ETH zürich

ETH Library

Structural and Mechanistic Insights into Eukaryotic N-Glycan Biosynthesis and Human Vitamin B12 Transport

Doctoral Thesis

Author(s):
Bloch, Joël S.

Publication date:
2020

Permanent link:
https://doi.org/https://doi.org/10.3929/ethz-b-000409204

Rights / license:
In Copyright - Non-Commercial Use Permitted

This page was generated automatically upon download from the ETH Zurich Research Collection.
For more information, please consult the Terms of use.

DISS. ETH NO. 26652

STRUCTURAL AND MECHANISTIC INSIGHTS INTO EUKARYOTIC N-GLYCAN BIOSYNTHESIS AND HUMAN VITAMIN B12 TRANSPORT

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

Presented by

JOËL SHARON BLOCH

MSc ETH in Biology, Major in Structural Biology and Biophysics, ETH Zurich

Born on 02.03.1989

Citizen of
Zürich ZH and of Germany

Accepted on the recommendation of

Prof. Dr. K. Locher, examiner
Prof. Dr. R. Glockshuber, co-examiner
Prof. Dr. R Dutzler, co-examiner

2020

Summary

Structural biology allows us to shed light onto the smallest processes of life. The genetic information that is stored in the DNA of every living cell holds the building plans for proteins: These tiny molecules, which are often only at the size of a few nanometers, are responsible most of the biochemical functions in life. By means of X-ray crystallography and cryo-EM we are able to determine the structures of these proteins at atomic resolution, which makes it possible to rationalize the biochemical functions that are carried out by these proteins.

The first part of this thesis describes structural and mechanistic studies on transcobalamin mediated vitamin B12 uptake in human cells. In humans, vitamin B12 is essential for DNA synthesis and for certain metabolic processes. Humans are dependent on oral uptake of vitamin B12 from animal sources or from supplements. After ingestion, the vitamin is transported into the bloodstream, via a multistep pathway involving different carrier- and transport proteins. Transcobalamin is a soluble two-domain protein that is secreted into the bloodstream and specifically recognizes and binds vitamin B12 by clamping it between its two binding domains. The transcobalamin B12 complex is captured by the cell surface receptor CD320 and the complex is taken up into the cell by receptor mediated endocytosis.

As transcobalamin binds to vitamin B12 at a virtually irreversibly high affinity, studying its binding to different ligands without denaturing of the protein poses a great difficulty. We have heterologously expressed the beta-domain of transcobalamin and co-crystallized it in the apo form, as well as in complex with vitamin B12, the B12 related molecule cobinamide and synthetic antivitamin B12, a compound that cannot be further processed by the cellular enzymes and is designated for clinical studies in which it is used to artificially induce vitamin B12 deficiency in humans or animals. Using X-ray crystallography, we have obtained structures at high-resolution which helped to rationalize the binding modes of the respective substrates. In addition, we could prove biochemically, that the two domains of transcobalamin only assemble upon binding of vitamin B12.

The second part of the thesis describes the structural and mechanistic studies on the ER-luminal glucosyltransferase ALG6: N-glycosylation is one of the most common post translational modifications and is found in all domains of life. In eukaryotes a lipid-linked oligosaccharide is generated by the stepwise enzymatic addition of hexoses

onto a membrane embedded dolichylphosphate precursor molecule. In the lumen of the endoplasmic reticulum, the oligosaccharide is then transferred onto an acceptor peptide of a secreted- or membrane protein by oligosaccharyltransferase. The glycan helps in protein folding, quality control and after diversification in the Golgi as an identification tag of the respective cell or protein. The lipid-linked oligosaccharide biosynthesis starts in the cytoplasm and after the addition of one phosphate, two N-acetylglucosamine molecules and five mannoses, the lipid linked oligosaccharide is flipped into the ER-lumen. There the six membrane integral glycosyltransferases ALG3, ALG9, ALG12, ALG6, ALG8 and ALG10 add four additional mannoses and three glucoses in a step wise fashion, leading to a final Dol-PP-$GlcNAc_2Man_9Glc_3$ molecule. I have been studying the structure and mechanism of ALG6, which adds the first glucose to the LLO.

Building on the previously established chemoenzymatic synthesis of Dol-PP-$GlcNAc_2Man_5$, we used purified enzymes ALG3, ALG9 and ALG12 and synthetic substrate molecules to reconstitute a part the pathway of ER-luminal enzymatic LLO synthesis. We obtained pure Dol-PP-$GlcNAc_2Man_9$, which is the acceptor substrate of ALG6. In addition, we were able to generate every substrate intermediate of the enzymatic pathway. With the help of the Dol-PP-$GlcNAc_2Man_9$ and synthetic Dol-P-Glc donor substrate we could demonstrate the function of ALG6.

