

# Supporting Information

## Site specific protein O-glucosylation with bacterial toxins

Y. Sun, L. M. Willis, H. Batchelder and M. Nitz

## **1. General Methods**

### **1.1 General Materials and Chemical Reagents**

All amino acid building blocks for peptide synthesis and chemicals were purchased from Sigma Aldrich (Oakville, ON). Fmoc-rink amide-MBHA resins were purchased from Ana Spec (Fremont, CA). 1% Picrylsulfonic acid in DMF solution used for primary amine test was purchased from Sigma Aldrich (Oakville, ON). UDP-glucose was purchased from Sigma Aldrich (Oakville, ON). Organic solvents used for peptide synthesis and HPLC were purchased from Sigma Aldrich (Oakville, ON). Restriction enzymes and quick-stick ligase were purchased from New England Biolabs (Mississauga, Ontario). QIA mini prep kit was purchased from New England Biolabs (Mississauga, Ontario). Amylose resin was purchased from New England Biolabs (Mississauga, Ontario). Protease Inhibitor Cocktail pellet was purchased from Roche Diagnostics (Laval, Quebec).

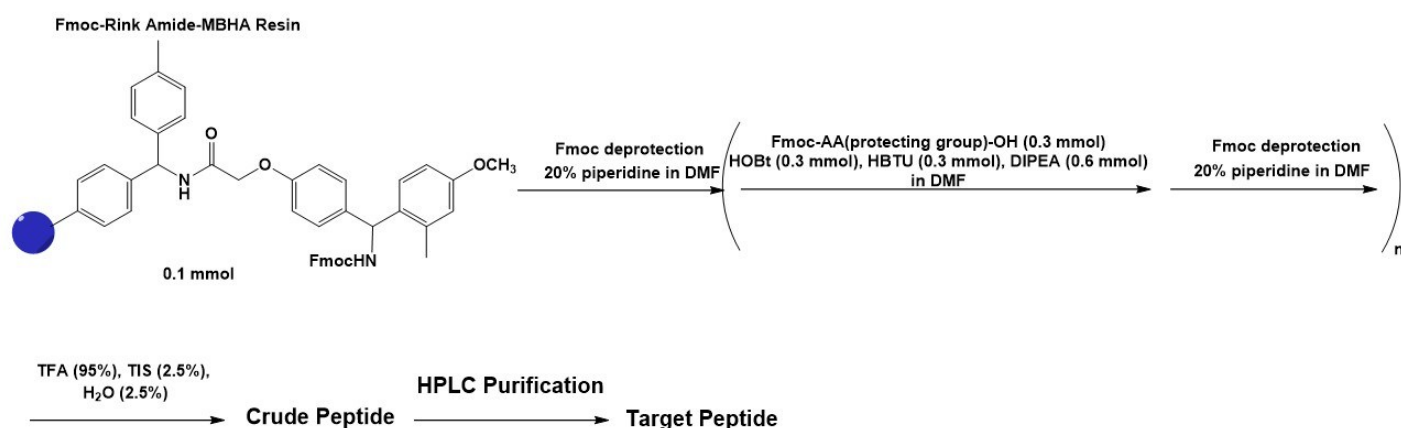
### **1.2 General Instruments**

The MALDI spectra were taken using a Waters Micromass MALDI micro MX™ (matrix-assisted laser/desorption/ionization time-of-flight mass spectrometer [MALDI-ToF MS]) using  $\alpha$ -cyano-4-hydroxycinnamic acid matrix. UV absorption was measured on either a Shimadzu UV-2401 PC spectrophotometer or a Nanodrop ND\_1000. Semi-preparative HPLC separations were performed on a Waters 1525 Binary HPLC pump and Waters 2487 dual  $\lambda$  absorption detector using a Waters XBridge™ Prep BEH130 C18 5  $\mu$ m (10  $\times$  250 mm) reverse phase column. Protein mass spectra were obtained from an ABI/Sciex QStar mass spectrometer fitted with a desalting column prior to direct introduction into the ESI source.  $^{19}\text{F}$  NMR spectra were recorded on an Agilent DD2-700 MHz spectrometer.

## 2. Solid-Phase Peptide Synthesis

### 2.1 Peptide synthesis (Scheme 1)

All peptides in this paper were synthesized on Fmoc-rink amide-MBHA resin using Fmoc-SPPS chemistry. Each peptide was synthesized on a scale of 0.1 mmol if not specified. Before synthesis, the resin was swollen in DMF overnight at r.t. Fmoc-L-Phe-OH, Fmoc-Pro-OH, Fmoc-Gly-OH, Fmoc-gly(OtBu)-OH, Fmoc-Tyr(tBu)-OH, Fmoc-Ile-OH, Fmoc-Thr(tBu)-OH, Fmoc-Ile-OH, Fmoc-Thr(tBu)-OH, Fmoc-Val-OH, Fmoc-Phe-OH, Fmoc-Asp((OtBu)-OH, Fmoc-Asn(Trt)-OH, Fmoc-Ser(tBu)-OH, Fmoc-Ala-OH, and Fmoc-L-Phe(4-F)-OH were used as building blocks. The coupling reaction was carried out in DMF (1.5 ml) containing Fmoc-L-amino acid (0.3 mmol), HBTU (0.3 mmol), HOBT (0.3 mmol), and DIPEA (0.6 mmol) for 1 h at r.t. Fmoc protecting group was removed by treating the resin with 20% piperidine in DMF (1.5 ml) for 15 min at r.t. TNBS test was used to confirm the completion of each Fmoc deprotection and coupling reaction. The final desired peptide was obtained by incubating the resin in the solution (95% TFA, 2.5% TIS, 2.5% H<sub>2</sub>O) for 2 h at r.t. with occasional inversion, followed by the RF-HPLC purification.



