Sequential Assignment of the Heat-labile Enterotoxin Backbone for Lipopolysaccharide Interaction Studies

by

Daniel Hatlem

Master Thesis in Biochemistry
Department of Biosciences
Faculty of Mathematics and Natural Sciences

University of Oslo

January 2015

2015
Sequential Assignment of the Heat-labile Enterotoxin Backbone for Lipopolysaccharide Interaction Studies
Daniel Hatlem
Print: Reprosentralen, Universitetet i Oslo

Acknowledgements

The work in this project was carried out in the laboratory of Professor Ute Krengel at the Department of Chemistry, University of Oslo.

I would first and foremost thank my main supervisor, Professor Ute Krengel, for recruiting me to her fantastic group shortly after we first met, and for providing me with an interesting and challenging master project. As my main supervisor, she has always shown great interest and support in my work.

I would also like to express my gratitude to my very competent co-supervisor, Dr. Per Eugen Kristiansen, who introduced me to the field of NMR and who pushed me through this project. Whenever there was a problem, he had a solution.

Furthermore, I would like to thank Julie Heggelund and Dr. Daniel Burschowksy. They have both been of invaluable help around the lab and have always been ready to answer all of my little questions about everything. Jan Anonsen also deserves a thank you for being a great help with the MS work and for putting up with my mislabeled samples.

Finally, I would like to thank all of the past and present members of the Protein Dungeon group. You are now too many to mention, but you know who you are. You have all made it into two fantastic years! A special thanks go to my family and friends, who are always of great support.

Oslo, January 2015
Daniel Hatlem

Abstract

Enterotoxigenic *E. coli* (ETEC) cause severe diarrhea, killing millions of people every year, mainly in the developing world. The main virulence factor of ETEC is the heat labile enterotoxin (LT), which is closely related to the cholera toxin (CT), the causative agent of cholera. LT consists of a catalytically active A-subunit sitting on top of the doughnut-shaped B-pentamer (LTB), which is responsible for binding the toxin to the epithelial cells in the intestinal lumen. The B-pentamer has two recognized binding sites. The primary binding site binds to the GM1 receptor, presented on the epithelial cell surface. More recently, a secondary binding site, which binds to blood group antigens, was discovered on LTB that was isolated from ETEC strains infecting humans (hLTB). It has also been shown that hLTB binds to lipopolysaccharides (LPS) presented in the outer bacterial membrane. This binding anchors LT to the bacterial surface and may explain the different toxin delivery mechanisms between LT and CT, where LT is delivered to the host cell while bound to outer membrane vesicles, whereas CT is secreted in soluble form. Some research has been conducted to characterize the binding properties of LPS to LT, but the binding properties of the LPS binding site are still elusive. The goal of this thesis is to study the interaction between LPS and hLTB by combining NMR spectroscopy and X-ray crystallography.

In this thesis, the resonance assignment of the hLTB backbone is presented. In the process, an improved production protocol of hLTB was developed, as well as adaptation of the bacterium *Vibrio* sp. 60 to 99% D_2O concentrations. The resonance assignment was verified by the mapping of the primary binding site using a known ligand, neolactotetraose. The interaction between hLTB and LPS was complicated due to the hydrophobic nature of the LPS lipid tail causing aggregation. A soluble LPS component, Kdo monosaccharide, was used in the hLTB binding studies. Using both the NMR assignment and a crystal structure that was solved to 2.2 Å, we could show that there is low-to-none affinity between hLTB and Kdo. Progress has been made in solubilizing LPS using different detergents, with clear notions on which way to continue. Extensive work has been done to prepare for binding studies between LPS and hLTB, paving the road for future explorations of the interactions of these important biomolecules.

Abstract (In Norwegian)

Enterotoksigen *E coli* (ETEC) foråsaker alvorlig diaré, og er ansvarlig for millioner av dødsfall hvert år, hovedsakelig i utviklingsland. Hovedvirulensfaktoren til ETEC er det varmelabile enterotoksinet (LT), som er nært beslektet til koleratoksinet (CT), årsaken til kolera. LT består av en katalytisk aktiv subenhet som sitter på toppen av en smultringformet B-pentamer (LTB), ansvarlig for å binde toksinet til epitelceller i tarmen. B-pentameren har to bindingseter. Det primære bindingsetet binder til GM1 gangliosidet, som finnes på overflaten til epitelcellene. Nylig har et sekundært bindingsete blitt oppdaget på LTB isolert fra ETEC-stammer som rammer mennesker (hLTB), og dette bindingsetet binder blodgruppeantigener. Det er også vist at hLTB binder til lipopolysakkarider (LPS), som finnes på bakterieoverflaten. Denne bindingen ankrer LT til bakterieoverflaten og kan forklare hvorfor LT blir levert til vertscellen bundet til bakterielle membranvesikler, mens CT slippes fritt ut i tarmen. Tidligere forskning har forsøkt å karakterisere bindingsegenskapene mellom LPS og LT, men denne gåten er fremdeles uløst. Målet med dette prosjektet er derfor å undersøke den spesifikke bindingen mellom LPS og hLTB ved å kombinere NMR-spektroskopi og røntgenkrystallografi.

I denne oppgaven presenteres resonansassigneringen av hLTBs "ryggrad". I prosessen har en forbedret protokoll av hLTB blitt utviklet og bakterien *Vibrio* sp. 60 har blitt tilpasset 99% D_2O. Resonansassigneringen ble bekreftet ved å analysere bindingen av en kjent ligand, neolaktotetraose (NEO) til hLTB. Interaksjonen mellom hLTB og LPS ble komplisert grunnet den hydrofobe lipidhalen til LPS, som foråsaket aggregering. En løselig LPS-komponent, monosakkaridet Kdo, ble brukt til hLTB bindingsstudier. Ved å kombinere både NMR-assigneringen og krystallstrukturen ble det vist at det er lav-til-ingen binding mellom Kdo og hLTB. Fremgang har blitt gjort for å øke løseligheten til LPS ved bruke forskjellige detergenter med en klar plan videre. Omfattende forberedende arbeid har blitt gjort for å studere bindingen mellom LPS og hLTB, og brøyter veiene videre for nye bindingstudier.

Contents

Abbreviations

General abbreviations

$^{1}H^{N}$ ..... amide-bound proton
BMRB ..... Biological Magnetic Resonance Data Bank
CSA ..... chemical shift anisotropy
$^{13}C'$ ..... carbonyl carbon
CMC ..... critical micelle concentration
CT ..... cholera toxin
CF ..... colonization factor
CTB ..... cholera toxin B-pentamer
D_2O ..... deuterium oxide
DPC ..... n-dodecylphosphocholine
DSS ..... 4,4-dimethyl-4-silapentane-1-sulfonic acid
ESRF ..... European Synchrotron Radiation Facility
EG ..... ethylene glycol
ETEC ..... enterotoxigenic *E. coli*
FID ..... free induction decay
GM1 ..... monosialotetrahexosylganglioside
GndHCl ..... guanidine hydrochloride
K_D ..... dissociation constant
H-D exchange ..... hydrogen-deuterium exchange
LPS ..... lipopolysaccharide
hLTB ..... human heat labile enterotoxin B-pentamer
HSQC ..... heteronuclear single quantum coherence spectroscopy
IPTG ..... isopropyl β-D-1-thiogalactopyranoside
Kdo ..... 3-deoxy-D-manno-octulosonic acid
LT ..... heat labile enterotoxin
LTB ..... heat-labile enterotoxin B-pentamer
MQ-H_2O ..... Milli-Q filtered and ion-exchanged water
NMR ..... nuclear magnetic resonance
NOE ..... nuclear overhauser enhancement/effect
MS ..... Mass spectrometry
OMV ..... outer membrane vesicle
PDB ..... Protein Data Bank
PAGE ..... polyacrylamide gel electrophoresis
PG ..... propylene glycol
r.f. ..... radio frequency
SDS ..... sodium dodecyl sulfate
ST ..... heat stable enterotoxin
TALOS ..... Torsion Angle Likeliness Obtained from Shift and Sequence Similarity

τ_c ..... correlation time
T_1 ..... longitudinal relaxation time constant
T_2 ..... transverse relaxation time constant
TROSY ..... transverse relaxation-optimized spectroscopy

1 Introduction

1.1 General background

1.1.1 Infectious diarrheal diseases and enterotoxigenic *E. coli*

Infectious diarrheal diseases affect billions of people every year and are annually responsible for the death of almost 1.5 million children under 5 years of age, the second leading cause of child death after pneumonia [1, 2]. The most common etiological agents include Rotavirus, *Shigella spp*, *Vibrio cholerae* and enterotoxigenic *E. coli* (ETEC), where ETEC is recognized as the most frequent bacterial cause [3]. ETEC is also the most common cause of traveler's diarrhea, which affect people who travel to regions where these diseases are endemic.

Until the second half of the 20th century, ETEC-induced diarrhea was often misdiagnosed as cholera or simply regarded as 'non-vibrio cholera' due to the very similar pathophysiology of the two diseases. It was only in the 1940s that it was suspected that a pathological *E. coli* strain was causing diarrhea, leading to much research without conclusive results as to the cause, as reviewed by Qadri et al [4]. The definite evidence of enterotoxin-producing *E. coli* came in 1967 when a team of clinical scientists from USA traveled to Kolkata, India, to conduct research on cholera [5]. They were soon side-tracked by patients admitted with cholera-like symptoms, but with high concentrations of *E. coli* in their stool. By injecting isolated *E. coli* cell cultures and cell culture filtrates into rabbit ileal loops, they established that enterotoxigenic *E. coli* was the etiological agent and that it released a toxin with very similar properties to the cholera toxin.

ETEC is mainly endemic in the developing world, where it is estimated to cause more than 840 million cases of diarrhea every year [4]. It usually spreads by fecal-oral transmission by contaminated drinking water and food, caused by poor sanitary conditions [6]. The disease has also been shown to cause epidemic outbreaks during floods [7]. Although ETEC affects people of all ages, the incident rate and morbidity is usually higher in children under the age of 5 [8]. Since ETEC frequently infect young children, causing malnutrition, stunted growth and death, it is recognized as a great burden on developing countries and the international community has initiated several programs to combat the disease with varying degrees of success.

The incubation time of ETEC is normally 1-3 days and the illness typically has a duration of 4-5 days, but could last much longer [9]. After infection, ETEC colonize the intestine and release enterotoxins, which cause symptoms that include diarrhea, fever, abdominal cramps and general malaise. The disease can usually be treated simply by replenishing fluids and electrolytes, but in more serious cases, antimicrobials and intravenous rehydration might be necessary [1, 4]. Treatment has been improved by the development

and distribution of oral rehydration solution (ORS) and educating local health workers about the disease [10, 11]. Prophylactic initiatives have been initiated in many developing countries and involve ensuring clean drinking water, improving sanitary conditions and ensuring proper nutrition [1]. Since it has been established that those affected by ETEC or *V. cholerae* have developed antibodies against certain virulence factors, several vaccines have been developed over the years, although with varying success [12, 13].

1.1.2 Enterotoxins and other virulence factors

ETEC represents a ubiquitous *E. coli* strain that express many distinct antigens encoded on either chromosomal DNA, or extrachromosomal plasmids. ETEC virulence plasmids encode two different enterotoxins, more than 25 different colonization factors (CFs) and other adhesins necessary for infection [14].

The main virulence factors of ETEC are the enterotoxins, called the heat-labile (LT) and the heat-stable (ST) enterotoxin. ST is a small peptide that activates guanylyl cyclase and is toxic for animals [15]. As the name implies, it retains its structure and function even when boiled. The heat-labile enterotoxin on the other hand, forms a hexameric complex that is irreversibly denatured when incubated above 65 °C (See section 3.4.1). A second heat-labile enterotoxin called LT-II is occasionally isolated from ETEC. LT-II share a high sequence identity with LT-I, but the structures differ significantly and LT-II will not be discussed further in this text [16]. LT (LT-I) is further described in section (1.1.3). ETEC is commonly classified by expression of ST, LT or both. Around 20 % of all ETEC isolated strains express LT and 80% express ST, 50% express both [17].

The colonization factors expressed in ETEC are proteinaceous filaments (fimbriae and fibrills) that allow the bacteria to adhere to the intestinal mucosa via highly specific interaction with epithelial glycoconjugates and play an important role in ETEC pathogenesis [18]. The genes that encode CFs are often flanked by insertion sequences and transposons, allowing different ETEC strains to adapt quickly to new conditions and evade the host immune system.

ETEC also expresses a wide variety of O-antigens, which are oligosaccharides associated with cell wall lipopolysaccharides (LPS). More than 78 different O-antigens have been identified in ETEC strains, and O-antigen serotyping has been shown to be a worthwhile way to identify pathogenic *E. coli* [19]. 34 different H-serogroups have been identified in ETEC. Like CFs, these are flagellar antigroups, but their pathophysiological importance is disputed [19].

The heterogeneities mentioned above underline the high diversity of ETEC, and coupled with the fact that ETEC is not detected during rutine stool analysis could explain why many cases are suspected to go unreported worldwide [9]. The fact that the genes encoding the virulence factors are located on extrachromosomal plasmids allows the dis-

ease to quickly adapt to new conditions by horizontal gene transfer, and might explain the quick emergence of resistant strains that has hampered the development of broad spectrum, long lasting vaccines [8, 12, 13, 20].

