses of YSDSPSTST and YSDSP(*O*-GlcNAc-Ser)TST, the Fmoc solid phase peptide synthesis (SPPS) method was employed. This method was introduced by Carpino in 1972 [7] and further applied by Atherton in 1978 [8, 9]. The main advantage of Fmoc is that it can be cleaved under very mild basic conditions (usually 20% piperidine in DMF). Because one of our target peptides YSDSP(*O*-GlcNAc-Ser)TST could easily loose *O*-GlcNAc under very strongly acidic conditions, the Fmoc SPPS method was an obvious choice. We employed Fmoc-Thr(*t*Bu)-Wang Resin as solid support in the whole SPPS. Actually, Fmoc-Thr(*t*Bu)-SASRIN is a better choice as solid support because it is much easier to cleave the peptide from this resin. We did not use it as the solid support due to its high cost.

The whole SPPS route for YSDSPSTST **1** is shown in Scheme 2.1. Fmoc-Thr(*t*Bu)-Wang Resin was added to the wells of an Advanced ChemTech 348 automated synthesizer, suspended and washed with THF then with DMF for 3-4 times. This wash procedure was used between every deprotection and coupling step. Then the Fmoc group was deprotected through the addition of 20% piperidine in DMF. The coupling reaction employed a standard HBTU coupling

procedure. The coupling reactions initially involved adding the Fmoc protected amino acids (5 equiv., 0.5 mol/L in DMF), followed by HBTU (5.0 equiv., 0.5 mol/L in DMF) and DIPEA (10.0 equiv., 0.5 mol/L in DMF). Then the reaction block was shaken for 2 h. Following filtration, the resin was washed by the previously motioned procedure prior to the start of the next cycle (total 8 cycles). After SPPS, YSDSPSTST **1** was cleaved from Wang resin by treatment with 95% TFA, 2.5% TIS and 2.5% water at 0 °C. The mixture temperature was allowed to slowly rise to room temperature and stirring was continued overnight. Then the resin was put into cold diethyl ether and a solid precipitated. This solid was washed with diethyl ether and lyophilized, yielding crude YSDSPSTST **1**, which was further purified by HPLC.

For YSDSP(*O*-GlcNAc-Ser)TST **2** synthesis, we employed the same synthetic strategy (Fmoc SPPS method). As the building block for the SPPS of **2**, the acetyl-protected *β*-D-GlcNAc-Ser analogue **3** was prepared by a previously reported procedure [10] through a coupling reaction between glycosyl chloride and protected serine in the presence of mercury bromide as mediating reagent and 1, 2-dichloroethane as solvent (Scheme 2.2). Then the same procedure as the synthesis of YSDSPSTST was applied on synthesizing the glycopeptide and the acetyl protected *O*-GlcNAcylated peptide **4** was obtained (Scheme 2.3). It is also worth mentioning here that when **4** was cleaved from the resin in the presence of TFA/TIS/water, we shortened the reaction time to 1.5 h to prevent the removal of *O*-GlcNAc from **4** due to the instability of **4** under strongly acidic conditions. It

was found that if the cleavage time was less than 1 h, some protecting groups (*t*Bu) were not removed; and if the cleavage time was more than 3 h, β-elimination reaction happened on compound **4** and as a result, by-product **5** was obtained. After protected glycopeptide **4** was purified by HPLC, it needed to be deacetylated to yield the desired substrate YSDSP(*O*-GlcNAc-Ser)TST **2**. Sodium methoxide in methanol was used to deacetylate **4**. Different conditions (Table 2.1) were used and the reactions monitored by HPLC (for HPLC conditions, please see Table 2.2). We found that if compound **4** was deacetylated under condition 1 in Table 2.1, it mainly experienced beta-elimination reaction at the same time to give both the desired compound **2** (t_R = 15.32 min) and undesired compound **5** (t_R = 16.72 min) in the ratio 1:3 (mainly undesired compound **5**) (Figure 2.2A). Then less basic conditions (condition 2 in Table 2.1) were selected to try to convert **4** to **2** and the HPLC result (Figure 2.2B) showed that the ratio of compound **2** (t_R = 15.37 min) to compound **5** (t_R = 16.75 min) increased from 1:3 to 3:1, but the staring material (compound **4**, t_R = 22.02 min) did not disappear completely. We assumed that the ratio of **2** to **5** mainly depended on the reaction time and slightly depended on the NaOMe concentration. Thus, much more basic conditions and very short reaction time (condition 3 in Table 2.2) were applied to this deacetylation and the HPLC result (Figure 2.2C) showed that the ratio of **2** (t_R = 15.35 min) to **5** (t_R = 16.83 min) was greatly improved to 49:1 and the starting material **4** had disappeared completely. After HPLC purification of crude **2**, the desired substrate **2** was readily obtained.

2.3 HPLC Purification of Substrates

After SPPS, the substrates needed to be further purified by reverse phase HPLC. Solvent A (0.1% HCOOH in CH_3CN, less polar) and solvent B (0.1% HCOOH in H_2O, polar) are the two solvents used to separate the desired peptides from other impurities. Although TFA (Trifluoroacetic acid) is the most commonly acidic modifier used in peptide-HPLC separation, 0.1% formic acid was added into both H_2O and ACN, because TFA would interfere with the mass spectral quality. Since ACN is less polar than H_2O, we can increase the ACN percentage to elute the peptide binding to the reverse phase HPLC column.

For purification of YSDSPSTST **1**, we optimized a HPLC method (Table 2.3) to separate it from the crude product. As shown in Table 2.3, from 0 min to 24 min, the ACN gradient was increased linearly from 1% to 9% and the desired peptide eluted out at 16.67 min (Figure 2.3A); From 25 min to 30 min, 100% ACN (Table 2.3) was used to wash out all the impurities remaining in the column; from 31 min to 40 min, the initial conditions were employed again to equilibrate the column for the next injection. To accelerate the separation rate, HPLC conditions with a faster flow rate (Table 2.4) were used and the HPLC chromatogram (Figure 2.3B) showed that the desired peptide (t_R = 8.65 min) was readily purified. But when we increased the flow rate to 6 mL/min to purify YSDSPSTST, the HPLC chromatogram showed that the desired peptide could not be purified at a satisfactory level due to the loss of peak resolution. After collection of the desired fractions, the liquid fractions were lyophilized on a

lyophilizer equipped with a centrifuge, yielding the pure YSDSPSTST **1**. Then we used CapLC to analyze it and the CapLC chromatogram (Figure 2.3C) showed that only one component (t_R = 6.82 min) was detected by CapLC. During the purification, ESI-MS was used to identify the desired peptide. The mass spectrum (Figure 2.3D) showed that the strong peak (*m/z* 944.47) was protonated peptide ion (MW = 943.38), and the peak at 966.49 was due to the sodiated peptide ion.

Intermediate acetyl protected glycopeptide **4** contains more acetyl protecting groups and thus is a less polar compound than TSDSPSTST **1**. Thus, compared with the HPLC method (Table 2.4) to purify YSDSPSTST, a less polar solvent (Table 2.5) was employed to purify compound **4**. As shown in Table 2.5, from 0 min to 12 min, the ACN was linearly increased from 1% to 30% and the desired acetylated glycopeptide **4** eluted at 9.48 min (Figure 2.4A). After the collection of the desired fraction (9.48 min peak), the liquid fractions were lyophilized to yield the pure YSDSP(*O*-GlcNAc-OAc_3-Ser)TST **4**. Then we used CapLC to analyze it and the CapLC chromatogram (Figure 2.4B) showed that only one component (t_R = 6.90 min) was detected by CapLC. Also it was characterized by MS (Figure 2.4C).

To optimize the HPLC separation method for monitoring the deacetylation reaction from **4** to **2**, many methods were attempted. After deacetylation of **4** under condition 1 in Table 2.1, the product was first analyzed by HPLC under the conditions shown in Table 2.6. The resulting HPLC chromatogram (Figure 2.5A) showed that the two products YSDSP(*O*-GlcNAc-Ser)TST **2** (t_R = 9.87 min) and

YSDSP(2-aminopropenoic acid)TST **5** (t_R = 10.03 min) overlapped with each other. To separate these two compounds, a more polar HPLC solvent (Table 2.7) was employed and the resulting HPLC chromatogram (Figure 2.5B) showed that compound **2** (t_R = 13.15 min) and compound **5** (t_R = 14.02 min) had almost been separated from each other completely. Then this HPLC method was further modified to give the HPLC method in Table 2.2, which could separate compound **2** (t_R = 15.32 min) and compound **5** (t_R = 16.72 min) totally (see Figure 2.5C). After HPLC purification of compound **2** and compound **5** by the same HPLC method (Table 2.2), the purified **2** and **5** were characterized by MS (Figure 2.6C for **2** and Figure 2.7B for **5**). Then we used CapLC to analyze them and the CapLC chromatograms (Figure 2.6B for **2** and Figure 2.7A for **5**) showed that for compound **2**, only one component (t_R = 6.97 min) was detected by CapLC and for compound **5**, two components (t_R = 7.85 min for unknown impurity and t_R = 8.45 min for compound **5**, about 1: 20) were detected by CapLC.

2.4 Conclusions

The target enzyme substrates, YSDSPSTST **1** for OGT and YSDSP(*O*-GlcNAc-Ser)TST **2** for NCOAT, were prepared by the Fmoc solid phase peptide synthesis method. The crude substrates were purified by preparative reversed phase HPLC. The purified substrates were analyzed and characterized by CapLC and ESI-MS.

2.5 Experimental

2.5.1 General Procedures

Peptide syntheses were performed on an Advanced ChemTech 348 automated synthesizer generously made available by Dr. Valliant at McMaster University and purifications of peptides were performed on a Waters Delta 600 HPLC with a semi-preparative column (Sunfire^{TM} prep C_{18} 5 μm 10 × 100 mm column). All CapLC analyses were carried out on a Waters Delta CapLC with a Waters NanoEase^{TM} column (Atlantis^{TM} dC_{18} 5 μm 300 μm × 150 mm column). The solvent for HPLC was A 0.1% HCOOH in CH_3CN, B 0.1% HCOOH in water, and the flow rate was 3 mL/min; the solvent for CapLC was A 0.1% HCOOH/1% CH_3CN/H_2O, B 0.1% HCOOH/CH_3CN, and the flow rate was 5 μL/min.

Proton NMR and ^{13}C NMR spectra were collected on Brucker AV 200 and AV 500 spectrometers and were referenced to internal tetramethylsilane ($SiMe_4$). Chemical shifts for ^{1}H NMR spectra are reported as δ in units of parts per million (ppm) downfield from $SiMe_4$ (δ 0.0) and relative to the signal of DMSO-d_6 (δ 2.5000). Multiplicities were given as: s (singlet); brs (broad singlet); d (doublet); t (triplet); q (quartet); dd (doublet of doublets); dt (doublet of triplets); dtq (doublet of triplets of quartets); or m (multiplet). The number of protons (n) for a given resonance is indicated by nH. Coupling constants are reported as a *J* value in Hz. Carbon nuclear magnetic resonance spectra (^{13}C NMR) are reported as δ in units

of parts per million (ppm) downfield from $SiMe_4$ (δ 0.0) and relative to the signal of DMSO-d_6 (δ 39.5).

ESI-MS analyses were carried out on a Micromass Quattro Ultima (LC-ESI/APCI Triple Quadrupole Mass Spectrometer) and HRMS determinations were performed on a Micromass Global Ultima (MALDI/CapLC-ESI Quadrupole Time of Flight Mass Spectrometer).

Fmoc-Thr(*t*Bu)-Wang Resin (100-200 mesh, 0.52-0.60 mmol/g) was purchased from Genemed Synthesis Inc and all protected amino acids for peptide synthesis were obtained from Advanced ChemTech Inc. The other chemicals were purchased from Sigma-Aldrich. 1, 2-dichloroethane was heated at reflux over CaH_2. The other reagents and solvents were used without further purification. Thin layer chromatography (TLC) was performed on pre-coated Merck 60 F_{254} plates and the spots were examined with UV light and sulfuric acid- ethanol spray. All experiments were carried out in dried glassware. Column chromatography was carried out by using silica gel (100- 200 mesh).

2.5.2 YSDSPSTST (1)

Solid phase peptide synthesis was carried on an Advanced ChemTech 348 automated synthesizer. Fmoc-Thr(*t*Bu)-Wang resin (100-200 mesh, 0.52-0.60 mmol/g, 2 × 200 mg) was used as the solid support resin and Fmoc protected amino acids (Tyr 0.92 g in 4 mL DMF, Ser 1.44 g in 10 mL DMF, Asp 0.82 g in 4 mL DMF, Pro 0.67 g in 4 mL DMF and Thr 1.19 g in 6 mL DMF) were used as the starting materials. Side-chain protection used *t*Bu for all amino acids.

Throughout the synthesis, Fmoc removal was achieved with piperidine in DMF (15 mL in 85 mL DMF). Fmoc-amino acids were coupled for a total of 30 h in the presence of HBTU (3.03 g in 16 mL DMF)/DIPEA (4.14 g in 16 mL DMF) at room temperature. After finishing solid phase peptide synthesis, the peptide resin (2 × 200 mg) was transferred to a glass bottle. Peptide YSDSPSTST was cleaved from the resin with concurrent removal of the *t*Bu side-chain protecting groups by treatment with TFA-TIS-H_2O (2 × 2 mL, 95% TFA, 2.5% TIS, 2.5% H_2O) for 24 h at 25 °C. Then the peptide solution was added to a centrifuge tube filled with cold diethyl ether (7 mL). After shaking this tube for half a minute, some precipitate came out and the solution was centrifuged at 3000 rpm and -5 °C for 2 min. The upper solution was removed and the solid was washed with cold diethyl ether (4 × 7 mL) and lyophilized, yielding a white solid. Then this solid was further purified by reversed phase HPLC with a semi-preparative column (t_R = 8.65 min, 1% A/99% B to 9% A/91% B in 12 min, the flow rate is 3.0 mL/min). The desired fractions were pooled and lyophilized to get the target peptide (40 mg) as white powder. **ESI-MS *m/z*:** 944.5 $[M+H]^+$, 966.5 $[M+Na]^+$, 942.5 $[M-H]^-$; **HRMS *m/z* $[M+H]^+$:** obsd 944.3853, calcd 944.3849 for $C_{38}H_{58}N_9O_{19}$; **CapLC analysis**: 90% A/10% B, retention time is 6.82 min and only one peak was shown on CapLC chromatogram.

2.5.3 *N*-(9-Fluorenylmethoxycarbonyl)-(2-acetamido-2-deoxy-3, 4, 6-tri-*O*-acetyl-β-D-glucopyranosyl)-L-Serine (3)

A mixture of 2-acetamido-2-deoxy-3,4,6-tri-*O*-acetyl-α-D-glucopyranosyl chloride (181 mg, 0.50 mmol) and *N*-(9-fluorenylmethoxycarbonyl)-L-serine (82 mg, 0.25 mmol) in 1, 2-dichloroethane (1.5 mL) was heat at reflux with mercuric bromide (198 mg, 0.55 mmol) for 9 h. Then the reaction mixture was washed with a solution of 3% EDTA in water and the organic layer dried over $MgSO_4$ and concentrated. The residue was eluted from a column of silica gel with $CHCl_3$: MeOH: AcOH (16:1:0.1 and then 8:1:0.1) to give the product *O*-GlcNAc-Fmoc-Ser (83 mg, 0.125 mmol, 50%) as an amorphous solid. $\mathbf{R_f}$ = 0.6 ($CHCl_3$: MeOH: AcOH = 8: 1: 0.1); **^{1}H NMR δ_H ($CDCl_3$/CD_3OD, 500 MHz)**: 7.78 (2H, d, *J* 7.3 Hz, Fmoc Ph), 7.64 (2H, d, *J* 7.6 Hz, Fmoc Ph), 7.43-7.31 (4H, m, Fmoc Ph), 5.20 (1H, t, $J_{2,3}$ 9.8 Hz, H-3), 5.02 (1H, t, $J_{3,4}$ 9.8 Hz, H-4), 4.63 (1H, d, $J_{1,2}$ 8.6 Hz, H-1), 4.49 (1H, dd, *J* 10.6 Hz, *J* 6.5 Hz, CH_{2a}Fmoc), 4.43-4.41 (1H, m, CH Ser), 4.39 (1H, dd, *J* 10.6 Hz, *J* 6.8 Hz, CH_{2b}Fmoc), 4.27-4.17 (2H, m, H-6, CH Fmoc), 4.19 (1H, dd, *J* 10.9 Hz, *J* 4.0 Hz, CH_{2a} Ser), 4.12 (1H, dd, $J_{5,6'}$ 2.1 Hz, $J_{6,6'}$ 12.2 Hz, H-6'), 3.89 (1H, dd, *J* 10.9 Hz, *J* 3.3 Hz, CH_{2b} Ser), 3.84 (1H, dd, $J_{1,2}$ 8.6 Hz, $J_{2,3}$ 9.8 Hz, H-2), 3.69 (1-H, ddd, $J_{4,5}$ 9.1 Hz, $J_{5,6}$ 4.5 Hz, $J_{5,6'}$ 2.1 Hz, H-5), 2.08-1.86 (12H, 4s, $COCH_3$); **^{13}C NMR δ_c (DMSO-d_6, 200 MHz):** 171.25, 170.11, 169.71, 169.61, 169.32, 155.98, 143.85, 143.75, 140.74, 127.69, 127.13, 125.28, 120.15, 100.28, 72.54, 70.87, 69.02, 68.47, 65.88, 61.78, 54.10, 53.07, 46.61, 22.63, 20.43, 20.33; **ESI-MS *m/z***: 657.2 $[M+H]^+$, 674.3 $[M+NH_4]^+$, 679.3

$[M+Na]^+$, 695.2 $[M+K]^+$, 655.3 $[M-H]^-$; **HRMS *m/z* $[M+H]^+$:** obsd 657.2269, calcd 657.2296 for $C_{32}H_{37}N_2O_{13}$.

2.5.4 YSDSP(*O*-GlcNAc-OAc_3-Ser)TST (4)

Fmoc-Thr(*t*Bu)-Wang resin (100-200 mesh, 0.52-0.60 mmol/g, 2 × 200 mg) was used as the solid support resin and Fmoc protected amino acids (Tyr 0.92 g in 4 mL DMF, Ser 1.44 g in 10 mL DMF, Pro 0.67 g in 4 mL DMF, Asp 0.82 g in 4 mL DMF, Thr 1.19 g in 6 mL DMF, compound **3** 2 × 200 mg in 2 × 0.5 mL DMF) were used as the starting materials. Side-chain protection used was *t*Bu for all amino acids. Fmoc removal was achieved with piperidine in DMF (15 mL in 85 mL DMF). Fmoc-amino acids were coupled for a total of 30 h in the presence of HBTU (3.03 g in 16 mL DMF)/DIPEA (4.14 g in 16 mL DMF) at room temperature. After finishing SPPS, the glycopeptide resin (2 × 200 mg) was transferred to a glass bottle. *O*-acetylated glycopeptide was cleaved from the resin with concurrent removal of the *t*Bu side-chain protecting groups by treatment with 2 × 2 mL TFA-TIS-H_2O (95% TFA, 2.5% TIS, 2.5% H_2O) for 1.5 h at 25 °C. Then the glycopeptide solution was added to a centrifuge tube filled with about 7 mL cold diethyl ether. After shaking this tube for half a minute, some precipitate came out and the solution was centrifuged for 2 min. The upper solution was removed and the solid was washed with cold diethyl ether (4 × 7 mL). This solid was lyophilized overnight and then purified by reversed phase HPLC with a semi-preparative column (t_R = 9.48 min, from 1% A/99% B to 30% A/70% B in 12 min, the flow rate is 3.0 mL/min). The correct fractions were pooled and lyophilized to

get about 35 mg of acetylated glycopeptide. **ESI-MS *m/z*:** 1273.6 $[M+H]^+$, 1295.6 $[M+Na]^+$, 1311.6 $[M+K]^+$, 1271.6 $[M-H]^-$; **HRMS *m/z* $[M+H]^+$:** obsd 1273.4922, calcd 1273.4960 for $C_{52}H_{77}N_{10}O_{27}$; **CapLC analysis**: 80% A/20% B, retention time is 6.90 min and only one peak was shown on CapLC chromatogram.

2.5.5 YSDSP(*O*-GlcNAc-Ser)TST (2)

The acetylated glycopeptide **4** (25 mg) was transferred to a glass bottle and sodium methoxide in methanol (200 μM, 2 mL) was added. This *O*-deacetylation reaction was monitored by reversed phase HPLC and ESI-MS. After 2 min (if the reaction time was longer than 10 minutes, some side products YSDSP(2-aminopropenoic acid)TST **5** would be obtained), the reaction proceeded completely and some dry ice was added until pH 6 was reached. Then the solvent was removed in vacuo and the residue was purified by semi-preparative reversed phase HPLC (t_R = 15.35 min, from 1% A/99% B to 1% A/99% B in 5 min, then 1% A/99% B to 20% A/80% B in 20 min, the flow rate is 3.0 mL/min). The correct fractions were pooled and lyophilized to get about 20 mg (yield: 89%) glycopeptide **2**. **ESI-MS *m/z*:** 1147.2 $[M+H]^+$, 1145.2 $[M-H]^-$; **HRMS *m/z* $[M+H]^+$:** obsd 1147.4673, calcd 1147.4643 for $C_{46}H_{71}N_{10}O_{24}$; **CapLC analysis**: 90% A/10% B, retention time is 6.94 min and only one peak was shown on CapLC chromatogram.

2.5.6 YSDSP(2-aminopropenoic acid)TST (5)

The acetylated glycopeptide **4** (5 mg) was transferred to a glass bottle and sodium methoxide in methanol (200 μM, 1 mL) was added. After 2 h, the reaction proceeded completely and some dry ice was added until pH 6 was reached. Then the solvent was removed in vacuo and the residue was purified by semi-preparative reversed phase HPLC (t_R = 16.72 min, from 1% A/99% B to 1% A/99% B in 5 min, then 1% A/99% B to 20% A/80% B in 20 min, the flow rate is 3.0 mL/min). The correct fractions were pooled and lyophilized to obtain beta-elimination product **5** (about 2 mg). **ESI-MS *m/z*:** 926.3 $[M+H]^+$, 924.5 $[M-H]^-$; **HRMS *m/z* $[M+H]^+$:** obsd 926.3738, calcd 926.3743 for $C_{38}H_{56}N_9O_{18}$; **CapLC analysis**: 90% A/10% B, retention time is 8.45 min and two peaks (t_R = 7.85 min for unknown impurity and t_R = 8.45 min for compound **5**, about 1: 20) were shown on CapLC chromatogram.

