

DISS. ETH NO. 19586

**N-linked protein glycosylation in *Escherichia coli* -  
A system for basic and applied research**

A dissertation submitted to

ETH ZURICH

for the degree of  
Doctor of Sciences

presented by

CHRISTIAN ANDREAS LIZAK

M.Sc. Biochemistry, Technische Universität München

born on November 6, 1979  
citizen of Germany

accepted on the recommendation of

Prof. Dr. Markus Aebi, examiner  
Prof. Dr. Andreas Plückthun, co-examiner  
Prof. Dr. Dario Neri, co-examiner  
PD. Dr. Martin Bachmann, co-examiner

2011

## Summary

This thesis focuses on the analysis of the N-linked protein glycosylation pathway of bacteria. By taking advantage of glycosylation competent *Escherichia coli* cells, distinct steps in the biosynthesis of glycoproteins have been investigated.

Asparagine-linked glycosylation is a post-translational modification where a preassembled oligosaccharide is covalently attached to proteins that contain a conserved sequence motif (Asn-Xaa-Ser/Thr). This kind of modification affects more than 50% of all eukaryotic proteins and is implicated in various processes such as protein folding and quality control or host-pathogen interactions. Studies of prokaryotes have revealed that N-linked protein glycosylation frequently occurs in archaea, but it also exists in some bacteria.

Chapter 1 describes the main steps in the biosynthesis of glycoproteins in eukaryotes and bacteria. A comparison of both systems allows the identification of general principles of N-linked protein glycosylation, but also highlights the differences arising from their physiological function. With regard to a comparative analysis of oligosaccharyltransferase (OST), the central enzyme in this pathway, general aspects about substrate specificity and catalysis are reviewed, and an evolutionary development of OST is suggested.

N-linked protein glycosylation in bacteria is well-described for the human pathogen *Campylobacter jejuni*, but other bacteria encoding a general N-linked glycosylation system have been reported. In chapter 2, the N-glycosylation system of the non-pathogenic rumen symbiont *Wolinella succinogenes* is characterized. The transfer of this N-glycosylation system into *E. coli* shows the biosynthesis of a glycan with significant structural deviations from the heptasaccharide of *C. jejuni*. This glycan increases the repertoire of oligosaccharides synthesized in *E. coli*, and further analyses of this novel carbohydrate structure might provide insights into the natural function of N-glycans in bacteria.

Chapter 3 implements the exploitation of the protein glycosylation system of *C. jejuni* as a strategy to produce glycoproteins in *E. coli*. Studies on a glycoengineered single-chain antibody fragment show the capability of the system for quantitative production of homogeneous glycoprotein. The analysis reveals an improved stability and an increased solubility of the glycosylated antibody fragment and suggests bacterial N-linked glycosylation as a possibility to specifically alter the biophysical properties of proteins.

The glycosylation reaction is catalyzed by OST, a hetero-oligomeric membrane protein complex in eukaryotes, where STT3 represents the central, catalytic enzyme. Homologues of STT3 can be found in bacteria and archaea, and the study of these single-subunit OSTs

allows addressing the mechanism of N-linked glycosylation in a less complex system. Chapter 4 illustrates a strategy for expression, purification and crystallization of single-subunit OSTs from bacteria. In chapter 5 the X-ray structure of a bacterial OST, the PglB protein from *Campylobacter lari*, in complex with an acceptor peptide is described. The structure provides the molecular basis of sequon recognition and reveals a catalytic site that is formed by the transmembrane domain of the protein. The results suggest a mechanism for amide nitrogen activation and provide opportunities for engineering of N-glycosylation in bacteria.

## Zusammenfassung

Die vorliegende Arbeit befasst sich mit der Analyse der N-Glykosylierung in Bakterien. Zur gezielten Untersuchung bestimmter Schritte in der Biosynthese von Glykoproteinen wurden *Escherichia coli* Zellen verwendet, die in der Lage sind N-Glykosylierung zu betreiben.

