BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net **Hypothesis**

Volume 6(5)

Analysis of glycosylation motifs and glycosyltransferases in Bacteria and Archaea

Syed Tabish, Abbas Raza, Arshan Nasir, Sadia Zafar, Habib Bokhari*

Department of Biosciences, COMSATS Institute of Information Technology, Park Road, Chak Shahzad, Islamabad, Pakistan; Habib Bokhari - Email: habib@comsats.edu.pk; Phone: +92-300-5127684; Fax: 0092-051-4442805; *Corresponding author

Received March 03, 2011; Accepted March 18, 2011, Published May 26, 2011

Abstract:
The process of glycosylation has been studied extensively in prokaryotes but many questions still remain unanswered. Glycosyltransferase is the enzyme which mediates glycosylation and has its preference for the target glycosylation sites as well as for the type of glycosylation i.e. N-linked and O-linked glycosylation. In this study we carried out the bioinformatics analysis of one of the key enzymes of *pgl* locus from *Campylobacter jejuni*, known as PglB, which is distributed widely in bacteria and AglB from archaea. Relatively little sequence similarity was observed in the archaeal AglB(s) as compared to those of the bacterial PglB(s). In addition we tried to the answer the question of as to why not all the sequins Asp-X-Ser/Thr have an equal opportunity to be glycosylated by looking at the influence of the neighboring amino acids but no significant conserved pattern of the flanking sites could be identified. The software tool was developed to predict the potential glycosylation sites in autotransporter protein, the virulence factors of gram negative bacteria, and our results revealed that the frequency of glycosylation sites was higher in adhesins (a subclass of autotransporters) relative to the other classes of autotransporters.

Keywords: Glycosylation, Bioinformatics, PglB, *Campylobacter jejuni*, Glycosyltransferase, Phylogenetics.

Background:

Glycosylation is the process of addition of a carbohydrate moiety to a protein molecule. Addition of a carbohydrate moiety to the side chain of an amino acid residue affects the physicochemical properties of that protein. Glycosylation process alters the different properties of proteins like proteolytic resistance, protein solubility, stability, local structure and immunogenicity **[1]**. Two important types of glycosylation are O-linked and N- linked Glycosylation. N-linked glycosylation is a co-translational process involving the transfer of a precursor oligosaccharide to asparagine residues in a sequon Asn-X-Ser/Thr (N-X-S/T), where X is any amino acid other than Proline **[2]**. This is however, not a specific consensus, since not all such sequins are modified in the cell. O-linked glycosylation involves the post-translational transfer of an oligosaccharide to a serine or threonine residue **[3]**. A large number of proteins contain the N-X-S/T motif and thus could be potential glycoproteins **[4]**. A protein can have a number of these tri-peptide motifs but only a few of them or selected ones get glycosylated due to structural constraints. Glycosylation was earlier thought to be restricted only to eukaryotes but now both O- and N-linked glycosylation pathways have been studied in detail in prokaryotes particularly in *Campylobacter jejuni* **[5]**.

Protein glycosylation is widespread in prokaryotes, with more than 70 bacterial glycoproteins reported so far [6]. Most of these are surface or secreted proteins that affect how bacteria interact with their environment, for instance, by influencing cell-cell interactions, surface adhesions, or by evasion of immune response **[6, 7]**. The well-known glycoproteins of *E. coli* present several similarities: (a) are secreted as autotransporters, which represent a branch of the type V secretion pathway. Autotransporter proteins belong to the family of outer membrane or secreted proteins. They are involved in virulence functions such as adhesion, aggregation, invasion, biofilm formation and toxication. (b) Have nearly identical N-terminal 19-amino-acid repeats; (c) are glycosylated by the addition of heptoses mediated by single glycosyltransferases that are functionally interchangeable; and (d) are versatile virulence factors mediating bacterial autoaggregation and biofilm formation as well as adhesion to and invasion of mammalian cells. Because of these similarities, enzymes like AIDA-I, TibA, and Ag43 have been named self-associating autotransporters (SAAT) **[8]**. In this study we predicted the number of glycosylation sites in a large group of autotransporter proteins of the bacterial pathogens. This may have implications in terms of their virulence and hence overall pathogenecity of the host bacteria possessing them.