References

1 L.K. Krepprel, M.A. Blomberg and G.W. Hart, *J. Biol. Chem.* **272** (1997) 9308-9315

2 W.A. Lubas, D.W. Frank, M. Krause and J.A. Hanover, *J. Biol. Chem.* **272** (1997) 9316-9324

3 W.A. Lubas and J.A. Hanover, *J. Biol. Chem.* **275** (2000) 10983-10988

4 L.K. Krepprel and G.W. Hart, *J. Biol. Chem.* **274** (1999) 32015-32022

5 N.E. Zachara and G.W. Hart, *Chem. Rev.* **102** (2002) 431-438

6 R.S. Haltiwanger, G.D. Holt and G.W. Hart, *J. Biol. Chem.* **265** (1990) 2563-2568

7 L. A. Carpino and G. Y. Han, *J. Org. Chem.* **37** (1972) 3404-3409

8 E. Atherton, H. Fox, D. Harkiss, C. J. Logan, R. C. Sherppard and B. J. Williams, *J. Chem. Soc. Chem. Commun.* **13** (1978) 537-539

9 E. Atherton, H. Fox, D. Harkiss and R. C. Sherppard, *J. Chem. Soc. Chem. Commun.* **13** (1978) 539-540

10 I. Carvalho, S. L. Scheuerl, K. P. R. Kartha and R. A. Field, *Carbohydr. Res.* **338** (2003) 1039-1043

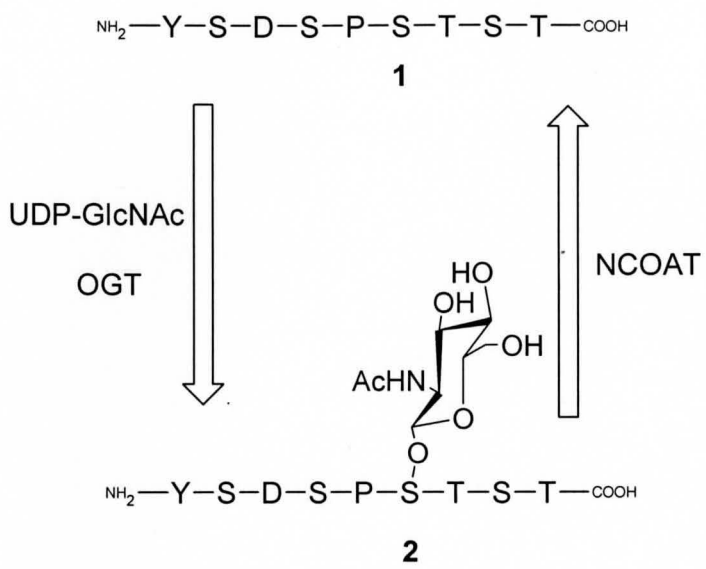
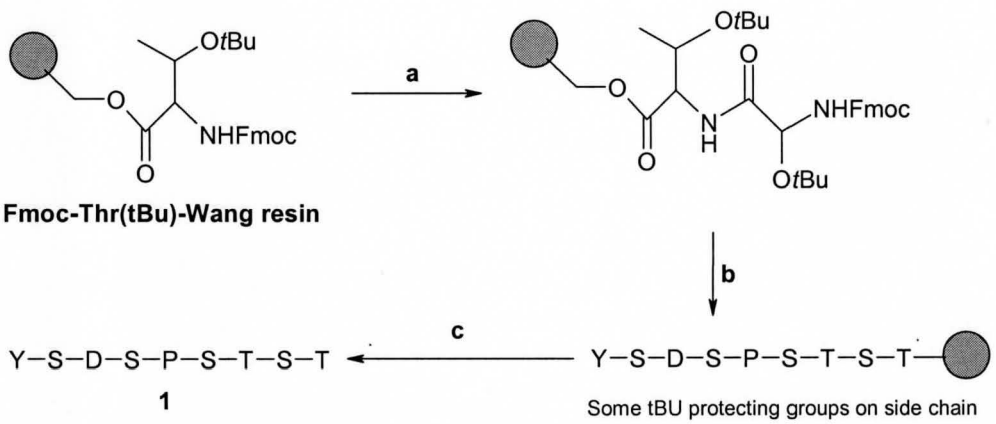
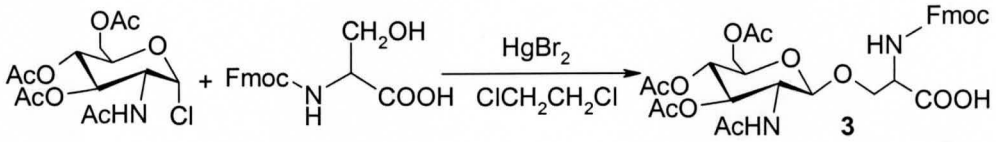


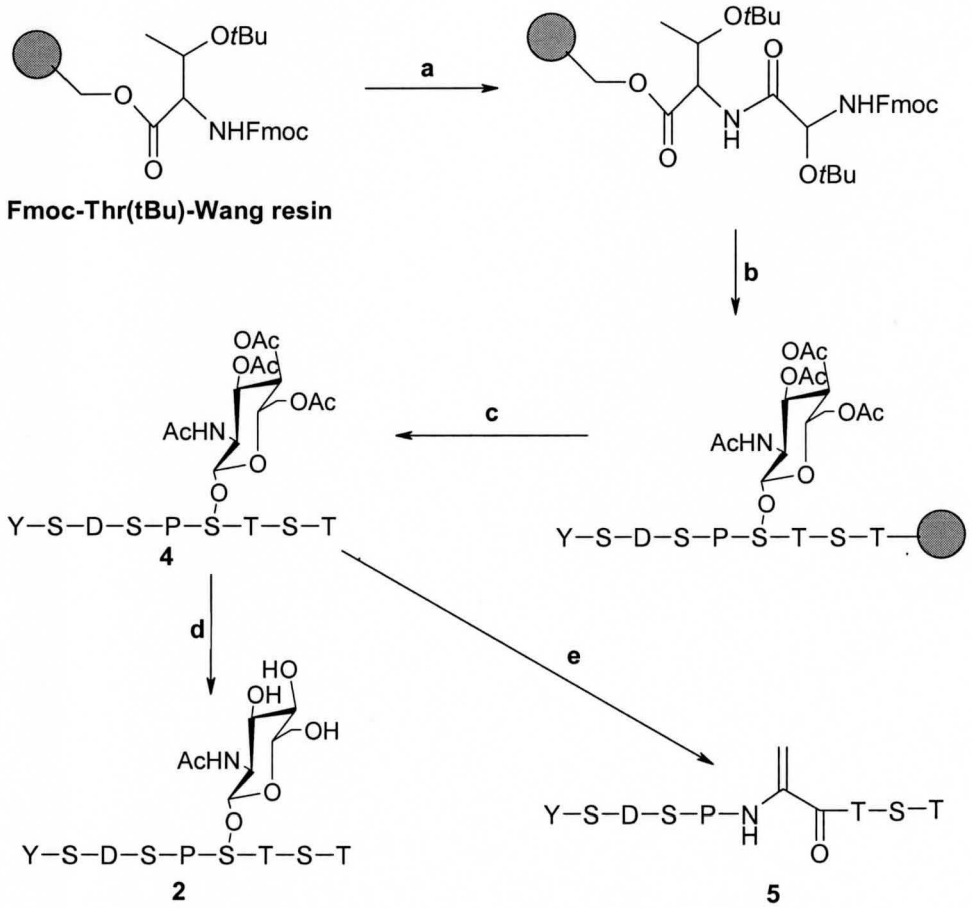
Figure 2.1 The *O*-GlcNAcylation cycle of an optimized enzyme substrate YSDSPSTST



Scheme 2.1 The SPPS route for YSDSPSTST. **(a)** 1. 20% piperidine; 2. Wash with THF then with DMF; 3. Fmoc-Ser(*t*Bu)-OH (5.0 equiv), HBTU (5.0 equiv), DIPEA (10.0 equiv); 4. Wash with THF then with DMF; The whole step **a** is cycle 1; **(b)** Cycle 2: Fmoc-Thr(*t*Bu)-OH; Cycle 3: Fmoc-Ser(*t*Bu)-OH; Cycle 4: Fmoc-Pro-OH; Cycle 5: Fmoc-Ser(*t*Bu)-OH; Cycle 6: Fmoc-Asp(*t*Bu)-OH; Cycle 7: Fmoc-Ser(*t*Bu)-OH; Cycle 8: Fmoc-Tyr(*t*Bu)-OH; Wash with THF then with DMF; **(c)** 95% TFA, 2.5% TIS, 2.5% Water.



Scheme 2.2 Synthesis of the building block *N*-(9-Fluorenylmethoxycarbonyl)-(2-acetamido-2-deoxy-3, 4, 6-tri-*O*-acetyl-β-D-glucopyranosyl)-L-Serine **3**



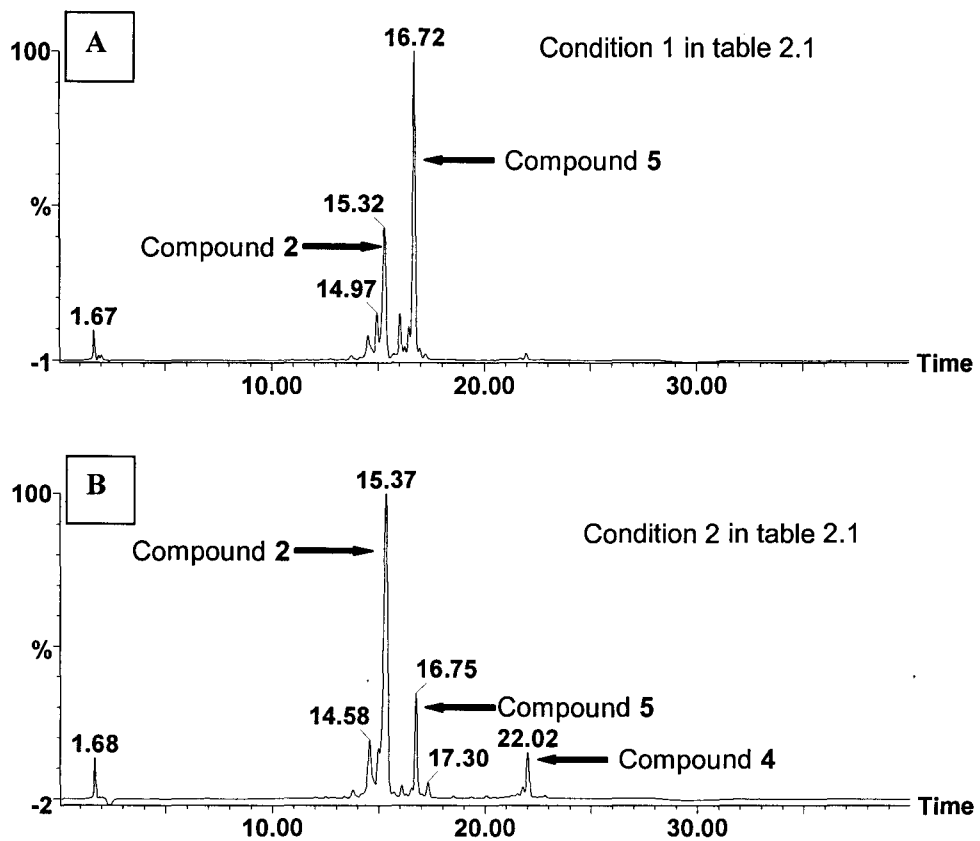
Scheme 2.3 The SPPS route for YSDSP(*O*-GlcNAc-Ser)TST. **(a)** 1. 20% piperidine; 2. Wash with THF then with DMF; 3. Fmoc-Ser(*t*Bu)-OH (5.0 equiv), HBTU (5.0 equiv), DIPEA (10.0 equiv); 4 Wash with THF then with DMF; The whole step **a** is cycle 1; **(b)** Cycle 2: Fmoc-Thr(*t*Bu)-OH; Cycle 3: Building block **3**; Cycle 4: Fmoc-Pro-OH; Cycle 5: Fmoc-Ser(*t*Bu)-OH; Cycle 6: Fmoc-Asp(*t*Bu)-OH; Cycle 7: Fmoc-Ser(*t*Bu)-OH; Cycle 8: Fmoc-Tyr(*t*Bu)-OH; Wash with THF then with DMF; **(c)** 95% TFA, 2.5% TIS, 2.5% Water; **(d)** NaOMe in MeOH (200 μM) 2 min; **(e)** NaOMe in MeOH (20 μM) 2 h.

	Concentrations (NaOMe in MeOH)	Reaction Time
Condition 1	20 μM	2 h
Condition 2	6 μM	2 h
Condition 3	200 μM	2 min

Table 2.1 The conditions for deacetylation of **4**

Time (min)	A%	B%	Flow rate (mL/min)
0	1	99	3 mL/min
5	1	99	3 mL/min
25	20	80	3 mL/min
40	1	99	3 mL/min

Table 2.2 The conditions for HPLC for monitoring the deacetylation from **4** to **2.** Scan wavelength is from 200nm-310nm. A: 0.1% HCOOH in CH_3CN; B: 0.1% HCOOH in water.



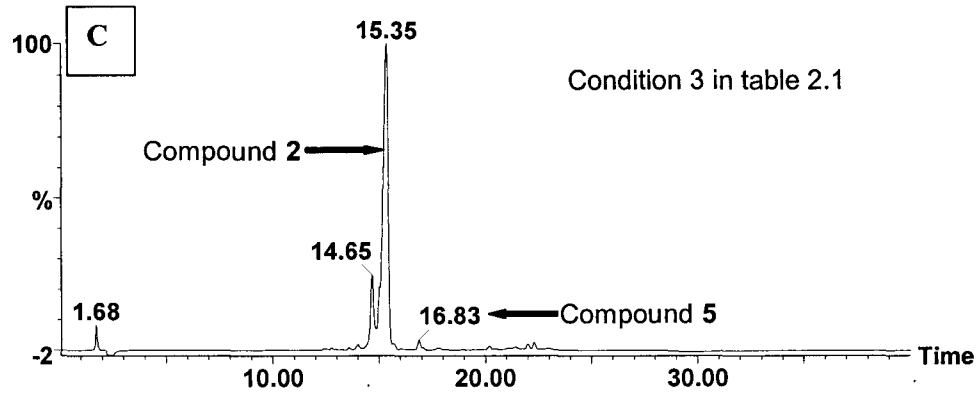


Figure 2.2 HPLC chromatograms for monitoring the deacetylation of **4** to **2**. These HPLC chromatograms used different reaction conditions (Table 2.1) and used the HPLC conditions in table 2.2. **A)** used reaction conditions 1 in Table 2.1; **B)** used reaction conditions 2 in Table 2.1; **C)** used reaction conditions 3 in Table 2.1

Time (min)	A%	B%
0	1	99
24	9	91
25	100	0
30	100	0
31	1	99
40	1	99

Table 2.3 HPLC conditions for separation of YSDSPSTST **1**. Flow rate: 1.4 mL/min; scan wavelength: 200 nm-310 nm.

Time (min)	A%	B%
0	1	99
12	9	91
13	100	0
15	100	0
16	1	99
20	1	99

Table 2.4 HPLC conditions for separation of YSDSPSTST **1**. Flow rate: 3 mL/min; scan wavelength: 200 nm-310 nm.

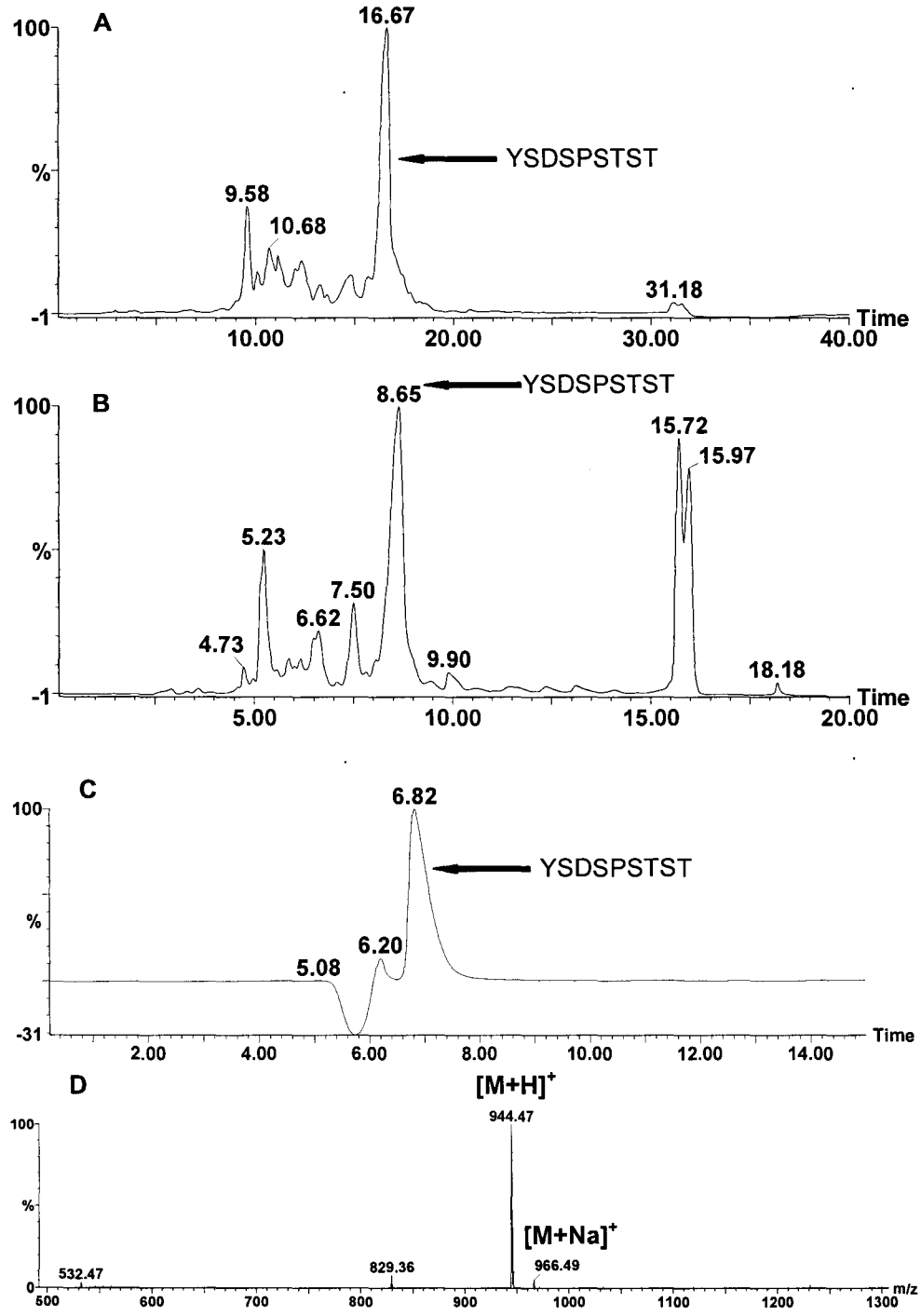
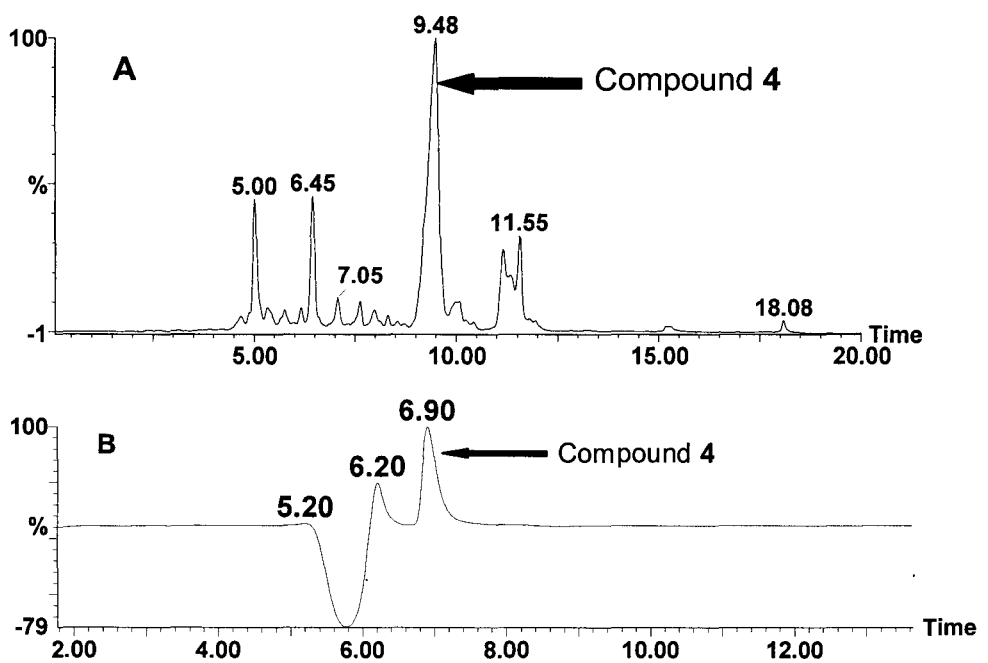

A
100
%
-1
16.67
YSDSPSTST
9.58
10.68
31.18
Time
10.00
20.00
30.00
40.00
B
100
%
-1
8.65
YSDSPSTST
15.72
15.97
5.23
7.50
6.62
4.73
9.90
18.18
Time
5.00
10.00
15.00
20.00
C
100
%
-31
6.82
YSDSPSTST
6.20
5.08
Time
2.00
4.00
6.00
8.00
10.00
12.00
14.00
D
[M+H]+
944.47
100
%
0
[M+Na]+
532.47
829.36
966.49
m/z
500
600
700
800
900
1000
1100
1200
1300

Figure 2.3 HPLC chromatograms, CapLC chromatogram and MS of YSDSPSTST. **A)** HPLC chromatogram of crude YSDSPSTST under HPLC condition in Table 2.3; **B)** HPLC chromatogram of crude YSDSPSTST under HPLC condition in Table 2.4; **C)** CapLC chromatogram of purified YSDSPSTST, CapLC condition 90% A/10% B (A 0.1% HCOOH/1% CH_3CN/H_2O, B 0.1% HCOOH/ CH_3CN), flow rate: 5 μL/min, scan wavelength: 195 nm-320 nm; **D)** MS of purified YSDSPSTST.

Time (min)	A%	B%
0	1	99
12	30	70
13	100	0
15	100	0
16	1	99
20	1	99

Table 2.5 HPLC condition for separation of YSDSP(*O*-GlcNAc-OAc_3-Ser)TST **4**. Flow rate: 3 mL/min; scan wavelength: 200 nm-310 nm.