1.1.3 LT structure and function

LT is a heterohexamer belonging to the AB_5 toxin family, which includes cholera toxin (CT), pertussis toxin and shiga toxin. LT consists of an A-subunit (LTA) inserted in a doughnut-shaped homopentameric B-pentamer (LTB). The structure of hLTB was first solved in 1991 by Sixma et al [21] and is presented in Figure 1. Even though ETEC infect several animals, LT is usually only found in strains that infect humans (hLT) and pigs (pLT). LT is structurally and functionally very similar to the cholera toxin (CT), which is secreted from *V. cholerae*. hLTB shares a sequence identity of approximately 84% with the cholera toxin B-subunit (CTB).

During infection, the bacteria colonize the intestine and secrete the LT into the lumen where the B-subunit anchors the complex to the epithelial cell surface and induce endocytosis. When the toxin is internalized into the cell, the catalytically active A-subunit dissociates from the complex and initiates a mechanism that cause an efflux of H_2O and salts over the membrane, causing diarrhea [22].

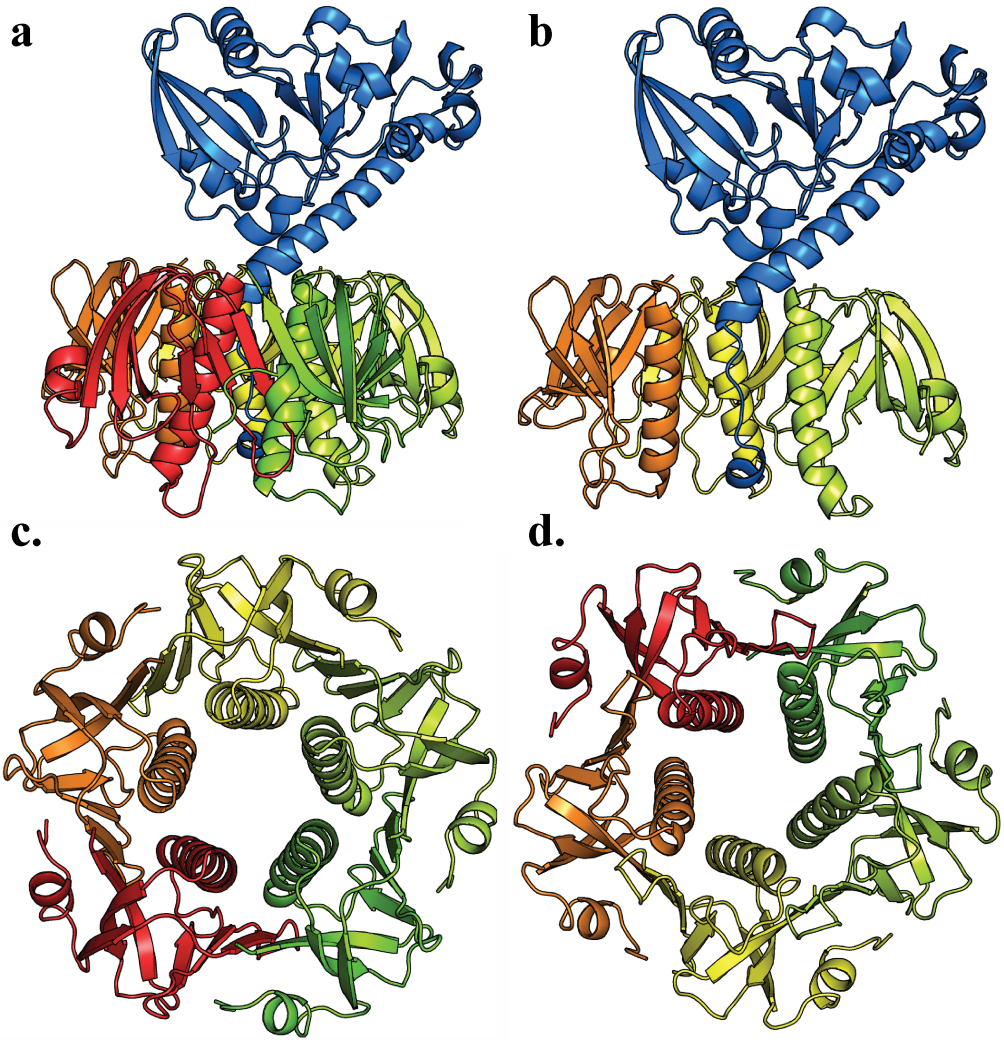


Figure 1: **a.** Side view of LT. The A-subunit (blue) inserted into the B-pentamer **b.** Side view of LT with two B-subunits removed to reveal the inside of the B-pentamer **c.** Top view of the B-pentamer **d.** Bottom view of the B-pentamer (PDB ID: 1LTS)

The mature A-subunit is a 240 amino acid protein of 27.5 kDa, which exhibits ADP-ribosylase activity. It is translated with a signal sequence that directs it to the periplasm after production [23], where it is cleaved off [24]. The protein subsequently folds spontaneously into a semi-stable tertiary structure. Bacterial disulfide oxidoreductase A (DsbA) forms a disulfide bond between Cys187 and Cys199 before LTA is partially activated by a proteolytic cleavage between residue 194 and 195. Cleavage forms the two LTA fragments LTA-1 (22 kDa) and LTA-2 (5.5 kDa), which are held together by the the disulfide bond [25]. LTA-2 is the part that associates with LTB.

The B pentamer (58.5 kDa) consists of 5 identical subunits of 103 residues (11.7 kDa) resulting in a 5-fold symmetry. Like LTA, it is also encoded with a signal sequence directing it to the periplasm, where the signal sequence is cleaved off [24, 26]. Cys9 and Cys86 are oxidized by DsbA allowing it to fold spontaneously into its tertiary structure. The B-subunits have been shown to spontaneously oligomerize into a B-pentamer, but only when Pro93 is in the *cis* configuration. It is thought that peptidyl prolyl *cis*-isomerase is responsible for catalytic isomerization of Pro93 [27].

The C-terminus of the A-subunit (LTA-2) was shown to modulate the assembly process of the holotoxin and the oligomerization of the B-pentamer is enhanced 3-fold compared to the absence of LTA [28, 29]. The exact mechanism is unknown, but it is assumed that 2–4 B-subunits form an intermediate product that recruits LTA, which further stabilize the complex and recruit the remaining B-subunits [28]. B-subunits can oligomerize into B-pentamers spontaneously and an excess of A-subunits is expressed with ca 1-2 LTAs produced per 5 B subunits. From an energy saving viewpoint this makes sense, as a surplus of LTA can be broken down and recycled, while the very stable LTB can be secreted without being bound to LTA (As reviewed by Mudrak et al [30]).

The assembled holotoxins are secreted from the bacteria over the outer membrane via the type II secretion complex (T2SS) [31]. Both ETEC and *Vibrio cholerae* share this complex and it has been shown that the B-pentamers of both toxins are the secretion substrates [30]. In contrast with CT, more than 95% of LT is secreted bound to lipopolysaccharides (LPS) in the outer membrane and delivered to the epithelial cells associated with outer membrane vesicles (OMVs) [32].

After the complex is secreted from the bacteria into the intestinal lumen, the B-pentamer adheres the toxin to the epithelial cells by binding to the GM1 ganglioside receptor on the epithelial cell surface [33]. The complex is internalized into the cell by endocytosis (Figure 2) and transported to the endoplasmatic reticulum (ER) via retrograde trafficking. Once in the ER, the disulfide bond that connects LTA-1 and LTA-2 is reduced, allowing LTA-1 to dissociate from the B-pentamer and translocate into the cytosol where it folds into its active state. Once active, LTA-1 ribosylates and activates

the GTP-binding $G_s\alpha$-protein, initiating the cAMP-dependent pathway that ends with the opening of Cystic fibrosis transmembrane conductance regulator (CFTR) i.e the G_s-protein activates the adenylate complex (AC), which activates protein kinase A (PKA) through production of cAMP. PKA phosphorylates CFTR which results in the secretion response that is the primary event responsible for causing diarrhea [34, 35].

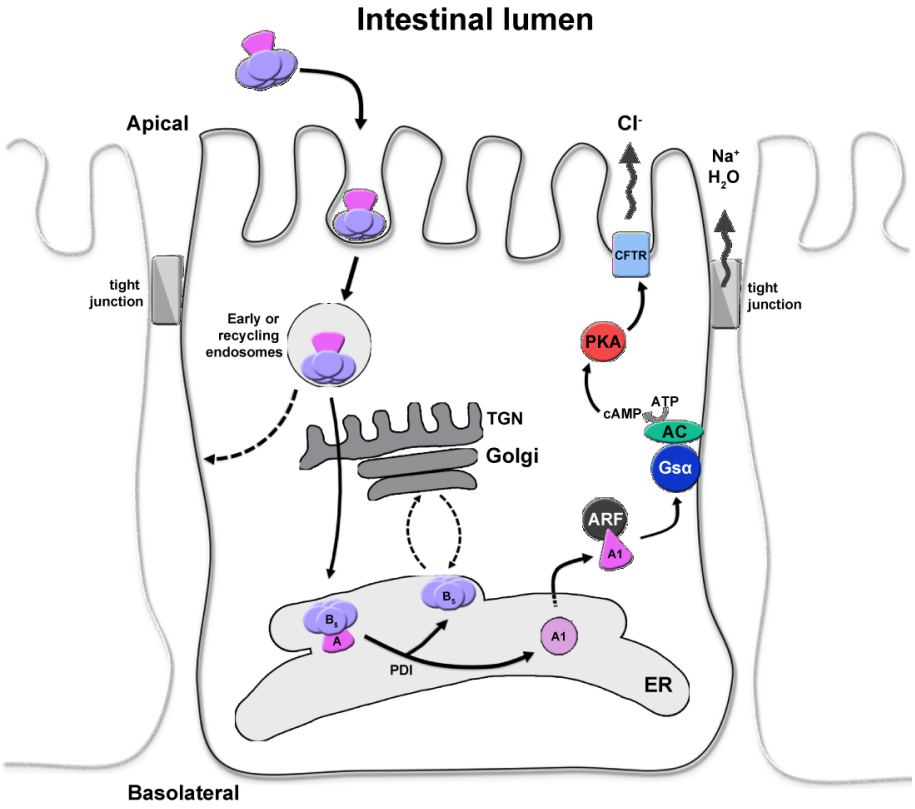


Figure 2: Endocytosis and subsequent catalytic activity of CT/LT (Adapted from figure by Julie Heggelund)

1.1.4 Toxin gene regulation

The genes encoding for LT are located on a ETEC virulence plasmid called pEnt and were sequenced in the early 80s [36, 37]. LTA and LTB are encoded by the two genes *eltA* and *eltB* (earlier called *toxA* and *toxB*) that overlap with one nucleotide at the distal end of *eltA* and proximal end of *eltB*. They are a part of the same two-gene operon, *eltAB*, and are transcribed as one mRNA strand [38]. Sequence comparison between *eltAB* and *ctxAB*, the corresponding CT operon, shows that they share a nucleotide identity of 77.9% suggesting that they share a common ancestor [39].

Expression of LT is shown to be regulated by environmental factors, and, more interestingly, conditions similar to those found in the intestine generally upregulate the expression

rate of LT. The main regulator of *eltAB* is the histone-like nucleoid-structuring protein (H-NS) that binds to a downstream regulatory element from the *eltAB* promotor and inhibits transcription. The inhibitory activity of H-NS is strongly affected by increasing temperatures and exhibits the lowest inhibition at 37 °C [40, 41]. Expression is also affected by osmolarity and pH, albeit to a lesser degree than temperature. The mechanisms for these conditions are unknown, but it has been shown that high osmolarity by NaCl and glucose increase LT expression while acidic pH is inhibiting [42, 43]. This is consistent with the high salt levels and neutral pH that is found in the intestine. Recently, it was also shown that H-NS also regulate the transcriptions of secretion system II that is necessary for secretion of LT [44].

Another factor that inhibits LT expression is the presence of short carbon-chains [45]. Such chains are found in the colon as short-chain fatty acids and might act as a signal for ETEC to stop LT-production when it has passed the intestine [46]. Higher cAMP levels also suppress LT production. This could function as a feedback mechanism as cAMP leaks out of the epithelieal cells after LTA activation and is detected by exogenous cAMP receptors on the bacterial surface. The factors that regulate LT expression can vary drastically between different ETEC strains [47].

1.1.5 Ligand binding to LTB

The B-pentamer is responsible for anchoring the holotoxin to the target cells receptor, GM1, and has also been shown to interact with several other molecules. At the present time, two types of saccharide binding sites have been identified on the B-subunit; the primary binding site, and the more recently identified secondary binding site. Figure 3 shows the binding sites and their binding substrates. A third, putative binding site binding to LPS is hypothesized to overlap with the secondary binding site [48].

Primary binding site

The main receptor for LTB and CTB is the GM1-ganglioside (Galβ3GalNAcβ4(NeuAcα3)-Galβ4GlcCer), which binds to the primary binding site, located in the lower interphase between the subunits (Figure 3) [49, 50]. GM1 is presented on the surface of the intestinal epithelial cells and upon binding to the B-pentamer, anchors the toxin to the surface prior to internalization. GM1 binding has been shown to be essential for the virulence of the toxins [51].

The interaction between LT and GM1 is characterized by a a large surface interaction and exceptionally high affinity (K_D = 49 nM [52]). Structural studies show that the terminal Galβ3 of GM1 is mainly responsible for the interaction by burrowing into a cleft in the binding site, where it interacts with residues Glu51, Gln61, Ans90, Lys91 and Trp88 (Figure 4 [53]). The sialic acid residue supports the binding by interacting with Arg13, Ile58 and Glu11 and Glu33, where the latter belongs to the neighboring subunit.