Glycosylation in Bacteria:

Studying glycosylation in relatively less-complicated bacterial systems, such as mucosal associated pathogens, provides the opportunity to exploit glycoprotein biosynthetic pathways. For example, *C. jejuni* has been established as an excellent model for an N-linked glycosylation pathway in bacteria, with the activities of the characterized *pgl* (protein glycosylation) gene cluster assembling and transferring a known heptasaccharide from a membrane-anchored undecaprenylpyrophosphate-linked donor to an asparagines residue in proteins at the classic Asn-X-Ser/Thr motif **[1]**. There is strong evidence for the presence of a conserved glycosylation operon known as *pgl* in *Campylobacter jejuni* and many other bacteria. Proteins encoded by the *pgl* locus are capable of carrying out functions ranging from the synthesis of structural components to the functional molecules, i.e. carbohydrate moieties and enzymes respectively, involved in the cascade of glycosylation. Glycosylation pathway in *C. jejuni* is encoded by the *pgl* gene cluster. One protein from this cluster, PglB is considered to be the oligosaccharyl transferase (OST) due to its homology with the Sttp3 protein, which is a subunit of yeast OST complex. N-linked glycosylation is a very common post-translational modification in eukaryotes **[6]**. PglB and Sttp3 both have a conserved signature motif 'WWDYG' which has been shown to be essential for activity in vivo.

PglB comprises of 10-12 predicted transmembrane domains and a small C-terminal periplasmic domain. It has also been shown that unlike the eukaryotic OSTs, PglB is capable of transferring the heptasaccharide to the asparagine side chain of fully folded acceptor proteins *in vitro* as well as in periplasm **[9]**. The optimal glycosylation consensus sequence for PglB is 'DQNAT' although additional binding determinants and local peptide confirmations are also likely to affect glycosylation efficiency in full-length proteins **[10]**. Furthermore, PglB has substrate flexibility and can accept multiple peptide substrates. In contrast to the eukaryotic N-linked glycosylation, the enzymes from *C. jejuni* can be readily over expressed in functional form in *E. coli*, which makes them more amenable to biochemical studies and hence in better understanding of the functioning of the enzyme **[11]**.

Glycosylation in Archaea:

AglB is a homologue of the bacterial PglB in Archaea. The understanding of genetic pathways for the assembly and attachment of N-linked glycans in bacterial systems far outweighs the knowledge of comparable processes in archaea. Recent characterization of a novel trisaccharide [b-ManpNAcA6Thr-(1-4)-b-GlcpNAc3NAcA-(1-3)-b-GlcpNAc] N-linked to asparagine residues in *Methanococcus voltae* flagellin and S-layer proteins affords new opportunities to investigate N-linked glycosylation pathways in archaea **[10]**.

Methodology:

Protein sequence of *pglB* gene from *Campylobacter jejuni* (Accession number: Q9S4V7) was retrieved from UniProt **[12]** and then Basic Local Alignment Search Tool (BLAST) **[13]** was used to search for its homologs in archaeal and bacterial proteomes. First PglB was searched against the bacterial resource using BLOSUM62 as the scoring matrix and at an expect threshold of 0.1 with search optimized to report only the best 100 hits. Out of the 63 hits, 20 hits were short listed for further analysis based on the following parameters, which are their functional description (i.e. glycosylation), percentage identity and the presence of motif WWDYG.