**Scheme S1.** General peptide synthesis procedure

## 2.2 Peptide purification

The crude peptides were dissolved in DMF and purified by semi-preparative RP-HPLC (XBridge Prep BEH 130 C 18, 10×250 mm). The purification condition of each peptide is shown in Table 1 and Table 2. HPLC fractions corresponding to the desire peptide were confirmed by MALDI. The peptide concentration was quantified by the UV absorption in 0.02 M phosphate buffer (6M guanidine hydrochloride, pH6.5) using an extinction coefficient that was calculated from ProtParam tool. Peptides were stored in DMSO at -20 °C.

**Table S1.** HPLC gradient

Method	Flow rate (mL/min)	H <sub>2</sub> O+0.1% TFA (%)	MeCN+0.1% TFA (%)	Time (min)
A	3.0	90→30	10→70	0→50
B	3.0	90→30	10→70	0→30
C	2.0	90→30	10→70	0→30
D	2.0	90→40	10→60	0→50
		40	60	50→70
E	2.0	90→50	10→50	0→100
F	3.0	90→30	10→70	0→30
		30→20	70→80	30→60

**Table S2.** Peptide HPLC purification methods

Name	Sequence	Method	Name	Sequence	Method
1	FVDEYDPTIEDSY	D	11	YIPTVFDNA	F
2	FPGEYIPTVF*DNY	D	12	YIPTVFDAY	F
3	PGEYIPTVF*DNY	D	13	YIPTVFANY	A
4	PTVF*D	D	14	YIPTVADNY	F
5	IPTVF*D	D	15	YIPTAFDNY	F
6	YIPTVF*D	C	16	YIATVFDNY	A
7	EYIPTVF*D	B	17	YAPTVDNY	A
8	YIPTVF*DNY	A	18	AIPTVFDNY	A
9	YIPTVFDN	E	19	YAPTVDAY	A
10	YIPTVFDNY	A			

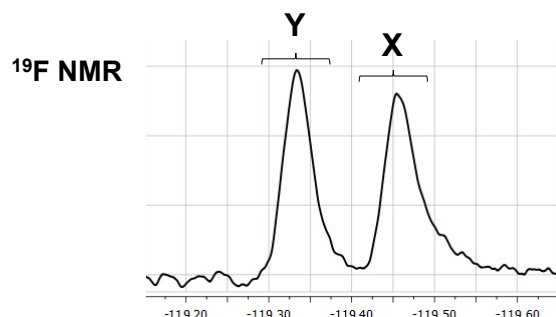
### 3. MALDI Kinetics

#### **3.1 Peptide glucosylation reaction**

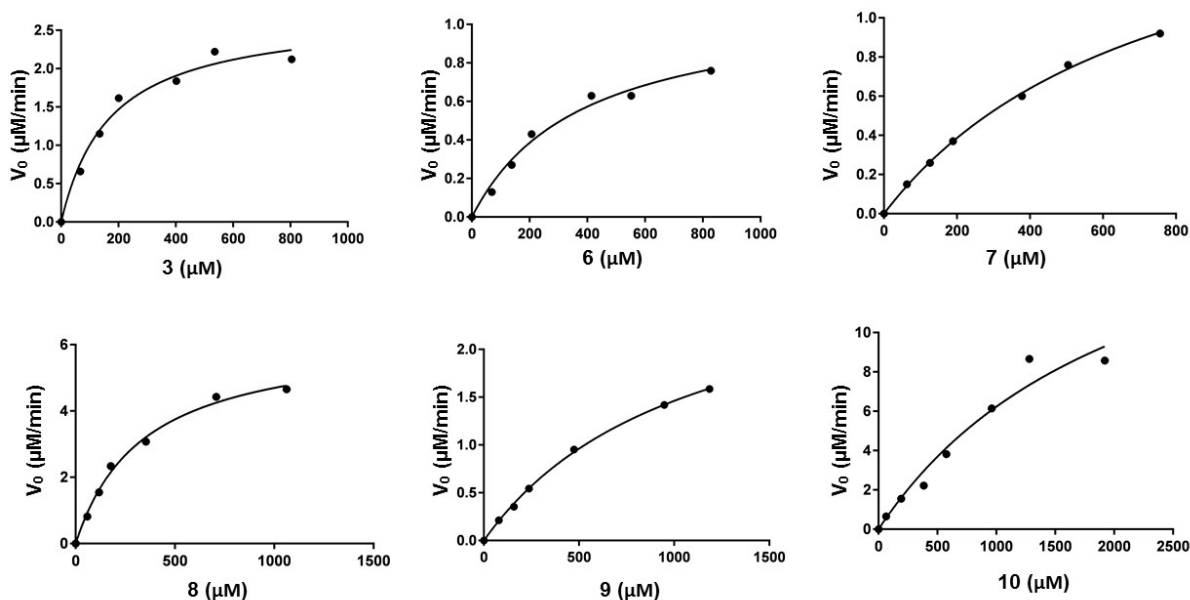
Peptide (various concentrations, maximum DMSO concentration 20%, Enzyme activity was not substantially affected by DMSO at this concentration) was incubated with Toxin B (2  $\mu$ M, or 5  $\mu$ M, or 10  $\mu$ M) and an excess UDP-glucose (10 mM) buffer (50 mM HEPES, 100 mM KCl, 2 mM  $MgCl_2$ , 1 mM  $MnCl_2$ , pH 7.5) at 37°C for 4 hours. Aliquots of the reaction mixture (10  $\mu$ L) were taken at 15 min, 30 min, 1hr, 2hr, and 4 hr time points. Immediately after removing the samples they were loaded and desalted on a C18 ZipTip using an 50/50 acetonitrile/water eluent. A sample (~2  $\mu$ L) was then spotted onto a crystallized CHCA matrix spot on the MALDI plate. The averaged MALDI spectrum was obtained by collecting 60 to 100 shots across the sample.