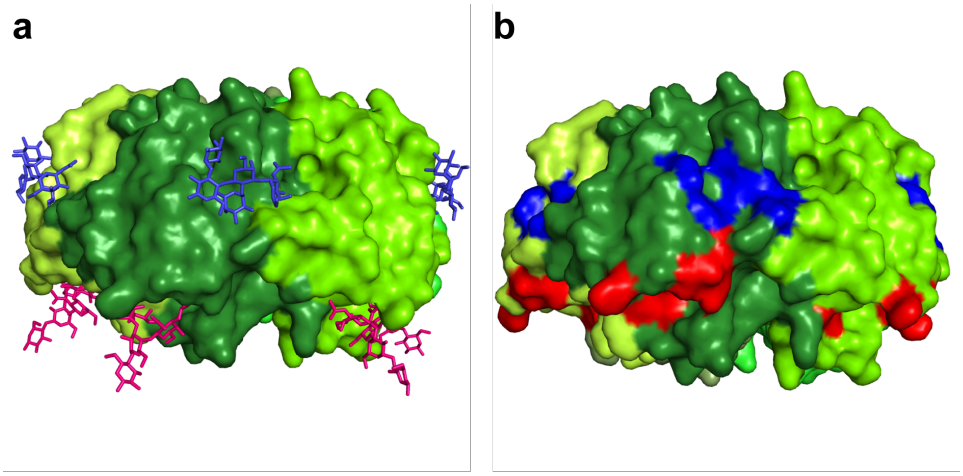


Figure 3: **a.** Surface representation of LT bound to GM1 (red) and blood A determinant (blue). **b.** Surface representation of residues involved in binding Prepared with PyMol (PDB ID: 2O2L, 2XRQ)

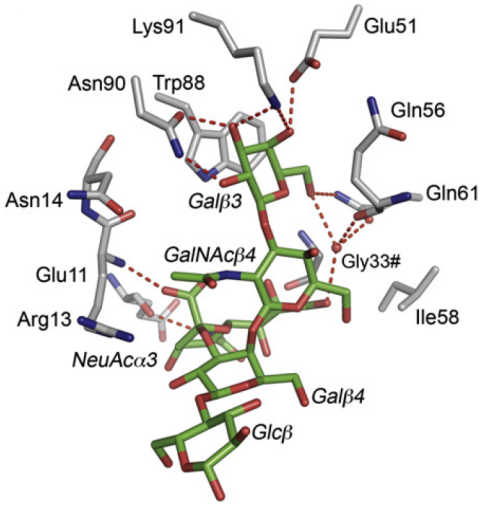


Figure 4: GM1 binding to pLTB (From Holmner et.al (2011)[53])

Even though the CTB primary binding site only binds to GM1, LTB has been shown to be more promiscuous by also binding, to GM_2, GD_2, GD_1b and glycoconjugates that carry N-acetyllactoseamine epitopes, such as neolactotetraosylceramide (NEO), although significantly weaker than GM1 [53, 54, 55].

Blood group binding

The secondary binding site was discovered in 2004 in a hybrid between LT and CT, and is shown to bind both blood group antigen A [56]. The site is distinct from the primary binding site and located at the interface of two subunits, but like the primary binding site, one subunit is mostly responsible for ligand interaction. Blood group A determinant was

successfully co-crystallized with LT in 2007, showing the detailed interaction between a wild type toxin and A-antigen (GalNAcα3(Fucα2)Galβ4(Fucα3)Glcβ) [57].

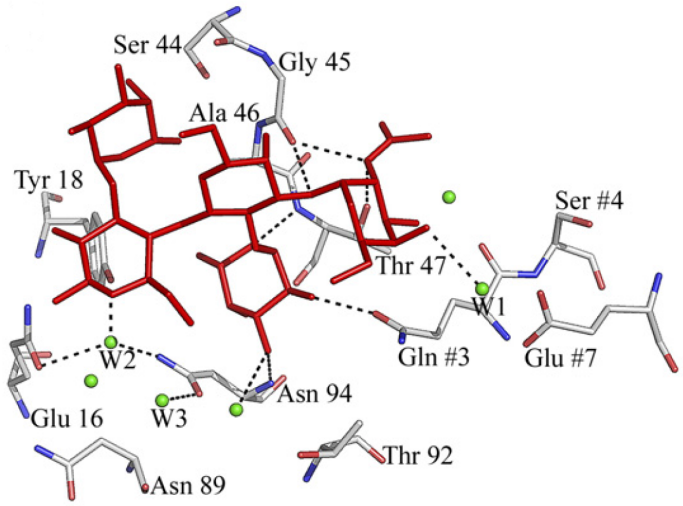


Figure 5: Blood group A determinant binding site (From Holmner et al (2007)[57]

The ABO-blood type determining ABH-antigens are glycoconjugates that are associated with the surface of red blood cells. The different antigens are composed of the base H-antigen ((Fucα2)Galβ4(Fucα3)GlcβNAc), associated with blood type O with an added terminal N-acetylgalactosamine or galactose for the A- and B- antigen, respectively [58] (Figure 5). Interestingly, most people also express blood group antigens in saliva and on epithelial cells in the digestive tract and intestines [59]. Based on the binding studies, it is therefore believed that blood antigen A and B, expressed on the mucosal layer of the intestinal cells, interfere with binding to GM1 and prevent or slow down toxin internalization into the cell. This may explain the observed blood group dependency of cholera, where the symptoms in people with these blood types are less severe than for those with blood type O [60, 61]. It is hypothesized that blood group antigen may interfer by preventing the toxin from reaching the main receptor [55]. Although cholera exhibits a blood group dependency, this is not the case for LT. The reason is unknown, but it is possible that LPS binding to an overlapping binding site outcompetes blood group binding [57].

Lipopolysaccharide

Lipopolysaccharides (LPS) are a major component of the outer monolayer of the outer membrane of gram negative bacteria and are essential for bacterial growth. LPS forms a hydrophilic layer around the bacteria that protects it against harsh conditions caused by extreme pH, detergents or chemical attack. LPS consists of 3 different parts, Lipid A, core oligosaccharide and the O-antigen (Figure 6).

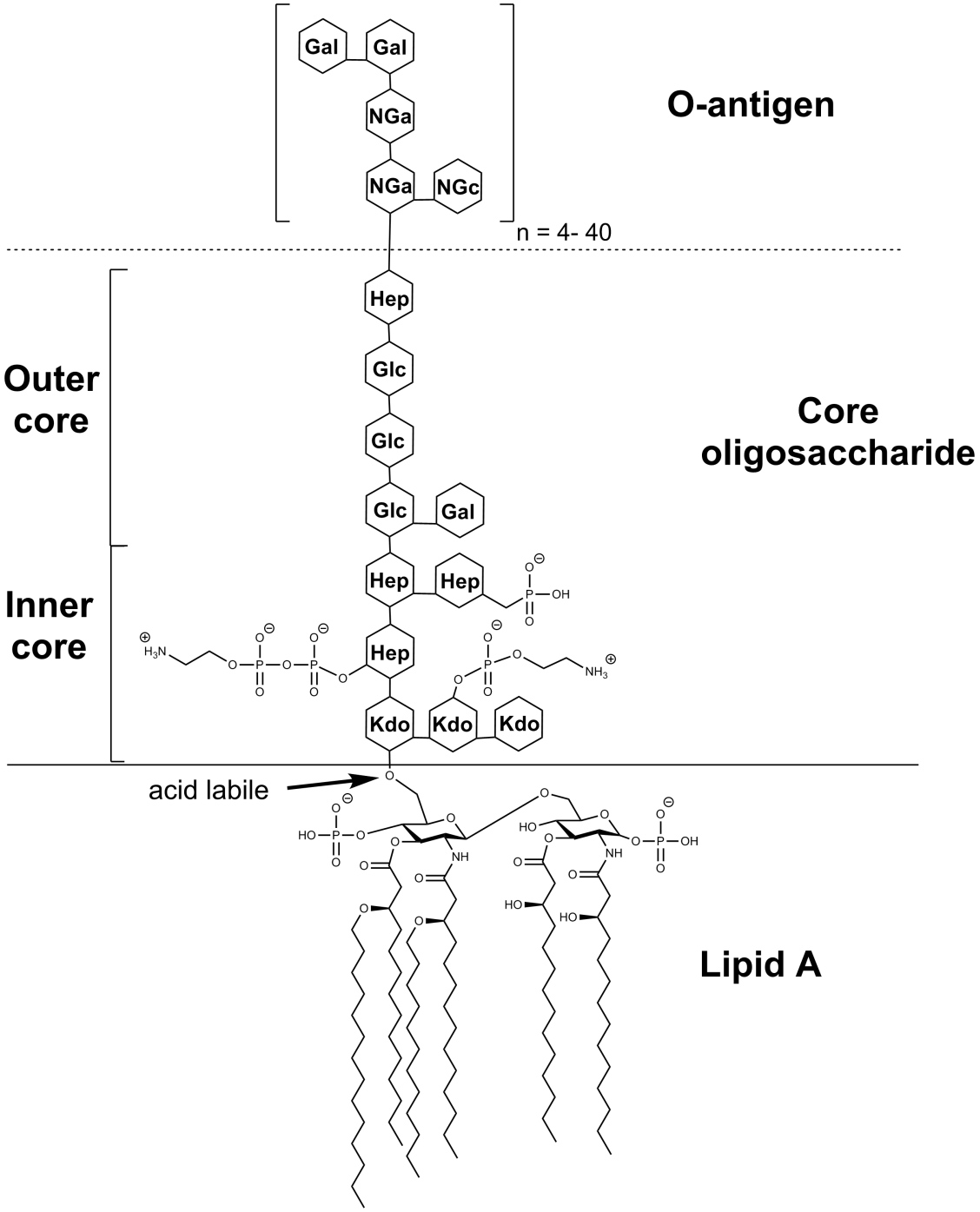


Figure 6: Schematic figure of *E. coli* K-12 LPS. NGa = GalNAc, NGc = GlcNAc.

Lipid A functions as the membrane anchor for the saccharide structure and consists of 6 12C–14C hydroxyacyl chains covalently linked to two β(1-6) linked glycoamines. The core oligosaccharide is covalently bound to Lipid A and is comprised of 10 branched sugar moieties, divided into the inner- and outer core. The proximal sugar is a 3-deoxy-D-manno-octulosonic acid (Kdo), which is the only sugar required for bacterial growth. Other sugar moieties include L-glycero-D-manno-heptose, glucose, galactose and N-acetyl-glucosamine [62]. The O-antigen is bound to the distal core oligosaccharide and is the most variable element of LPS. It consists of up to 50 repeats of small oligosaccharides, usually tetrasaccharides, forming the hydrophilic, protective layer around the bacteria [63]. The O-antigen has several functions, including protection against harsh conditions and aiding interbacterial recognition. The high variety of the O-antigen can also prevent host antibody recognition of the bacteria. LPS without O-antigen occur naturally in all bacterial strains and is called LPS-Ra, or rough LPS. LPS with only Kdo sugars is called LPS-Re or, deep rough LPS [63].

It has been established that LT, but not CT, is secreted in association with the bacterial outer membrane by binding to LPS [64, 32]. However, CT has been shown to bind to *E. coli* LPS, and LT introduced into *Vibrio* sp. 60 on an exogenous plasmid is secreted in soluble form, suggesting that both CT and LT can bind to *E. coli* LPS, but not *V. cholerae* LPS [65]. It has been shown that the smallest LPS structure recognized by LT is LPS-Re, carrying only the Kdo-sugar. The difference in binding may be caused by phosphorylation of the Kdo sugars on *Vibrio* LPS, preventing CT binding and resulting in secretion of soluble toxin [65, 48]. The difference in binding caused by a difference in LPS also suggest that both LT and CT share a common LPS binding site. At the present time, no crystal structures of LT bound to LPS or other topical epitopes have been studied, but point mutation studies show that residues Gln3, Thr47, Tyr18 and A46 are required for binding, indicating that LPS binds to a distinct binding site that overlaps with the secondary binding site [48]. This binding may explain why ETEC is not blood group dependent, as opposed to cholera. It has also been shown that LT can bind both LPS and GM1 concurrently, suggesting that the two binding sites are not overlapping or in competition [65].

The role of LPS binding to LT still remains elusive, but several possible functions are suggested. Since LT covers the ETEC vesicle, it assists the targeting of the epithelial cell and mediates the internalization of the whole vesicle [66]. ETEC vesicles carry many pathogens that would not gain entry to the cell in soluble form, this mechanism could aid the pathogenesis of ETEC by delivering the pathogens in one binding event [67]. Additionally, the binding between LT and LPS has a high affinity, and the OMV may protect LT against hydrolysis. Concurrent binding of LT to both LPS and GM1 may assist CFs greatly in adhering ETEC to the intestinal wall, giving ETEC an advantage in colonization of the intestine [13, 68].

1.2 Aims and approach for this thesis

For the past 20 years, much progress has been made in the understanding of the binding properties of the B-pentamers with regards to GM1 and blood group antigens. However, the binding between LPS and LT remains elusive. This interaction is thought to be important for ETEC pathogenesis and plays an important role in different toxin delivery mechanisms that are observed between LT and CT. By elucidating the binding properties between LPS and hLTB, a greater understanding of the role of LPS binding on toxin delivery, pathogenesis and perhaps blood-group dependence of the diseases may be achieved.

The main objective of this thesis was to study the interaction between hLTB and LPS by combining two powerful methods in structural biology, X-ray crystallography and nuclear magnetic resonance (NMR). This was achieved by assigning the hLTB backbone for titration studies and co-crystallizing the LPS core sugar, Kdo, which has been shown to bind to hLTB independently of the remaining LPS. The experimental work was divided into four parts:

The aim of the first part was to develop a suitable expression and labeling protocol for hLTB. This was necessary to obtain high yields of useable hLTB for the NMR experiments.