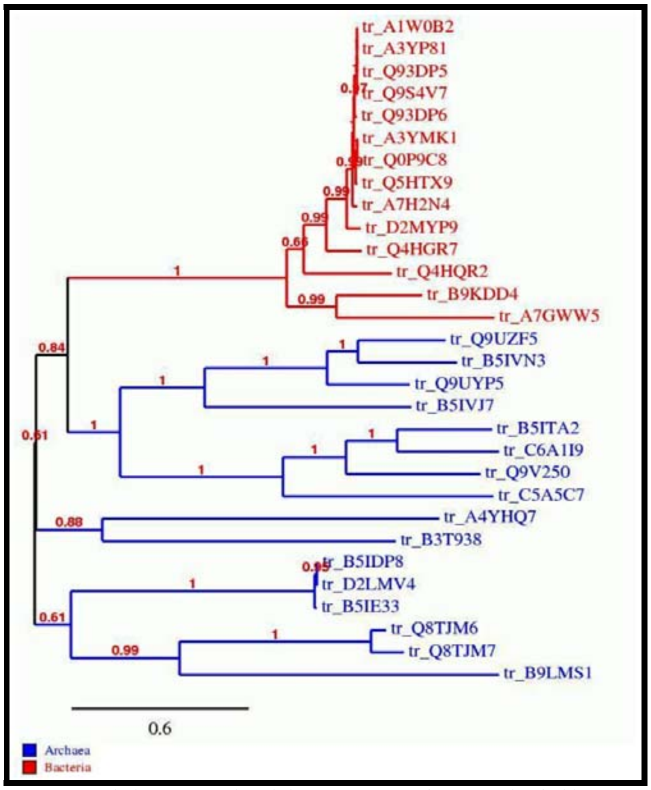


Figure 1: Phylogram produced by Mr. Bayes for a total of thirty (30) glycosyltransferases from archaea (16) and *C. jejuni* (14)

Search for PglB homologs against the archaeal database yielded 37 hits and 16 were short-listed based on the above-mentioned parameters. Search against the archaeal database was carried out using adjusted parameters such as setting the expect threshold to 10.0 in order to accommodate more results and search optimized to report only the best 100 hits. Selected sequences were given as input to generate the Multiple Sequence alignment by MUSCLE **[14]**. Alignment file generated by MUSCLE was given as input to the BioNJ **[15]** program, ran at 1000 bootstraps, to compute a distance tree to represent the phylogenetic relationships between bacterial and archaeal enzymes. Trees were visualized using TreeDyn **[16]**. Mr. Bayes was also used to further confirm the results **[17]** **(Figure 1)**. To test protein models for the prediction of glycosylation sites, seven hundred autotransporter proteins were retrieved from NCBI **[18]**. These proteins were classified into five main classes on the basis of their functional domains using Pfam **[19]**. These classes are adhesins, lipolytic, proteases, toxins, and SPATES.

Prediction of Glycosylation Sites:

For N-linked and O-linked glycosylation, a signal peptide is needed in the target protein. We used two online glycosylation site prediction servers i.e. NetOGlyc 3.1 for the prediction of O-linked glycosylation sites **[20]**, and NetNGlyc1.0 **[21]** for the prediction of N-linked glycosylation sites. It was also used for the prokaryotic proteins, due to the reason that N-glycosylation machinery in prokaryotes resembles to that of eukaryotes **[22]**.

Development of Database:

The database of PglB protein sequence was developed in MySQL. MySQL is currently the most popular open source database server in existence. On top of that, it is very commonly used in conjunction with PHP scripts to create powerful and dynamic server-side applications.

Software Development:

A tool for finding specific motifs in glycosyltransferases was developed using PHP as the scripting language in Macromedia Dream weaver. This tool is named Oligosaccharyl Transferase Prediction Tool (OTPT) version 1.0 and it can check for the presence of specific motifs in query sequences.

Results and Discussion:

Protein sequences of PglB from *C. jejuni* and its archaeal homologs were stored in a database which carries information regarding the accession numbers, length of proteins, their names and records of sequence/motif of five amino acids essential for N-linked glycosylation. All the protein sequences contain common motifs with few exceptions in archaeal proteins where there was a single amino acid change from Y→N, Y→F, Y→Q and Y→W. OTPT can detect such motifs.