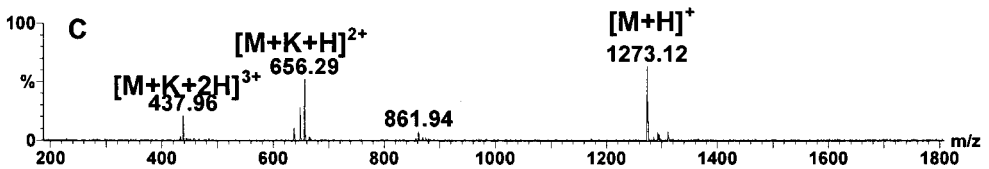


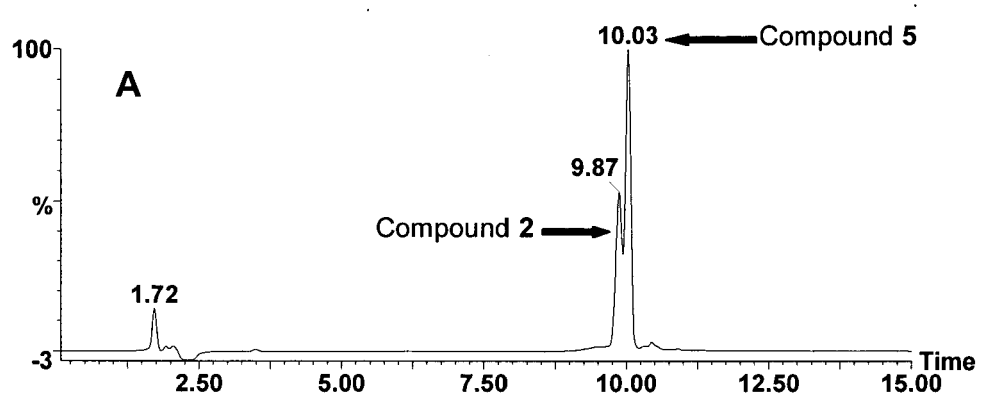
Figure 2.4 HPLC chromatogram, CapLC chromatogram and MS of YSDSP(*O*-GlcNAc-OAc_3-Ser)TST **4**. **A)** HPLC chromatogram of crude **4** under HPLC condition in Table 2.5; **B)** CapLC chromatogram of purified **4**, CapLC condition 80% A/20% B (A 0.1% HCOOH/1% CH_3CN/H_2O, B 0.1% HCOOH/ CH_3CN), flow rate: 5 μL/min, scan wavelength: 195 nm-320 nm; **C)** MS of purified **4**.

Time (min)	A%	B%
0	1	99
5	1	99
15	99	1

Table 2.6 A HPLC method to analyze the result of deacetylation of **4** to **2**. Flow rate: 3 mL/min. scan wavelength: 200 nm-310 nm

Time (min)	A%	B%	Flow rate (mL/min)
0	1	99	3 mL/min
5	1	99	3 mL/min
25	35	65	3 mL/min
40	1	99	3 mL/min

Table 2.7 A HPLC method to analyze the result of deacetylation of **4** to **2**. scan wavelength: 200 nm-310 nm



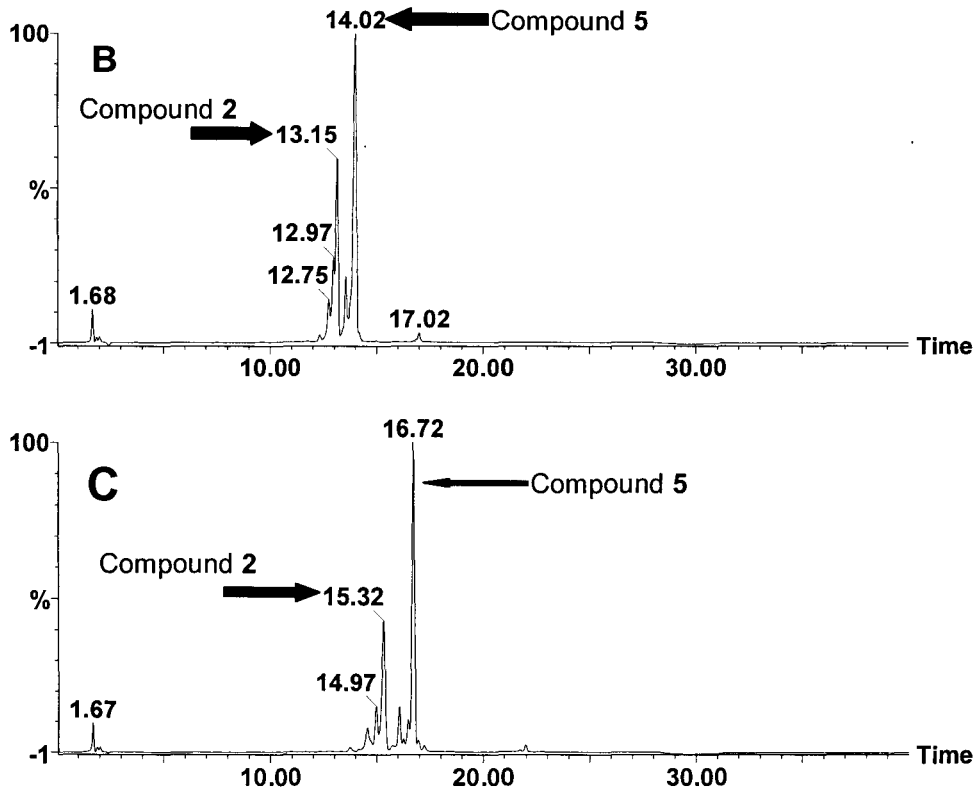
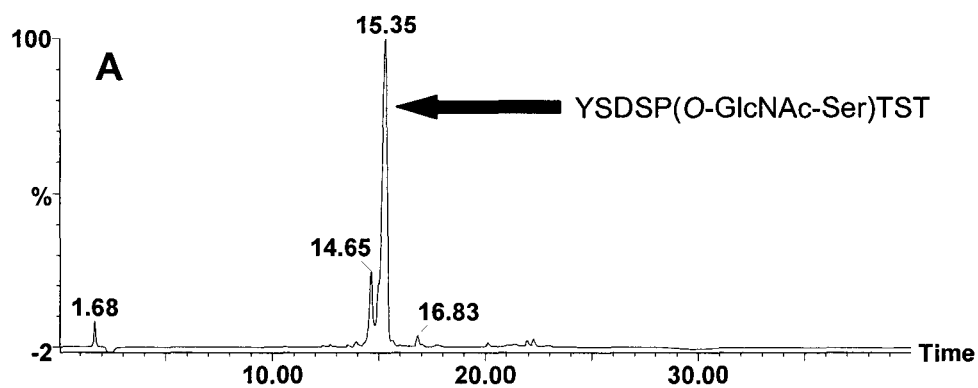


Figure 2.5 HPLC chromatograms of deacetylation of **4** to **2** under the condition 1 in Table 2.1. **A)** used HPLC analytical condition in Table 2.6; **B)** used HPLC analytical condition in Table 2.7; **C)** used HPLC analytical condition in Table 2.2.



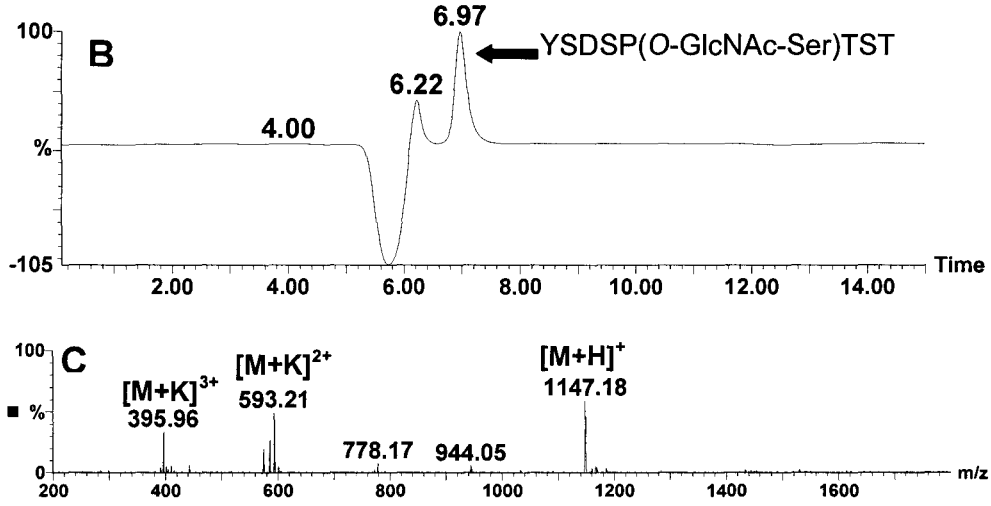


Figure 2.6 HPLC chromatogram, CapLC chromatogram and MS of YSDSP(*O*-GlcNAc-Ser)TST **2**. **A)** HPLC chromatogram of crude **2** under HPLC condition in Table 2.2; **B)** CapLC chromatogram of purified **2**, CapLC condition 90% A/10% B (A 0.1% HCOOH/1% CH_3CN/H_2O, B 0.1% HCOOH/CH_3CN), flow rate: 5 μL/min, scan wavelength: 195 nm-320 nm; **C)** MS of purified **2**.

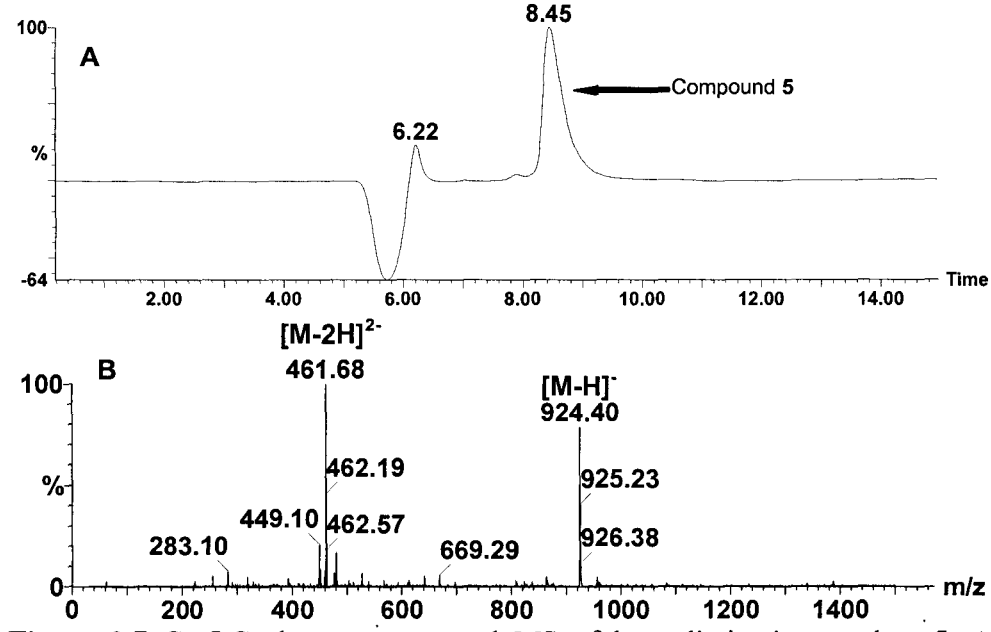


Figure 2.7 CapLC chromatogram and MS of beta-elimination product **5**. **A)** CapLC chromatogram of purified **5**, CapLC condition 90% A/10% B (A 0.1% HCOOH/1% CH_3CN/H_2O, B 0.1% HCOOH/CH_3CN), flow rate: 5 μL/min, scan wavelength: 195 nm-320 nm; **B)** MS of purified **5**.

Chapter 3

Development of CapLC and MS Methods for *O*-GlcNAcylation Study

3.1 Introduction

Of all protein PTMs, glycosylation is by far the most common, and is a target for proteomic research. As chapter 1 has shown, glycosylation affects various important cellular processes and this PTM of proteins is believed to play key roles in controlling many diseases [1]. However, to date, glycosylation analysis still remains a challenging task. MS, in combination with modern separation technologies such as CapLC, is one of the most powerful and versatile tools for studying this kind of PTM of proteins [2-5]. Thus, in this chapter, several CapLC and MS methods for studying *O*-GlcNAcylation were developed.

3.2 Peptide Sequencing by MS/MS

After SPPS and HPLC purifications of the enzyme substrates (YSDSPSTST **1** and YSDSP(*O*-GlcNAc-Ser)TST **2**), and beta-elimination peptide **5**, their sequences and structure needed to be studied. As is well known, tandem mass spectrometry (MS/MS) can be used to provide structural information about a compound by breaking it into daughter ions and identifying these resulting daughter ions. Peptide sequencing is one of the most useful applications

of tandem mass spectrometry in peptide studies. Thus, here, a MS/MS method was employed to determine the structural information and sequences of compounds **1**, **2**, and **5** (For detailed information about SPPS and purifications of these three peptides, please see chapter 2).

Analytical samples were prepared by dissolving the solids **1**, **2**, and **5** in HPLC grade solvent (CH_3CN: H_2O = 1: 1 with 0.5 mM NH_4OAc) to the concentration of 5 pmol/μL. MS/MS experiments were performed on the Micromass Global Q-TOF Ultima at McMaster regional center for mass spectrometry.

For YSDSPSTST **1**, its mass spectrum had already been obtained (see Figure 2.3D) and the ion at *m/z* 944.49 was ascribed to protonated molecular ion peak. Next we employed a daughter ion scanning experiment in which *m/z* 944.49 was broken into many fragments and these fragments were detected as shown on the MS/MS spectrum (Figure 3.1A). As Figure 3.1A shows, there was only one intense peak (*m/z* 492.23), which was due to the y_5 fragment ion of *m/z* 944.49. This means the peptide bond between the Ser residue and the Pro residue is much more easily broken than any other peptide bonds. To get more fragment information, Figure 3.1A was magnified by 24 times (Figure 3.1B). Table 3.1 lists all theoretical a, b, c, x, y, z fragment ions of *m/z* 944.49. The ions marked in red were all fragment ions that were present in the magnified MS/MS spectrum of *m/z* 944.49. As a result, most of the a, b, c, x, y, z fragment ions of *m/z* 944.49 were detected and the MS/MS data matched the sequence of YSDSPSTST **1** very well.

Then the same experiment was performed on beta-elimination product **5**. As Figure 3.2A shows, only one intense peak (*m/z* 472.23, y_5 fragment ion) was found, which was very similar to YSDSPSTST (Figure 3.1A). To get more fragment information, Figure 3.2A was magnified by 6 times (Figure 3.2B). Table 3.2 lists all theoretical a, b, c, x, y, z fragment ions of *m/z* 926.37. The ions marked in red were all fragment ions that were present in the magnified MS/MS spectrum of *m/z* 926.37. As a result, most of the a, b, c, x, y, z fragment ions of *m/z* 926.37 were found and the MS/MS data matched the sequence of beta-elimination product **5** very well.

Then glycopeptide YSDSP(*O*-GlcNAc-Ser)TST **2** was analyzed with the same MS/MS procedure. The MS/MS spectrum (Figure 3.3) showed that the protonated glycopeptide ion (*m/z* 1147.45) was mainly fragmented into *m/z* 944.40, which was ascribed to the protonated YSPSDSTST ion, from removal of *O*-GlcNAc sugar ring, and *m/z* 492.24, which was due to the y_5 fragment ion of the protonated YSPSDSTST ion. It could be concluded that the bond between the sugar and the peptide was much more easily broken than any other bonds. If an MS/MS experiment was further performed on fragment ion *m/z* 944.40, the same spectrum as Figure 3.1 was obtained. From this MS/MS data, the *O*-GlcNAcylation site could not be determined because it was very easy for compound **2** to loose the *O*-GlcNAc sugar ring. The *O*-GlcNAcylation site mapping is discussed in the next paragraph.

3.3 Glycosylation Site Mapping by β-Elimination and MS/MS

Since the MS/MS method described above could not map the *O*-GlcNAcylation site of compound **2** directly, a new assay was developed to solve this problem. Treatment of glycopeptide **2** with the very strong base sodium methoxide in methanol could afford beta-elimination product **5** (Scheme 3.1) [6]. Then we could use MS/MS to detect which residue undergoes the beta-elimination reaction, thus determining the *O*-GlcNAcylation site.

The glycopeptide **2** (100 μg) was transferred to a centrifuge tube (1 mL) and sodium methoxide in methanol (200 μM, 100 μL) was added. After 2 h, the reaction proceeded completely and dry ice was added until a pH of 6 was reached. Then the solution was analyzed on the Micromass Global Q-TOF Ultima at McMaster regional center for mass spectrometry.

As Figure 3.4 shows, *m/z* 926.40 was the protonated product ion after the beta-elimination reaction from **2** (MW = 1146.46 g/mol) to **5** (MW = 925.39 g/mol) and the main fragment daughter ion was *m/z* 474.23, which was the y_5 fragment ion. To search for the beta-elimination site, we amplified Figure 3.4 from *m/z* 490 to *m/z* 730 by 10 times to give Figure 3.5 and in Figure 3.5, *m/z* 550.22 and *m/z* 619.26 were assigned to the b_5 fragment ion and the b_6 fragment ion respectively. The difference in mass (*m/z* 69.04) between the b_5 fragment ion and the b_6 fragment ion corresponded to a 2-aminopropenoic acid residue. This experimental result indicated the 2-aminopropenoic position, thus determining the *O*-GlcNAcylation site of glycopeptide **2**. This beta-elimination method in

combination with MS/MS could thus be a powerful tool to map the *O*-GlcNAcylation sites of proteins after trypsin digestion.

3.4 Standard Curves of Enzyme Substrates for CapLC

In this project, CapLC was used as a major method to study *O*-GlcNAcylation. Determining the standard curves of enzyme substrates (**1** and **2**) would facilitate the quantitative study of *O*-GlcNAcylation.

Compounds **1**, **2** and **5** were dissolved in HPLC grade water to a certain concentration respectively. All CapLC analyses were carried out on a Waters Delta CapLC with a Waters NanoEase^{TM} column (Atlantis^{TM} dC_{18} 5 μm 300 μm × 150 mm column) and analytical conditions employed 90% A/10% B (A 0.1% HCOOH/1% CH_3CN/H_2O, B 0.1% HCOOH/CH_3CN), flow rate: 5 μL/min, scan wavelength: 195 nm-320 nm. Representative CapLC chromatograms of **1**, **2**, and **5** are shown in Figure 2.3C, Figure 2.6B and Figure 2.7A respectively. The relationship between the injected quantities (pmol as unit) of compounds with their UV response (area under UV response peaks) gave CapLC standard curves that are shown in Figure 3.6, Figure 3.7 and Figure 3.8. It is also noteworthy that when the injected quantities of all three compounds were reduced to 0.4 pmol, these three compounds could not be discriminated from the noise of the CapLC machine. Thus, it could be concluded that the limit of detection for CapLC machine was around 0.5 pmol, which would assist in the design of the CapLC assays.

3.5 CapLC-MALDI MS Analysis of Protein Digests

After proteins are modified (such as by glycosylation and photoaffinity labeling), they need to be analyzed to determine the modification sites. It is very difficult to detect the proteins directly by MS due to their high molecular weight. Usually, they need to be digested to fragments by enzymes such as trypsin before analysis. These fragments can then be analyzed by CapLC-MALDI MS to map the modified sites of the proteins. Here, α-crystallin was chosen as the research target, because one of its residues could be *O*-GlcNAcylated. We wanted to see if we could use CapLC-MALDI MS to detect this *O*-GlcNAcylation site. Also, developing a CapLC-MALDI MS analytical assay is useful for studying OGT structural information in the near future.

α-Crystallin was purchased from Sigma-Aldrich and it was digested by the following procedure: A solution of 40 μg α-crystallin, 25 μL (0.48 g/mL urea, 410 mM NH_4HCO_3) water solution, and 125 μL water was incubated at 65 °C for 15 min to unfold the protein. After this mixture was cooled to room temperature, 10 μL (0.1 g/L) trypsin solution was added. Then the mixture was incubated at 58 °C for 1 h. Then 2.6 μL (0.1 g/L) trypsin solution was added. Finally, the mixture was incubated at 37 °C for 12 h. The resulting solution was stored at -80 °C until analyzed by CapLC-MALDI MS.

CapLC analysis was performed on a Waters Delta CapLC with a Waters NanoEase™ column (Atlantis™ dC_{18} 5 μm, 300 μm × 150 mm column) and analytical conditions employed 90% A/10% B (A 0.1% HCOOH/1%

CH_3CN/H_2O, B 0.1% HCOOH/CH_3CN), flow rate: 5 μL/min, scan wavelength: 195 nm-320 nm. The CapLC samples were collected with a Waters LC-MALDIprep sample collect module. Each spot collected 1 min CapLC eluent fractions and each inject volume was 2 μL for CapLC.

MALDI MS analyses were carried out on a Micromass TofSpec 2E at McMaster regional center for mass spectrometry. The MALDI matrix employed 0.5 mM α-CHCA solution (in CH3CN: 0.1% TFA in H_2O = 1: 1). Each spot contained 2 μL of this matrix solution.

The CapLC analytical conditions are shown in Table 3.3, which was optimized previously and is suitable for analyzing most protein digests. The CapLC chromatogram of α-crystallin trypsin digests is shown in Figure 3.9 and the detected fragments of α-crystallin trypsin digests by MALDI MS are listed in Table 3.4, which were also assigned to the corresponding peaks on the CapLC trace (Figure 3.9). These experimental results showed that most fragments in the α-crystallin trypsin digests could be detected by CapLC-MALDI MS, but the *O*-GlcNAcylation site (162 Ser residue could be *O*-GlcNAcylated, it was T19-20, *m/z* of protonated modified fragment T19-20 should be 1844.89) could not be found in the MALDI mass spectra. This result could be due to one or more of the following three reasons: 1) a small amount of α-crystallin was *O*-GlcNAcylated, so that it could not be detected by MALDI MS; 2) *O*-GlcNAcylated peptides tend to loose the sugar under the MALDI MS operation conditions, thus making it very difficult to detect them; 3) the MALDI MS instrument could not recognise *O*-

GlcNAcylated peptides in the presence of a large excess of nonglycosylated ones due to the MALDI instrument suppression effect. Although the *O*-GlcNAcylation site of α-crystallin could not be confirmed by this CapLC-MALDI MS assay due to a lot of difficulties, this assay still could analyze most trypsin digests of α-crystallin, which will help us study the structure information of OGT in the near future.

3.6 Suppression Effect Study of MALDI MS for Detecting Glycopeptides

The yield of *O*-GlcNAcylation of our substrate YSDSPSTST **1** is about 1%-5%. Thus, if we wanted to detect the *O*-GlcNAcylated peptide **2** directly by MS, we need to know the limit of detection of glycopeptide by MALDI MS and also need to overcome the difficulties the MALDI instrument suppression effect brought. In this section, the limit of detection of glycopeptide by MALDI MS was measured and the MALDI instrument suppression effect was studied.

Compounds **1** and **2** were dissolved in HPLC grade water to a certain concentration. MALDI MS analyses were carried out on a Micromass TofSpec 2E at McMaster regional center for mass spectrometry. The MALDI matrix employed 0.5 mM α-CHCA solution (in CH3CN: 0.1% TFA in H_2O = 1: 1). Each spot contained 2 μL of this matrix solution.