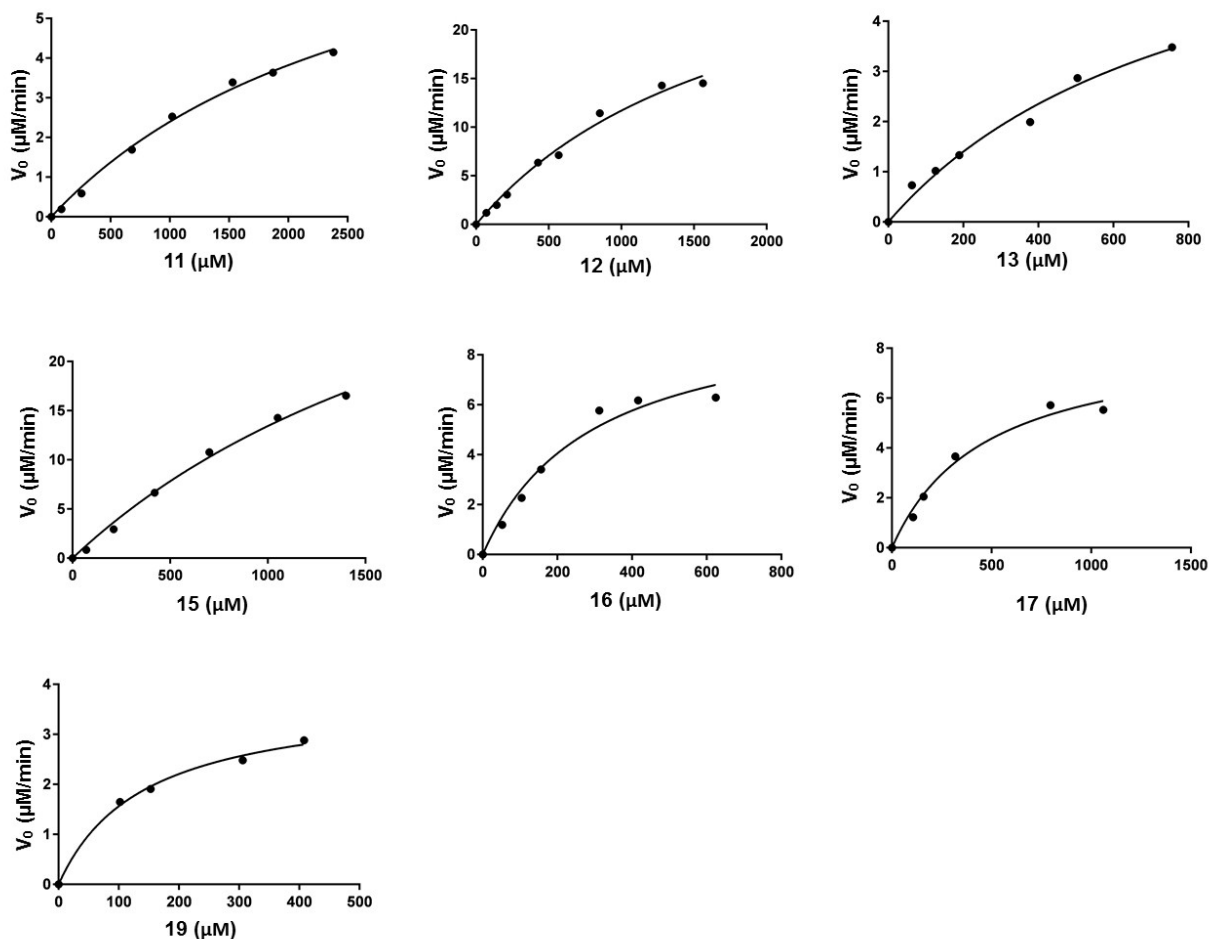
#### **3.2 Michaelis Menten fitting of peptide substrates**

The glycosylation rate was calculated by the formula  $\% \text{ of glycosylation} = \frac{Y}{X + Y}$  (Figure S1). From velocity vs peptide concentration plots the  $K_m$  and  $k_{cat}$  values were determined by fitting with Graph Pad Prism 6 (Figure S2).



**Figure S1.** <sup>19</sup>F NMR spectrum of partially glycosylated Peptide **2**. Peptide **2** (100  $\mu$ M) was incubated with Toxin B (10  $\mu$ M) and UDP-Glc (10 mM) in reaction buffer (50 mM HEPES, 100 mM KCl, 2 mM MgCl<sub>2</sub>, 1 mM MnCl<sub>2</sub>, pH 7.5) at 37° C for 60 min. The sample was boiled and clarified by centrifugation. After addition of 5% D<sub>2</sub>O the samples were analyzed by <sup>19</sup>F NMR. Y denotes the glycosylate, X denotes the native **2**.