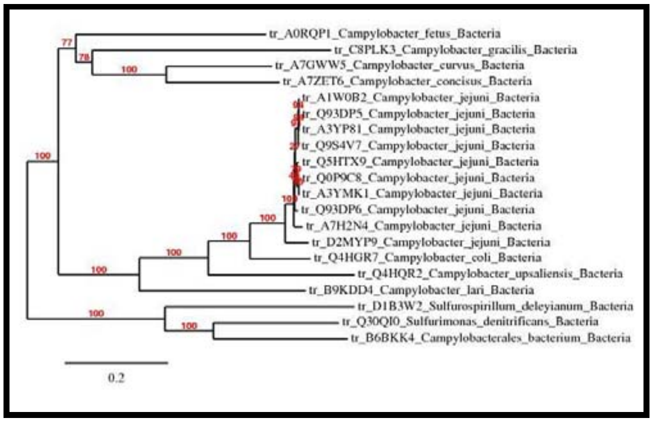

77
tr_A0RQP1_Campylobacter_fetus_Bacteria
78
tr_C8PLK3_Campylobacter_gracilis_Bacteria
100
tr_A7GWW5_Campylobacter_curvus_Bacteria
tr_A7ZET6_Campylobacter_concisus_Bacteria
tr_A1W0B2_Campylobacter_jejuni_Bacteria
tr_Q93DP5_Campylobacter_jejuni_Bacteria
tr_A3YP81_Campylobacter_jejuni_Bacteria
tr_Q9S4V7_Campylobacter_jejuni_Bacteria
tr_Q5HTX9_Campylobacter_jejuni_Bacteria
tr_Q0P9C8_Campylobacter_jejuni_Bacteria
tr_A3YMK1_Campylobacter_jejuni_Bacteria
100
tr_Q93DP6_Campylobacter_jejuni_Bacteria
100
tr_A7H2N4_Campylobacter_jejuni_Bacteria
100
tr_D2MYP9_Campylobacter_jejuni_Bacteria
100
tr_Q4HGR7_Campylobacter_coli_Bacteria
100
tr_Q4HQR2_Campylobacter_upsaliensis_Bacteria
100
tr_B9KDD4_Campylobacter_lari_Bacteria
100
tr_D1B3W2_Sulfurospirillum_deleyianum_Bacteria
100
tr_Q30QI0_Sulfurimonas_denitrificans_Bacteria
tr_B6BKK4_Campylobacterales_bacterium_Bacteria
0.2

Figure 2: Phylogram of bacterial glycosyltransferases drawn by BioNJ run at 1000 bootstraps. Close clustering of bacterial enzymes could be seen.

Phylogenetic Analysis of PglB Sequences:

First objective of this study was to analyze the evolutionary relationships among the OSTs from archaeal and bacterial Proteins. Twenty bacterial and sixteen archaeal enzymes were used for phylogenetic studies. The bacterial tree shows close grouping of bacterial enzymes **(Figure 2)** but the archaeal tree shows clustering of archaeal enzymes into four sub-clusters probably indicative of relatively low level of sequence homology but still sharing common functionality **(Figure 3)**. All 36 sequences were then used to make the final phylogenetic tree. BioNJ was used to draw the phylogenetic tree (ran at 1000 bootstraps). Trees produced were then visualized using TreeDyn and represented in a phylogram **(Figure 4)**. Phylogenetic tree differentiates the bacterial and archaeal enzymes in two separate clusters.

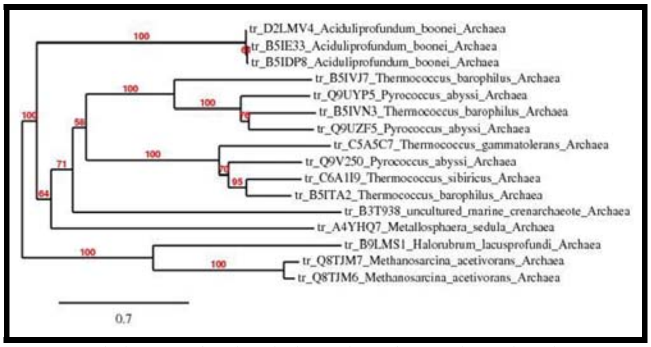

tr_D2LMV4_Aciduliprofundum_boonei_Archaea
tr_B5IE33_Aciduliprofundum_boonei_Archaea
tr_B5IDP8_Aciduliprofundum_boonei_Archaea
tr_B5IVJ7_Thermococcus_barophilus_Archaea
tr_Q9UYP5_Pyrococcus_abyssi_Archaea
tr_B5IVN3_Thermococcus_barophilus_Archaea
tr_Q9UZF5_Pyrococcus_abyssi_Archaea
tr_C5A5C7_Thermococcus_gammatolerans_Archaea
tr_Q9V250_Pyrococcus_abyssi_Archaea
tr_C6A1I9_Thermococcus_sibiricus_Archaea
tr_B5ITA2_Thermococcus_barophilus_Archaea
tr_B3T938_uncultured_marine_crenarchaeote_Archaea
tr_A4YHQ7_Metallosphaera_sedula_Archaea
tr_B9LMS1_Halorubrum_lacusprofundi_Archaea
tr_Q8TJM7_Methanosarcina_acetivorans_Archaea
tr_Q8TJM6_Methanosarcina_acetivorans_Archaea
100
100
100
100
58
71
64
100
70
95
100
100
0.7

Figure 3: Phylogram of archaeal glycosyltransferases drawn by BioNJ run at 1000 bootstraps. Four clusters could be seen among the archaeal enzymes probably suggestive of low level of sequence homology within.