Figure 3.10 shows that when 5 pmol of glycopeptide **2** was spotted, it was easily detected (*m/z* 1169.58 was sodiated glycopeptide ion in Figure 3.10A);

when the quantity of **2** was deceased to 1 pmol, it still could be detected (Figure 3.10B); but when the quantity of **2** was further deceased to 0.1 pmol, compound **2** could not be detected by MALDI MS (Figure 3.10C). This means that the limit of detection of **2** by MALDI MS should be 0.1 pmol-1 pmol.

Then we employed a mixture of **1** and **2** to study the MALDI instrument suppression effect. In this experiment, the same quantity of **2** (5 pmol) and different amounts (10 pmol, 100 pmol and 500 pmol) of **1** were used. The results showed that when the ratio of **1** to **2** was 2: 1 (10 pmol: 5 pmol), glycopeptide **2** could be detected obviously (Figure 3.11A); when the ratio of **1** to **2** was increased to 20: 1 (100 pmol: 5 pmol), the intensity of the sodiated ion peak of glycopeptide **2** (*m/z* 1169.61) was greatly reduced (Figure 3.11B); when the ratio of **1** to **2** was further reduced to 100: 1 (500 pmol: 5 pmol), glycopeptide **2** could not be detected by MALDI MS (Figure 3.11C). From these experimental results, it could be concluded that the MALDI instrument suppression effect could greatly affect the detection of glycopeptides by MALDI MS and that we should be careful of this problem when we designed the assay for studying *O*-GlcNAcylation of peptides.

3.7 Ion Mobility Study of 1, 2 and 5

Ion mobility spectrometry (IMS) is an analysis method which measures the mobility (a parameter that is dependent of ion mass, shape and size) of ions in the gas phase. IMS couples extremely good sensitivity with fast operation speed and operation simplicity. IMS coupled with a mass spectrometer can make it

possible to obtain both size and mass information of unknown chemicals simultaneously [7]. All these features have allowed IMS to develop as a research tool for the analysis of proteomics and metabolomics. In this section, some preliminary ion mobility studies of compounds **1**, **2** and **5** are introduced.

Analytical samples were prepared by dissolving the solids **1**, **2**, and **5** in the HPLC grade solvent (CH_3CN: H_2O = 1: 1 with 0.5 mM NH_4OAc) to the concentration of 10 pmol/μL. IMS experiments were performed on an ion mobility spectrometer at York University.

Table 3.5 lists all ion mobility values for many ions of compounds **1**, **2** and **5**. Table 3.6, Table 3.7 and Table 3.8 show the ion mobility changes with charge state transformation (from 1 to 2), Na effect (Na replacing H) and mode switch (from positive mode to negative mode). As Table 3.6 shown, when the charge state was increased from 1 to 2, $[M+2H]^{2+}$ for compound **1** and $[M+Na-3H]^{2-}$ for compound **2** showed a small increase in mobility (taking into consideration the 2% experimental error, percentage change has a derived 2.8% error, and has to be greater than 2.8%). As Table 3.7 shows, only $[M+Na+H]^{2+}$for **2** and $[M+Na-3H]^{2-}$for **2** showed a slight increase in mobility compared to their corresponding protonated ions. As Table 3.8 shows, compared with their corresponding positive ions, all negative ions for compound **1** and compound **5** showed more decrease in mobility than all negative ions of compound **2**. The ion mobility values of compound **2** were greater than those of compound **1** and compound **5**, due to the greater molecular weight of **2** than that of **1** and **5**. Thus,

IMS could be considered as a tool to separate compound **1** and compound **2** due to the big difference between their ion mobility, since it is very hard to separate them by HPLC and CapLC (the retention time for these two compounds is very close).

3.8 Conclusions

Compounds **1**, **2** and **5** were characterized by MS/MS methods. The *O*-GlcNAcylation site of **2** was confirmed by beta-elimination and MS/MS. Several CapLC and MS methods were developed for studying *O*-GlcNAcylation.

References

1 N.E. Zachara and G.W. Hart, *Chem. Rev.* **102** (2002) 431-438

2 H. Geyer and R. Geyer, *Biochim. Biophys. Acta* **1764** (2006) 1853-1869

3 M. Wuhrer, A. M. Deelder and C. H. Hokke, *J. Chromatogr. B.* **825** (2005) 124-133

4 M. Mann and O. N. Jensen, *Nat. Biotechnol.* **21** (2003) 255-261

5 W. Morelle, K. Canis, F. Chirat, V. Faid and J. C. Michalski, *Proteomics* **6** (2006) 3993-4015

6 K. Vosseller, K.C. Hansen, R.J. Chalkley, J.C. Trinidad, L. Wells, G.W. Hart and A.L. Burlingame, *Preteomics* **5** (2005) 388-398.

7 C. S. Creaser, J. R. Griffiths, C. J. Bramwell, S. Noreen, C. A. Hill and C. L. P. Thomas, *Analyst* **129** (2004) 984-994

8 A. P. Snyder, Interpreting Protein Mass Spectra, Oxford University Press (NY), 2000, 160-161

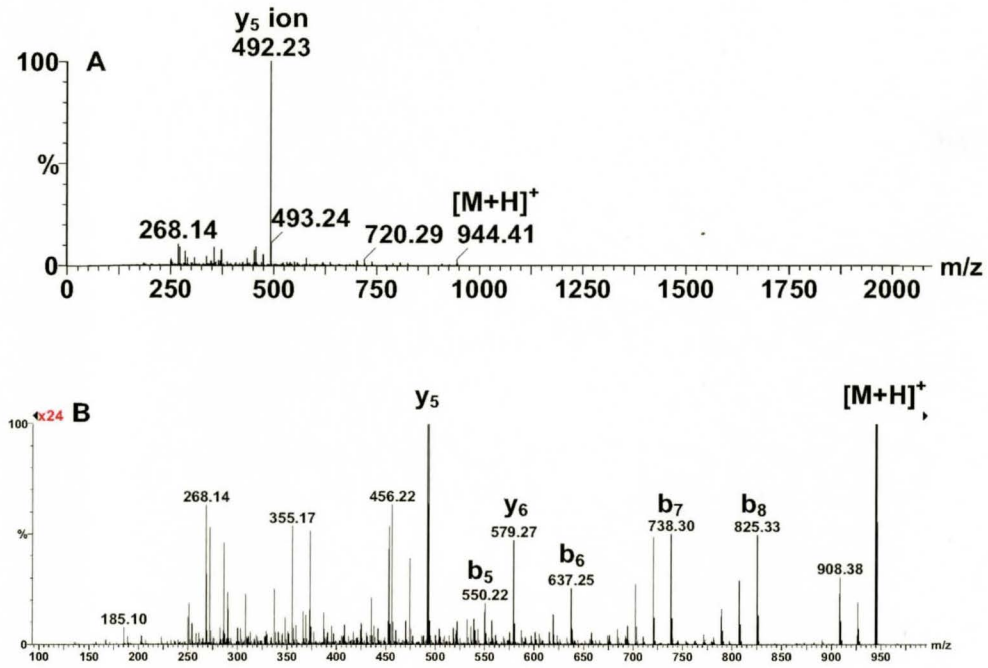


Figure 3.1 MS/MS spectra of *m/z* 944.41 (YSDSPSTST **1**). **A)** MS/MS spectrum of *m/z* 944.41; **B)** expanded MS/MS spectrum of *m/z* 944.41

136.08 (a_1)	223.11 (a_2)	338.14 (a_3)	425.17 (a_4)	552.22 (a_5)	609.25 (a_6)	710.30 (a_7)	797.33 (a_8)
164.07 (b_1)	251.10 (b_2)	366.13 (b_3)	453.16 (b_4)	550.21 (b_5)	637.25 (b_6)	738.29 (b_7)	825.33 (b_8)
179.08 (c_1)	266.11 (c_2)	381.14 (c_3)	468.17 (c_4)	565.23 (c_5)	652.26 (c_6)	753.31 (c_7)	840.34 (c_8)
148.06 (x_1)	235.09 (x_2)	336.14 (x_3)	423.17 (x_4)	520.23 (x_5)	607.26 (x_6)	722.28 (x_7)	809.32 (x_8)
120.07 (y_1)	207.10 (y_2)	308.15 (y_3)	395.18 (y_4)	492.23 (y_5)	579.26 (y_6)	694.29 (y_7)	781.32 (y_8)
105.06 (z_1)	192.09 (z_2)	293.13 (z_3)	380.17 (z_4)	477.22 (z_5)	564.25 (z_6)	679.28 (z_7)	766.31 (z_8)

Table 3.1 Theoretical a, b, c, x, y, z fragments of *m/z* 944.41. The ions identified in red were the fragments that were detected in the MS/MS spectrum (Figure 3.1B). For nomenclature of the fragments, see reference 8

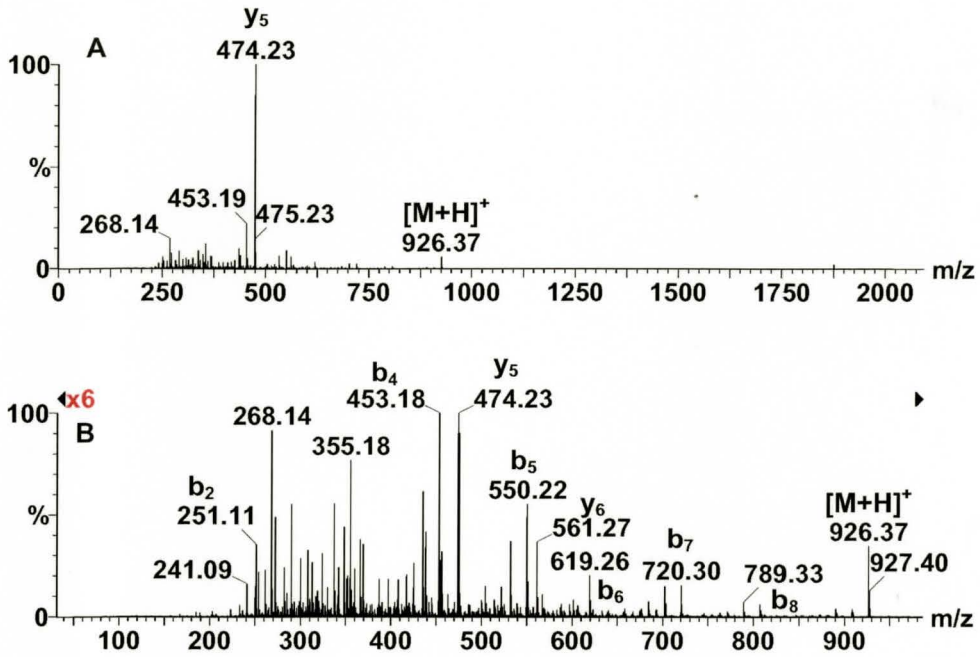


Figure 3.2 MS/MS spectra of *m/z* 926.37 (beta-elimination product **5**). **A)** MS/MS spectrum of *m/z* 926.37; **B)** expanded MS/MS spectrum of *m/z* 926.37

136.08 (a_1)	223.11 (a_2)	338.14 (a_3)	425.17 (a_4)	552.22 (a_5)	591.24 (a_6)	692.29 (a_7)	779.32 (a_8)
164.07 (b_1)	251.10 (b_2)	366.13 (b_3)	453.16 (b_4)	550.21 (b_5)	619.24 (b_6)	720.28 (b_7)	807.32 (b_8)
179.08 (c_1)	266.11 (c_2)	381.14 (c_3)	468.17 (c_4)	565.23 (c_5)	634.25 (c_6)	735.29 (c_7)	822.33 (c_8)
148.06 (x_1)	235.09 (x_2)	336.14 (x_3)	405.16 (x_4)	502.21 (x_5)	589.25 (x_6)	704.27 (x_7)	791.31 (x_8)
120.07 (y_1)	207.10 (y_2)	308.15 (y_3)	377.17 (y_4)	474.22 (y_5)	561.25 (y_6)	676.28 (y_7)	763.31 (y_8)
105.06 (z_1)	192.09 (z_2)	293.13 (z_3)	362.16 (z_4)	459.21 (z_5)	546.24 (z_6)	661.27 (z_7)	748.30 (z_8)

Table 3.2 Theoretical a, b, c, x, y, z fragments of *m/z* 926.37. The ions identified in red were the fragments that were detected in the MS/MS spectrum (Figure 3.2B)

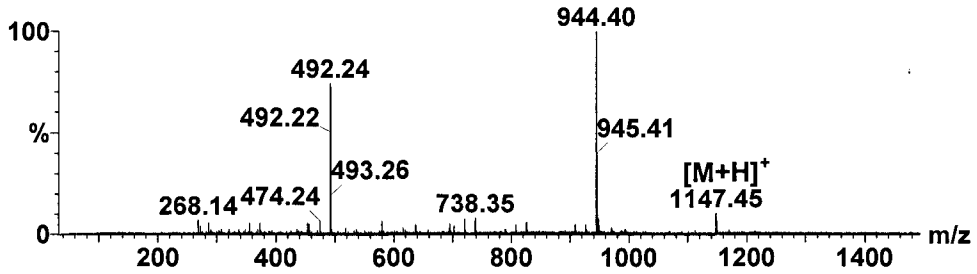
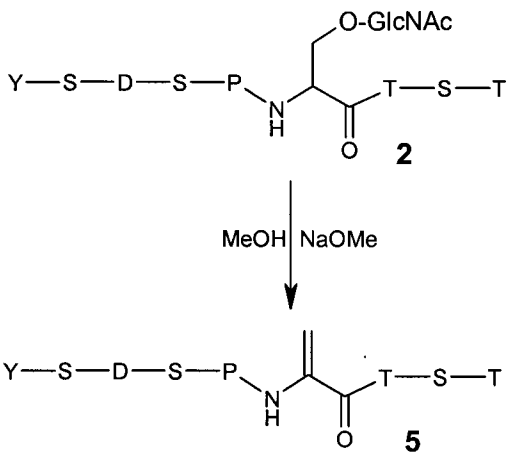


Figure 3.3 MS/MS spectrum of *m/z* 1147.45 (glycopeptide **2**)



Scheme 3.1 Beta-elimination reaction from **2** to **5** under strongly basic condition

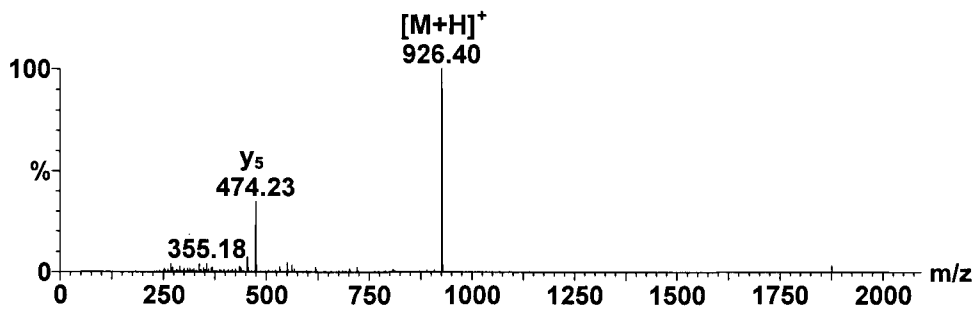


Figure 3.4 MS/MS spectrum of protonated beta-elimination product ion *m/z* 926.40

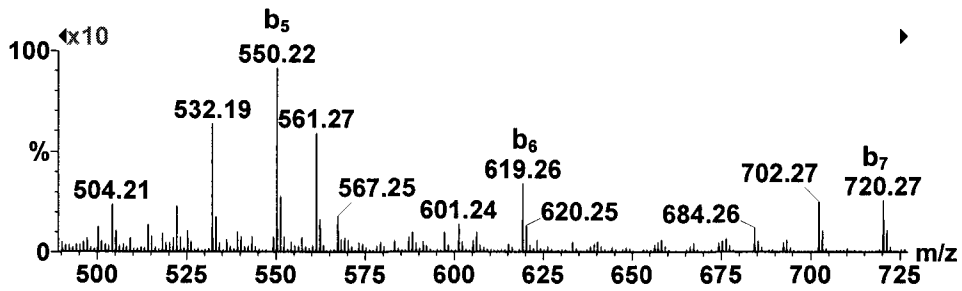


Figure 3.5 Partial MS/MS spectrum of protonated beta-elimination product ion *m/z* 926.37 (From *m/z* 490 to *m/z* 730)

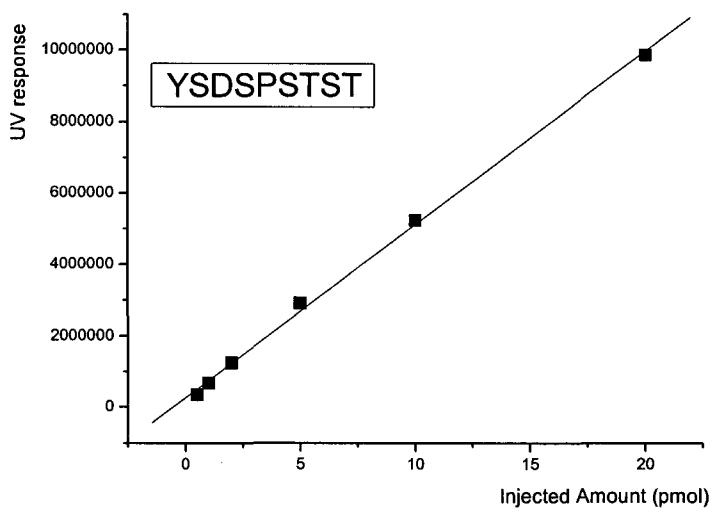


Figure 3.6 Standard curve of YSDSPSTST for CapLC. (Y = 474724.09X + 265170.42, R = 0.99914, A = 265170.42 ± 94474.56, B = 474724.09 ± 10049.63 for Y = BX+A). Each data point is the average of 3 determinations. Error bars are not shown as the σ for each set of 3 was less than 1%

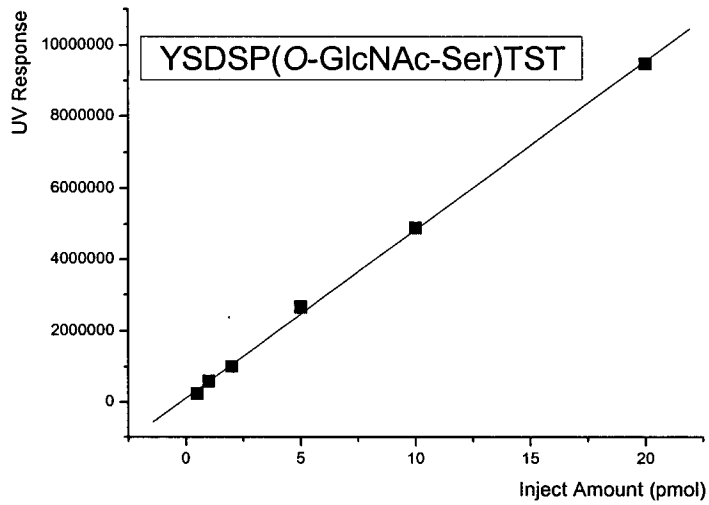


Figure 3.7 Standard curve of YSDSP(*O*-GlcNAc-Ser)TST for CapLC. (Y = 471124.16X + 115936.47, R = 0.99951, A = 115936.47 ± 69491.29, B = 471124.16 ± 7392.06 for Y = BX+A). Each data point is the average of 3 determinations. Error bars are not shown as the σ for each set of 3 was less than 1%

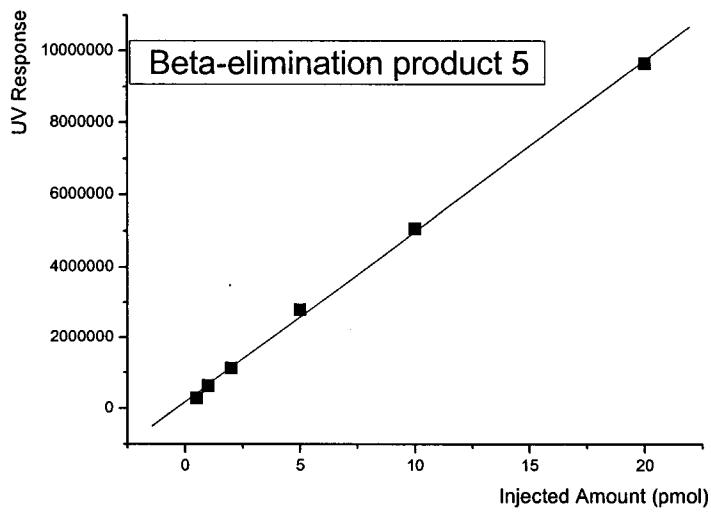


Figure 3.8 Standard curve of beta-elimination product **5** for CapLC. (Y = 477559.33X + 189832.88, R = 0.99933, A = 189832.88 ± 73421.32, B = 477559.33X ± 8512.22 for Y = BX+A). Each data point is the average of 3 determinations. Error bars are not shown as the σ for each set of 3 was less than 1%

Time (min)	A%	B%
0	2	98
26	54	46
36	54	46
40	2	98
60	2	98

Table 3.3 CapLC condition for analyzing α-crystallin trypsin digests. (A 0.1% HCOOH/1% CH_3CN/H_2O, B 0.1% HCOOH/CH_3CN), flow rate: 5 µL/min, scan wavelength: 195 nm-320 nm

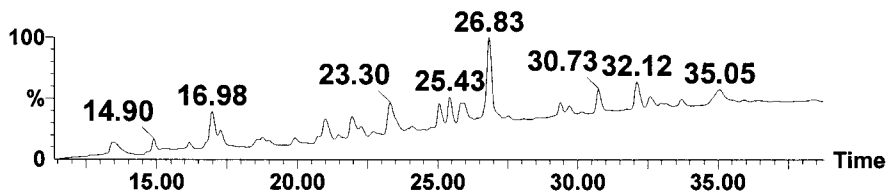


Figure 3.9 CapLC chromatogram of α-crystallin trypsin digests. Using the analytical conditions in Table 3.3

Fragment	*m/z* of $[M+H]^+$ on mass spectrum	CapLC Retention time (min)
T3	1037.54	23.30
T4	3364.65	35.05
T5-6	1806.97	32.12
T6	1175.63	25.40
T8-9	1223.70	25.43
T9	980.58	16.98
T10	1172.60	25.40
T11	1300.65	26.83
T13	1090.49	22.15
T16-17	2944.42	33.72
T18	1224.60	25.92
T19-20	1641.84	30.73
T20	1018.47	21.13

Table 3.4 The detected fragments of α-crystallin trypsin digests by MALDI MS. CapLC retention time corresponds to the CapLC peaks on Figure 3.9