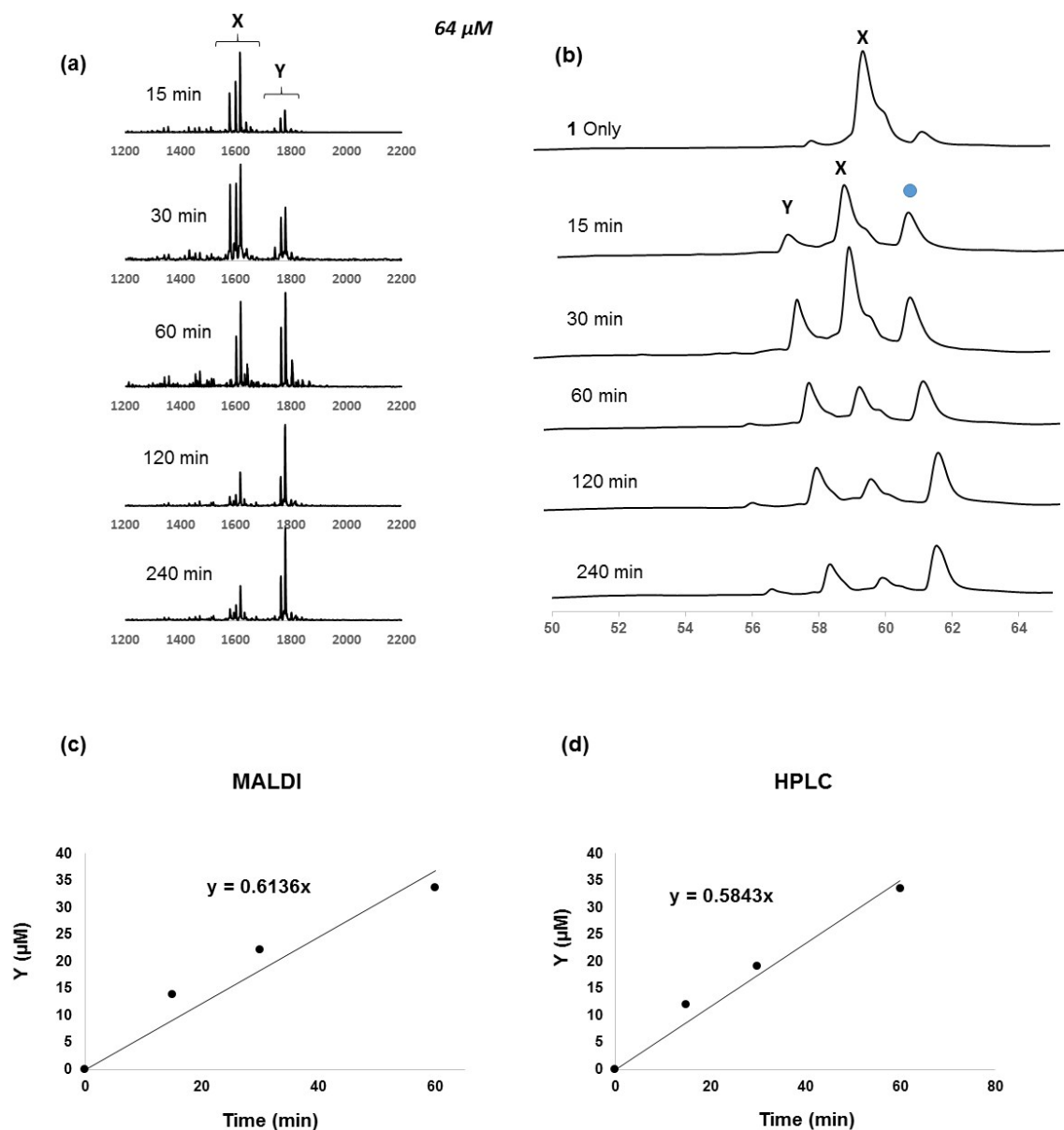
**Figure S2.** Fitting of glycosylation rates to Michaelis Menten parameters of peptides **3-19**