The aim of the second part was to perform an backbone assignment of hLTB by using triple resonance experiments. This would allow us to study the interactions between hLTB and different ligands by performing NMR-titration experiments.

The aim of the third part was to study the interaction between hLTB and LPS by performing NMR-titration experiments with LPS-Ra and Kdo from the core sugar complex.

The aim of the fourth part was to co-crystallize hLTB with Kdo in order to solve the high resolution structure of the interaction.

1.3 NMR-spectroscopy

1.3.1 General background

Spin, or spin angular momentum, is an intrinsic property of many subatomic particles. All particles that have a spin quantum number $\neq 0$ have a spin angular momentum and a magnetic moment [69] (M_i). The nuclear spin, denoted I by convention, is formed by combining the spin number of protons and neutrons. In liquid state NMR, we are primarily interested in nuclei with I = 1/2 [70]. ^{1}H is the most important nuclei with I = 1/2, being a part of water and all organic molecules, and has a special place in NMR. A selection of nuclei and their properties are presented in Table 1.

Table 1: A selection of nuclear isotopes and their properties at 14.095 T

Isotope	**Spin**	**Natural abundance (%)**	**Gyromagnetic ratio ($\gamma/10^6 s^{-1}T^{-1}$)**	**Resonance frequency at 14.095 T (MHz)**
^{1}H	1/2	~100	42.577	600.130
^{2}H	1	0.015	6.536	92.124
^{12}C	0	98.9	-	-
^{13}C	1/2	1.1	10.708	150.903
^{14}N	1	99.6	-	-
^{15}N	1/2	0.37	-4.137	60.834

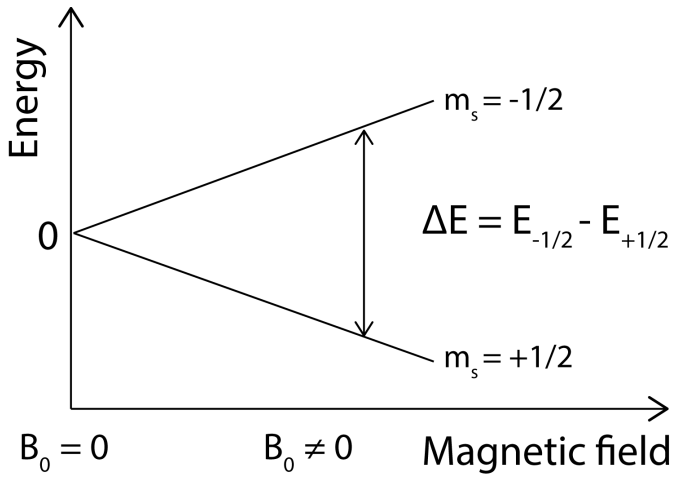


Figure 7: Zeeman splitting of nuclei 1/2 in a magnetic field

Nuclei with I = 1/2 have two spin states that are degenerate in absence of a magnetic field. However, when the nuclei are inserted into a magnetic field, the spin states break apart into 2 separated energy states [70]. This effect is called Zeeman splitting, and is illustrated for a nuclei with spin 1/2 in Figure 7. The two energy levels represent low energy spins that are parallel (N^+, or 'spin-up') and high energy spins that are antiparallel (N^-,

or 'spin-down') with the external field [69]. The magnetic moment of the nucleus is aligned with the spin axis of the nuclei. When the sample is inserted into a magnetic field, the magnetic moments will precess around the field, defined as the Z-axis [70]. The precession frequency of a particular nuclei is called the Larmor frequency and is proportional with the external magnetic field.

$$\omega^0 = -\gamma B_0 \tag{1}$$

where ω is the frequency of the nuclei in Hz and γ is the gyromagnetic ratio of the nuclei. The distribution of spin population in a nucleus with spin 1/2 is defined by the Boltzman factor for a two-state system [70]

$$\frac{N^+}{N^-} = e^{\frac{\Delta E}{kT}} \tag{2}$$

where where N_i is the number of spins in state i, T is the absolute temperature and k is Boltzmann's constant. ΔE is the energy difference between the two spin states and is defined in a magnetic field as

$$\Delta E = |\hbar \gamma B_0| \tag{3}$$

where $\hbar$ is the reduced Planck constant, and B_0 is the external magnetic field. By combining equation (2) and (3),

$$\frac{N^+}{N^-} = e^{\frac{|\hbar \gamma B_0|}{kT}} \tag{4}$$

the difference in distribution increases with stronger magnetic field and larger gyromagnetic ratio. This difference in distribution between spin-up and spin-down generates a net magnetic moment, which is the basis for the signal that is detected in NMR [69]. This is also the reason why ^{1}H, which has a comparatively large γ, provide an extraordinary strong signal in NMR [71].

The 1D NMR experiment

When a sample is inserted into a magnetic field the magnetic moments of the nuclei will align with the field, forming a net magnetic moment. However, the different spins are incoherent, or out of phase with each other and there is no detectable signal. In order to observe a signal, the sample is excited by applying a short, powerful radio frequency (r.f.) pulse that rotates the spins into the transverse XY-plane, forming a coherent signal (Figure 8) [69].

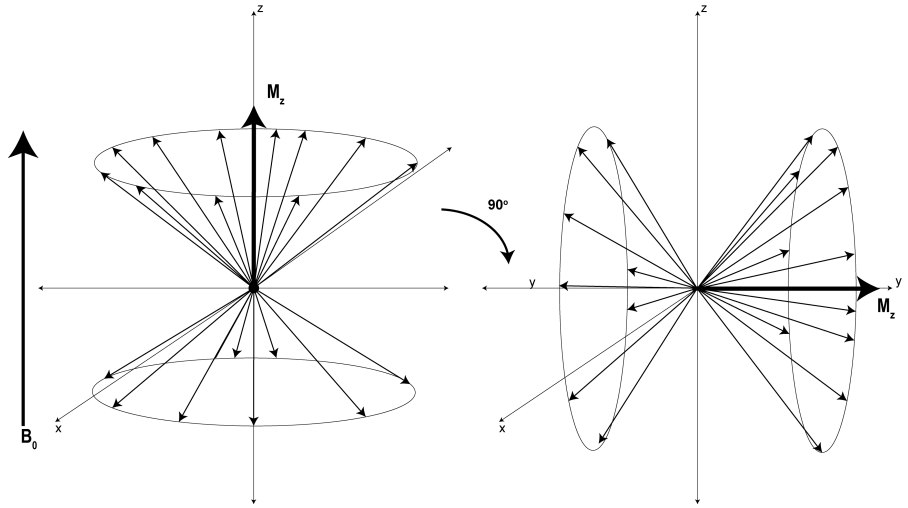


Figure 8: The net magnetization that is aligned with the external field is excited by a 90 degree pulse and rotated into the transverse plane

After the r.f.-pulse is turned off, the net magnetic moment will continue to precess around the Z-axis in phase, forming a measurable signal by induction. The sum of these signals form a complex wave called the free induction decay (FID), which is detected by a coil in the NMR instrument. In order to interpret the FID, it is transformed from the time domain to the frequency domain by a Fourier transformation (Figure 9).

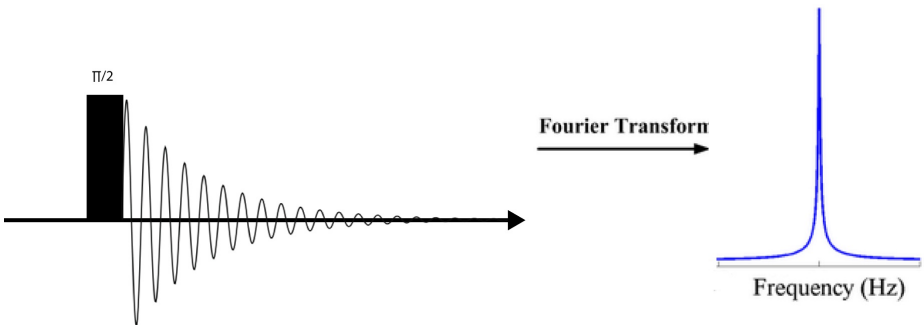


Figure 9: A simple NMR-experiment. The black box represents the 90 degree pulse that rotates the magnetization into the XY-plane. The FID is collected directly afterwards and converted into the frequency domain by a fourier transformation

Signal to noise

The signal to noise of an NMR-experiment is given by:

$$S/N \propto n\gamma_e\sqrt{\gamma_s^3 B_0^3 t} \tag{5}$$

where γ_e and γ_s represent the gyromagnetic ratios of the excitation and detection nuclei, respectively [69]. The n represents the amount of sample, B_0 the field strength of

the magnet and t the number of scans. The signal to noise of an experiment can therefore be increased by increasing the number of scans or using a stronger magnet.

Chemical shift

Although two nuclei of the same isotope should have the same resonance frequency, this frequency can differ slightly because of an effect called chemical shift [71]. Chemical shift is caused by a number of reasons, but the most important one is the different chemical environment caused by the different electron distribution in the molecule [70]. The electrons induce a small field in either the opposite or same direction as the external field, causing a shielding or deshielding effect that affects the local magnetic field. Nuclei will therefore resonate with slightly different frequencies depending on the local electronic distribution. Different functional groups have characteristic chemical shifts. The chemical shift is fundamental in NMR to distinguish functional groups and in studying the chemical environment of nuclei [71].

The ppm scale

Since the chemical shift is field dependent, it is impractical to compare spectra collected at different field strengths. All chemical shifts are therefore referenced in the ppm scale:

$$\delta = \frac{v_o - v_r}{v_r} 10^6 \tag{6}$$

where δ is the chemical shift value in ppm, v_o is the observed frequency of the nuclei and v_r is the reference frequency of the nucleus [69] v_r is observed by adding an internal standard to the sample. An common internal standard for ^{1}H is 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) [72].

^{15}N, ^{13}C and ^{1}H in amino acids have characteristic chemical shift values, depending on the structure of the amino acid. Typical chemical shift values for $^{13}C_\alpha$ and $^{13}C_\beta$ in the 20 common amino acids are presented in Figure 10.

J-coupling

J-coupling (also called scalar coupling) is an effect where covalently bound nuclei interact with each other through the electrons that form chemical bonds [70, 71]. The size of the J-coupling constant between two nuclei, A and B is measured in Hz and denoted $^nJ_{AB}$, where n is the number of bonds between the spins. The typical coupling constants in the protein backbone are independent of the type of amino acid (Figure 11). This is of importance in backbone protein NMR experiments since it allows selective transfer of the magnetization to spins in the preceding or same residue [73, 74] (Section 1.3.4).

Dipole coupling

Dipole coupling is an interaction between two spins that depends on the distance between

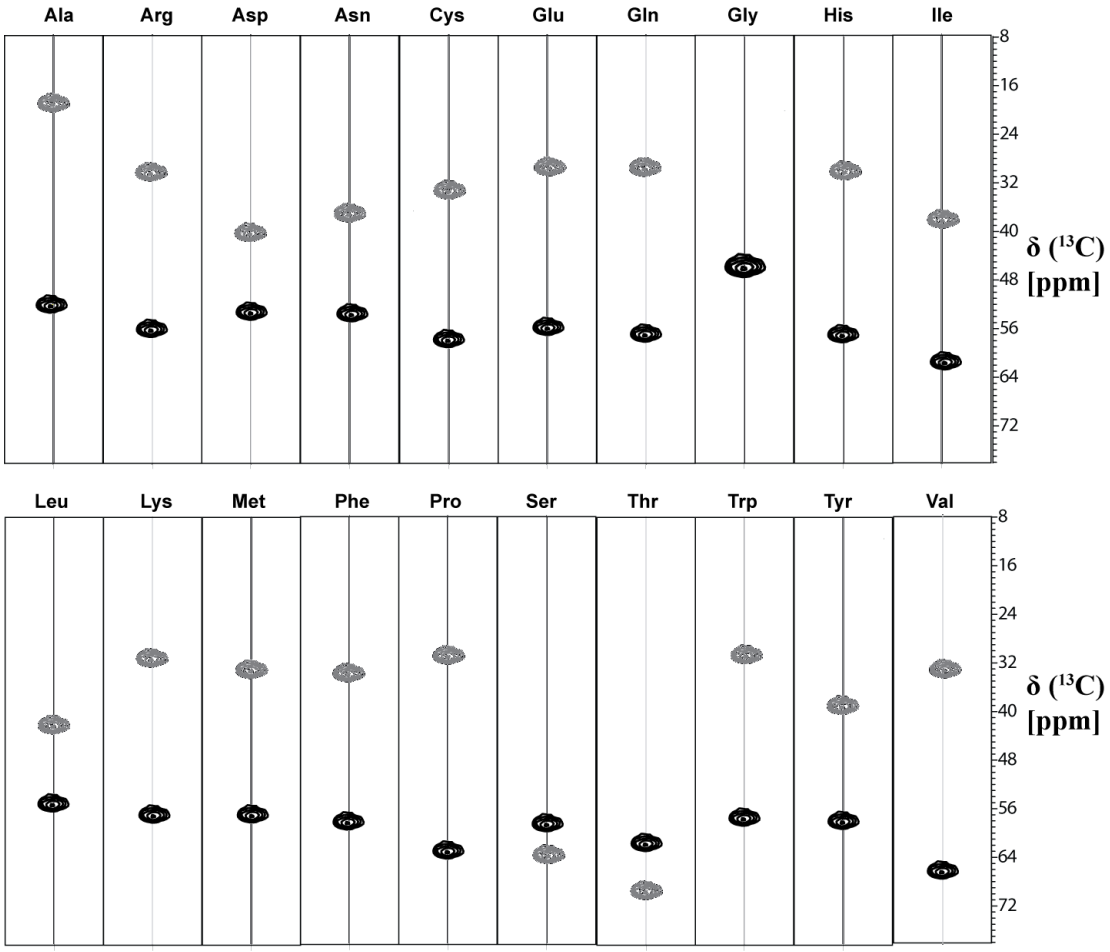


Figure 10: Typical chemical shift values of $^{13}C_\alpha$ (black) and $^{13}C_\beta$ (grey) in the different amino acids (The chemical shift values used to make this figure were collected from BMRB)