We reconstituted ALG6 into lipid nanodisc and with the help of a Fab binder from a synthetic library we were able to obtain a structure of ALG6 at 3.0 Å resolution by cryo-EM. An additional structure at 3.9 Å was determined with the protein reconstituted in detergent and with the synthetic donor-substrate analog Dol_{25}-P-Glc bound to it. The structure of the substrate-bound enzyme in combination with site directed mutagenesis helped to identify the catalytic residue Asp69, which is likely to act as a general base in ALG6 and which hints to a possible catalytic mechanism. ALG6 is a member of the GT-C superfamily. The GT-C superfamily categorizes membrane integral glycosyltransferases which require isoprenoid-inked sugar donor substrates. Structurally, the GT-C superfamily is only poorly defined. By comparing our structure of ALG6 to other published structures of GT-C enzymes, we were able to identify common motifs in GT-C enzymes and we observed an evolutionarily conserved modular design of GT-C enzymes.

Our described method for the chemo-enzymatic synthesis of eukaryotic LLOs and LLO intermediates provides and unprecedent tool to study enzymes involved in LLO synthesis as well as oligosaccharyltransferases.

Our structural findings not only explain the function of ALG6 and give therefore insight into ER-based LLO biosynthesis, but also, they help to rationalize the architecture and mechanism of GT-C enzymes in general.

With ALG6 being 64 kDa in size and the complexed Fab 48 kDa, the structure of ALG6:6AG9-Fab complex at 3.0 Å resolution, is at this time reportedly the highest resolved structure of a membrane-bound protein complex resolved by cryo-EM. We therefore also demonstrate the usefulness of Fab technology for the study of small membrane proteins and describe a useful tool for further studies of e.g. small membrane integral glycosyltransferases.

Zusammenfassung

Strukturbiologie hilft uns die kleinsten Prozesse des Lebens zu erklären. Die genetische Information, die in der DNS, jeder lebenden Zelle aufbewahrt wird enthält die Baupläne für Proteine. Diese Makromoleküle, sind für die meisten biochemischen Funktionen in der belebten Welt verantwortlich. Röntgenkristallographie und Kryo-Elektronenmikroskopie ermöglichen es uns, die dreidimensionalen Strukturen dieser Proteine mit atomarer Auflösung aufzuklären. Dies ermöglicht es, die biochemischen Funktionen dieser Proteine zu erklären.

Der erste Teil dieser Dissertation beschreibt strukturelle und mechanistische Studien zur Transcobalamin-abhängigen Aufnahme von Vitamin B12 in menschlichen Zellen. Vitamin B12 ist essenziell für die DNS Synthese und für gewisse metabolische Prozesse im Menschen. Menschen, sind auf die Aufnahme von Vitamin B12 aus tierischen Quellen oder aus Nahrungsergänzungsmitteln angewiesen. Nach der Nahrungsaufnahme gelangt das Vitamin B12 über mehrere Schritte, durch verschiedene Transportproteine in die Blutbahn. Transcobalamin ist ein lösliches Zweidomänenprotein, welches in die Blutbahn sekretiert wird. Es bindet spezifisch an Vitamin B12 indem es das Vitamin, zwischen den beiden Proteindomänen einklemmt. Der Komplex von Transcobalamin und Vitamin B12 wird dann von dem Zelloberflächenrezeptorprotein CD320 gebunden und mit Hilfe von Clathrin-vermittelter Endozytose in die Zelle aufgenommen.

Da Transcobalamin mit einer nahezu unendlich hohen Affinität an Vitamin B12 bindet, stellt das Studieren der Bindungsinteraktionen von Transcobalamin mit anderen Liganden, ohne dafür Transcobalamin denaturieren zu müssen, eine grosse Hürde dar. Um diese Hürde zu Überschreiten haben wir die beta-Domäne von Transcobalamin heterolog exprimiert und in der Abwesenheit eines Liganden, sowie als Komplex mit Vitamin B12, dem verwandten Molekül Cobinamid, sowie synthetischem Antivitamin B12 kristallisiert. Antivitamin B12 ist eine synthetisch veränderte Form von Vitamin B12, welche von der Zellmaschinerie nicht weiterverarbeitet werden kann. Es dient dazu für klinische Studien in Patienten oder Tieren kurzzeitig einen künstlichen Vitamin B12 Mangel zu erzeugen. Mit Hilfe von hochauflösender Röntgenkristallographie konnten wir die Strukturen der jeweiligen Komplexe ermitteln, welche uns Aufschluss über die jeweiligen Substratbindungsmodi von Transcobalamin gaben. Zusätzlich gelang es uns biochemisch aufzuzeigen, dass

sich die beiden Domänen von Transcobalamin nur durch die Zugabe von Vitamin B12 zusammenfügen.