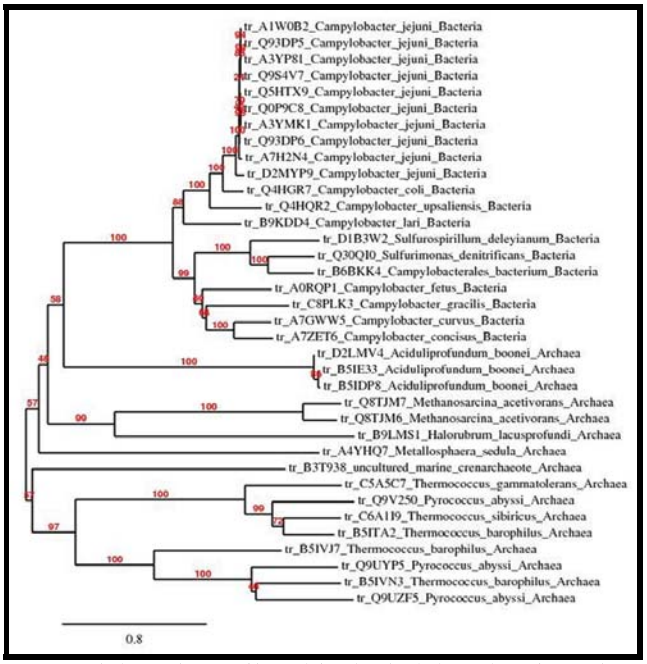

tr_A1W0B2_Campylobacter_jejuni_Bacteria
tr_Q93DP5_Campylobacter_jejuni_Bacteria
tr_A3YP81_Campylobacter_jejuni_Bacteria
tr_Q9S4V7_Campylobacter_jejuni_Bacteria
tr_Q5HTX9_Campylobacter_jejuni_Bacteria
tr_Q0P9C8_Campylobacter_jejuni_Bacteria
tr_A3YMK1_Campylobacter_jejuni_Bacteria
tr_Q93DP6_Campylobacter_jejuni_Bacteria
tr_A7H2N4_Campylobacter_jejuni_Bacteria
tr_D2MYP9_Campylobacter_jejuni_Bacteria
tr_Q4HGR7_Campylobacter_coli_Bacteria
tr_Q4HQR2_Campylobacter_upsaliensis_Bacteria
tr_B9KDD4_Campylobacter_lari_Bacteria
tr_D1B3W2_Sulfurospirillum_deleyianum_Bacteria
tr_Q30QI0_Sulfurimonas_denitrificans_Bacteria
tr_B6BKK4_Campylobacterales_bacterium_Bacteria
tr_A0RQP1_Campylobacter_fetus_Bacteria
tr_C8PLK3_Campylobacter_gracilis_Bacteria
tr_A7GWW5_Campylobacter_curvus_Bacteria
tr_A7ZET6_Campylobacter_concisus_Bacteria
tr_D2LMV4_Aciduliprofundum_boonei_Archaea
tr_B5IE33_Aciduliprofundum_boonei_Archaea
tr_B5IDP8_Aciduliprofundum_boonei_Archaea
tr_Q8TJM7_Methanosarcina_acetivorans_Archaea
tr_Q8TJM6_Methanosarcina_acetivorans_Archaea
tr_B9LMS1_Halorubrum_lacusprofundi_Archaea
tr_A4YHQ7_Metallosphaera_sedula_Archaea
tr_B3T938_uncultured_marine_crenarchaeote_Archaea
tr_C5A5C7_Thermococcus_gammatolerans_Archaea
tr_Q9V250_Pyrococcus_abyssi_Archaea
tr_C6A1I9_Thermococcus_sibiricus_Archaea
tr_B5ITA2_Thermococcus_barophilus_Archaea
tr_B5IVJ7_Thermococcus_barophilus_Archaea
tr_Q9UYP5_Pyrococcus_abyssi_Archaea
tr_B5IVN3_Thermococcus_barophilus_Archaea
tr_Q9UZF5_Pyrococcus_abyssi_Archaea
100
100
100
100
100
88
100
99
100
58
100
46
57
99
100
97
100
99
100
100
0.8