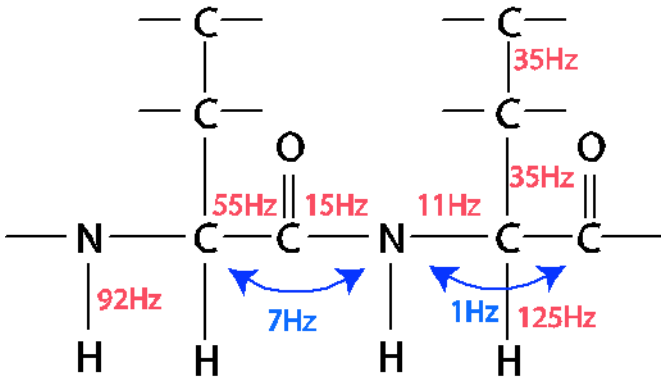


Figure 11: Different J-coupling constants between different nuclei in a protein backbone. Red number indicate ^{1}J and blue ^{2}J coupling constants (Borrowed from Per Eugen Kristiansen)

them [71]. The magnitude of the the dipole interaction between nuclei j and k is given by the equation:

$$b_{j,k} = -\frac{\mu_0}{4\pi}\frac{\gamma_j \gamma_k \hbar}{r_{j,k}^3} \tag{7}$$

where γ_j and γ_k are the gyromagnetic ratios of nuclei j and k, and $\mu_0 = 4\pi \times 10^{-7}$ H/m [69]. This means that the dipole-interaction is most pronounced between nuclei with high gyromagnetic ratio, but decrease rapidly when the distance between the nuclei increases. This means that nuclei directly bound to protons experience strong dipole interaction.

Dipole coupling cause cross relaxation (see below) between nuclei, and is a fundamental aspect in the nuclear overhauser effect/enhancement (NOE) [70]. NOE spectroscpy (NOESY) use this effect to cross correlate spins through space to gain information about the distance between nuclei. The intensity of the observed cross peaks are highly dependent on the distance and can usually only be observed up to ~5 Å [71].

Relaxation

Relaxation time is a term that describes the process when a system reaches thermal equilibrium after it has been perturbed. Two types of relaxation are commonly recognized in NMR: longitudinal relaxation and transverse relaxation, denoted by the time constants T_1 and T_2 respectively [70]. When a sample is in equilibrium in the absence of a magnetic field, the magnetic moment of the different nuclei can point in any direction and the sum of all spins are 0. The longitudinal relaxation time, T_1, describes the time needed for 1/e of the sample to reach an equilibrium when it is inserted into a magnetic field or after it has been perturbed by an r.f. pulse [69]. Although the direction of the nuclear magnetization is independent of the orientation of the molecule, the tumbling of the molecule change the relative position of the other electrons and nuclei around the nucleus. The relative movement of the magnetic particles compared to the nucleus causes fluctuations of the magnetization as the molecule tumbles and is called chemical shift anisotropy (CSA). CSA is usually not an issue for small molecules, since the effect is averaged out because of their rapid rotation (given by correlation time, τ_c) [71]. However, larger molecules rotate slower, causing the effects of CSA and dipole-coupling to modulate the precession frequency of the nuclei, leading to decay of the signal coherence [69]. This is the cause of transverse relaxation, which causes peak broadening and loss of signal. To counteract the problem of transverse relaxation in large molecules, experiments that use the transverse relaxation-optimized spectroscopy (TROSY) technique are used [75].

1.3.2 Transverse relaxation-optimized spectroscopy (TROSY)

When molecules become sufficiently large, the transverse relaxation rate increase because of decreased mobility and longer correlation time. This a common problem in biomolecular triple-resonance experiments, but can be counteracted by using TROSY-experiments [75].

The ^{15}N and ^{13}C nuclei in the protein is affected by dipole coupling from bound protons and CSA from surrounding electrons, however when the spin of the bound electron is in the N^- state, the dipole and CSA fields oppose, and cancel each other out. TROSY-experiments take advantage of this effect by selecting signals of nuclei that cross correlated with protons in the N^- state. Since dipole-coupling is constant, and CSA is field-dependent, the TROSY effect is more effective at higher fields [69]. TROSY spectra have sharper peaks than their non-TROSY counterparts, but are generally less sensitive. All experiments used in this project use TROSY.

1.3.3 ^{15}N Heteronuclear Single Quantum Correlation (^{15}N-HSQC)

HSQC spectroscopy is a 2 dimensional heteronuclear NMR-technique that transfers the magnetization from amide protons to the nitrogens by a ^{1}J INEPT (Insensitive Nuclei Enhanced by Polarization Transfer) transfer [76] (Figure 12).

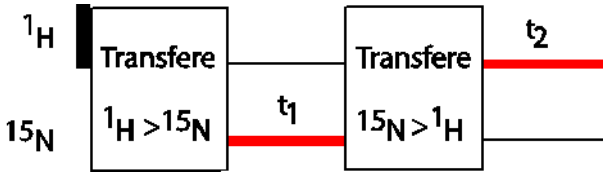


Figure 12: A schematic of the polarization transfer steps in a HSQC experiment (Adapted from figure by Per Eugen Kristiansen).

The HSQC experiment starts by exciting the protons by a 90 degree pulse. The magnetization is transferred to ^{15}N by an $^1J_{HN}$ INEPT transfer [76], where the magnetization is allowed to evolve for a given time period, t_1. After the evolution period, the magnetization is transferred back to the proton by a $^1J_{NH}$ INEPT transfer and measured (t_2). Since only a certain point of the signal can be transferred from the first nucleus, several experiments must be repeated with increments in the t_1 evolution period to obtain a 2D spectrum. The signal from the different ^{15}N time points is then transformed by a 2D fourier transformation to form the indirect dimension.

This spectrum shows peaks in the ^{15}N-^{1}H-dimension, where each peak represents the backbone amide of one amino acid. Some peaks can also belong to side chain groups belonging to Gln, Asn, Arg and Trp [72].
The TROSY version of ^{15}N-HSQC is originally called [^{1}H-^{15}N]-TROSY [77, 78]. In this thesis [^{1}H-^{15}N]-TROSY will be referred to as TROSY-HSQC.

1.3.4 3D experiments for backbone assignment

Triple resonance 3D experiments are used for sequential backbone assignment by correlating the backbone amides with ^{13}C carbons in the same or preceding residue [74, 79, 80, 81,

82, 83]. 3D backbone spectra can be thought of as a normal ^{15}N-HSQC spectrum where the third, ^{13}C, dimension is lying perpendicular to the ^{15}N-^{1}H-plane. The ^{13}C peaks will then appear perpendicular to their correlated amide-peak, and together they are called spin systems. It is important that the protein is uniformly ^{13}C and ^{15}N labeled in these experiments [73].

Triple resonance backbone experiments transfer the magnetization from the amide to nearby ^{13}C by ^{1}J couplings. The difference in $^{1}J_{NC'}$ and $^{1}J_{NC\alpha}$ J-coupling constants (Figure 11) in the the protein backbone allows the experiment to selectively transfer the magnetization via CO (^{13}C') specifically to ^{13}C in the preceding residue (^{13}Ci-1), or to ^{13}C in both the preceding and ^{13}C in the same residue (^{13}Ci) by transferring via $^{13}C_\alpha$ and ^{13}C' [74]. An example of the magnetization transfer pathway for HNCACB [82] is presented in Figure 13. Table 2 lists the 6 common backbone experiments and shows which carbons are observed in the different experiments. The listed experiments are available as TROSY versions [77].

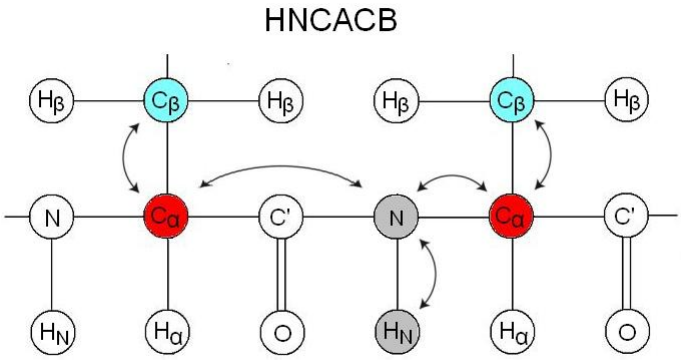


Figure 13: Magnetization transfer pathway for the HNCACB experiment. The magnetization is transferred from $^{1}H^{N}$ via N to $^{13}C_\alpha$ and $^{13}C_\beta$, and back again. This experiment observed the $^{13}C_\alpha$ and $^{13}C_\beta$ in both the preceding an same reside. The corresponding HN(CO)CACB transfer the magnetization via ^{13}C' to observe only the $^{13}C_\alpha$i-1 and $^{13}C_\beta$i-1. (Source: Wikimedia Commons)

Experiments that observe the same carbons on the preceding and same residue, are used in tandem with a different experiment that only observe the carbons in the preceding residue. HNCACB and HN(CO)CACB are the most common spectra used to sequentially assign the backbone amides of the protein. The HN(CO)CACB correlates the amides with the preceding $^{13}C_\alpha$i-1 and $^{13}C_\beta$i-1. The HNCACB correlates the backbone amides with $^{13}C_\alpha$ and $^{13}C_\beta$ in both preceding and same residue. By comparing a spin system in the two spectra, the $^{13}C_\alpha$ and $^{13}C_\beta$ can be identified (Figure 14) When all $^{13}C_\alpha$i $^{13}C_\alpha$i-1, $^{13}C_\beta$i and $^{13}C_\beta$i-1 are identified, the different spin systems can be linked together by comparing the ^{13}Ci-1 values with the ^{13}C values in the different spin systems (Figure 15).

When several spin systems are linked, they can be assigned to the protein sequence

Table 2: Different backbone experiments and their observed nuclei. • indicate observed signal, ∘ indicate observed, oppositely phased signal. Magnetization is transferred via, but not observed, nuclei in parenthesis.

Experiment	Residue					
	n-1			n		
	CO	Ca	Cb	CO	Ca	Cb
HNCO	•					
HN(CA)CO	•			•		
HN(CO)CA		•				
HNCA		•			•	
HN(CO)CACB		•	•			
HNCACB		•	∘		•	∘

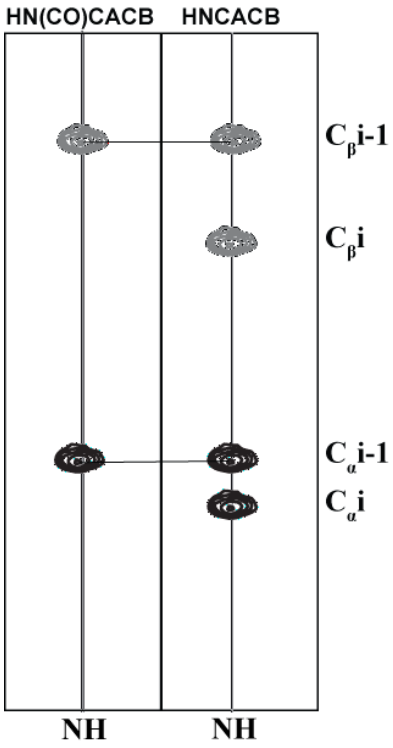


Figure 14: Comparison of a spin system in HN(CO)CACB and HNCACB. HN(CO)CACB correlate the amide ^{1}H and ^{15}N to $^{13}C_\beta$-1 and $^{13}C_\alpha$-1. The two other peaks that show up in HNCACB must therefore be $^{13}C_\alpha$ and $^{13}C_\beta$ correlated to that amide. In practice, the spectra are compared in a computer program (Cara) that allows the user to manually mark the different peaks as ^{13}Ci-1 or ^{13}Ci

by comparing the chemical shift values and other characteristics associated with certain amino acids (Figure 10) e.g. glycine does not have a $^{13}C_\beta$ peak. The chain in Figure 15 has the sequence X-G-X-A-S/T. When a chain of 3 or more candidates are linked, and it contains one or more characteristic residues, it can be tentatively assigned to the sequence. From then on, the chain is extended by linking more spin systems until it meets another chain, or until a discrepancy is encountered. In that case, it has to be assessed whether

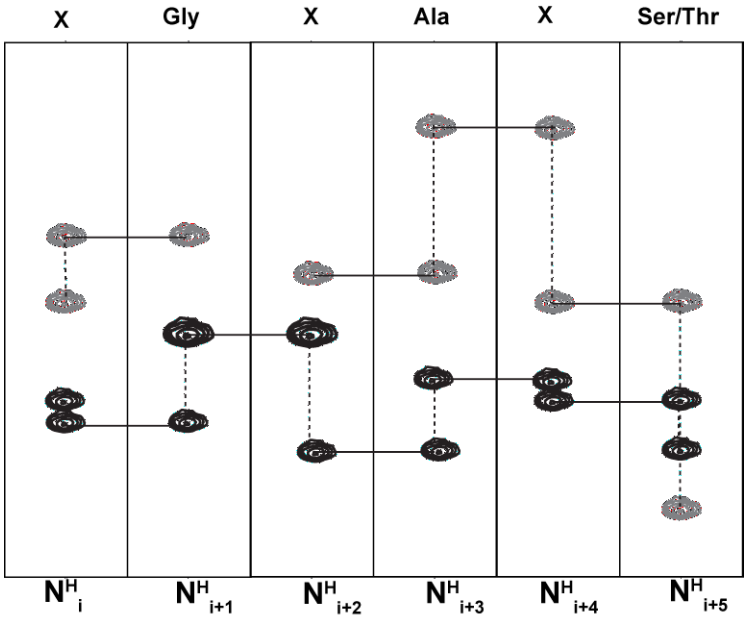


Figure 15: Six spin systems are linked together by comparing the different ^{13}C values. $^{13}C_\alpha$ is in black and $^{13}C_\beta$ in grey.

the assignment so far, or the linking, is incorrect.