## 4. HPLC Kinetics

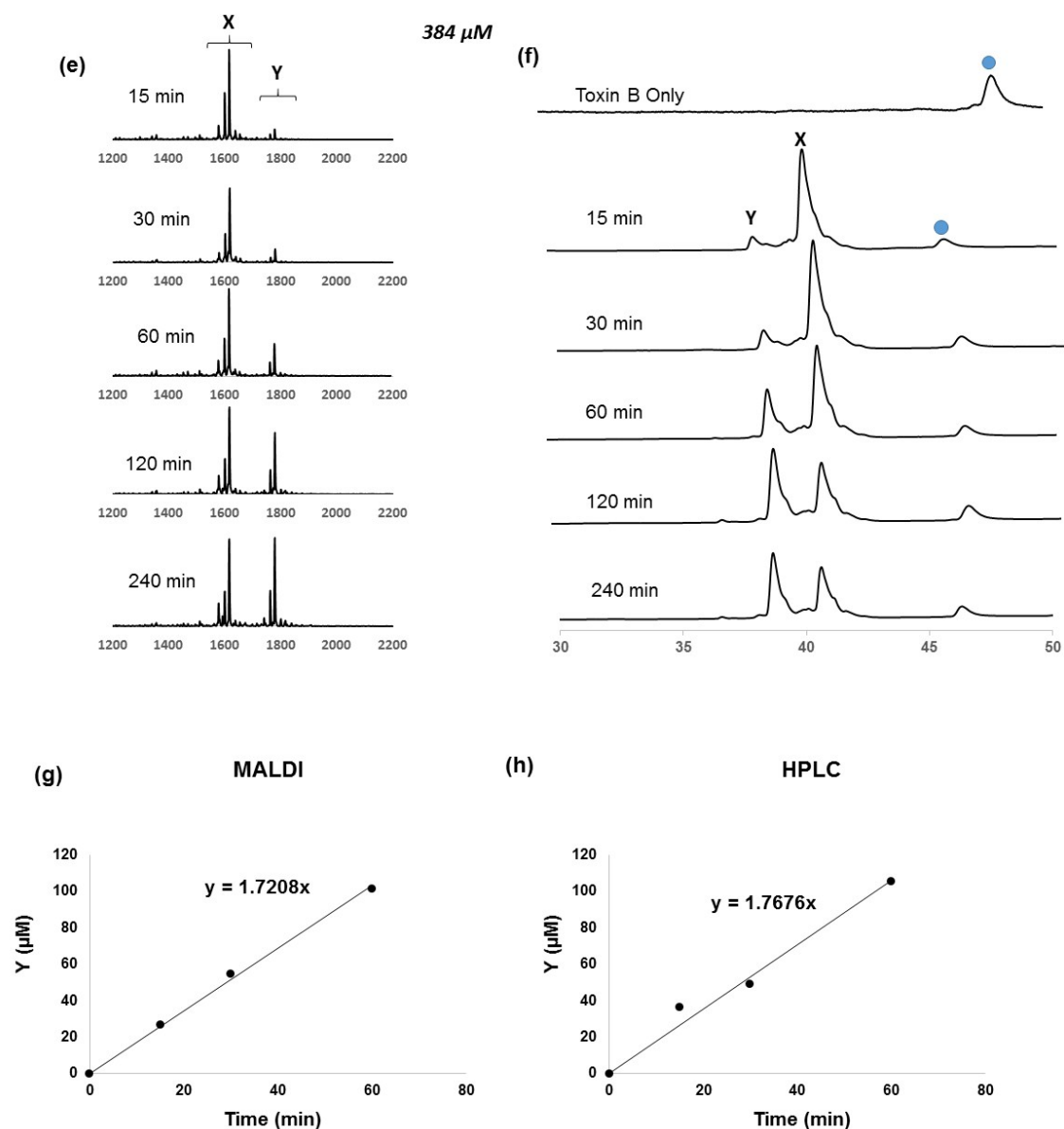
### 4.1 Glucosylation conditions

Peptide **2** (64  $\mu\text{M}$  or 384  $\mu\text{M}$ ) was incubated with Toxin B (10  $\mu\text{M}$ ) and an excess UDP-glucose (10 mM) in reaction buffer (50 mM HEPES, 100 mM KCl, 2 mM  $\text{MgCl}_2$ , 1 mM  $\text{MnCl}_2$ , pH 7.5) at 37° C for 15 mins, 30 mins, 1 hr, 2 hr and 4 hrs. The total volume of each sample was 100  $\mu\text{L}$ . Each sample was diluted with 100  $\mu\text{L}$   $\text{H}_2\text{O}$  and the entire solution was injected into RP-HPLC (XBridge Prep BEH 130 C 18, 10×250 mm) and elution was followed at 215 nm to determine ratio of starting peptide to glucosylated

peptide for kinetic analysis. The HPLC gradient used was Method A in Table S2. The HPLC results of **2** glucosylation was shown in Figure S3. The % of glycosylation was calculated by  $\% \text{ of glycosylation} = \frac{Y}{X + Y}$ . The initial velocities calculated from MALDI and HPLC are agree within 10% (Figure S3 c & d, and g & h).







**Figure S3.** Comparison between MALDI and HPLC kinetics. (a) MALDI spectra of **2** glucosylation at 64  $\mu\text{M}$ ; (b) HPLC spectra of **2** glucosylation at 64  $\mu\text{M}$ ; (c) Initial velocity of **2** (64  $\mu\text{M}$ ) calculated by MALDI results; (d) Initial velocity of **1** (64  $\mu\text{M}$ ) calculated by HPLC results; (e) MALDI spectra of **2** glucosylation at 384  $\mu\text{M}$ ; (f) HPLC spectra of **2** glucosylation at 384  $\mu\text{M}$ ; (g) Initial velocity of **2** (384  $\mu\text{M}$ ) calculated by MALDI results; (h)

Initial velocity of **2** (384  $\mu$ M) calculated by HPLC results; X: native **2**; Y: glucosylated **2**;  
blue dot: Toxin B

## **5. MBP- tagged protein constructs**

### **5.1 Cloning**

#### ***Construction of plasmids encoding N-terminal MBP-tag fusions.***

The Q5 site directed mutagenesis kit (New England Biolabs) was used to introduce a stop codon into the vector pCW*malE*-thrombin using the primers 5'-TATCACCAAGTAATGAGTGGCGGACATA and 5'-CGAGTCTGCGCGTCTTTC. The MalE protein expressed from the resulting plasmid contains the sequence KDAQTRITK at the C-terminal end. This plasmid was used as the parent for insertion of sequences encoding the peptide YIPTVFDAY immediately after the ATG start codon, using the Q5 SDM kit according to manufacturer's instructions. Primers were as follows:

5'- TTTTGATGCATACAAAATCGAAGAAGGTAACTG and 5'-  
ACGGTCGGGATGTACATCTGACCTCCTAAGCATC for YIPTVFDAY.

Plasmids were verified by sequencing and then transferred to *E. coli* BL21 for expression.

#### ***Construction of plasmids encoding C-terminal MBP-tag fusions.***

Oligonucleotides were designed so that complimentary pairs contained overhanging sequence that could anneal with *Nde*II/*Sal*I digested pCW*malE*-thrombin (REF). The sequences were as follows: 5'-

TATGTACGCCCCGACCGTTTTTGATGCATACTAATAG and 5'-

TCGACTATTAGTATGCATCAAAAACGGTCGGGGCGTACA for YAPTVFDAY, 5'-

TATGTACATCCCGACCGTTTTTGGATGCATACTAATAG and 5'-

TCGACTATTAGTATGCATCAAAAACGGTCGGGATGTACA for YIPTVFDAY, and 5'-

TATGTACGCCCCGACCGTTTTTGGATACTAATAATAG and 5'-

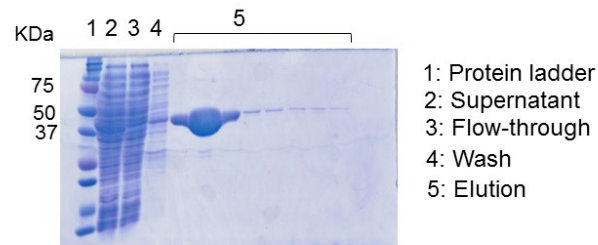
TCGACTATTAGTAGTTATCAAAAACGGTCGGGGCGTACA for YAPTVDNY.

Pairs of oligos were annealed at 95°C for 30 sec followed by 25°C for 15 min, then phosphorylated using polynucleotide kinase (New England Biolabs) as per manufacturer's instructions. Inserts were ligated with *Nde*I/*Sall*-cut pCW*malE*-thrombin and used to transform *E. coli* DH5α by electroporation. Plasmids were verified by sequencing and then transferred to *E. coli* BL21 for expression.

## **5.2 Expression and purification**

*E. coli* BL 21 (DE3) transformant containing the pCW-tagged MBP plasmid was grown in LB broth containing 100 µg/ml ampicillin at 37°C. When the OD<sub>600</sub> reached ~0.6, IPTG (0.5 mM) was added to induce the protein expression. The cells were then allowed to grow at 16 °C for 20 hours. The cells were harvested by centrifugation at 4000 rpm at 4°C for 40 min. The pellet was resuspended in 20 mM Tris HCl pH 8.0, 200 mM NaCl containing a Protease Inhibitor Cocktail pellet (Roche Diagnostics, Laval, Quebec) and sonicate on ice. The lysate was centrifuged at 12000 rpm for 40 min at 4°C to remove the cell debris. The supernatant was applied to amylose resin which was equilibrated with washing buffer (20 mM TrisHCl pH 8, 200 mM NaCl). The column was washed with 10 column volumes of washing buffer and MBP-tagged proteins eluted with elution buffer (20 mM TrisHCl pH 8, 200 mM NaCl, and 10 mM maltose). The fractions containing desired protein were detected by SDS-PAGE (Figure S4) and the purified proteins were

combined, purified through P10 column, and further concentrated by Amicon Ultra. Protein concentration was determined by UV absorption based on the theoretical extinction coefficient. Final protein was stored in -20 °C.