Die Asparagin gebundene Glykosylierung stellt eine posttranslationale Modifikation dar, in der ein vorgefertigtes Oligosaccharid kovalent an ein Protein gebunden wird, das über ein konserviertes Sequenzmotiv (Asn-X-Ser/Thr) verfügt. Diese Art der Modifikation betrifft über 50% aller eukaryotischen Proteine. Sie hat Auswirkungen auf eine Vielzahl von zellulären Prozessen, wie die Faltung und Qualitätskontrolle von Proteinen, oder die Interaktion zwischen Pathogenen und deren Wirtszellen. Untersuchungen an Prokaryoten haben gezeigt, dass die N-gebundene Glykosylierung in Archaeen weit verbreitet ist, aber dass sie auch in einigen Bakterien auftritt.

Kapitel 1 beschreibt die grundlegenden Schritte der Biosynthese von Glykoproteinen in Eukaryoten und in Bakterien. Ein Vergleich dieser beiden Systeme lässt die generellen Prinzipien der N-gebunden Protein Glykosylierung erkennen. Er hebt aber auch diejenigen Unterschiede hervor, die sich aus den unterschiedlichen physiologischen Funktionen ergeben. Eine Analyse der Oligosaccharyltransferase (OST), dem zentralen Enzym in diesem Prozess, liefert einen Überblick über die generellen Aspekte der Substrat Spezifität, sowie der katalytischen Reaktion. Zudem wird eine evolutionäre Entwicklung der OST vorgeschlagen.

Die N-Glykosylierung in Bakterien ist für das menschliche Pathogen *Campylobacter jejuni* umfassend beschrieben. Es gibt aber auch Berichte über andere Bakterien, die ein generelles System zur N-Glykosylierung in ihrem Genom tragen. In Kapitel 2 wird das N-Glykosylierungssystem des apathogenen Symbionten *Wolinella succinogenes*, der im Pansen von Rindern lebt, charakterisiert. Eine Übertragung dieses Systems in *E. coli* Zellen zeigt die Biosynthese eines N-Glykans, das signifikante, strukturelle Unterschiede zu dem Heptasaccharid von *C. jejuni* aufweist. Dieser Zucker erhöht das Repertoire an N-Glykanen, die in *E. coli* synthetisiert werden können. Eine weiterführende Analyse dieser neuen Kohlenhydrat Struktur kann zudem Einblicke in die natürliche Funktion von N-Glykanen in Bakterien liefern.

Kapitel 3 beschreibt eine Strategie, wie das System der Protein Glykosylierung von *C. jejuni* für die Herstellung von Glykoproteinen in *E. coli* verwendet werden kann. Untersuchungen an einem, für die Glykosylierung überarbeitetem, "single-chain" Antikörper Fragment zeigen, dass das System in der Lage ist, quantitative Mengen an homogen glykosyliertem Protein zu

produzieren. Eine Analyse des glykosylierten Antikörper Fragments lässt eine verbesserte Stabilität sowie eine erhöhte Löslichkeit erkennen. Die bakterielle N-Glykosylierung stellt somit eine Möglichkeit dar, die biophysikalischen Eigenschaften von Proteinen gezielt zu verändern.

Die Glykosylierungsreaktion wird von OST katalysiert. In Eukaryoten ist diese ein heterooligomerer Komplex von Membranproteinen, in dem STT3 die zentrale, katalytische Enzymeinheit darstellt. Homologe zu STT3 existieren in Bakterien und Archaeen. Untersuchungen dieser OSTs, die aus einer einzigen Untereinheit bestehen, ermöglichen es, sich mit dem Mechanismus der N-Glykosylierung in einem weniger komplexen System zu befassen. Kapitel 4 zeigt eine Strategie für die Expression, die Reinigung und das Kristallisieren von bakteriellen OSTs auf. In Kapitel 5 wird die Röntgenkristallstruktur einer bakteriellen OST, dem PglB Protein aus *Campylobacter lari*, im Komplex mit einem Akzeptor Peptid beschrieben. Diese Struktur liefert das molekulare Grundverständnis für die Erkennung von Sequenzmotiven in Glykosylierungsstellen und offenbart das katalytische Zentrum, welches von der Transmembrandomäne des Proteins gebildet wird. Aus diesen Erkenntnissen kann man einen Mechanismus für die Aktivierung des Amid Stickstoffs ableiten. Weiterhin liefert diese Struktur eine Möglichkeit, die N-Glykosylierung in Bakterien gezielt zu verändern.