Figure 4: Phylogenetic tree of bacterial and archael glycosyltransferases drawn by BioNJ at 1000 bootstraps. Bacterial enzymes at the top and archaeal enzymes at the bottom.

Retrieval of Autotransporter Proteins:

About seven hundred autotransporter protein sequences were retrieved and classified by Pfam and then used to check for the presence of glycosylation sites. In total, 366 autotransporter proteins were classified, out of which 298 were Adhesins, 37 IgA1 proteases, 11 Lipolytic, 16 SPATES and 4 Toxins. AIDA-I precursor, outer membrane autotransporter, lipase/esterase; EstA are some examples of these autotransporters. Our results indicate that among the autotransporters, members of adhesins subfamily are more frequently glycosylated compared to members of other subfamilies in the group. Although there is a clear bias in the number of adhesin proteins being the largest group of the five classes of autotransporters but they are the only group in which potential N-glycosylation sites could be detected (109 in total). No N-glycosylation sites could be detected in rest of the four classes of autotransporters. From the results of glycosylation site prediction, we inferred a contradictory result that many of the autotransporter proteins without conventional signal peptide are also glycosylated **(Table 1 see Supplementary material)**.

Glycosylation Site Prediction Results:

O-glycosylation occurs more frequently in autotransporters. The reason is that, autotransporter proteins are membrane proteins and they are present on the surface. It is observed relatively more in adhesins and lipolytic class of autotransporter proteins. While the results of NetNGlyc show that there is a low level of N-glycosylation in autotransporter proteins because mostly it lacks the signal peptide for N-glycosylation except for the adhesins class **(Table 1)**. This higher number is probably because the conserved motif for N-Linked glycosylation occurs more frequently in adhesins than in any other class of autotransporter proteins. Graph shows occurrence of both types of glycosylation in adhesins **(Figure 5)**. Furthermore, the archaeal and bacterial glycosyltransferases were aligned using MUSCLE in order to search for a conserved pattern of amino acids in the flanking sequence of N-X-S/T motif but it is clear from the output of MUSCLE that no such significant pattern could be identified which is common across bacteria and archaea except for a conserved tyrosine residue **(Figure 6)**.

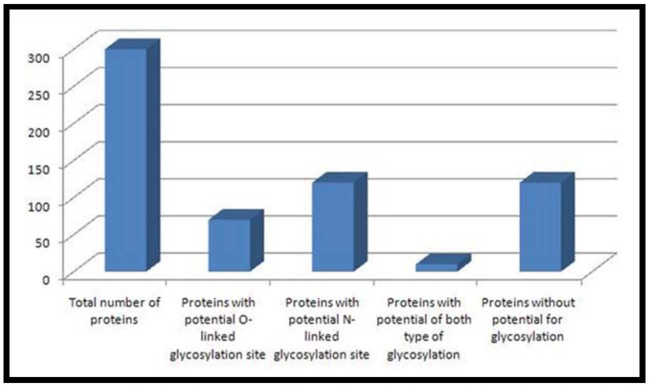
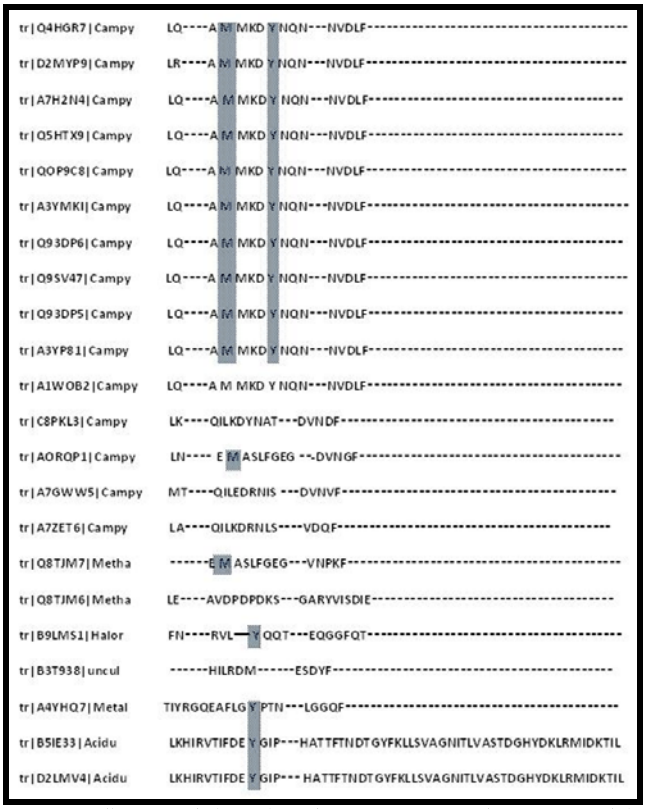