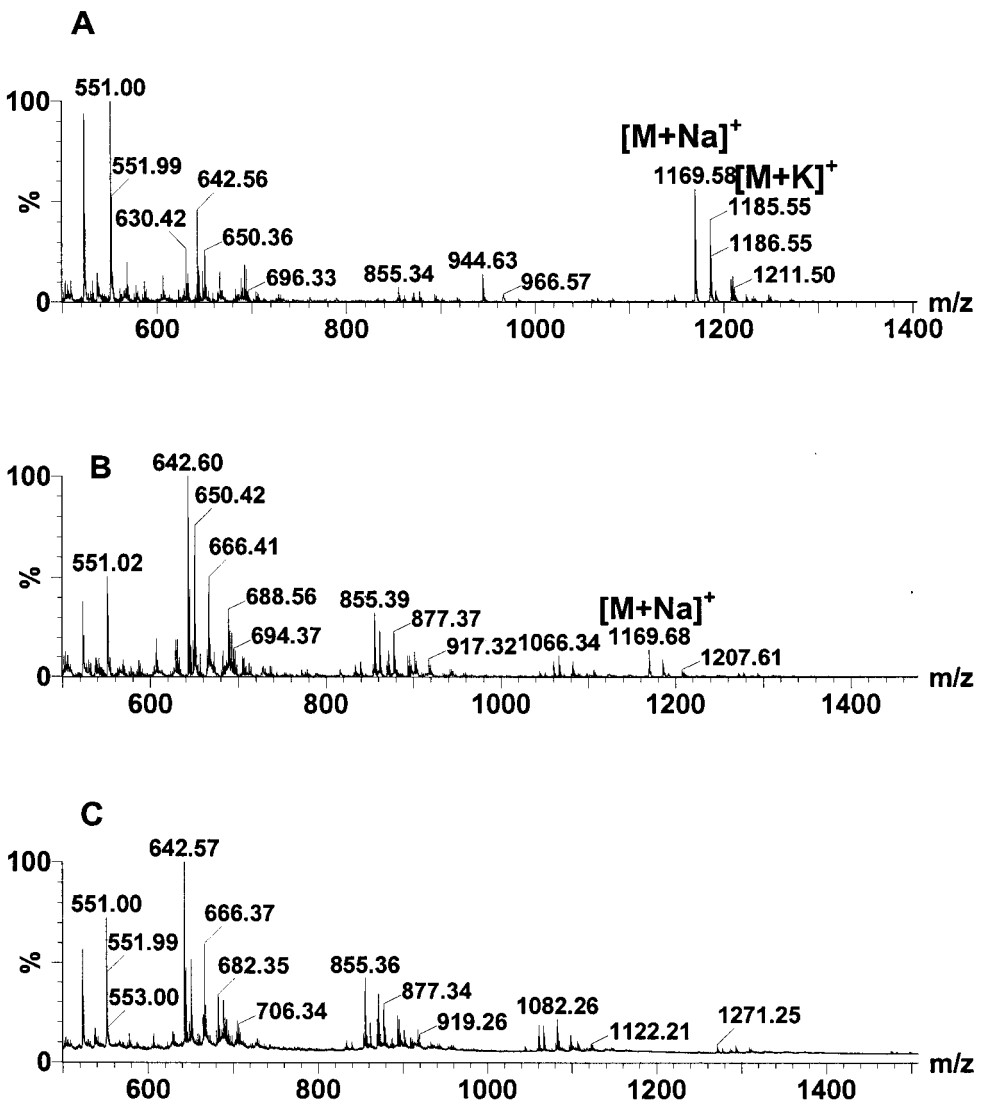


Figure 3.10 MALDI mass spectra of glycopeptide **2** with different quantities. **A)** 5 pmol; **B)** 1 pmol; **C)** 0.1 pmol

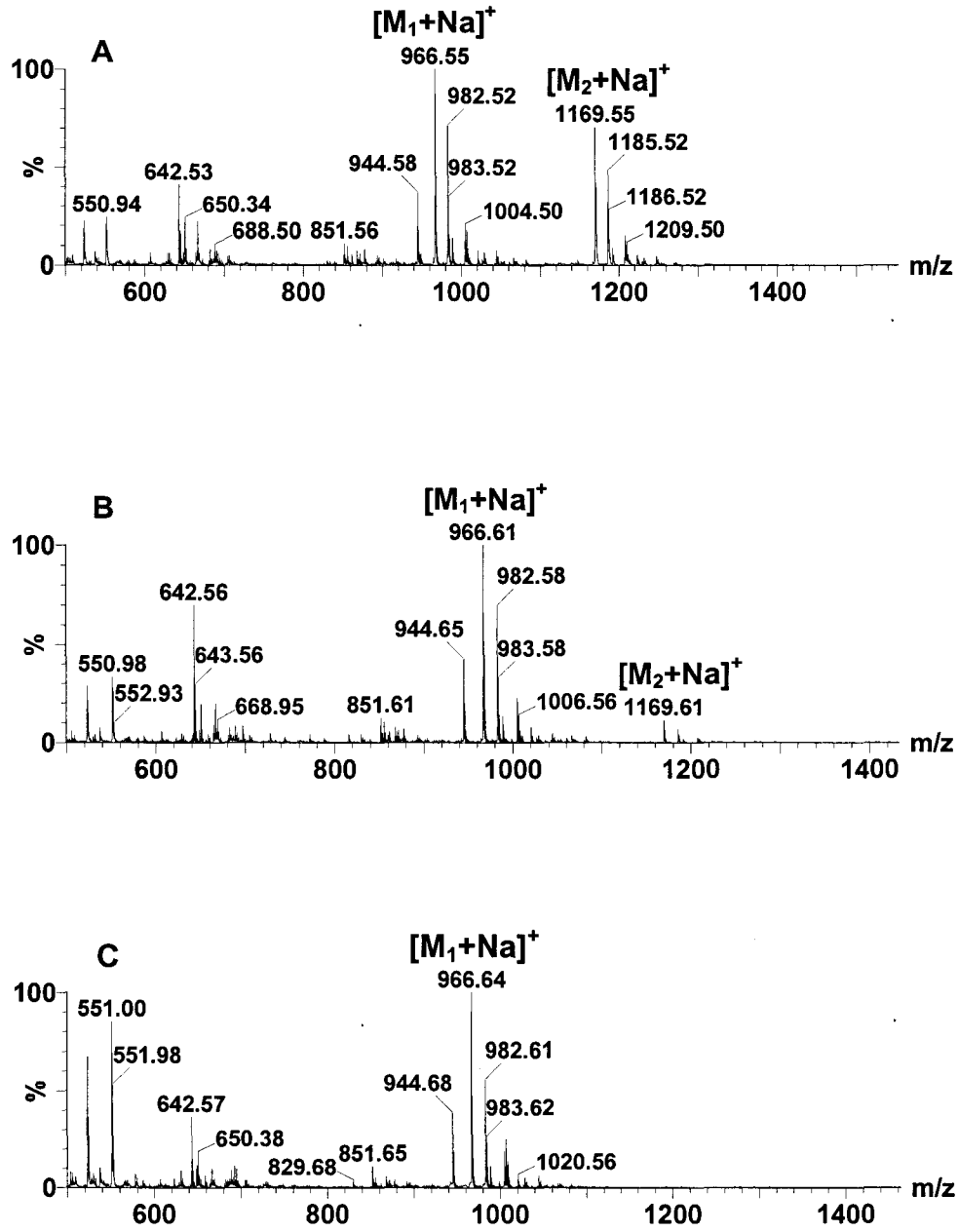


Figure 3.11 MALDI mass spectra of peptide **1** and glycopeptide **2** with different ratio. **A)** 10 pmol **1** and 5 pmol **2**; **B)** 100 pmol **1** and 5 pmol **2**; **C)** 500 pmol **1** and 5 pmol **2**; M_1 for YSDSPSTST **1** and M_2 for glycopeptide **2**

	Compound **1**	Compound **2**	Compound **5**
$[M+H]^+$	202	235	207
$[M+2H]^{2+}$	209	235	208
$[M-H]^-$	196	233	196
$[M-2H]^{2-}$	196	233	195
$[M+Na]^+$	207	239	209
$[M+Na+H]^{2+}$	212	243	213
$[M+Na-2H]^-$	201	232	200
$[M+Na-3H]^{2-}$	196	241	195

Table 3.5 Ion mobility values for all ions of **1**, **2**, and **5**. Units are Ω/A^2 for ion mobility values.

	Compound **1**	Compound **2**	Compound **5**
$[M+H]^+$ to $[M+2H]^{2+}$	3.5%	0	0.5%
$[M-H]^-$ to $[M-2H]^{2-}$	0	0	-0.5%
$[M+Na]^+$ to $[M+Na+H]^{2+}$	2.4%	1.7%	1.9%
$[M+Na-2H]^-$ to $[M+Na-3H]^{2-}$	-2.5%	3.9%	-2.5%

Table 3.6 The changes of ion mobility values from 1 charge to 2 charges.

	Compound **1**	Compound **2**	Compound **5**
$[M+H]^+$ to $[M+Na]^+$	2.5%	1.7%	1.0%
$[M-H]^-$ to $[M+Na-2H]^-$	2.6%	0.4%	2.0%
$[M+2H]^{2+}$to $[M+Na+H]^{2+}$	1.4%	3.4%	2.4%
$[M-2H]^{2-}$to $[M+Na-3H]^{2-}$	0	3.4%	0

Table 3.7 The changes of ion mobility values from protonated ions to sodiated ions.

	Compound **1**	Compound **2**	Compound **5**
$[M+H]^+$ to $[M-H]^-$	-3.0%	-0.9%	-5.3%
$[M+2H]^{2+}$to $[M-2H]^{2-}$	-6.2%	-0.9%	-6.3%
$[M+Na]^+$to $[M+Na-2H]^-$	-2.9%	-2.9%	-4.3%
$[M+Na+H]^{2+}$to $[M+Na-3H]^{2-}$	-7.5%	-0.8%	-8.5%

Table 3.8 The changes of ion mobility values from positive ions to negative ions.

Chapter 4

Synthesis of *S*-GlcNAc-Serine as the Building Block for Preparing *S*-GlcNAcylated Peptides

4.1 Introduction

Many nuclear and cytoplasmic proteins are covalently modified by the attachment of a single monosaccharide *N*-acetylglucosamine (GlcNAc) to serine or threonine, a reaction catalyzed by the enzyme *O*-GlcNAc transferase (OGT) [1]. On the other hand, *O*-GlcNAc can be removed from modified proteins and returned to their unmodified state by a family 84 glycoside hydrolase known as *O*-GlcNAcase [2-4].

More than 200 proteins to date have been shown to be *O*-GlcNAc-modified [5] and these proteins with this kind of modification play a role in a number of processes of fundamental importance to cell survival, including cell cycle regulation and cell division [6-9], import and export through the nuclear pore [7, 10], transcription [11-15], translation [15], and others. Unlike the many phosphatases which exist in a cell, *O*-GlcNAcase is the only nuclear and cytoplasmic enzyme known to be capable of removing *O*-GlcNAc residues and it resides within a larger protein, termed nuclear and cytoplasmic O-GlcNAcase and acetyltransferase (NCOAT), which contains a C-terminal acetyltransferase domain [16].

As a result of the biological importance of *O*-GlcNAcase, small molecule inhibitors have received more and more attention [17-21], both as probes for discovering the role of *O*-GlcNAcase in biological processes [22], and in potential therapeutic applications [19, 23]. Thus, our preliminary interest mainly focuses on designing and preparing inhibitors of NCOAT, particularly those that function at the domain for catalyzing the removing *O*-GlcNAc from modified proteins.

4.2 The Design of NCOAT Inhibitors

In chapter 2, the optimized enzyme substrate YSDSP(*O*-GlcNAc-Ser)TST **2** was introduced. Here its sulfur substituted analogue YSDSP(*S*-GlcNAc-Ser)TST (Figure 4.1) was chosen as a candidate inhibitor of NCOAT, because thioglycosides are known to be glycosidase resistant [24]. If this *S*-GlcNAcylated peptide substrate shows a potent inhibitory effect on NCOAT, then it can be further modified by photoaffinity reactive groups such as benzophenone and azide or by radioactive labeling such as in tritiated analogues. These compounds then become potential photoaffinity labeled or radioactive probes, which can be used for studying the role of *O*-GlcNAcylation in biological processes and the structure of NCOAT. Before solid phase synthesis of this *S*-GlcNAcylated peptide, its required building block *N*-Fmoc-S-GlcNAc-OAc_3-Serine (Figure 4.2) needs to be prepared. Thus, in this chapter, a synthetic route was introduced for preparing this building block.

4.3 The Synthesis of *N*-Fmoc-*S*-GlcNAc-OAc_3-Serine

First, the synthetic strategy for preparing *N*-Fmoc-*O*-GlcNAc-OAc_3-Serine **3** (Scheme 2.2) [25] was modified to synthesize the building block *N*-Fmoc-*S*-GlcNAc-OAc_3-Serine. However, the coupling reaction of glycosyl chloride and *N*-Fmoc protected cysteine in the presence of mercury bromide as mediating reagent and 1, 2-dichloroethane as solvent gave an unexpected product which was shown to be the dimeric *N*-Fmoc protected cysteine containing a disulfide linkage. This was probably due to the oxidation of the mercapto groups to the disulfide linkage, instead of generating the desired building block *N*-Fmoc-*S*-GlcNAc-OAc_3-Serine (Scheme 4.1). Then a multi-step synthetic route (Scheme 4.2) was used for preparing this building block. As Scheme 4.2 shows, the iodo alanine derivative **8** [26] was obtained by a series of simple modifications of L-Serine. Compound **7** was prepared by reacting the amino group and hydroxyl group of L-Serine with Fmoc and OTs respectively. An S_N2 reaction between NaI (the nucleophilic reagent) with compound **7** (OTs as a leaving group) was used to synthesize compound **8**. On the other hand, *N*-acetylglucosamine chloride was converted to the known 1-isothiourea derivative **9** [27, 28], which was precipitated during the reaction and readily isolated (about 83% overall yield from GlcNAc) by simple filtration after the reaction mixture was cooled. Finally, coupling of **8** and **9** in DMSO-H_2O at 0 °C gave pure product *S*-GlcNAc-OAc_3-Ser-COOBn **10** [29]. According to the previous literature [29], the building block

N-Fmoc-*S*-GlcNAc-OAc_3-Serine would be obtained after hydrogenation of **10** with H_2 and Pd/C in methanol.

After completion of the synthesis of **10**, a recent paper [30] was published that implied that NCOAT could remove *S*-GlcNAc from *S*-GlcNAc glycopeptides or *S*-GlcNAc proteins (Figure 4.3), which means that YSDSP(*S*-GlcNAc-Ser)TST may not be a good candidate as an inhibitor of NCOAT. Also, a site-directed mutagenesis study confirms that Asp174 and Asp 175 are the key catalytic residues of NCOAT protein [31]. The above developments made OGT a priority for my research goal. However, this synthetic method provides a facile route to *N*-Fmoc-*S*-GlcNAc-OAc_3-Serine, which still can serve as a building block for preparing *S*-GlcNAcylated peptides in general.

4.4 Conclusions

N-Fmoc-*S*-GlcNAc-OAc_3-Serine as the building block for solid phase synthesis of *S*-GlcNAcylated peptides was prepared by a facile synthetic method. Although recent research developments with NCOAT suggest that *S*-GlcNAcylated peptides might be not good inhibitor candidates for NCOAT, they can still be used in other research fields.

4.5 Experimental

4.5.1 General Procedures

Proton NMR and ^{13}C NMR spectra were collected on Brucker AV 200 and AV 500 spectrometers and were referenced to internal tetramethylsilane ($SiMe_4$).

Chemical shifts for ^{1}H NMR spectra are reported as δ in units of parts per million (ppm) downfield from $SiMe_4$ (δ 0.0) and relative to the signal of chloroform-*d* (δ 7.2600) and DMSO-d_6 (δ 2.5000). Multiplicities are given as: s (singlet); brs (broad singlet); d (doublet); t (triplet); q (quartet); dd (doublet of doublets); dt (doublet of triplets); dtq (doublet of triplets of quartets); or m (multiplet). The number of protons (n) for a given resonance is indicated by nH. Coupling constants are reported as a *J* value in Hz. Carbon nuclear magnetic resonance spectra (^{13}C NMR) are reported as δ in units of parts per million (ppm) downfield from $SiMe_4$ (δ 0.0) and relative to the signal of chloroform-*d* (δ 77) and DMSO-d_6 (δ 39.5).

ESI-MS analyses were carried out on a Micromass Quattro Ultima (LC-ESI/APCI Triple Quadrupole Mass Spectrometer) and HRMS determinations were performed on a Micromass Global Ultima (MALDI/CapLC-ESI Quadrupole Time of Flight Mass Spectrometer).

All chemicals were purchased from Sigma-Aldrich. Dichloromethane was heated at reflux over CaH_2 for 4 h and distilled; pyridine and triethylamine were heated at reflux overnight over potassium hydroxide and distilled; acetone was heated at reflux over anhydrous K_2CO_3 for 2 h and distilled. The other reagents and solvents were used without further purification.

Thin layer chromatography (TLC) was performed on pre-coated Merck 60 F_{254} plates and the spots were examined with UV light and sulfuric acid-ethanol

spray. All experiments were carried out in dried glassware. Column chromatography was carried out by using silica gel (100-200 mesh).

4.5.2 *N*-Fmoc-L-Serine Benzyl Ester (6)

A solution of L-serine benzyl ester hydrochloride (928 mg, 4 mmol), Et_3N (0.84 mL, 6 mmol), and 9-fluorenylmethyl *N*-succinimidyl carbonate (1348 mg, 4 mmol) in CH_2Cl_2 (20 mL) was stirred at room temperature for 2 h. The mixture was washed with 0.1 N HCl (20 mL) and water (2 × 20 mL). The organic phase was dried over anhydrous Na_2SO_4 overnight, and then the solution was evaporated in vacuo to give a syrup. CH_2Cl_2: hexane (1: 1, 7 mL) was used to recrystallize the residue to give the product as a white solid (1.70 g, 4 mmol). The yield is 100%. $\mathbf{R_f}$ = 0.35 (CH_2Cl_2: MeOH = 20: 1); **^{1}H NMR δ_H (DMSO-d^6, 200 MHz)**: 7.89-7.13 (14H, m, Fmoc-Ph, Bn-Ph, NH-Ser), 5.13 (2H, s, OCH_2Bn), 5.01 (1H, t, *J* 5.8 Hz, OH), 4.38-4.16 (4H, m, CH-Fmoc, CH_2-Fmoc, CH-Ser), 3.70 (2H, t, *J* 5.3 Hz, CH_2-Ser); **^{13}C NMR δ_c (DMSO-d_6, 200 MHz):** 170.73, 156.14, 143.77, 140.74, 135.98, 128.39, 127.96, 127.62, 127.10, 125.26, 120.14, 72.66, 65.88, 61.24, 56.86, 46.60; **ESI-MS *m/z*:** 418.3 $[M+H]^+$, 435.4 $[M+NH_4]^+$, 440.3 $[M+Na]^+$, 456.3 $[M+K]^+$; **HRMS *m/z* $[M+H]^+$:** obsd 418.1671, calcd 418.1654 for $C_{25}H_{24}NO_5$.

4.5.3 *N*-Fmoc-*O*-Tosyl-L-Serine Benzyl Ester (7)

N-Fmoc-L-serine benzyl ester **6** (1.5 g, 3.60 mmol) was dissolved in pyridine (15 mL) and cooled to 0 °C. Then tosyl chloride (1.37 g, 7.19 mmol) was

added to the solution. The reaction mixture was stirred overnight, then the solution was poured into ice-water (40 mL) and the aqueous mixture was extracted with ethyl acetate (60 mL). The organic layer was washed with 1 M citric acid (3 × 20 mL), saturated $NaHCO_3$ (2 × 20 mL), NaCl solution (20 mL), 1 M HCl (2 × 20 mL), saturated $NaHCO_3$ (2 × 20 mL) and NaCl solution (3 × 20 mL). The organic extracts were dried over anhydrous Na_2SO_4, filtered and concentrated under reduced pressure. The crude product was recrystallized from ethanol-petroleum ether (22 mL) to give the product (1.89 g, 3.31 mmol) as a white powder. The yield is 92%. $\mathbf{R_f}$ = 0.4 (Hexane: EtOAc = 7: 3); **^{1}H NMR δ_H ($CDCl_3$, 200 MHz)**: 7.80-7.22 (17H, m, Fmoc-Ph, Bn-Ph, Tosyl-Ph), 5.72 (1H, d, *J* 7.8 Hz NH-Ser), 5.17 (2H, s, OCH_2Bn), 4.63 (1H, d, *J* 7.8 Hz, CH-Ser), 4.50-4.14 (5H, m, CH-Fmoc, CH_2-Fmoc, CH_2-Ser), 2.34 (3H, s, CH_3-Tosyl); **^{13}C NMR δ_c ($CDCl_3$, 200 MHz):** 168.07, 155.49, 145.17, 143.60, 143.50, 141.18, 134.58, 132.05, 129.88, 128.61, 128.28, 127.92, 127.75, 127.08, 125.09, 119.96, 69.00, 68.01, 67.45, 53.43, 46.86, 21.52; **ESI-MS *m/z*:** 572.3 $[M+H]^+$, 589.3 $[M+NH_4]^+$, 594.3 $[M+Na]^+$, 610.3 $[M+K]^+$; **HRMS *m/z* $[M+H]^+$:** obsd 572.1738, calcd 572.1743 for $C_{32}H_{29}NO_7S$.

4.5.4 Iodo-*N*-Fmoc-L-Alanine Benzyl Ester (8)

A solution of *N*-Fmoc-*O*-Tosyl-L-Serine Benzyl Ester 7 (1 g, 1.75 mmol) and NaI (1.5 g, 10 mmol) in acetone (20 mL) was stirred under N_2 for 24 h at room temperature. Then the mixture was evaporated to give a syrup and the syrup

was diluted in EtOAc (40 mL). The solution was washed with $Na_2S_2O_3$ (2 × 20 mL), and then saturated NaCl solution (2 × 20 mL). The organic layer was dried over anhydrous Na_2SO_4 and evaporated to give the crude product. This material was recrystallized from ethanol-petroleum ether (20 mL) to give white crystals (0.9 g, 1.71 mmol) as the product. The yield is 98%. $\mathbf{R}_f$ = 0.6 (Hexane: EtOAc = 7: 3); **^{1}H NMR δ_H ($CDCl_3$, 200 MHz)**: 7.80-7.26 (13H, m, Fmoc-Ph, Bn-Ph), 5.73 (1H, d, *J* 7.2 Hz NH-Ser), 5.24 (2H, s, OCH_2Bn), 4.63 (1H, dt, *J* 7.2 Hz, *J* 3.6 Hz, CH-Ser), 4.42 (2H, d, *J* 3.0 Hz, CH_2-Fmoc), 4.26 (1H, t, *J* 3.0 Hz, CH-Fmoc), 3.62 (2H, d, *J* 3.6Hz, CH_2-Ser); **^{13}C NMR δ_c ($CDCl_3$, 200 MHz):** 168.09, 155.30, 143.54, 141.23, 134.61, 128.63 (4H), 127.72 (3H), 127.06 (3H), 125.08 (2H), 119.98 (3H), 68.13, 67.33, 54.01, 46.99, 7.39; **ESI-MS *m/z*:** 528.2 $[M+H]^+$, 545.2 $[M+NH_4]^+$, 550.2 $[M+Na]^+$, 566.2 $[M+K]^+$; **HRMS *m/z* $[M+H]^+$:** obsd 528.0671, calcd 528.0672 for $C_{25}H_{23}NO_4I$.