1.3.5 Secondary structure prediction

Chemical shift values provide important information about the local chemical environment. The chemical shifts for backbone atoms are strongly dependent on the ϕ and ψ torsion angles of the residues, and by analyzing these chemical shifts, the secondary structure of the protein can be accurately predicted [84]. TALOS+ (Torsion Angle Likeliness Obtained from Shift and Sequence Similarity) is a software that compares input chemical shift values and sequence to a protein databank containing high resolution crystal structures and their observed chemical shifts [85]. The program calculates ϕ and ψ torsion angle restraints by analyzing the chemical shifts of ^{13}C', $^{13}C_\alpha$, $^{13}C_\beta$, ^{15}NH, ^{15}N and $^{1}H^\alpha$, allowing prediction of secondary structures. TALOS+ consistently predicts 88% the dihedral angles, but this value decreases when fewer chemical shift values are present.

1.3.6 Titration experiments

NMR titration experiments is conducted by stepwise adding ligand to an ^{15}N-labeled hLTB solution while collecting a series of TROSY-HSQC spectra at increasing ligand concentrations [72]. Upon ligand binding, the chemical environment of the residues that are involved in the interaction will change, causing a perturbation of the corresponding amide peaks. The binding constant can be calculated by measuring the peak movements as a response to increasing ligand concentrations. If the different amide peaks have been

assigned to the sequence, the residues involved in binding can also be identified [72]. If a structure of the protein is available, it is possible to map the interaction residues onto the structure to localize the binding site.

1.4 X-ray crystallography

X-ray crystallography is a technique in structural biology that allows us to solve structure of biomolecules to atomic resolution. The molecule of interest is crystallized and exposed to an intensive X-ray beam that is scattered into a diffraction pattern and detected. By analyzing the pattern, an electron density map that can be used to construct a model of the molecule. Biomolecular X-ray crystallography is a well established field, and the books by McPherson [86, 87], Blow [88] and Rupp [89] provide a good introduction to this method.

Protein crystals are highly ordered aggregates of proteins with the same orientation that are repeated indefinitely in all three dimension. The smallest repeating unit of a crystal is called the unit cell and may contain one or more molecules, which often are arranged symmetrically. Crystals therefore serve two important functions in X-ray diffraction:

1. The signal is amplified by every molecule in the crystal
2. By locking the molecules in the same orientations, crystals only allow reflections from certain angles and thus sharp diffraction spots to be generated.

When an X-ray beam hits a crystal, it is diffracted by the electrons that surrounds the atoms in the molecule. The waves interfere with each other and form constructive interference at certain angles. Constructive interference is observed by the detector as spots defined by Bragg's law:

$$2dsin\theta = n\lambda \tag{8}$$

where d is the distance between sets of lattice planes, θ is the angle of the incident and outgoing beam and λ is the wavelength. By measuring the angles and intensity of the spots it is possible to treat them as wave functions (called structure factors) with a given direction and wavelength that is defined by the location and intensity of the spot. However, the diffraction pattern provides no information about the phases of the structure factors. This is called the phase problem and it can be solved by molecular replacement. Phases of an already solved structure are applied to the structure factors. By combining the structure factors by Fourier transformation and integrating the wave in three dimensions, an electron density map of the unit cell can be constructed. When the protein sequence is known, the structure can be imposed on the electron density to construct the model. Today, all of these steps are computer-aided and the most time-consuming step in X-ray crystallography is to obtain the crystals.

Crystallization is a complex process that occurs when a molecule is brought into a supersaturated solution and molecules starts forming ordered aggregates (Figure 16). The free energy difference between remaining in solution and forming aggregates decreases and the molecule starts forming small intermediate aggregates that are dissolved as quickly as they are formed. If the concentration is further increased (Figure 16b), the aggregates can form for long enough to allow the nucleus of the aggregate remain stable. When a crystal nucleus has formed, it becomes more energy favourable for the remaining molecules in solution to join the crystal. As the crystal grows it recruits molecules from the solution until the concentration reaches the solubility maximum of the molecule (Figure 16c)

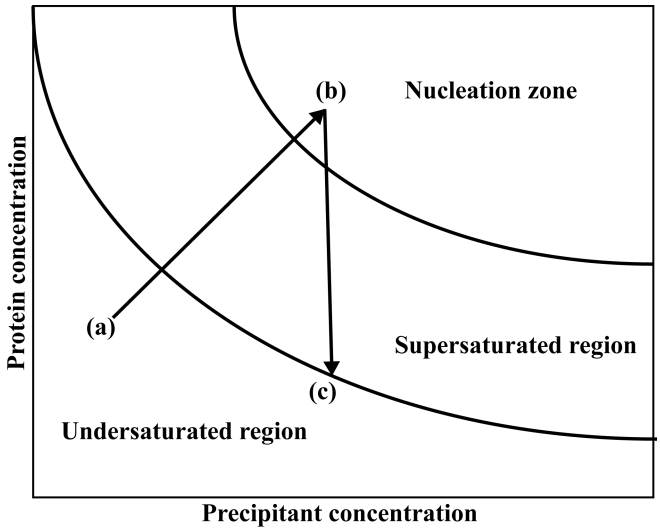


Figure 16: Phase diagram illustrating the process of crystallization. **a.** The molecules are in equilibrium in a undersaturated point. **b.** The concentration is brought into the nucleation zone, forming stable nuclei. **c.** When a stable crystal is formed, it recruits protein from the solution, thereby lowering the concentration until an equilibrium is reached at the supersaturation phase transition.

In practice, the most common crystallization technique is vapor diffusion where a protein-containing drop is stored in a small enclosure with a well/reservoir solution with higher osmolarity/salt concentration (Figure 17). Because of the different solvent potential between the two solutions, the well solution will attract more vapor in an osmosis-like manner and the drop will lose solvent, thereby increasing the concentration of protein and solvent.

Protein crystallization requires very pure samples (>95%) and in addition to the precipitant, the solubility of proteins depends on many other factors including pH, ionic strength, temperature and overall solvent composition. At present time it is not possible to predict which conditions will lead to crystallization and the process of obtaining crystals are still based on trial and error by preparing crystallization solutions with varying conditions. It can be a daunting task to vary every condition in the solution, which is why

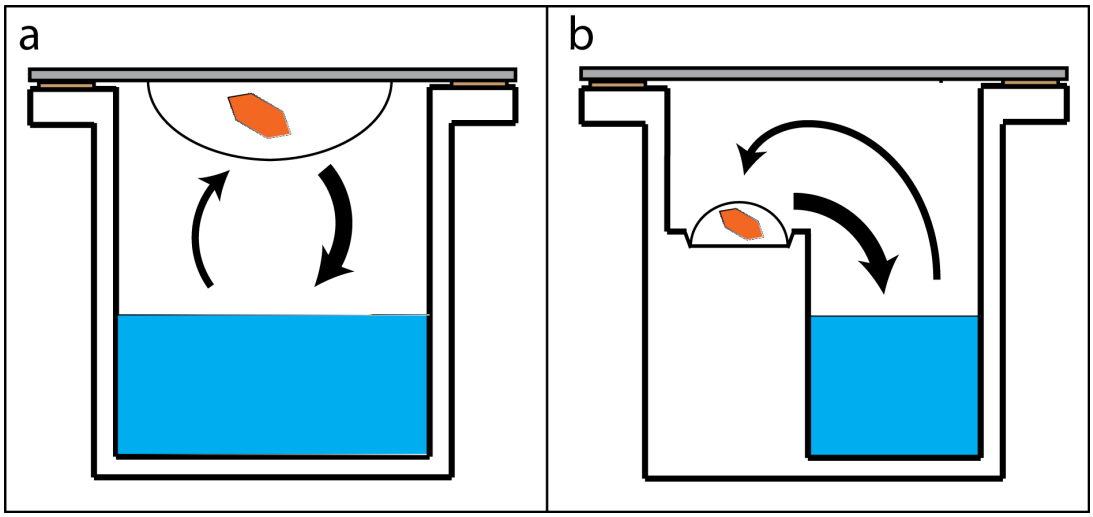


Figure 17: Two common vapor diffusion techniques. **a.** hanging drop. **b.** sitting drop.

commercial screening kits have been developed to speed up the process. These kits are developed based on conditions that most often yield crystals and different kits vary different sets of conditions i.e different precipitants, pH, divalent cations, different cryoprotectants and different detergents (for membrane proteins). When a crystal is obtained, the conditions are usually optimized to increase the diffracting properties. A common optimization technique is to vary two and two conditions, like pH and precipitant concentration, in a 2D-grid.

2 Methods

A list of materials, chemicals and buffers used in this project can be found in appendix B.

2.1 Production and purification of hLTB

The hLTB pentamer was produced by over-expression of the hLTB encoding gene ExtB cloned into a pMMB66EH (Appendix F) and transformed into the non-pathogenic marine bacterium *Vibrio* sp. 60. ExtB is regulated by a *lac*-operon and hLTB is produced by IPTG-induced overexpression.

2.1.1 Production of hLTB

The following protocol for production and purification of hLTB is also the basis for production and purification of isotopically labeled hLTB.

Bacterial cultures were grown in M9 minimal media (Appendix D). Normal culture sizes were 1 L for production of unlabeled protein and 0.2 L for production of isotopically labeled protein, grown in 2 L and 0.5 L baffled Erlenmeyer flasks, respectively. Precultures were 1/10 the volume of the main culture. A preculture was inoculated with frozen stock *Vibrio* sp. 60 and grown overnight in a baffled Erlenmeyer flask at 30 °C at 125 rpm. OD_{600nm} was measured and the culture centrifuged at 8000 $\times$ g for 10 minutes at 4 °C in 50 mL tubes. The supernatant was discarded and the bacterial pellet resuspended in the main culture to a starting OD_{600nm} of 0.1. Protein expression was induced in early exponential growth phase at OD_{600nm} of 0.4 by adding 1 M IPTG to a final concentration of 1 mM. hLTB was harvested by centrifuging the culture for 45 minutes at 10 000 $\times$ g in 50 mL falcon tubes and saving the supernatant. The bacterial pellet was discarded.

Frozen stocks were prepared by centrifuging 50 mL mature culture, collecting the pellet and adding 50% glycerol to a final concentration of 12.5%. The stocks were flash frozen in liquid nitrogen and stored at -80 °C.

2.1.2 Purification of hLTB

After the application of the supernatant to a column of immobilized galactose, the column was washed 3 times with 20 mL PBS-buffer, and hLTB was eluted twice with 10 mL 300 mM galactose elution buffer and collected in 2 fractions. The 20 mL eluate was dialysed 3 times against 2 L buffer for 3 hours, 6 hours and overnight. The dialysis buffer was PBS-, Tris- or NMR-buffer depending on further use. After dialysis, the protein was concentrated by centrifuging in a 20 mL Vivaspin column at 6000 $\times$ g, and the final concentration measured by spectrophotometry at 280 nm by using the theoretical extinction coeficient of hLTB, 0.97 M^{-1} cm^{-1} for hLTB assuming reduced Cys residues (Calculated

by ExPASy (web.expasy.org/protparam)).

Samples for SDS-PAGE were collected during the process for analysis. 15 µL of each sample were mixed with 5 µL 4x SDS-PAGE gel loading buffer and heated to 80 °C for 10 minutes. 15 µL of each sample solution were applied to a 4-12% Bis-Tris gel and run for 35 minutes at 200 V. The gel was stained with Coomasie staining buffer for 2-5 hours and destained overnight with MQ-H_2O.

2.2 Production of ^{15}N-labeled and ^{15}N-^{13}C-labeled hLTB

2.2.1 Production of ^{15}N-hLTB

For the production of ^{15}N-hLTB, $^{14}NH_4Cl$ was replaced with $^{15}NH_4Cl$ for the preparation of the growth media. Otherwise, the protocol remained unchanged.

2.2.2 Production of ^{15}N-^{13}C-hLTB

2.2.2.1 ^{15}N-^{13}C-hLTB production with high cell density

A 0.8 L cell culture was prepared in unlabeled M9-medium and grown according to the protocol in section 2.1. At $OD_{600nm} = 0.6$, the cells were pelleted by centrifugation for 10 minutes at 8000 $\times$ g, washed with 1X M9-salt solution to remove all nitrogen and carbon sources and centrifuged again. The pellet was resuspended in 1/4 volume (200 mL) isotopically labeled M9 minimum media and incubated for 30 minutes before induction. The protein was harvested and purified after 5 hours of expression. Isotopically labeled M9 media was prepared by substituting normal glucose and $^{14}NH_4Cl$ with ^{13}C-glucose and $^{15}NH_4Cl$.