Figure 5: Graph showing potential number of glycosylation sites in adhesin class of autotransporter proteins



```
tr|Q4HGR7|Campy    LQ----A M MKD Y NQN---NVDLF----------------------------------------
tr|D2MYP9|Campy    LR----A M MKD Y NQN---NVDLF----------------------------------------
tr|A7H2N4|Campy    LQ----A M MKD Y NQN---NVDLF----------------------------------------
tr|Q5HTX9|Campy    LQ----A M MKD Y NQN---NVDLF----------------------------------------
tr|QOP9C8|Campy    LQ----A M MKD Y NQN---NVDLF----------------------------------------
tr|A3YMKI|Campy    LQ----A M MKD Y NQN---NVDLF----------------------------------------
tr|Q93DP6|Campy    LQ----A M MKD Y NQN---NVDLF----------------------------------------
tr|Q9SV47|Campy    LQ----A M MKD Y NQN---NVDLF----------------------------------------
tr|Q93DP5|Campy    LQ----A M MKD Y NQN---NVDLF----------------------------------------
tr|A3YP81|Campy    LQ----A M MKD Y NQN---NVDLF----------------------------------------
tr|A1WOB2|Campy    LQ----A M MKD Y NQN---NVDLF----------------------------------------
tr|C8PKL3|Campy    LK----QILKDYNAT---DVNDF----------------------------------------
tr|AORQP1|Campy    LN---- E M ASLFGEG --DVNGF----------------------------------------
tr|A7GWW5|Campy    MT----QILEDRNIS ---DVNVF----------------------------------------
tr|A7ZET6|Campy    LA----QILKDRNLS----VDQF----------------------------------------
tr|Q8TJM7|Metha    ------E M ASLFGEG---VNPKF----------------------------------------
tr|Q8TJM6|Metha    LE----AVDPDPDKS---GARYVISDIE----------------------------------------
tr|B9LMS1|Halor    FN----RVL— Y QQT---EQGGFQT----------------------------------------
tr|B3T938|uncul    ------HILRDM------ESDYF----------------------------------------
tr|A4YHQ7|Metal    TIYRGQEAFLG Y PTN---LGGQF----------------------------------------
tr|B5IE33|Acidu    LKHIRVTIFDE Y GIP---HATTFTNDTGYFKLLSVAGNITLVASTDGHYDKLRMIDKTIL
tr|D2LMV4|Acidu    LKHIRVTIFDE Y GIP--- HATTFTNDTGYFKLLSVAGNITLVASTDGHYDKLRMIDKTIL
```

Figure 6: Alignment of glycosyltransferases. No conserved pattern around the N-X-S/T motif detected except for a conserved tyrosine residue.

Tool Development:

As explained earlier, WWDYG is a specific motif which is important for the proper functioning of OSTs. Software for the identification of this conserved motif in glycosyltransferases was developed using PHP. We tested a total of 90 proteins of Archaea and Bacteria to detect motifs (NXS/T and WWDYG) in protein sequences. Results were saved into the database of PglB.