4.5.5 2-(2-acetamido-3, 4, 6-tri-O-acetyl-2-deoxy-β-D-glucopyranosyl)-2-thiopseudourea hydrochloride (9)

2-acetamido-2-deoxy-3,4,6-tri-*O*-acetyl-α-D-glucopyranosyl chloride (2 g, 5.52 mmol) and thiourea (2 g, 26.31 mmol) in acetone (35 mL) were heated at reflux for 30 min. The solution, clear initially, crystallized during the heating period. The solution was cooled to 0 °C, filtered and the product was washed with acetone to give crude product (1.91 g). This crude product was recrystallized in MeOH: acetone (1: 1, 20 mL) to obtain the product (1.85 g, 4.57 mmol) as white

needles. The yield is 83%. **^{1}H NMR δ_H (DMSO-d_6, 200 MHz)**: 9.40 (3H, s, NH, NH_2), 8.48 (1H, d, *J* 9.4 Hz NHAc), 5.75 (1H, d, *J* 10.4 Hz, H-1), 5.09 (1H, dd, *J* 9.2 Hz, *J* 9.8 Hz, H-3), 4.88 (1H, dd, *J* 9.2 Hz, *J* 9.8 Hz, H-4), 4.27-3.90 (4H, m, H-2, H-5, H-6_a, H-6_b), 1.97 (3H, s, OAc), 1.94 (3H, s, OAc), 1.89 (3H, s, OAc), 1.71 (3H, s, OAc); **^{13}C NMR δ_c (DMSO-d_6, 200 MHz):** 170.07, 169.89, 169.61, 169.31, 167.37, 80.62, 74.69, 72.82, 67.87, 61.48, 51.25, 22.57, 20.62, 20.45, 20.43; **ESI-MS *m/z*:** ESI-MS 406.3 $[M+H]^+$; **HRMS *m/z* $[M+H]^+$:** obsd 406.1271, calcd 406.1284 for $C_{15}H_{24}N_3O_8S$.

4.5.6 *S*-GlcNAc-OAc$_3$-Ser-COOBn (10)

To a cooled (about 0 °C) and stirred solution of potassium carbonate (187.4 mg, 1.36 mmol) and sodium metabisulphite (208 mg, 1.09 mmol) in water (3 mL) was added compound **9** (599 mg, 0.51 mmol) in DMSO (3 mL) and compound **8** (629.6 mg, 1.10 mmol) in DMSO (5 mL). The reaction mixture was stirred for 1.5 h at 0 °C, and then the mixture was poured into ice-water (20 mL). The products were extracted with $CHCl_3$ (20 mL) and the organic layer was washed with 0.5 M $KHSO_4$ (10 mL), water (10 mL), and saturated NaCl solution (10 mL). The organic layer was dried over anhydrous Na_2SO_4 and concentrated in vacuo. Then the crude product was purified using a silica gel column (first toluene: EtOAc = 5: 1, Then 1: 1) to give the product (540 mg, 0.71 mmol) as a light yellow powder. The yield is 65%. $\mathbf{R}_f$ = 0.5 (CH_2Cl_2: MeOH = 20: 1); **^{1}H NMR δ_H ($CDCl_3$, 500 MHz)**: 7.96-7.30 (13H, m, aromatic), 5.81 (1H, d, *J* 8.7 Hz, NH Ser),

5.35 (1H, d, *J* 8.4 Hz, NHAc), 5.10 (1H, dd, *J* 9.8 Hz, *J* 9.6 Hz, H-3), 4.68 (1H, dd, *J* 9.8 Hz, *J* 9.6 Hz, H-4), 4.57 (1H, d, *J* 10.5 Hz, H-1), 4.55-4.38 (5H, m, SCH_2Ph, CH_2-Fmoc, CH-Ser), 4.28-4.19 (3H, m, H-6, CH-Fmoc, CH_{2a}-Ser), 4.08 (1H, dd, H-6′), 3.85 (1H, dd, *J* 10.8 Hz, *J* 2.7 Hz, CH_{2b}-Ser), 3.7 (1H, t, *J* 10.0 Hz, H-2), 3.64-3.59 (1H, m, H-5), 2.04-1.75 (12H, 4s, $COCH_3$); **^{13}C NMR δ_c ($CDCl_3$, 200 MHz):** 170.94, 170.62, 170.33, 169.19, 169.10, 143.72, 141.20, 134.97, 128.61, 128.29, 127.70, 127.09, 124.97, 124.93, 119.96, 83.63, 75.99, 73.58, 68.13, 67.52, 67.10, 62.06, 53.45, 52.86, 47.01, 32.03, 23.10, 20.55(4C); **ESI-MS *m/z*:** 763.3 $[M+H]^+$, 780.3 $[M+NH_4]^+$, 785.3 $[M+Na]^+$, 801.2 $[M+K]^+$; **HRMS *m/z* $[M+H]^+$:** obsd 763.2505, calcd 763.2537 for $C_{39}H_{43}N_2O_{12}S$.

References

1 L. K. Krepprel, M. A. Blomberg and G. W. Hart, *J. Biol. Chem.* **272** (1998) 9308-9315

2 L. Wells, Y. Gao, J. A. Mahoney, K. Vosseller, C. Chen, A. Rosen and G. W. Hart, *J. Biol. Chem.* **277** (2002) 1755-1761

3 Y. Gao, L. Wells, F. I. Comer, G. J. Parker and G. W. Hart, *J. Biol. Chem.* **276** (2001) 9838-9845

4 D. L. Dong and G. W. Hart, *J. Biol. Chem.* **269** (1994) 19321-19330

5 R. Gupta, H. Birch, K. Rapachi, S. Brunak and J. E. Hansen, *Nucleic Acids Res.* **27** (1999) 370-372

6 R. A. Haltiwanger, W. G. Kelly, E. P. Roquemore, M. A. Blomberg, L. Y. Dong, L. Kreppel, T. Y. Chou and G. W. Hart, *Biochem. Soc. Trans.* **20** (1992) 264-269

7 G. W. Hart, R. S. Haltiwanger, G. D. Holt and W. G. Kelly, *Ciba Found Symp.* **145** (1989) 102-112

8 K. P. Kearse and G. W. Hart, *Arch. Biochem. Biophys.* **290** (1991) 543-548

9 C. F. Chou and M. B. Omary, *J. Biol. Chem.* **268** (1993) 4465-4472

10 R. S. Haltiwanger, M. A. Blomberg and G. W. Hart, *J. Biol. Chem.* **267** (1992) 9005-9013

11 F. I. Comer and G. W. Hart, *Biochim. Biophys. Acta* **1473** (1999) 161-171

12 X. Yang, F. Zhang and J. E. Kudlow, *Cell* **110** (2002) 69-80

13 X. Cheng and G. W. Hart, *J. Biol. Chem.* **276** (2001) 10570-10575

14 X. Yang, K. Su, M. D. Roos, Q. Chang, A. J. Paterson and J. E. Kudlow, *Proc. Natl. Acad. Sci. U. S. A.* **98** (2001) 6611-6616

15 S. P. Jackson and R. Tjian, *Cell* **55** (1988) 125-133

16 C. Toleman, A. J. Paterson, T. R. Whisenhunt and J. E. Kudlow, *J. Biol. Chem.* **279** (2004) 53665-53673

17 G. Legler, E. Lullau, E. Kappes and F. Kastenholz, *Biochim. Biophys. Acta* **1080** (1991) 89-95

18 M. Horsch, L. Hoesch, A. Vasella and D. M. Rast, *Eur. J. Biochem.* **197** (1991) 815-818

19 J. Liu, A. R. Shikhman, M. K. Lotz and C. H. Wong, *Chem. Biol.* **8** (2001) 701-711

20 S. Knapp, D. Vocadlo, Z. N. Gao, B. Kirk, J. P. Lou and S. G. Withers, *J. Am. Chem. Soc.* **118** (1996) 6804-6805

21 M. S. Macauley, G. E. Whitworth, A. W. Debowski, D. Chin and D. J. Vocadlo, *J. Biol. Chem.* **280** (2005) 25313-25322

22 P. Lalegerie, G. Legler and J. M. Yon, *Biochimie* **64** (1982) 977-1000

23 M. B. Tropak, S. P. Reid, M. Guiral, S. G. Withers and D. Mahuran, *J. Biol. Chem.* **179** (2004) 13478-13487

24 Z. J. Witczak and D. Boryczewski, *Biorg. Med. Chem. Lett.* **8** (1998) 3265-3268

25 I. Carvalho, S. L. Scheuerl, K. P. R. Kartha and R. A. Field, *Carbohydr. Res.* **338** (2003) 1039-1043

26 R. F. W. Jackson, N. Wishart, A. Wood, K. James and M. J. Wythes, *J. Org. Chem.* **57** (1992) 3397-3404

27 D. Horton and M. L. Wolfrom, *J. Org. Chem.* **27** (1962) 1794-1800

28 S. Chipowsky and Y. C. Lee, *Carbohydr. Res.* **31** (1973) 339-346

29 Y. Ohnishi, M. Ichikawa and Y. Ichikawa, *Biorg. Med. Chem. Lett.* **10** (2000) 1289-1291

30 M. S. Macauley, K. A. Stubbs and D. J. Vocadlo, *J. Am. Chem. Soc.* **127** (2005) 17202-17203

31 N. Cetinbas, M. S. Stubbs, K. A. Drapala and D. J. Vocadlo, *Biochemistry* **45** (2006) 3835-3844

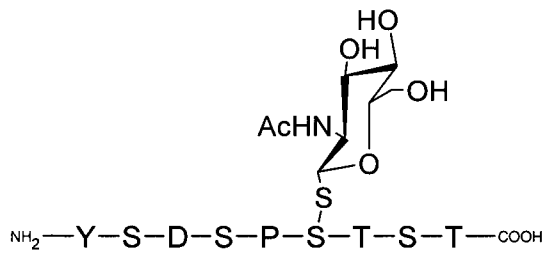


Figure 4.1 The structure of designed inhibitor for NCOAT

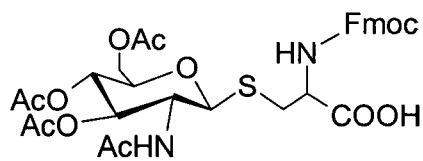


Figure 4.2 The structure of building block *N*-Fmoc-S-GlcNAc-OAc_3-Serine

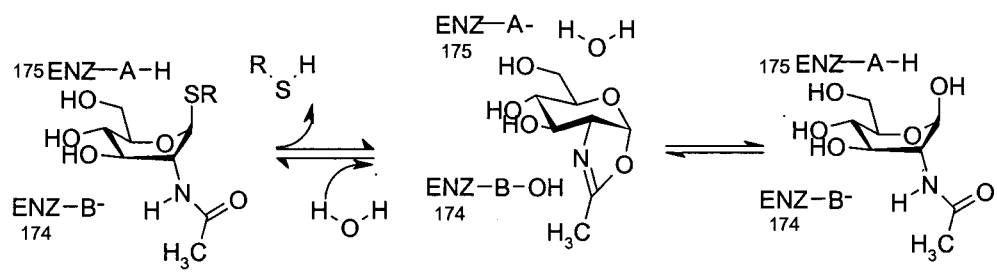
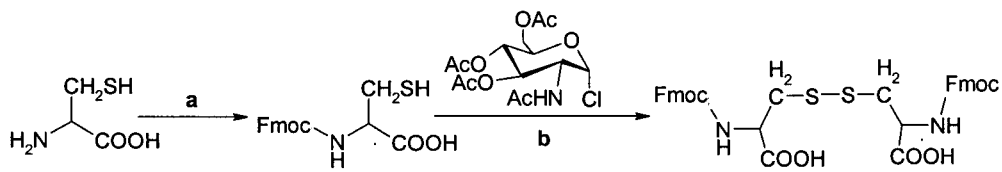
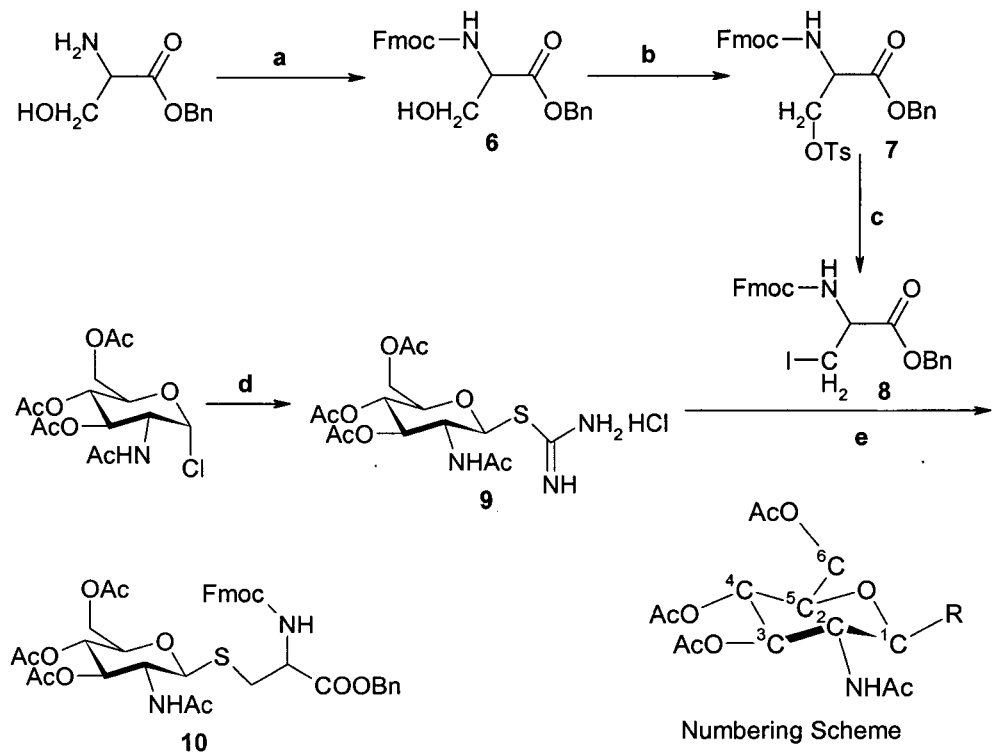


Figure 4.3 *O*-GlcNAcase catalyzes cleavage of thioglycosides without general acid catalysis. Asp 174 and Asp 175 are the catalytic sites of *O*-GlcNAcase. This picture was taken from reference 30



Scheme 4.1 The preliminary synthetic route for the building block shown in Figure 4.2. **(a)** Fmoc-OSu, CH_2Cl_2, Et_3N, rt, 83%; **(b)** $HgBr_2$, 1, 2-dichloroethane, reflux, 50%



Scheme 4.2 Synthesis of building block **10**. **(a)** Fmoc-OSu, CH_2Cl_2, Et_3N, rt, 100%; **(b)** pyridine, TsCl, 0 °C, 92% ; **(c)** NaI, acetone, rt, 99%; **(d)** thiourea, acetone, 57 °C, 83%; **(e)** K_2CO_3, $Na_2S_2O_5$, DMSO, H_2O, 0 °C, 65%. These sugar numbers were defined for ^1H NMR data

Chapter 5

Design and Syntheses of Two Photoaffinity Labelled Probes for Mapping the Active Sites of sOGT

5.1 Introduction

O-GlcNAcylation, a kind of glycosylation, is catalyzed by a speicfic enzyme known as *O*-GlcNAc transferase (OGT), which transfers N-acetylglucosamine (GlcNAc) from UDP-GlcNAc to selected serine or threonine residues of many nuclear and cytoplasmic proteins [1-4]. This post-translational modification affects many cellular processes, such as transcription, translation and protein-protein interactions [5] and OGT is also believed to play important roles in these processes [6-12]. Studying the active site of OGT can undoubtedly provide a deeper understanding of the biological functions of *O*-GlcNAcylation. However, to date, the crystal structure of the OGT catalytic domain has still not been determined. Although the crystal structure of MurG, which has a similar catalytic domain to OGT, has been reported (Figure 5.1) [13, 14], the known inhibitors of MurG do not inhibit OGT. This means that the active sites of these two proteins should be quite distinct. Thus, mapping the active site of OGT is still a challenging task. Recently, our group started a project to prepare compounds and develop methods to map the active site of OGT. sOGT is an isoform of OGT, containing the catalytic domain

and a shorter tetratricopeptide repeat (TPR) domain than other isoforms of OGT (Figure 5.1) [15]. Because it is much easier to express significant quantities of sOGT, here we used sOGT as the research goal instead of OGT. In this chapter, two potent photoaffinity labelled probes for mapping the active sites of sOGT were designed and prepared.

5.2 Design of Two Photoaffinity Labelled Probes

Phenyl 5-chloro-2-oxo-3-hydrobenzoxazole-3-carboxylate (Figure 5.2), **11**, was reported as a potent inhibitor of sOGT by Walker and co-workers [16]. Based on this core structure, we used photoaffinity labeled groups to modify this inhibitor in the side chain to prepare the potential photoaffinity labeled probes, **12** and **13**, which will be used for mapping the active sites of sOGT. Here, two photoreactive groups, benzophenone **14** and azide **15**, were selected (Figure 5.3), as they are considered to be the most reliable and have demonstrated utility in several enzyme active site studies [17-22]. The selected photoaffinity groups react with the target proteins (here, sOGT) through different reactive intermediates: a radical (benzophenone) and a nitrene (aromatic azide) [18, 19].

The use of different photoreactive groups in different side chains increases the chance of successful cross-linking to the target protein, since it cannot be predicted in advance which group will be most readily attached to sOGT. Upon protein digestion, the fragments are to be analyzed by CapLC and MALDI MS to detemine the peptide sequences to which the photoaffinity labeled probes are bound and the photoreactive groups are attached. Correlating this information to the

structural data available, this could subsequently reveal where the active site regions of sOGT are.

5.3 Syntheses of Two Photoaffinity Labelled Probes

The reported inhibitor of sOGT, **11**, was synthesized as a standard compound to evaluate the new assay for measuring the IC_{50} values of the sOGT inhibitors which will be introduced in sections below. Compound **11** was prepared by a coupling reaction of chlorzoxazone at *N*-3 with the acylating agent phenyl chloroformate (Scheme 5.1) [23]. The same approach was also applied for the synthesis of benzophenonyl 5-chloro-2-oxo-3-hydrobenzoxazole-3-carboxylate **12** and 5-azido-2-oxo-3-hydrobenzoxazole-3-carboxylate **13**. The acylating agent benzophenonyl chloroformate **16** was synthesized by treating 4-hydroxybenzophenone with bis(trichloromethyl) carbonate under very strongly basic conditions [24], since it was not commercially available. It is important to control the exact ratio of 4-hydroxybenzophenone to bis(trichloromethyl) carbonate at 3: 1, otherwise some by-products such as dibenzophenonyl carbonate were obtained and they would cause problems during separation. Without any further purification of **16**, coupling between the acylating agent **16** and chlorzoxazone gave **12**, the photoaffinity labeled analogue of **11** containing a benzophenone group (Scheme 5.1) [23].

Compound **13** was synthesized by a multistep synthetic route (Scheme 5.2). As an important intermediate, azidozoxazone **19** was prepared through a series of simple modifications of commericailly available 2-amino-4-nitrophenol.

A cyclic carbamate **17** was first formed between both the amino and phenol groups of 2-amino-4-nitrophenol with 1,1′-carbonyl diimidazole in THF [25]. Hydrogenation of the aromatic nitro group of **17** gave **18**. After the aromatic amino group of **18** was converted into an azide by treatment with triflyl azide, triethyl amine, and $CuSO_4$ as catalyst [26], the intermediate **19** was obtained smoothly. Finally, the coupling reaction between **19** and the acylating agent phenyl chloroformate was carried out at room temperature to give the photoaffinity labeled analogue of **11** containing an aromatic azide group, **13**, successfully [23]. It is also worth mentioning that **11**, **12**, and **13** could not be purified by silica gel column chromatography in our hands because silica gel resulted in loss of the acyl group.

5.4 IC_{50} Measurements of Two Photoaffinity Labelled Probes

To evaluate the inhibitor performance of the two photoaffinity labelled probes for mapping the active site of sOGT, a new assay was developed (Scheme 5.3). The substrate peptide YSTSPSTST **1** could be transiently *O*-GlcNAcylated by sOGT, which transferred GlcNAc from UDP-GlcNAc to a specific serine residue [27]. The resulting mixture was analyzed by CapLC (Figure 5.4A) and MALDI MS (Figure 5.4B). Figure 5.4 showed that the *O*-GlcNAcylated product YSTSP(*O*-GlcNAc-Ser)TST **2** could be obtained (Figure 5.4 B) through this enzyme catalytic reaction, however, with the CapLC analytical method of 93% A/7% B, the peptide and glycopeptide could not be separated (both are present in the 9.70 min peak in Figure 5.4A). To optimize the CapLC analytical conditions

in order to separate peptide **1** and glycopeptide **2**, a series CapLC analytical methods (Table 5.1, Table 5.2 and Table 5.3) were evaluated to separate a mixture of **1** and **2** (50 pmol: 50 pmol). The CapLC results (Figure 5.5A, Figure 5.5B and Figure 5.5C) showed that **2** could not be separated from **1** completely, no matter which CapLC analytical method was employed. Since glycopeptide **2** could not be directly detected in the presence of a large amount of **1** by CapLC, the *O*-GlcNAcylated product YSTSP(*O*-GlcNAc-Ser)TST **2** was further treated with very strong base (NaOMe) to promote the beta-elimination reaction, leading to the formation of **5** (Figure 5.6D and Scheme 5.3) [28]. The inhibitors of sOGT could inhibit the *O*-GlcNAcylation of **1** to **2**. Treatment of the mixture (**1**, sOGT, UDP-GlcNAc) with more concentrated inhibitor initially led to the formation of less **2**. As a result, less **5** was obtained (Figure 5.6A, Figure 5.6B and Figure 5.6C). The quantity of beta-elimination product **5** could be measured by CapLC. According to the relationship between the inhibitor concentration and the quantity of the final product **5**, the IC_{50} curves could be drawn for each inhibitor (Figure 5.7, Figure 5.8 and Figure 5.9) and the IC_{50} values could be calculated by the software Grafit 5. First, the IC_{50} value of the reported sOGT inhibitor **11** was measured by this new assay and found to be 29 ± 2 μM, which agreed well with the previously reported value of compound **11** [16]. Since this new assay was reliable for measuring the IC_{50} values of sOGT inhibitors, we applied this assay to the IC_{50} value measurement of the two photoaffinity labelled compounds, **12** and **13**. The IC_{50} values of **12** and **13** for sOGT were found to be 71 ± 1 μM and 36 ± 3 μM,

which means that both **12** and **13** are effective inhibitors of sOGT and can be useful tools to probe the active site of sOGT.