2.2.2.2 ^{15}N-^{13}C-hLTB production with long term expression

The expression rate of hLTB over time was determined by inducing a 1.0 L unlabeled culture and allowing it to express for 40 hours. 50 mL samples were collected every 2 hours for the first 28 hours and at 40 hours. Each sample was centrifuged and 40 mL of each supernatant were applied to an ÄKTA galactose affinity column and washed with 20 mL PBS-buffer. The hLTB was eluted with 0.6 M galactose elution buffer. The resulting elution peaks were integrated by the ÄKTA Unicorn control software and plotted against time. The optical density was also measured at all time points. This experiment was repeated with a 58 hour expression with longer time point intervals and no OD_{600nm} measurements.

To examine whether the addition of glucose resulted in higher yield, two 200 mL unlabeled cultures were grown as described in section 2.1 and allowed to express hLTB for 48 hours and 1 g glucose was added to one culture after 24 hours.

2.2.2.3 Final production of ^{15}N-^{13}C-labeled hLTB

A 200 mL culture was prepared with $^{15}NH_4Cl$ and ^{13}C-glucose as nitrogen and carbon source and grown according to the protocol in 2.1. After induction, the bacteria were allowed to express for 72 hours with addition of 1 g ^{13}C-glucose after 24 hours.

2.3 D_2O training of *Vibrio* sp. 60

All cultures were either 2.5 or 5.0 mL, grown in glass culture tubes at 30 °C, 165 rpm. The tubes were autoclaved and dried in an oven at 120 °C for 20 minutes to remove any residual H_2O prior to inoculation. D_2O containing M9 minimal medium was prepared simply by exchanging H_2O with D_2O. At higher (>87%) levels of D_2O, the component solutions were prepared with D_2O as well. Maximum D_2O concentration in the media was restricted to 99% as the commercial BME-vitamin mix was prepared with H_2O.

For the adaption, a starting culture with 0% D_2O was grown overnight to an OD_{600nm} of 0.6. 1/10 of the volume was then transferred directly to the following culture with higher D_2O concentration. The next culture was then grown overnight and so on. All cultures were pelleted and frozen in 12.5% glycerol stocks after transfer. The transfer and growth steps for the final procedure are listed in Table 3.

Table 3: Transfer and growth steps for D_2O adaption of *Vibrio* sp. 60

Culture#	**D_2O (%)**	**Volume (mL)**	**OD_{600nm} start**	**OD_{600nm} end**
1	0	5	0	0.5
2	30	5	0.04	0.5
3	60	5	0.04	0.5
4	85	5	0.28	0.6
5	86	5	~0.1	1.8
6	87	5	0.12	1.7
7	88	5	0.14	1.7
8	89	5	~0.1	1.8
9	90	2.5	~0.1	1.2
10	91	2.5	~0.1	1.2
11	92	2.5	~0.1	0.9
12	93	2.5	~0.1	0.85
13	94	2.5	~0.1	0.47
14	95	2.5	0.05	0.16
15	96	2.5	~0.1	1.2
16	97	2.5	~0.1	1.4
17	98	2.5	~0.1	1.21
18	99	2.5	~0.1	0.98

2.4 Production of ^{15}N-^{2}H-hLTB

^{15}N-^{2}H-labeled hLTB was produced to estimate the expression yield in D_2O before production of more expensive ^{15}N-^{13}C-^{2}H-hLTB. A 200 mL M9 minimal media culture with $^{15}NH_4Cl$ and unlabeled glucose was prepared with 99% D_2O, inoculated with D_2O adapted *Vibrio* sp. 60, induced at OD_{600nm} 0.4 and allowed to express for 72 hours. 0.5% (1 g) non-labeled glucose was added after 24 hours of expression. Purification as in section 2.1.

2.5 Production of ^{15}N-^{13}C-^{2}H-hLTB

To produce triple labeled hLTB with high isotope incorporation, the culture and preculture were prepared with $^{15}NH_4Cl$, ^{13}C-^{2}H-glucose and 99% D_2O. The cultures were inoculated and grown as in section 2.4 and expressed for 72 hours. 0.5% (w/v) perdeuterated ^{13}C-glucose was added after 24 hours. The protein was purified according to section 2.1, but not concentrated. One 200 mL and two 150 mL cultures were grown in total.

2.6 Hydrogen-deuterium exchange

2.6.1 Heat treatment

The approximate melting temperature of hLTB was determined to 65-70 °C by heating a ^{15}N-labeled hLTB sample in 5 °C increments until denaturation. TROSY-HSQC spectra were collected between every increment to monitor if the structure was properly folded.

H-D-exchange was performed on ^{15}N-^{2}H-hLTB by heating the sample to 65 °C for 50 minutes, cooled to 35 °C and subsequently collecting a TROSY-HSQC spectrum to monitor the number of peaks.

2.6.2 Chemical denaturation and refolding

2.6.2.1 Denaturation

Two batches of ^{15}N-^{13}C-^{2}H-hLTB denatured and refolded. 8.5 M guanidine hydrochloride (GndCl) was added to the 20 mL ^{15}N-^{13}C-^{2}H-hLTB eluate to a final concentration of 6 M guanidine hydrochloride, 125 µg/mL hLTB. 1.43 M mercaptoethanol was added to a final concentration of 5 mM and the final pH adjusted to 8.0 by drop wise adding 1 M NaOH. The solution was left at room temperature overnight to allow complete H-D-exchange.

2.6.2.2 Refolding

The denatured protein was refolded by transferring the solution to a 3500 MWCO dialysis bag and dialysed at room temperature according to the steps in Table 4. Unfolded protein was collected by centrifuging the sample at 8000 × g, 20 °C for 10 minutes and resolubilized in 6 M GndHCl. The supernatant, containing the refolded, soluble protein, was saved for concentration by ultrafiltration.

Table 4: Dialysis steps for refolding hLTB. Glutathione was added in a 1:9 ratio of reduced and oxidized form. The first step was performed without stirring to allow slow equilibration.

Step	Volume	Content after equilibration	Duration
1	0.5 L	0.5 M GndHCl, 3 mM glutathione, NMR-buffer pH 8.0	Overnight
2	2.0 L	1 mM glutathione,NMR-buffer pH 8.0	Overnight
3	2.0 L	NMR-buffer pH 6.5	3 hours
4	2.0 L	NMR-buffer pH 6.5	6 hours
5	2.0 L	NMR-buffer pH 6.5, 0.02% NaN_3	Overnight

2.7 Preparation of LPS incorporated micelles

Based on protocol by Avanti polar lipids [90]. 2.0 mg LPS and 10.5 mg DPC in was partially dissolved in 2 mL 1:1 methanol:MQ-H_2O and mixed until the solution did not turn more clear. The solution was centrifuged at 16 000 $\times$ g for 10 minutes to remove the precipitate and the supernatant transferred to a 10 mL round bottom flask. The solvent was removed by rotary evaporation overnight at room temperature. The remaining lipid/detergent film was dissolved 1 mL 1X NMR buffer while keeping the temperature above 50 °C. The solution was aged overnight at above 50 °C on the rotavapor. The solution was sonicated for 60 minutes and centrifuges at 10 000 $\times$ g for 10 minutes and the supernatant was immediately titrated to 0.1 mM ^{15}N-hLTB in 57 mM (2%) DPC.

2.8 Hydrolysis of the core oligosaccharide

A 400 µL solution of 12.5 mg/mL (2.84 mM) LPS, 2% DPC was adjusted to pH 2.3 by adding 1 M HCl. The solution was heat treated in a PCR machine at 99 °C (maximum temperature) for 2 hours. The pH was adjusted to 6.5 by adding 1 M NaOH and the sample was centrifuged at 16 000 $\times$ g for 20 minutes to remove any precipitate. The supernatant was collected and titrated to 0.1 mM ^{15}N-hLTB.

2.9 NMR-spectroscopy

2.9.1 Sample preparation

All NMR samples were dialysed against 1X NMR-buffer and prepared with 10% D_2O to provide lock signal, and 0.2 mM DSS as internal chemical shift reference. Additional NaN_3 was added to samples with long acquisition time to a final concentration of 0.04%.

2.9.2 Data acquisition and processing

All spectra were acquired on a Bruker Avance II 600 Mhz NMR-spectrometer at 25 or 35 °C. ^{1}H chemical shifts were referenced to 0.2 mM DSS and ^{15}N- and ^{13}C-chemical shifts were indirectly referenced by chemical shift ratios from BMRB. The spectra were processed using either TopSpin 1.3 or 2.6.

Table 5: NMR sample preparation

Sample intention	Volume (μL)	Concentration (mM)	NMR-tube type
Titration	350-400	0.1-0.4	Wilmad 5mm, THIN WALL, 7", 600MHz
Assignment	300	1.0	Shigemi 5mm, D_2O-matched

2.9.3 Assignment

All peaks were picked and assigned in Cara 1.8.4.2 [91]. Chemical shift values were exported to Sparky web client after assignment for chemical shift indexing [84, 85]. All NMR-spectra figures were prepared and exported using Caras 'Print/preview'-function.

2.9.4 Titration experiments

NMR titration experiments were conducted by collecting a series of TROSY-HSQC spectra at increasing titrant concentrations.

Initial LPS titration experiment

The LPS titration experiment was carried out by stepwise adding 20 mg/mL (4.8 mM)) LPS in 1x NMR-buffer to a 0.38 mM ^{15}N-hLTB sample and TROSY-HSQC spectra in between at 25 °C. The additions are listed in table 6. The molecular weight (MW) of LPS-Ra was calculated to 4175 Da based on the structure of Raetz et al [62].

Table 6: LPS titration scheme.

Step	hLTB conc. (mM)	LPS conc. (mM)	[LPS]:[hLTB]
1	0.38	0.03	0.09
2	0.37	0.08	0.20
3	0.37	0.13	0.35
4	0.36	0.18	0.50
5	0.36	0.21	0.57
6	0.35	0.31	0.86
7	0.33	0.57	1.73

Titration LPS in DPC

The LPS titration experiment in DPC was carried out by stepwise adding 10 mg/mL (2.4 mM) LPS in 1X NMR-buffer in 285 mM (10%) perdeuterated DPC to a 0.38 mM hLTB sample in 171 mM (6%) perdeuterated DPC and recording TROSY-HSQC, and 1D spectra in between at 35 °C. The steps are listed in table 7. 570 mM (20%) DPC was added for step 7 and 8. The Sample was sonicated for 10 minutes at ~50 °C between every step.

Table 7: LPS in DPC titration scheme.

Step	hLTB conc.(mM)	LPS conc. (mM)	[LPS]:[hLTB]	DPC (%)
1	0.38	0.06	0.16	6.1
2	0.37	0.12	0.32	6.2
3	0.35	0.22	0.63	6.4
4	0.32	0.41	1.28	6.7
5	0.28	0.63	2.25	7.1
6	0.23	0.96	4.17	7.6
7	-	0.89	-	8.3
8	-	0.76	-	10.0

DPC titration to hLTB

The DPC titration experiment was carried out by stepwise adding 570 mM (20%) DPC in 1x NMR-buffer to a 0.1 mM ^{15}N-hLTB sample and recording TROSY-HSQC spectrum in between, at 35 °C. The additions are listed in table 8.

Table 8: DPC titration scheme.

Step	hLTB conc. (mM)	DPC conc. (mM)	[LPS]:[hLTB]
1	0.1	1.42	14.2
2	0.1	2.83	28.4
3	0.1	5.66	56.6

Kdo-titration

The Kdo titration experiment was carried out by stepwise adding 200 mM Kdo solution in 1X NMR buffer to a 400 µL 0.34 mM hLTB sample at 35 °C and recording an TROSY-HSQC in between. The additions are listed in table 9.

Table 9: KDO titration scheme.

Step	hLTB conc. (mM)	KDO conc. (mM)	Kdo:hLTB ratio
1	0.33	0.24	0.74
2	0.32	0.70	2.21
3	0.32	1.16	3.68
4	0.31	3.39	10.85
5	0.31	7.76	25.14
6	0.29	15.71	53.71
7	0.27	30.69	114.21
8	0.23	54.36	235.22

2.9.5 Calculation of chemical shift perturbations

The perturbation of each peak was measured in Cara [91] and the combined chemical shift change of ^{1}H and ^{15}N was calculated by the formula [92]

$$\Delta\delta = \sqrt{\Delta\delta(^{1}\mathrm{H})^2 + \frac{1}{5}(\Delta\delta(^{15}\mathrm{N}))^2}$$

2.10 Crystallization experiments

2.10.1 Crystallization

Crystallization conditions were initially screened using commercial protein crystallization kits with 96 different conditions (Table 10). The conditions were 0.60 μL sitting drops consisting of 1:1 and 1:3 protein:well solution on 96 well plates. The plates were sealed with tape and stored in low light at 20 °C. All experiments were set up with varying concentrations of hLTB and different ligands. The different concentrations and conditions are showed in Table 10.

Table 10: Screening conditions for hLTB with ligands

hLTB conc. (mg/mL)	**Buffer**	**Ligand**	**Ratio**	**Screen**
6.2 mg/mL	PBS	KDO	1:10	JCSG+, PGA, Morpheus
11.0 mg/mL	PBS	KDO	1:50	PGA, Morpheus
6.2 mg/mL	PBS	KDO	1:100	PGA, Morpheus
8.2 mg/mL	PBS	KDO	1:100	PGA
6.2 mg/mL	Tris	(None)	-	JCSG+, PGA, Morpheus
8.2 mg/mL	Tris	KDO	1:10	PGA, Morpheus

Crystal hits were further investigated and optimized by 24 well hanging drop plates with varying conditions around the specific hit, usually starting with pH-precipitant screenings, followed by other variations like additive screen and different buffers. 1 mL well solution was mixed directly in the different wells and mixed by vortexing the whole plate. The hanging drops were prepared by pipetting 2 μL well solution onto a silanized glass cover slide and then carefully pipetting 2 μL protein solution onto the first drop.