Conclusion:

In this study, we compared and analyzed the oligosaccharyl transferases from different species of bacteria and archaea. Phylogenetic analysis of the bacterial enzymes revealed close grouping of bacterial enzymes within bacteria whereas the archaeal enzymes reflected less sequence homology even though functionally they are same. It was also noticed that proteins from several species share a conserved motif essential for N-linked glycosylation especially some of the bacterial species i.e. Acceptor site "NXS/T" and Oligosaccharyl transferases "WWDYG". Software was developed for the efficient searching of these motifs in a given query sequence. Low sequence conservation was observed in the flanking sequences of these potential glycosylation sites (N-X-S/T) when multiple glycosyltransferases were aligned using MUSCLE. These protein sequences of PglB and its homologs were stored in a database i.e., "Database of PglBs". Prediction of glycosylation sites in autotransporter proteins was done using online tools. Although the presence of a signal peptide is necessary for glycosylation but even in the absence of signal peptide glycosylation sites were observed in autotransporters. Among the five classes of autotransporters, adhesins showed significant higher glycosylation frequency than the other classes of autotransporter proteins. More analyses on glycosyltransferases would help our understanding of these pathways and help direct new ways for further research. The outcome of the study may have implications in understanding the roles of the oligosaccharyl transferase and redefining the virulence potential of each adhesin.

Acknowledgements:
Prof. Brendan W. Wren, Head of the Dept of Pathogen Molecular Biology, LSHTM, for his invaluable guidance and Department of Biosciences for providing necessary facilities to carry out our work.

References:

[1] Gupta R & Brunak S. *Pac Symp Biocomput*. 2002 310-22 [PMID: 11928486]
[2] Yan A & Lennarz WJ. *J Biol Chem.* 2005 **280**(5): 3121 [PMID: 15590627]
[3] Van den Steen P *et al*. *Crit Rev Biochem Mol Biol*. 1998 **33**: 151 [PMID: 9673446]
[4] Apweiler R *et al*. *Biochim Biophys Acta.* 1999 **1473**: 4 [PMID: 10580125]
[5] Szymanski CM *et al*. *Trends Microbiol.* 2003 **11**: 233 [PMID: 12781527]
[6] Szymanski CM & Wren BW. *Nat Rev Microbiol*. 2005 **3**: 225 [PMID: 15738950]
[7] Schmidt MA *et al*. *Trends Microbiol.* 2003 **11**: 554 [PMID: 14659687]
[8] Klemm P *et al*. *Int J Med Microbiol.* 2006 **296**: 187 [PMID: 16600681]
[9] Kowarik M *et al*. *Science* 2006 **314**: 1148 [PMID: 17110579]
[10] Kowarik M *et al*. *EMBO J.* 2006 **25**: 1957 [PMID: 16619027]
[11] Glover KJ *et al*. *Chem Biol.* 2005 **12**: 1311 [PMID: 16356848]
[12] http://www.uniprot.org.
[13] http://www.expasy.org/tools/blast/
[14] http://www.phylogeny.fr/version2_cgi/one_task.cgi?task_type=muscle
[15] http://www.phylogeny.fr/version2_cgi/one_task.cgi?task_type=bionj
[16] http://www.phylogeny.fr/version2_cgi/one_task.cgi?task_type=treedyn
[17] http://www.phylogeny.fr/version2_cgi/one_task.cgi?task_type=mrbayes
[18] http://www.ncbi.nlm.nih.gov/
[19] http://pfam.sanger.ac.uk/
[20] http://www.cbs.dtu.dk/services/NetOGlyc/
[21] http://www.cbs.dtu.dk/services/NetNGlyc/
[22] Altschul SF *et al*. *J Mol Biol.* 1990 **215**: 403 [PMID: 2231712]

Edited by P Kangueane
Citation: Tabish *et al*. Bioinformation 6(5): 191-195 (2011)

Supplementary material:

Table 1: Results of glycosylation site predictions among autotransporter proteins.

Protein Family	Class of Protein	Total number of proteins	Proteins with potential O-linked glycosylation site	Proteins with potential N-linked glycosylation site	Proteins with potential of both type of glycosylation	Proteins without potential for glycosylation
Autotransporter Proteins	Adhesion	298	61	109	8	120
	IgA1-protease	37	3	0	7	27
	Lipolytic	11	8	0	1	2
	SPATES	16	1	0	0	15
	Toxins	4	1	0	0	3