5.5 Conclusions

Two photoaffinity labelled analogues of sOGT inhibitor **12** and **13** have been designed and prepared, and their IC_{50} values for sOGT have been measured by a new assay which was developed for measuring the IC_{50} values of sOGT inhibitors. The inhibitor performance of the two analogues showed that they both were potent inhibitors of sOGT and could be powerful photoaffinity labelled probes to map the active sites of sOGT in the near future.

5.6 Experimental

5.6.1 General Procedures

^1H NMR and ^{13}C NMR spectra were collected on Brucker AV 200 and AV 500 spectrometers and were referenced to internal tetramethylsilane ($SiMe_4$). Chemical shifts for ^1H NMR spectra are reported as δ in units of parts per million (ppm) downfield from $SiMe_4$ (δ 0.0) and relative to the signal of chloroform-*d* (δ 7.2600) and DMSO-d_6 (δ 2.5000). Multiplicities were given as: s (singlet); brs (broad singlet); d (doublet); t (triplet); q (quartet); dd (doublet of doublets); dt (doublet of triplets); dtq (doublet of triplets of quartets); or m (multiplet). The number of protons (n) for a given resonance is indicated by nH. Coupling constants are reported as a *J* value in Hz. Carbon nuclear magnetic resonance spectra (^{13}C NMR) are reported as δ in units of parts per million (ppm) downfield

from $SiMe_4$ (δ 0.0) and relative to the signal of chloroform-*d* (δ 77) and DMSO-d_6 (δ 39.5).

ESI-MS analyses were carried out on a Micromass Quattro Ultima (LC-ESI/APCI Triple Quadrupole Mass Spectrometer) and HRMS determinations were performed on a Micromass Global Ultima (MALDI/CapLC-ESI Quadrupole Time of Flight Mass Spectrometer).

All CapLC analyses were carried out on a Waters Delta CapLC.The solvent for CapLC was A 0.1% HCOOH/1% CH_3CN/H_2O, B 0.1% HCOOH/ CH_3CN, and the flow rate was 5 μL/min.

All chemicals were purchased from Sigma-Aldrich. Dichloromethane was heated at reflux over CaH_2 for 4 h and distilled; triethylamine was heated at reflux overnight over potassium hydroxide and distilled; THF was freshly distilled under nitrogen after sodium treatment. The other reagents and solvents were used without further purification.

Thin layer chromatography (TLC) was performed on pre-coated Merck 60 F_{254} plates and the spots were examined with UV light and sulfuric acid-ethanol spray. All experiments were carried out in dried glassware. Column chromatography was carried out by using silica gel (100-200 mesh).

5.6.2 Phenyl 5-chloro-2-oxo-3-hydrobenzoxazole-3-carboxylate (11)

To a solution of chlorzoxazone (1.6957 g, 10 mmol) in THF (10 mL) with triethylamine (1.443 mL, 10 mmol) was added a solution of phenyl chloroformate (1.26 mL, 10 mmol) in THF (10 mL). After stirring overnight at room temperature, the mixture was filtered and the solvent was removed in vacuo. The solid was recrystallized from ethanol (30 mL), giving pure product (1.7 g, 5.88 mmol, 59%) as white needle crystals. $\mathbf{R_f}$ = 0.7 (Hexane: EtOAc = 6: 1); **^{1}H NMR δ_H ($CDCl_3$, 500 MHz)**: 7.88 (1H, s), 7.53-7.42 (2H, m), 7.40-7.24 (4H, m), 7.20 (1H, d, *J* 4.8 Hz); **^{13}C NMR δ_c ($CDCl_3$, 500 MHz):** 149.77, 148.26, 147.59, 140.48, 130.39, 129.80, 127.88, 127.05, 125.45, 121.09, 115.44, 111.05; **ESI-MS *m/z*:** 290.0 $[M+H]^+$, 307.1 $[M+NH_4]^+$; **HRMS *m/z* $[M+Na]^+$:** obsd 312.0054, calcd 312.0040 for $C_{14}H_8NO_4ClNa$; **CapLC analysis**: 99% A/1% B to 1% A/99% B in 40 min (linear), retention time is 18.15 min (After subtracting the delay time of 13.40 min) and only one peak was shown on CapLC chromatogram.

5.6.3 Benzophenonyl 5-chloro-2-oxo-3-hydrobenzoxazole-3-carboxylate (12)

To a stirred solution of bis(trichloromethyl) carbonate (0.197 g, 0.67 mmol) dissolved in CH_2Cl_2 (20 mL) and cooled to 0 °C, was added dropwise a solution of 4-hydroxybenzophenone (0.396 g, 2 mmol) in THF/CH_2Cl_2 (1 mL/2 mL) and then a solution of NaOH (0.08 g, 2 mmol) in water (1 mL) was added.

During addition, the temperature of the mixture was kept at 0-5 °C. After the addition was completed, the temperature of mixture was allowed to slowly rise to room temperature and stirring was continued for 2 h. The water layer was removed. The organic phase was washed with cooled water (3 × 5 mL), dried over $CaCl_2$ and filtered. This solution (mainly compound **16**) was kept for use in the next reaction without any further purification. Then to a solution of chlorzoxazone (0.339 g, 2 mmol) in THF (3 mL) with Et_3N (0.29 mL, 2 mmol) was added the above solution. This mixture was stirred overnight at room temperature. Then the mixture was filtered and the solvent was removed in vacuo. The solid was recrystallized from EtOH (10 mL) to give the final product (339 mg, 0.86 mmol, 43%) as a white powder. $\mathbf{R_f}$ = 0.7 (Hexane: EtOAc = 6: 1); **^{1}H NMR δ_H ($CDCl_3$, 500 MHz):** 7.95 (2H, d, *J* 5.6 Hz), 7.90 (1H, s), 7.83 (2H, d, *J* 5.6 Hz), 7.62 (1H, t, *J* 4.8 Hz), 7.51 (2H, d, *J* 5.1 Hz), 7.47 (2H, d, *J* 5.9 Hz), 7.32 (1H, d, *J* 5.0 Hz), 7.21 (1H, d, *J* 5.0 Hz); **^{13}C NMR δ_c ($CDCl_3$, 500 MHz):** 195.13, 152.57, 148.08, 147.12, 140.51, 137.22, 136.32, 132.69, 131.85, 130.51, 129.96, 128.43, 127.70, 125.66, 121.11, 115.42, 111.16; **ESI-MS *m/z*:** 394.0 $[M+H]^+$, 415.9 $[M+Na]^+$; **HRMS *m/z* $[M+H]^+$:** obsd 394.0498, calcd 394.0482 for $C_{21}H_{13}NO_5Cl$; **CapLC analysis**: 99% A/1% B to 1% A/99% B in 40 min (linear), retention time is 20.18 min (After subtracting the delay time of 13.40 min) and only one peak was shown on CapLC chromatogram.

5.6.4 Nitrozoxazone (17)

1,1′-Carbonyl diimidazole (2.89 g, 17.8 mmol) was added to a solution of 2-amino-4-nitrophenol (2.5 g, 16.2 mmol) in dried THF (50 mL) at room temperature, and the mixture was stirred for 2 h at the same temperature. Dilute HCl solution (1 M, 25 mL) was added to the mixture and the aqueous phase was extracted with diethyl ether (60 mL). The organic solution was washed with saturated aqueous NaCl solution, and dried over $MgSO_4$. The solvent was removed in vacuo. The crude product was recrystallized from AcOEt-hexane (40 mL, 1: 1) to give the final product (2.83 g, 15.7 mmol, 97%) as dark yellow needles. $\mathbf{R_f}$ = 0.5 ($CHCl_3$: MeOH: AcOH = 8: 1: 0.1); **^{1}H NMR δ_H (DMSO-d_6, 200 MHz):** 11.91 (1H, brs, N-H), 7.95 (1H, dd, *J* 2.4 Hz, *J* 8.8 Hz, aromatic-H), 7.73 (1H, d, *J* 2.4 Hz, aromatic-H), 7.43 (1H, d, *J* 8.8 Hz, aromatic-H); **^{13}C NMR δ_c (DMSO-d_6, 200 MHz):** 154.09, 147.82, 143.67, 131.20, 118.59, 109.79, 104.95; **ESI-MS *m/z*:** 179.0 $[M-H]^-$; **HRMS *m/z* $[M-H]^-$:** obsd 179.0085, calcd 179.0093 for $C_7H_3N_2O_4$.

5.6.5 Aminozoxazone (18)

To nitrozoxazone **17** (2 g, 11.1 mmol) in MeOH (50 mL) solution, was added Pd/C (10wt%, 1 g). At the same time, hydrogen in a balloon (about 20 cm diameter) was used to pressurize the flask (100 mL) 6-7 times. The reaction mixture was stirred under hydrogen overnight at room temperature. The mixture was filtered and evaporated. The crude product was recrystallized from

MeOH/CH_2Cl_2 (20 mL, 1: 2) to give the product (1.5 g, 10 mmol, 90%) as yellow needles. $\mathbf{R_f}$ = 0.2 ($CHCl_3$: MeOH: AcOH = 8: 1: 0.1); **^{1}H NMR δ_H (DMSO-d_6, 200 MHz):** 11.21 (1H, brs, N-H), 6.88 (1H, d, *J* 8.4 Hz, aromatic-H), 6.34 (1H, d, *J* 2.2 Hz, aromatic-H), 6.24 (1H, dd, *J* 2.2 Hz, *J* 8.4 Hz, aromatic-H), 5.06 (2H, brs, NH_2); **^{13}C NMR δ_c (DMSO-d_6, 200 MHz):** 155.21, 145.34, 135.06, 130.98, 109.73, 107.16, 95.99; **ESI-MS *m/z*:** 149.0 $[M-H]^-$, 150.9 $[M+H]^+$, 172.9 $[M+Na]^+$; **HRMS *m/z* $[M+H]^+$:** obsd 151.0504, calcd 151.0508 for $C_7H_7N_2O_2$.

5.6.6 TfN_3 Synthetic Route

A solution of NaN_3 (3.5 g, 54 mmol) in water (8 mL) and CH_2Cl_2 (3 mL) was cooled to 0 °C in an ice-bath. To this vigorously stirred solution, Tf_2O (2.54 g, 1.51 mL, 9 mmol) was added dropwise through a syringe. After stirring at 0 °C for 2 h, the organic phase was separated and the aqueous phase was extracted with CH_2Cl_2 (3 mL). The combined organics were washed with saturated aqueous $NaHCO_3$ (10 mL) and saved for use in the next step reaction.

5.6.7 Azidozoxazone (19)

Aminozoxazone **18** (450 mg, 3 mmol) was dissolved in MeOH (2 mL) in a 50 mL round-bottomed flask. Et_3N (909 mg, 1.25 mL, 9 mmol) and a solution of $CuSO_4$ (24 mg, 0.15 mmol in 0.5 mL water) were added to the flask consecutively. After the solution was cooled to 0 °C, a freshly prepared CH_2Cl_2 solution of TfN_3 was added. After the addition was complete, the temperature of the mixture was allowed to slowly rise to room temperature and stirring was continued for 1.5 h.

The mixture was poured into saturated aqueous $NaHCO_3$ (30 mL) and extracted with CH_2Cl_2 (3 × 30 mL). The combined organics were washed with NaCl solution (30 mL), dried over Na_2SO_4, filtered and concentrated. The crude product was purified by silica-gel column chromatography ($CHCl_3$: MeOH = 25: 1) to obtain pure product (475.2 mg, 2.7 mmol, 90%) as a white powder. $\mathbf{R_f}$ = 0.6 ($CHCl_3$: MeOH: AcOH = 8: 1: 0.1); **^{1}H NMR δ_H (DMSO-d_6, 200 MHz):** 11.75 (1H, brs, N-H), 7.27 (1H, d, *J* 9.2 Hz, aromatic-H), 6.91 (1H, d, *J* 2.3 Hz, aromatic-H), 6.77 (1H, dd, *J* 2.3 Hz, *J* 9.2 Hz, aromatic-H); **^{13}C NMR δ_c (DMSO-d_6, 200 MHz):** 154.49, 140.65, 135.25, 131.61, 112.36, 110.55, 100.93; **ESI-MS *m/z*:** 175.0 $[M-H]^-$, 146.9 $[M-H-N_2]^-$; **HRMS *m/z* [M-H]$^-$:** obsd 175.0255, calcd 175.0256 for $C_7H_3N_4O_2$.

5.6.8 5-azido-2-oxo-3-hydrobenzoxazole-3-carboxylate (13)

To a solution of azidozoxazone **19** (176 mg, 1 mmol) in THF (5 mL) with Et_3N (144.3 μL, 1 mmol) solution, was added a solution of phenyl chloroformate (126 μL, 1 mmol) in THF (5 mL). After stirring overnight at room temperature, the mixture was filtered and the solvent was removed in vacuo. The crude product was recrystallized from EtOH (2 mL) to give the product (207 mg, 0.7 mmol, 70%) as white needles. $\mathbf{R_f}$ = 0.7 (Hexane: EtOAc = 6: 1); **^{1}H NMR δ_H ($CDCl_3$, 500 MHz)**: 7.60 (1H, s), 7.48 (2H, d, *J* 5.6 Hz), 7.37 (1H, t, *J* 4.8 Hz), 7.33 (2H, d, *J* 5.4 Hz), 7.25 (1H, d, *J* 5.2 Hz), 6.96 (1H, d, *J* 5.3 Hz); **^{13}C NMR δ_c ($CDCl_3$, 200 MHz):** 149.70, 148.51, 147.85, 138.93, 137.30, 129.80, 128.10, 127.05,

121.08, 115.82, 111.12, 106.12; **ESI-MS *m/z*:** 296.9 $[M+H]^+$, 314.0 $[M+NH_4]^+$, 319.0 $[M+Na]^+$; **HRMS *m/z* $[M+H]^+$:** obsd 297.0627, calcd 297.0624 for $C_{14}H_9N_4O_4$; **CapLC analysis**: 99% A/1% B to 1% A/99% B in 40 min (linear), retention time is 18.10 min (After subtracting the delay time of 13.40 min) and only one peak was shown on CapLC chromatogram.

5.6.9 Expression and Purification of sOGT

This work was done by another group member, Mr. Xi Wang. The expression of a soluble truncated variant of OGT was made possible using a pET28 based plasmid obtained as a kind gift from Dr. Walker at Harvard University. BL21 (DE3) cells were transformed and grown to an O.D. between 1.1 and 1.3 at 37 °C. The culture was cooled to 16 °C, and induced with 0.2 mM IPTG for 24 hours; the cells were pelleted then frozen at –80 °C. The pellets were resuspended in 50 mL of lysis buffer (50 mM Tris-HCl pH 7.5, 200 mM NaCl, 1 mM EDTA, 5 mM DTT, 1 mM PMSF) and a French press was used to lyse the cells. The lysate was then clarified by ultra centrifugation at 15,000 rpm and then loaded on a Hi-Trap nickel column (GE/Amersham Bioscience) for affinity purification relying on the encoded His-tag. The column was washed with ten column volumes of 50 mM imidazole, pH = 7.5, 250 mM NaCl, and the protein was eluted with 250 mM imidazole, pH = 7.5, 250 mM NaCl. About 15 mL of the fraction was collected and a Amicon Ultra-15 centrifuge with a 10 kDa cut off was used to concentrate the sOGT protein and exchange the buffer to 25 mM

Hepes, 10 mM $MgCl_2$, 1 mM EDTA, and 500 μM of tris(hydroxypropyl)phosphate. The aliquots were stored at –80 °C for further use.

5.6.10 Assay for Measuring IC_{50} Values of sOGT Inhibitors

This assay was performed in a 1.5 mL microcentrifuge tube by adding UDP-GlcNAc (1 μL, 1μg/μL), substrate peptide YSDSPSTST (20 μg), purified sOGT protein (1 μL, 1μg/μL), and a certain concentration of inhibitor (in 1 μL DMSO). Then TAB buffer (25 mM Hepes, 1 mM EDTA, and 10 mM $MgCl_2$) was added to a final volume of 50 μL. After incubation at 37 °C for 30 minutes, the reaction was stopped by addition of formic acid (450 μL, 50 mM). The reaction mixture was then loaded onto a 0.5 mL Sephadex (SP-25) column equilibrated in 50 mM formic acid. The column was washed with 50 mM formic acid and eluted with 2 mL of 0.5 M NaCl. The collected solution was lyophilized to give a white powder. This resulting solid was dissolved in 50 μL NaOMe in MeOH (1 μg/μL) and the solution was shaken at room temperature for 2 h. The solution was neutralized with dry ice to pH 7 and then lyophilized. The residual solid was dissolved in HPLC grade water (50 μL) and this solution was further analyzed by CapLC.

References

1 L. K. Krepprel, M. A. Blomberg and G. W. Hart, *J. Biol. Chem.* **272** (1997) 9308-9315

2 W. A. Lubas, D. W. Frank, M. Krause and J. A. Hanover, *J. Biol. Chem.* **272** (1997) 9316-9324

3 W. A. Lubas and J. A. Hanover, *J. Biol. Chem.* **275** (2000) 10983-10988

4 L. K. Krepprel and G. W. Hart, *J. Biol. Chem.* **274** (1999) 32015-32022

5 N. E. Zachara and G. W. Hart, *Chem. Rev.* **102** (2002) 431-438

6 R. Shafi, S. P. N. Lyer, L. G. Ellies N. O'Donnell, K. W. Marek, D. Chui and G. W. Hart, Proc. *Natl. Acad. Sci. USA* **97** (2000) 5735-5739

7 S. P. Jackson and R. Tjian, *Cell* **55** (1988) 125-133

8 M. D. Roos, K. H. Su, J. R. Baker and J. E. Kudlow, *Mol. Cell. Biol.* **17** (1997) 6472-6480

9 X. Y. Yang, K. H. Su, M. D. Roos, Q. Chang, A. J. Paterson and J. E. Kudlow, *Proc. Natl. Acad. Sci. USA* **98** (2001) 6611-6616

10 I. Han and J. E. Kudlow, *Mol. Cell. Biol.* **17** (1997) 2550-2558

11 K. H. Su, M. D. Roos, X. Y. Yang, I. Han, A. J. Paterson and J. E. Kudlow, *J. Biol. Chem.* **274** (1999) 15194-15202

12 K. H. Su, X. Y. Yang, M. D. Roos, A. J. Paterson and J. E. Kudlow, *Biochem. J.* **348** (2000) 281-289

13 S. Ha, D. Walker, Y. G. Shi and S. Walker, *Protein Sci.* **9** (2000) 1045-1052

14 Y. N. Hu, L. Chen, S. Ha, B. Gross, B. Falcone, D. Walker, M. Mokhtarzadeh and S. Walker, *Proc. Natl. Acad. Sci. USA* **100** (2003) 845-849

15 J. A. Hanover, S. Yu, W. B. Lubas, S. H. Shin, M. Ragano-Caracciola, J. Kochran and D. C. Love, *Arch. Biochem. Biophys.* **409** (2003) 287-297

16 B. J. Gross, B. C. Kraybill and S. Walker, *J. Am. Chem. Soc.* **127** (2005) 14588-14589

17 F. Kotzybahibert, I. Kapfer and M. Goeldner, *Angew. Chem., Int. Ed. Engl.* **34** (1995) 1296-1312

18 G. Dorman and G. D. Prestwich, *Trends Biotechnol.* **18** (2000) 64-77

19 Y. Hatanaka and Y. Sadakane, *Curr. Topics Med. Chem.* **2** (2002) 271-288

20 S. A. Fleming, *Tetrahedron* **51** (1995) 12479-12520

21 W. Davidson, G. A. McGibbon, P. W. White, C. Yoakim, J. L. Hopkins, I. Guse, D. M. Hambly, L. Frego, W. W. Ogilvie, P. Lavallee and J. Archambault, *Anal. Chem.* **76** (2004) 2095-2102

22 G. Kukolj, G. A. McGibbon, G. McKercher, M. Marquis, S. Lefebvre, L. Thauvette, J. Gauthier, S. Goulet, M. A. Poupart and P. L. Beaulieu, *J. Biol. Chem.* **280** (2005) 39260-39267

23 J. H. Musser, H. Jones, S. Sciortino, K. Bailey, S. M. Coutts, A. Khandwala, P. Sonninogoldman, M. Leibowitz, P. Wolf and E. S. Neiss, *J. Med. Chem.* **28** (1985) 1255-1259

24 H. Shi, W. Hu and Y. Sun, *J. Chem. Res.* **10** (2004) 708-709

25 K. Hiroya, S. Matsumoto, M. Ashikawa, H. Kida and T. Sakamoto, *Tetrahedron* **61** (2005) 12330-12338

26 Q. Liu and Y. Tor, *Org. Lett.* **5** (2003) 2571-2572

27 R. S. Haltiwanger, G. D. Holt and G. W. Hart, *J. Biol. Chem.* **265** (1990) 2563-2568

28 K. Vosseller, K. C. Hansen, R. J. Chalkley, J. C. Trinidad, L. Wells, G. W. Hart and A. L. Burlingame, *Preteomics* **5** (2005) 388-398

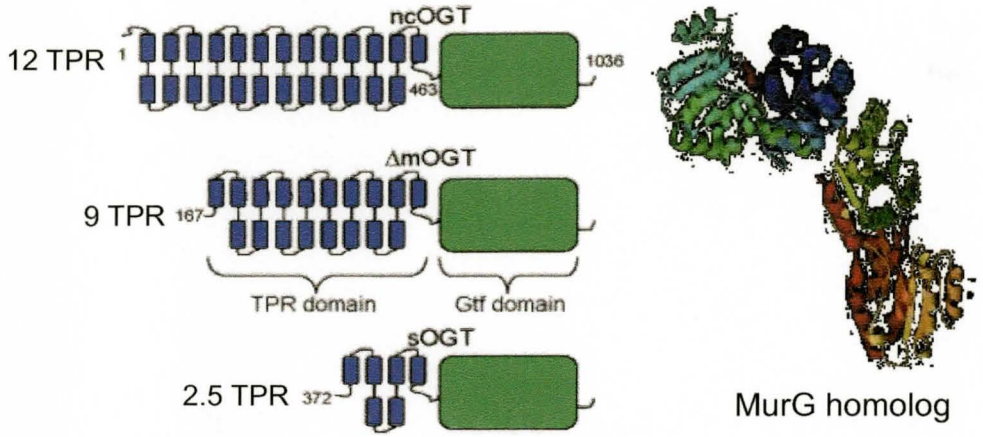


Figure 5.1 The structures of ncOGT, ΔmOGT, sOGT and MurG. ncOGT, ΔmOGT and sOGT are three different isoforms of OGT with different lengths of TPR domains. MurG belongs to the same family as OGT