**Figure S4.** SDS-PAGE of purified MBP-tagged protein (construct 2).

### **5.3 MBP-tagged protein sequences**

#### **MBP-12**

MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGP  
DIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIY  
NKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGK  
YDIKDVGVNDAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWS  
NIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEA  
VNKDKPLGAVALKSYYYELAKDPRIAATMENAQKGEIMPNIQMSAFWYAVRTAVINAA  
SGRQTVDEALKDAQTRITKGGGHIFNPRGSHMYIPTVFDAY

#### **12-MBP**

MYIPTVFDAYKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQ  
VAATGDGPDIIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPI  
AVEALSLIYNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGY  
AFKYENGKYDIKDVGVNDAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTI  
NGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYL  
LTDEGLEAVNKDKPLGAVALKSYYYELAKDPRIAATMENAQKGEIMPNIQMSAFWYA  
VRTAVINAASGRQTVDEALKDAQTRITK

#### **MBP-His**

MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGP  
DIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIY  
NKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGK  
YDIKDVGVNDAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWS

NIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEA  
VNKDKPLGAVALKSYYYEELAKDPRIAATMENAQKGEIMPNIPQM SAFWYAVRTAVINAA  
SGRQTVDEALKDAQTRITKHHHHHHH

#### MBP-19

MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGP  
DIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIY  
NKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGK  
YDIKDVGVNDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWS  
NIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEA  
VNKDKPLGAVALKSYYYEELAKDPRIAATMENAQKGEIMPNIPQM SAFWYAVRTAVINAA  
SGRQTVDEALKDAQTRITKGGGHIFNPRGSHMYAPTVDAY

#### MBP-17

MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGP  
DIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLIY  
NKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGK  
YDIKDVGVNDNAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWS  
NIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLEA  
VNKDKPLGAVALKSYYYEELAKDPRIAATMENAQKGEIMPNIPQM SAFWYAVRTAVIN  
AASGRQTVDEALKDAQTRITKGGGHIFNPRGSHMYAPTVDNY

Yellow: Thrombin-recognition site

Red: Peptide tag

### 5.4 Glucosylation of MBP-tagged proteins

The MBP-tagged protein (100  $\mu$ M) was incubated with Toxin B (10  $\mu$ M) and UDP-Glucose (10 mM) in reaction buffer (50 mM HEPES, 100 mM KCl, 2 mM MgCl<sub>2</sub>, 1 mM MnCl<sub>2</sub>, pH 7.5) at 37°C overnight. The resulted glucosylation reaction was monitored by ESI (Figure S5).

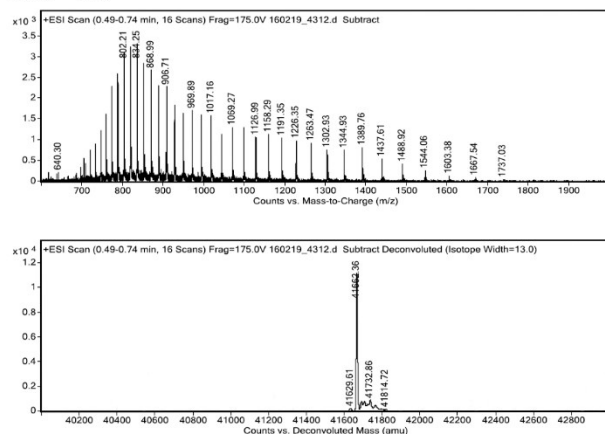
**Table S3.** ESI analysis of glycosylated MBP.

Protein	TcdB	Expected	Observed	%
	GTD	<i>m/z</i>	<i>m/z</i>	glycosylated

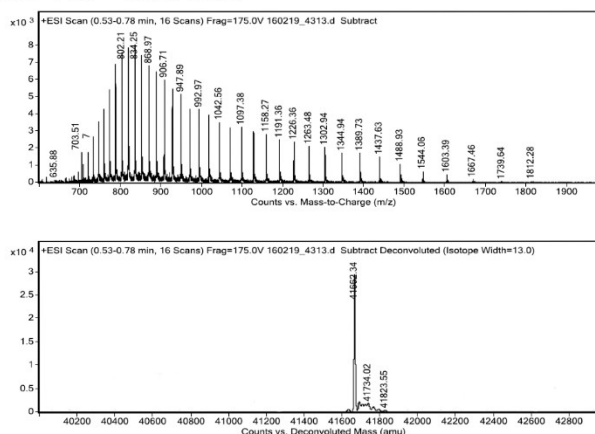
<b>MBP-His</b>	-	41661.3	41662.5	-
<b>MBP-His</b>	+	41661.3	41662.5	0
<b>MBP-12</b>	-	43257.2	43258.7	-
<b>MBP-12</b>	+	43419.2	43420.4	80/~95*
<b>MBP-19</b>	-	43215.1	43216.7	-
<b>MBP-19</b>	+	43377.1	43378.9	60
<b>MBP-17</b>	-	43258.1	43259.6	-
<b>MBP-17</b>	+	43420.1	43421.9	64
<b>12-MBP</b>	-	41908.7	41910.2	-
<b>12-MBP</b>	+	42070.1	4272.2	>95*