Der Zweite Teil dieser Dissertation beschreibt strukturelle und mechanistische Studien an der ER-luminalen Glukosyltransferase ALG6: N-Glykosilierung ist eine der am häufigsten vorkommenden posttranslationalen Proteinmodifikationen und kann in allen Domänen des Lebens gefunden werden. In Eukaryonten wird ein Lipid-gebundenes Oligosaccharid durch das enzymkatalysierte, schrittweise Hinzufügen von Hexosen an einen Membrangebundenen Dolichylphosphat-Vorläufer generiert. Im Lumen des Endoplasmatischen Retikulum (ER) wird das Oligosaccharid dann durch die Oligosaccharyltransferase an eine Asparagin-Seitenkette eines sekretorischen Proteins oder eines Membranproteins gehängt. Das N-Glykan hilft dem Protein bei der Proteinfaltung und ist an der Qualitätskontrolle der Faltung vor dem Weitertransport zum Golgi-Aparatus beteiligt. Nach einer ausgiebigen Diversifizierung durch eine Vielzahl von Glukosidasen und Glykosyltransferasen im Golgi erreicht das Glykan seine endgültige Zelltyp- und Glykoprotein-spezifische kovalente Struktur. Die Biosynthese des lipidgebundenen Oligosaccharids beginnt im Zytoplasma. Nach dem Anfügen einer Phosphatgruppe, zwei N-Acetylglucosamin Molekülen und fünf Mannosen an Dolicholphosphat wird das lipidgebundene Oligosaccharid ins Lumen des ER übergeführt. Von dort aus fügen die sechs membranintegralen Glykosyltransferasen ALG3, ALG9, ALG12, ALG6, ALG8 und ALG10 schrittweise, gesamthaft vier Mannosen und drei Glukosen an, was zum finalen Dol-PP-$GlcNAc_2Man_9Glc_3$ Molekül führt. In dieser Doktorarbeit habe ich mich mit der Struktur und dem enzymatischen Mechanismus von ALG6 befasst, welches die erste Glukose an das Oligosaccharid anfügt.

Aufbauend auf der bereits etablierten, chemoenzymatischen Synthese von Dol-PP-$GlcNAc_2Man_5$ haben wir mit Hilfe der aufgereinigten Enzyme ALG3, ALG9 und ALG12, sowie mit synthetisch hergestellen Substratmolekülen einen Teil des enzymatischen, ER-luminalen LLO-Synthese-Stoffwechselwegs nachgebaut. Es gelang uns reines Dol-PP-$GlcNAc_2Man_9$ aufzureinigen, welches das Substrat von ALG6 darstellt. Zusätzlich konnten wir alle Zwischenprodukte des Stoffwechselwegs herstellen. Mit Hilfe des Dol-PP-$GlcNAc_2Man_9$ sowie synthetisch hergestelltem Dol-P-Glc gelang es uns die Aktivität von ALG6 zu demonstrieren.

Wir betteten ALG6 in eine Nanodisc ein und mit Hilfe von eines Antikörper Fab-Fragments, das aus einer Synthetischen Bibliothek stammte, gelang es uns mit Kryo-

EM eine Struktur von ALG6 mit einer Auflösung von 3.0 Å zu ermitteln. Eine zusätzliche Struktur mit einer Auflösung von 3.9 Å wurde ermittelt. Hierbei wurde in Detergenz gelöstes, mit Fab komplexiertes ALG6 welches ebenfalls mit dem synthetisch hergestellten Donor-Substrat Dol-P-Glc komplexiert war verwendet. Die Struktur des Substrat-gebundenen Enzyms, sowie gezielte Mutagenese des *ALG6* Gens ermöglichten es, die katalytische Seitenkette Asp69, welche vermutlich als allgemeine Base fungiert, zu identifizieren. Dies gab uns ebenfalls Hinweise auf einen möglichen Reaktionsmechanismus von ALG6. ALG6 gehört zur GT-C Superfamilie Membran-integraler Glykosyltransferasen, welche Isoprenoid-gebundene Zucker als Donor-Substrate verwenden. Mit unseren Strukturen konnten wir gemeinsame Motive in GT-C Enzymen aufzeigen und haben eine evolutionär konservierte, modulare Architektur der GT-C Enzyme entdeckt.

Die von uns beschriebene Methode für die chemo-enzymatische Synthese von eukaryotischen LLOs und LLO-Zwischenprodukten ist ein noch einzigartiges Werkzeug zur Charakterisierung der in der LLO-Synthese involvierten Enzyme.

Die Strukturellen Einblicke in dieser Studie erklären nicht nur die Funktion von ALG6 und vermitteln dadurch Einsicht in die ER-basierte Biosynthese von LLOs, sondern tragen auch zum generellen Verständnis der Architektur und der Mechanismen von GT-C Enzymen bei.

Mit einer Masse von lediglich 64 kDa für ALG6 und 48 kDa für das Fab-Antikörper Fragment, stellt Struktur des ALG6:6AG9-Fab Komplexes mit einer Auflösung von 3.0 Å, die momentan am höchsten aufgelöste Kryo-EM Struktur eines membrangebundenen Komplexes dieser Grösse dar. Wir konnten somit die Nützlichkeit der Fab-Fragmente für die Strukturaufklärung von kleinen Membranproteinen demonstrieren. Solche Fab-Fragmente könnten sich für die Strukturaufklärungen von weiteren Membranproteinen, wie zum Beispiel kleinen, Membran-integralen Glykosyltransferasen als nützlich erweisen.