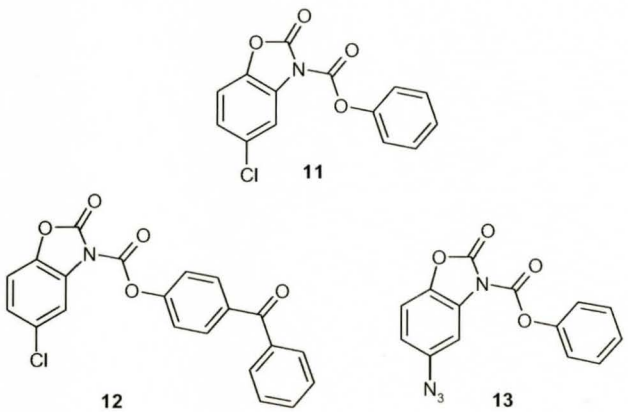


Figure 5.2 The structures of **11**, **12** and **13**. Compound **11** is a known inhbitor of sOGT; based on the core structure of **11**, two potential photoaffinity labelled probes for mapping the active sites of sOGT, **12** and **13**, were designed

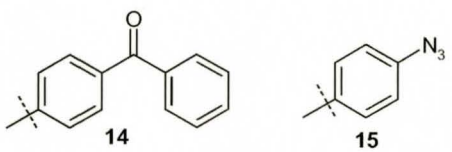
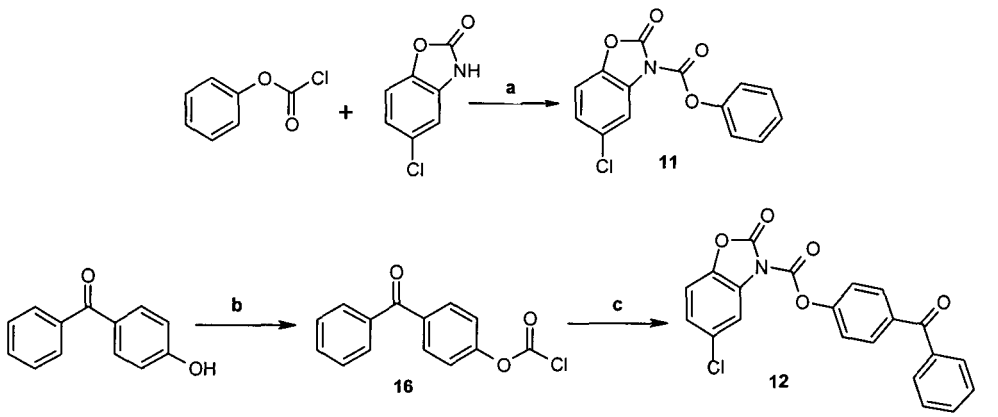
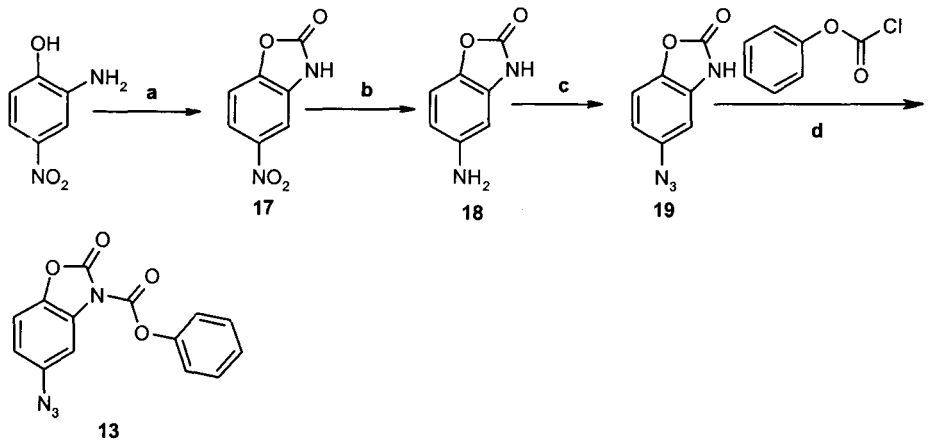


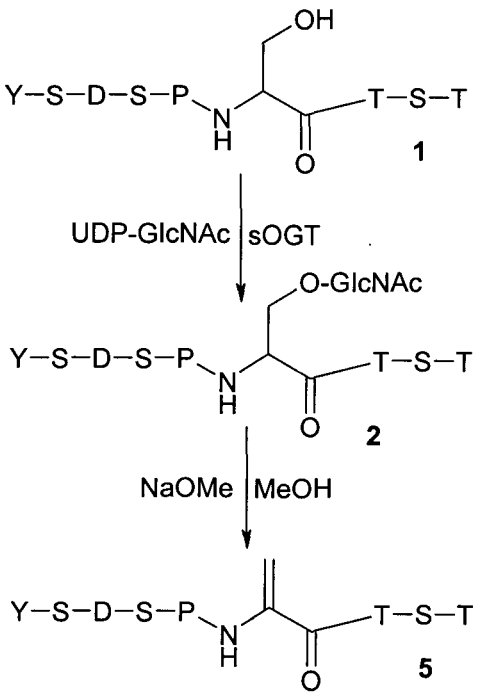
Figure 5.3 The selected photoaffinity groups



Scheme 5.1 Syntheses of **11** and **12**. **(a)** THF, Et_3N, rt, 59%; **(b)** Bis(trichloromethyl) carbonate, CH_2Cl_2, NaOH solution, 0 °C to rt; **(c)** chlorzoxazone, THF, Et_3N, rt, total yield for step b and step c is 43%



Scheme 5.2 Synthesis of **13**. **(a)** 1,1′-carbonyl diimidazole, THF, rt, 97%; **(b)** H_2, Pd/C, MeOH, rt, 90%; **(c)** MeOH, Et_3N, $CuSO_4$ solution, TfN_3 in CH_2Cl_2, 0 °C to rt, 90%; **(d)** THF, Et_3N, rt, 70%.



Scheme 5.3 sOGT catalyzed *O*-GlcNAcylation of peptide YSDSPSTST and beta-elimination of *O*-GlcNAcylated peptide YSDSP(*O*-GlcNAc-Ser)TST.

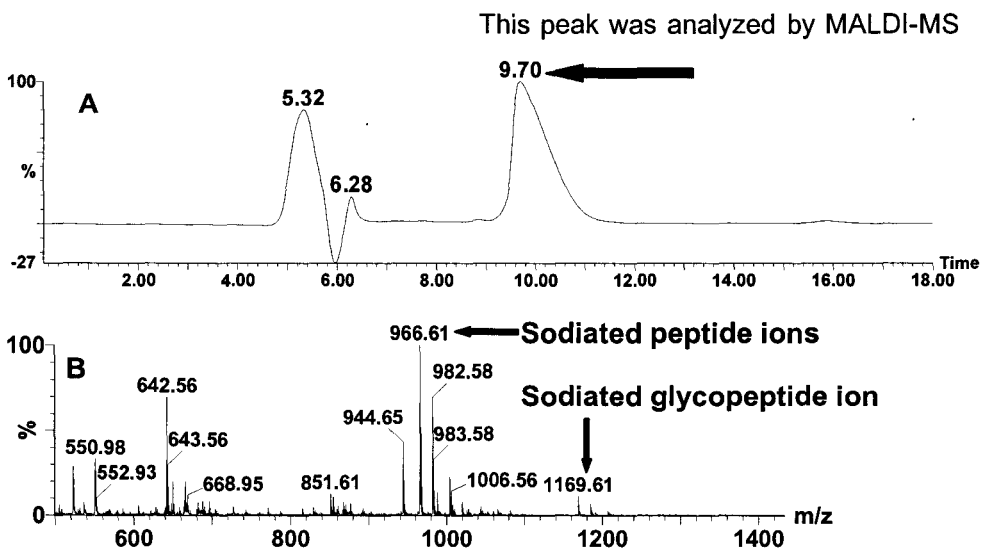


Figure 5.4 CapLC-MALDI MS analysis of *O*-GlcNAcylation of YSDSPSTST **1.** **A)** CapLC analysis of *O*-GlcNAcylation of **1**, 93% A/7% B, flow rate: 5 μL/min, scan wavelength: 195 nm-320 nm; **B)** MALDI MS of the peak (retention time is 9.70 min on Figure 5.4A)

Time (min)	A%	B%
0	96	4
30	92	8
31	96	4
40	96	4

Table 5.1 CapLC analytical method for separating **1** and **2** (the ratio of **1** to **2** is 1:1) (A 0.1% HCOOH/1% CH_3CN/H_2O, B 0.1% HCOOH/CH_3CN), flow rate: 5 μL/min, scan wavelength: 195 nm-320 nm

Time (min)	A%	B%
0	96	4
30	94	6
31	96	4
40	96	4

Table 5.2 CapLC analytical method for separating **1** and **2** (the ratio of **1** to **2** is 1:1) (A 0.1% HCOOH/1% CH_3CN/H_2O, B 0.1% HCOOH/CH_3CN), flow rate: 5 μL/min, scan wavelength: 195 nm-320 nm

Time (min)	A%	B%
0	96	4
30	95	5
31	96	4
40	96	4

Table 5.3 CapLC analytical method for separating **1** and **2** (the ratio of **1** to **2** is 1:1) (A 0.1% HCOOH/1% CH_3CN/H_2O, B 0.1% HCOOH/CH_3CN), flow rate: 5 μL/min, scan wavelength: 195 nm-320 nm

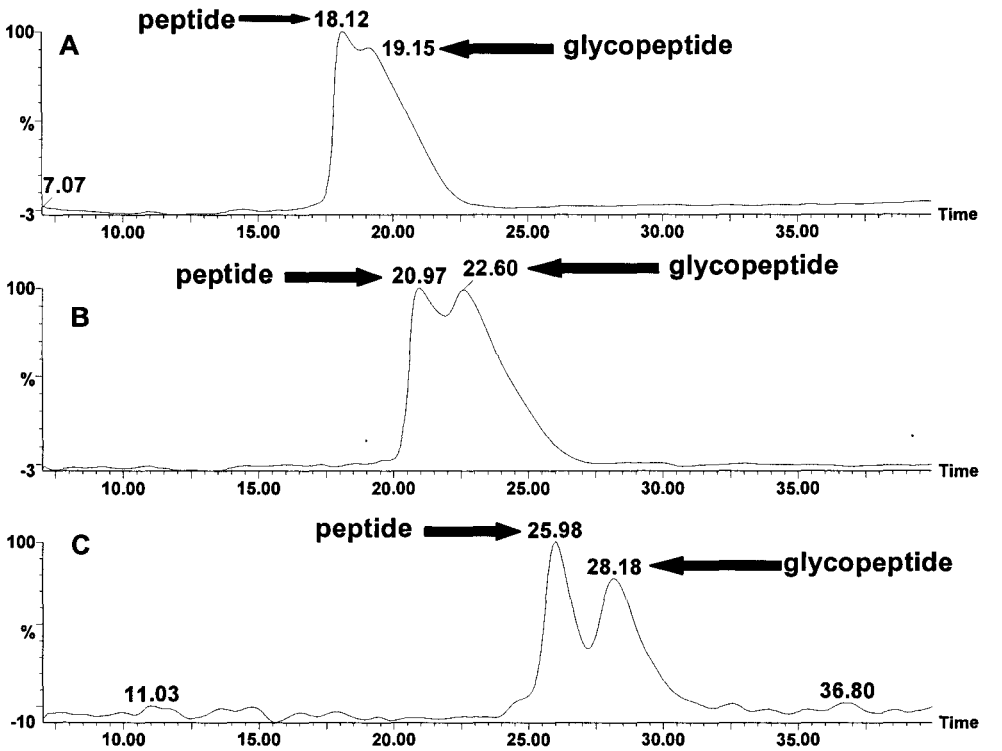
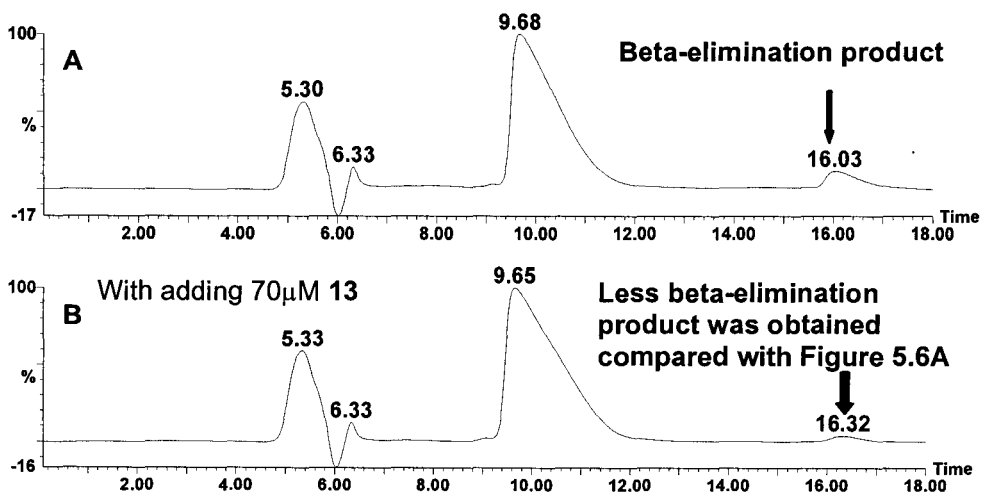


Figure 5.5 CapLC chromatograms of a mixture of peptide **1** and glycopeptide **2** (50 pmol: 50 pmol, 1: 1). **A)** The analytical method is shown in Table 5.1; **B)** the analytical method is shown in Table 5.2; **C)** the analytical method is shown in Table 5.3



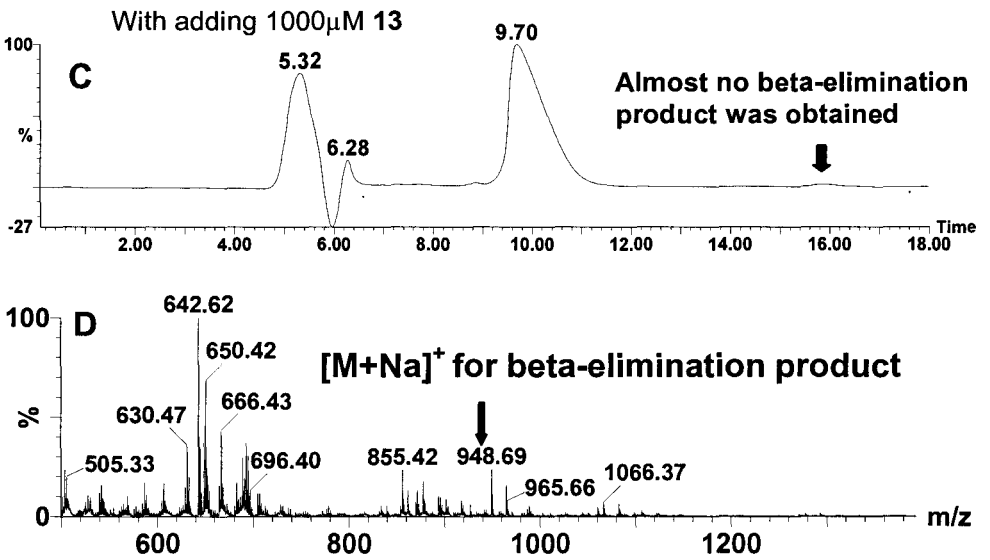


Figure 5.6 IC_{50} measurements of **11**, **12** and **13** with sOGT by CapLC-MALDI MS. **A)** CapLC chromatogram of the product after *O*-GlcNAcylation and beta-elimination treatment of YSDSPSTST **1** (without adding inhibitor); **B)** CapLC chromatogram of the product after *O*-GlcNAcylation and beta-elimination treatment of YSDSPSTST **1** (with adding 70 μM **13**); **C)** CapLC chromatogram of the product after *O*-GlcNAcylation and beta-elimination treatment of YSDSPSTST **1** (with adding 1000 μM **13**); **D)** MALDI MS of the peak (retention time is 16.03 min on Figure 5.6A);The CapLC analytical method was 93% A/7% B, flow rate: 5 μL/min, scan wavelength: 195 nm-320 nm

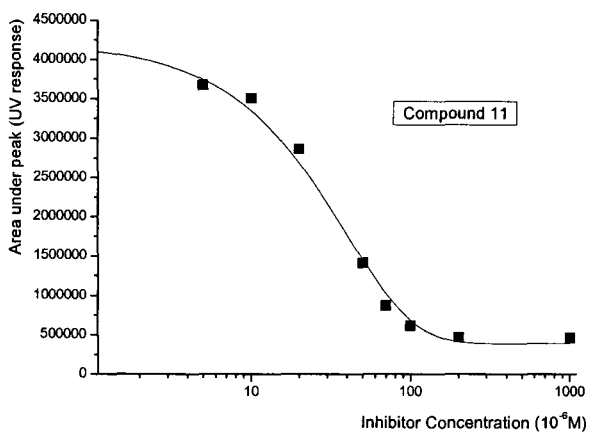


Figure 5.7 IC_{50} curves of compound **11** with sOGT. Each data point is the average of 3 determinations (from three repeated independent experiments). Error bars are not shown as the σ for each set of 3 was less than 1%

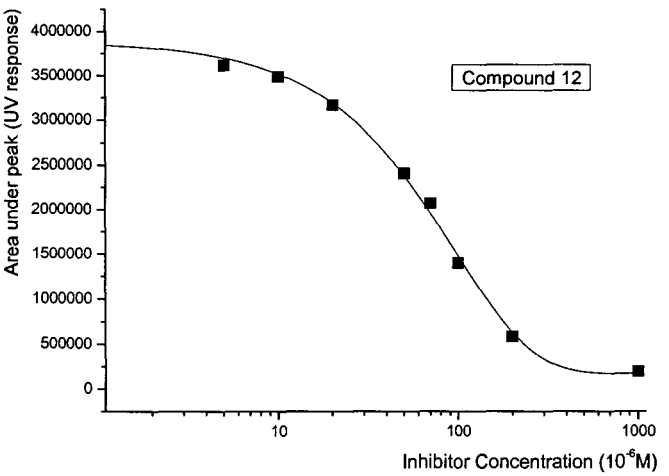


Figure 5.8 IC_{50} curves of compound **12** with sOGT. Each data point is the average of 3 determinations (from three repeated independent experiments). Error bars are not shown as the σ for each set of 3 was less than 1%

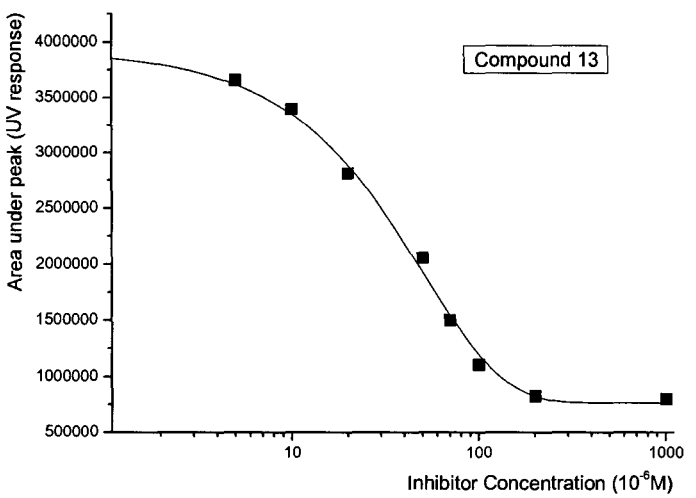


Figure 5.9 IC_{50} curves of compound **13** with sOGT. Each data point is the average of 3 determinations (from three repeated independent experiments). Error bars are not shown as the σ for each set of 3 was less than 1%

Chapter 6

Concluding Remarks

6.1 General Overview of the Thesis

A variety of important cellular functions rely on *O*-GlcNAcylation, a kind of post-translational glycosylation that modifies nuclear and cytoplasmic proteins on serine or threonine residues by the addition of a single sugar moiety β-*N*-acetylglucosamine (*O*-GlcNAc). Generally, two enzymes can catalyze the addition and removal of the *O*-GlcNAc *O*-glycosidic linkage. The *O*-GlcNAc transferase (OGT) catalyzes the transfer of the *N*-acetylglucosamine from uridine diphosphate *N*-acetylglucosamine to the hydroxyl group of specific Thr and Ser residues. On the other hand, hydrolysis of the sugar moiety from proteins is achieved via the bi-functional nuclear cytoplasmic *O*-GlcNAcase and acetyltransferase (NCOAT). Since *O*-GlcNAcylation plays very important roles in many cellular processes such as transcription, translation and protein-protein interaction, my research mainly focused on preparing compounds and developing methods to study the *O*-GlcNAcylation and discover the structural information of OGT and NCOAT.

In this thesis, the important enzyme substrates (YSDSPSTST for OGT and YSDSP(*O*-GlcNAc-Ser)TST for NCOAT) were prepared by solid phase synthesis and purified by preparative reverse phase HPLC. Their structure information and sequence were confirmed by ESI-MS and tandem MS (MS/MS). The building

block *S*-GlcNAc-Ac_3-Ser-COOBn for preparing *S*-GlcNAcylated peptides was prepared by a multistep procedure. Two potent photoaffinity labeled probes for mapping the active site of sOGT were designed and synthesized, and their IC_{50} values for sOGT were measured by CapLC and a beta-elimination method. Also, many analytical methods were developed for studying *O*-GlcNAcylation, including discovering the *O*-glycosylation site of peptides by MS/MS and beta-elimination methods, sequencing peptides by MS/MS, analyzing the protein digests by CapLC-MALDI MS and studying the ion mobility of peptides and glycopeptides by ion mobility spectrometry.

6.2 Outline of Future Work

Although the two photaffinity labeled probes were proven to have inhibitory effects on sOGT, they still need to be applied for mapping the active site of sOGT through photoaffinity labeling experiments, and by using CapLC-MALDI MS as the analytical tool. Thus, we are planning to use the prepared enzyme substrates, the developed method (photoaffinity labeling method) and the synthesized photoaffinity labeled probes to discover the active sites of sOGT. To date, we always used peptide YSDSPSTST and glycopeptide YSDSP(*O*-GlcNAc-Ser)TST as the enzyme substrates to study the *O*-GlcNAcylation, it would be interesting to extend these *O*-GlcNAcylation studies to a protein level. The protein beta-catenin could be selected as an enzyme substrate because it contains many modification sites. We can analyze *O*-GlcNAcylated beta-catenin trypsin digests by the CapLC-MALDI MS assay developed in this work. If CapLC-

MALDI MS can detect the *O*-GlcNAcylation of beta-catenin, we are also planning to use it as a powerful tool to study the interplay between *O*-GlcNAcylation and phosphorylation of different substrates such as YSDSPSTST, beta-catenin peptide (21mer peptide purchased from P.R.China) or beta-catenin protein.