\*rxns with alkaline phosphatase and 20mM UDP-Glc

(a)  
MBP-His

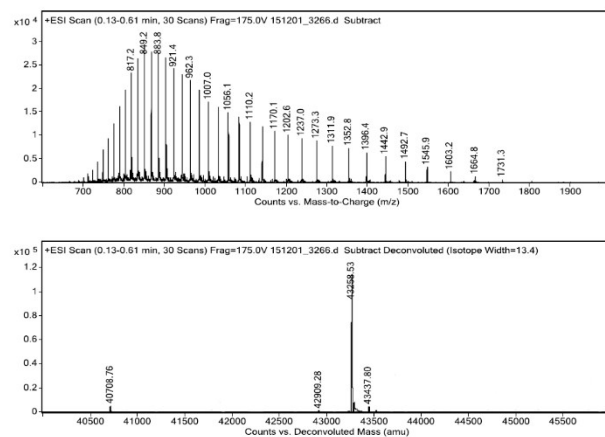


(b)  
MBP-His + TcdB GTD



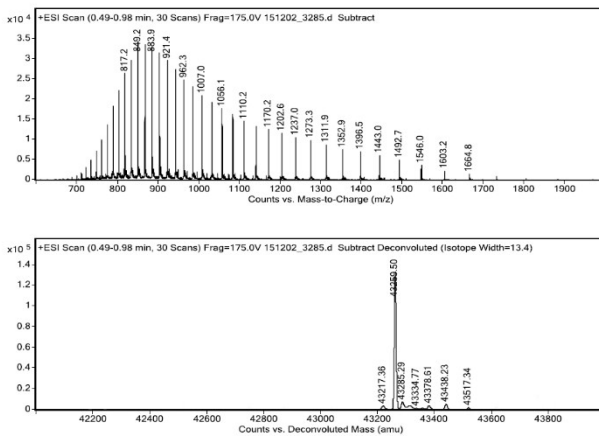
**Figure S5.** ESI analysis of glucosylated proteins. Proteins were incubated with UDP-Glc (10 mM) and Toxin B (10  $\mu$ M) in reaction buffer (50 mM HEPES, 100 mM KCl, 2 mM MgCl<sub>2</sub>, 1 mM MnCl<sub>2</sub>, pH 7.5) at 37°C.

(a)  
MBP-12

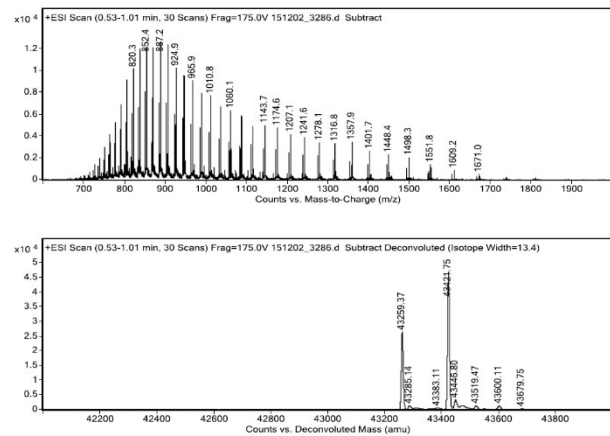


**Figure S6.** ESI analysis of glucosylated proteins. Proteins were incubated with UDP-Glc (10 mM) and Toxin B (10  $\mu$ M) in reaction buffer (50 mM HEPES, 100 mM KCl, 2 mM MgCl<sub>2</sub>, 1 mM MnCl<sub>2</sub>, pH 7.5) at 37°C.

(a)  
MBP-17

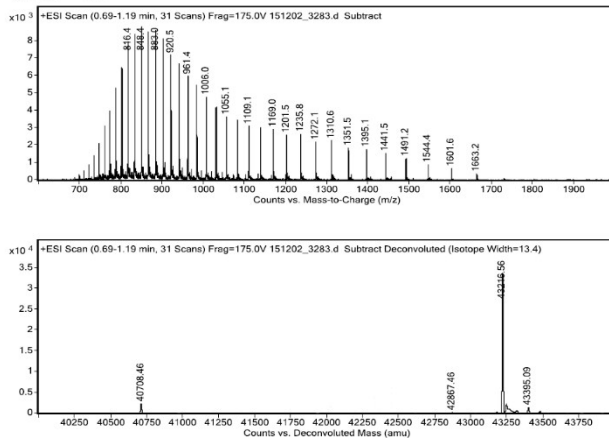


(b)  
MBP-17 + TcdB GTD

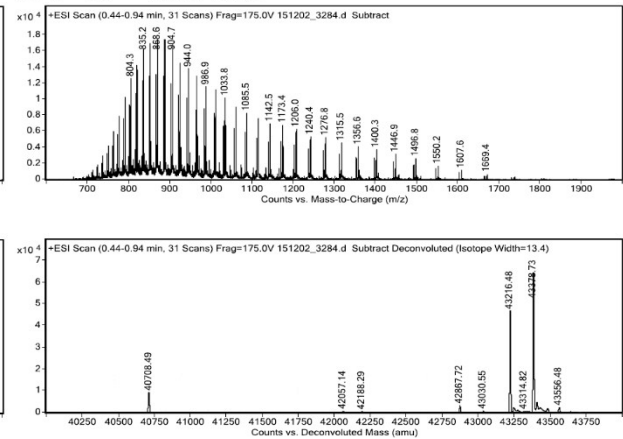


**Figure S7.** ESI analysis of glucosylated proteins. Proteins were incubated with UDP-Glc (10 mM) and Toxin B (10  $\mu$ M) in reaction buffer (50 mM HEPES, 100 mM KCl, 2 mM MgCl<sub>2</sub>, 1 mM MnCl<sub>2</sub>, pH 7.5) at 37°C.

(a)  
MBP-19



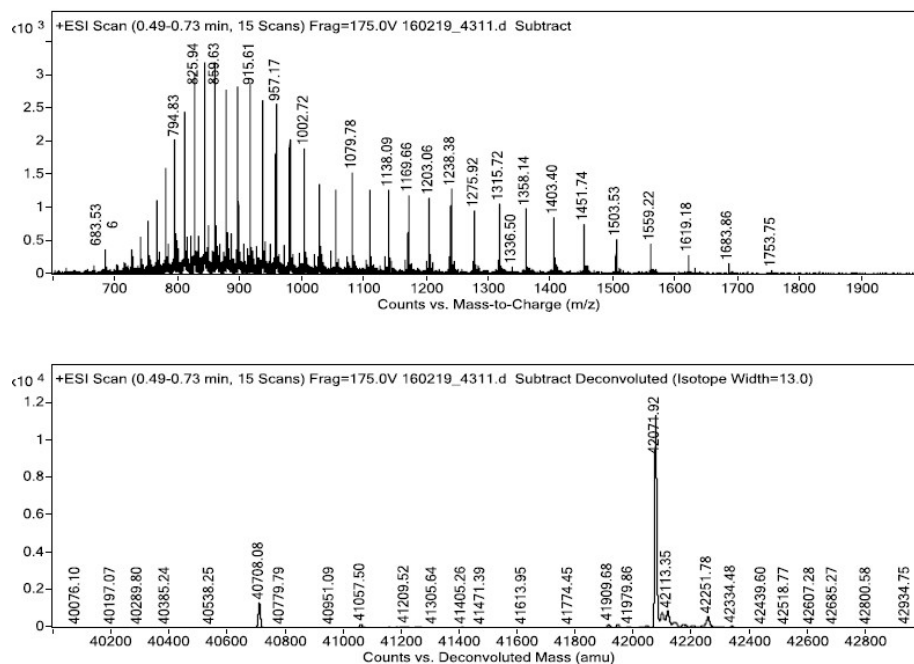
(b)  
MBP-19 + TcdB GTD



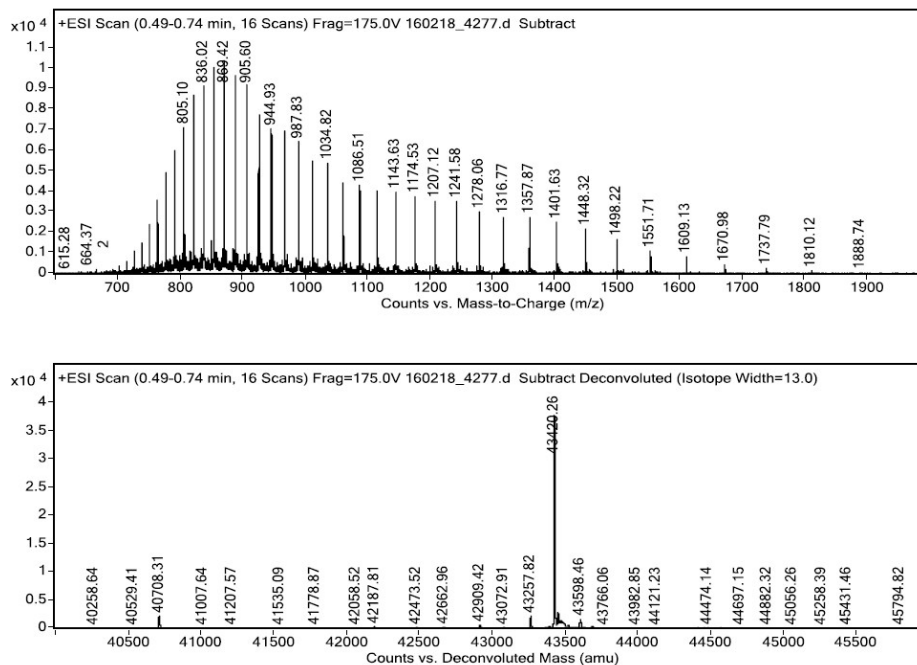


**Figure S8.** ESI analysis of glucosylated proteins. Proteins were incubated with UDP-Glc (10 mM) and Toxin B (10  $\mu$ M) in reaction buffer (50 mM HEPES, 100 mM KCl, 2 mM MgCl<sub>2</sub>, 1 mM MnCl<sub>2</sub>, pH 7.5) at 37°C.

## MBP-12



## 12-MBP



**Figure S9.** ESI analysis of glucosylated proteins. Proteins were incubated with UDP-Glc (20 mM), Toxin B (10  $\mu$ M) and Alkaline Phosphatase in reaction buffer (50 mM HEPES, 100 mM KCl, 2 mM  $MgCl_2$ , 1 mM  $MnCl_2$ , pH 7.5) at 37°C. Top MBP-12, Bottom 12-MBP

## 6. Toxin B Expression and Purification

*E.coli* BL 21 (DE3) transformant containing the Toxin B plasmid was grown in LB broth containing 50  $\mu$ g/ml ampicillin at 37°C. When the  $OD_{600}$  reached  $\sim 0.8$ , IPTG (0.5 mM) was added to induce the protein expression. The cells were then allowed to grow at 37°C for another 3 hours. The cells were harvested by centrifugation at 4000 rpm at 4°C for 40 min. The pellet was resuspended in 20 mM TrisHCl pH 8.0, 500 mM NaCl containing a Protease Inhibitor Cocktail pellet (Roche Diagnostics, Laval, Quebec) and sonicate on ice. The lysate was centrifuged at 12000 rpm for 40 min at 4°C to remove the cell debris. The supernatant was applied to Ni-NTA agrose resin which was equilibrated with equilibrating buffer (20 mM TrisHCl pH 8, 500 mM NaCl). The column was washed with

washing buffer (20 mM TrisHCL pH 8.0, 500 mM NaCl, and 20 mM imidazole) and Toxin B protein was eluted with elution buffer (20 mM TrisHCl pH 8, 500 mM NaCl, and 200 mM imidazole). The fractions containing desired protein were detected by SDS-PAGE and the purified protein was dialyzed against 50 mM HEPES, pH7.5 and was further concentrated by Amicon Ultra. Protein concentration was quantified by UV absorption according to the theoretical extinction coefficient. Final protein was stored at -20 °C in 50mM HEPES pH 7.5 and is stable for extended periods of time.