Streak seeding was performed by stroking a cat whisker through the donor drop and then stroking it through the new drop. Microseeding was performed by pipetting 2 μL of the crystal-containing drop and diluting it in 28 μL well solution. The crystals were crushed by adding a glass bead and vortexing for 10 s. 0.4 μL of the resulting solution was added to the new drop.

Crystals were directly picked in 100, 150 or 200 µm litho loops and briefly dipped in cryo solution containing well solution, cryoprotectant and equal amount of ligand. The crystal was then transferred directly into a cryostream at 100 K and stored in liquid nitrogen.

Table 11: Conditions for crystals used in data collection

hLTB conc. (mg/mL)	hLTB:Kdo	Conditions	Cryoprotectant
8.8 mg/mL 1x PBS	1:10	0.6 M Na_2HPO_4, K_2HPO_4, 0.1 M Tris, pH 7.5	20% (V/V) propylene glycol
7.4 mg/mL 1x PBS	1:50	0.5 M Na_2HPO_4, K_2HPO_4, 0.1 M Tris, pH 7.0	20% (V/V) propylene glycol

2.10.2 Data collection

Data was collected at 100 K at the European Synchrotron Radiation Facility (ESRF) on the following beamlines: ID23-1, ID29 and ID30A-1. Collection strategies were suggested by MxCube 2.0 and datasets were collected from all crystals diffracting with higher resolution than 5 Å.

Crystal conditions for data sets that were collected are presented in table 11.

2.10.3 Structure determination and refinement

The data sets were indexed, integrated and scaled by XDS (10.9.4) and merged with Aimless [93]. Phases were solved by molecular replacement with PhaserMR using the previously solved hLTB structure of Holmner et.al (2011) [53] (PDB ID:2O2L). The structure was refined by reiterations of model building in Coot and refinement in Refmac5. The mentioned programs, except for XDS, is a part of the CCP4 (6.4.0) software suite [94]. Figures were prepared with PyMol (1.3) and Coot.

3 Results and discussion

3.1 Production of hLTB

An expression system for hLTB was provided by Timothy Hirst [95]. The plasmid sequence is presented in the master thesis of Siri Trygsland [96].

hLTB is secreted from the bacteria after production [95] and can be harvested by pelleting the bacteria and collecting the supernatant. The protocol for production and purification of hLTB used in this work (Section 2.1), was originally based on a protocol by Heggelund et al (2012) [55]. However, Heggelunds original protocol contained three extra steps that involved precipitation of hLTB from the supernatant, resuspension and dialysis before further purification. By omitting these steps, we increased the overall yield by ~50 %. The yield was further improved by extending the expression time from overnight (~16 hours) to 48 hours (See section 3.2.3).

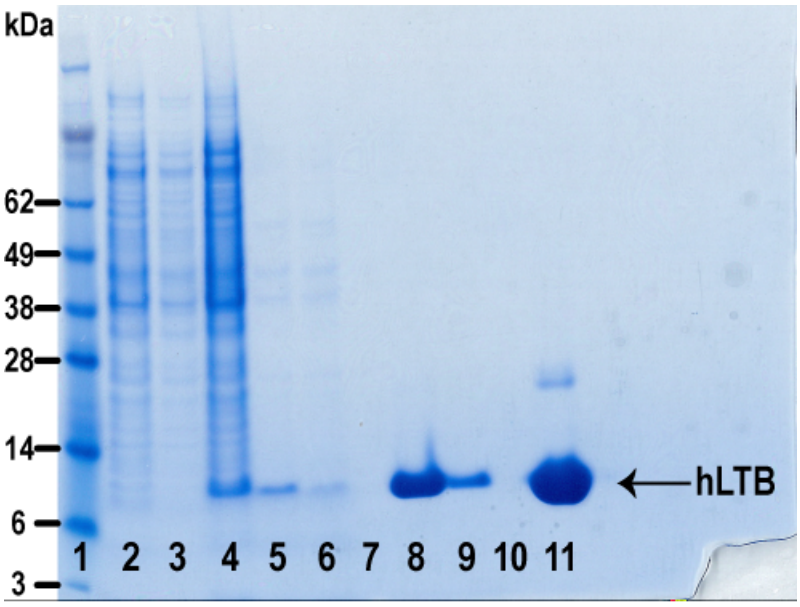


Figure 18: A representative 4-12% Bis-Tris NuPAGE gel of samples collected during production and purification of hLTB. Lane **1:** Seeblue Pre-Stained standard **2:** Preculture after overnight growth **3:** Main culture at induction point **4:** Main culture 48 hours after induction **5:** Supernatant at harvest **6:** Affinity column flowthrough **7:** Wash **8:** 1st eluate fraction **9:** 2nd eluate fraction **10:** (Empty) **11:** Protein after dialysis and concentration (4 µL 5mg/mL loaded)

Samples were collected during production and purification, and analyzed by SDS-PAGE. The theoretical mass of a hLTB subunit is 11.7 kDa and the expected band is visible between the 6 and 14 kDa marker. An unknown band is visible in the overloaded sample (Lane 11) at around 20 kDa (Figure 18). This band may be caused by an impurity,

but considering that it is only visible when the sample is overloaded, it must constitute a small part of the total protein amount. As the NMR and crystallography experiments were not affected in any noticeable way, no measures were taken to further improve the purification protocol. hLTB in the flow through was reapplied to the column to maximize the yield. A typical yield of this production method was around 9 mg hLTB per liter medium.

^{14}N and ^{12}C are the most abundant nitrogen and carbon isotopes, however, they do not have the properties commonly needed for liquid state NMR-spectroscopy (See section 1.3.1). hLTB was therefore isotopically labeled with ^{15}N and ^{13}C by expressing it in a minimal media with $^{15}NH_4Cl$ and ^{13}C-glucose as nitrogen and carbon source. M9 minimal media [97] was chosen for this purpose, and the recipe was modified by increasing the NaCl concentration to 15 g/L to accommodate growth of the marine *Vibrio* sp. 60 (Appendix D).

3.2 Production of ^{15}N-hLTB

^{15}N-labeled hLTB is required for ^{15}N-HSQC spectra for the titration experiments. ^{15}N-hLTB was produced by using $^{15}NH_4Cl$ as nitrogen source when preparing the cultures, and the ^{15}N-incorporation was determined to approximately 98.7% by MS (Appendix E).

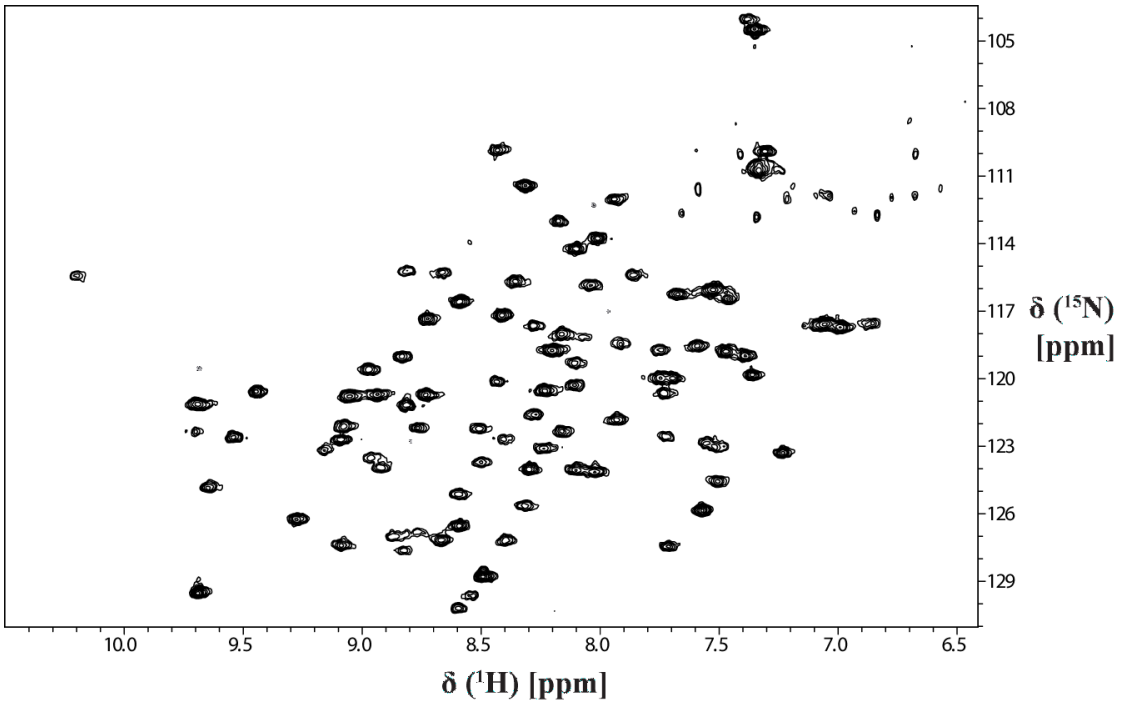


Figure 19: A TROSY-HSQC-spectrum of ^{15}N-hLTB.

A TROSY-HSQC spectrum of ^{15}N-labeled hLTB was collected to ensure that ^{15}N was

properly incorporated, showed in Figure 19. The different peaks represent the backbone amides of the different residues in hLTB. The small peaks in the top right corner are typical for Gln and Asn sidechain amides [72].

3.2.1 Production of ^{15}N-^{13}C-hLTB

^{15}N-^{13}C-labeled hLTB was necessary for triple resonance experiments for backbone assignment (See section 1.3), and is produced by using $^{15}NH_4Cl$ and ^{13}C-labeled glucose as nitrogen and carbon source when preparing the minimal media. However, ^{13}C-glucose is an expensive chemical, making it necessary to optimize the protein yield per amount of glucose used. Two different production methods were attempted to maximize hLTB yield per amount of glucose; short term expression in high density cell medium, and long term expression with addition of glucose. Approximately 4.8 mg hLTB was required for an NMR-sample for triple resonance spectroscopy.

3.2.2 Protocol A: ^{15}N-^{13}C-hLTB production with high cell density

Marley et al [98] developed a method to enhance the yield of ^{15}N-^{13}C-labeled protein by achieving high cell density in a small labeled medium prior to induction. The high cell density was achieved by growing the bacteria until exponential growth phase in a large, unlabeled culture and subsequently pellet, wash and transfer them to a labeled medium 1/4 the size. After a short recovery period to allow growth and production of labeled metabolites, hLTB production was induced by addition of IPTG. To prevent incorporation of ^{12}C due to unlabeled metabolites the expression was stopped after 5 hours when it was assumed that the ^{13}C-glucose was about to be exhausted. All cultures were prepared with $^{15}NH_4Cl$ to ensure highest possible ^{15}N incorporation.

One 800 mL ^{15}N-labeled culture was grown and transferred to a ^{13}C-^{15}N-labeled culture for expression. After a 30 minute recovery period, hLTB-production was induced by IPTG and allowed to express for 5 hours after which hLTB was purified as normal. Samples for SDS-PAGE were collected during the experiment (Figure 20).

Surprisingly, only 0.48 mg ^{15}N-^{13}C-hLTB was produced, or 2.4 mg hLTB per g ^{13}C-glucose used. This was only 1/3 of the normal yield after overnight expression and too low to produce the quantity needed for an NMR sample. It was initially thought that the low yield was caused by a low initial expression rate and the experiment was repeated with 9 hours expression time. However, this did not increase the ^{15}N-^{13}C-hLTB yield sufficiently.

The high cell density method was developed for high yield protein production in *E. coli* at 37 °C. It was suspected that the low hLTB yield could be caused by slower metabolic rate in *Vibrio* sp. 60, which is incubated at 30 °C, or simply by a slower expression rate in *Vibrio* sp. 60. Nevertheless, the low yield and slightly lower ^{13}C-incorporation than

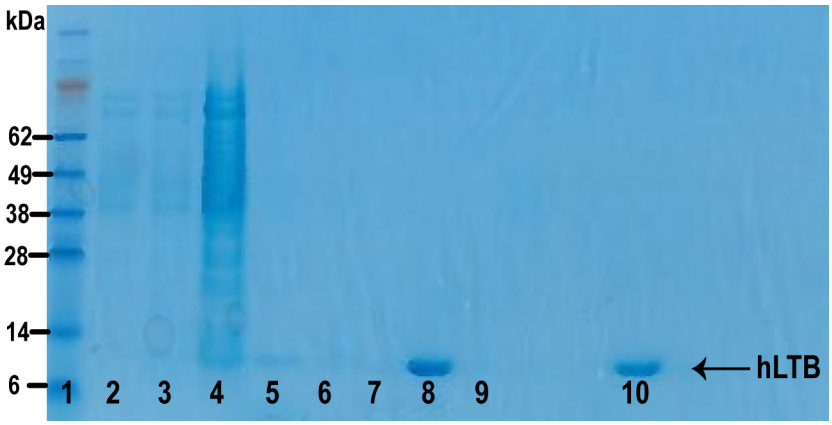


Figure 20: 4-12% Bis-Tris NuPAGE gel of production and purification of non-labeled hLTB with enhanced cell density. Lane **1:** Seeblue gel standard **2:** Preculture after overnight **3:** Main culture at volume reduction **4:** Main culture after volume reduction and 5 hours expression **5:** Supernatant at harvest **6:** Affinity column flowthrough **7:** Wash **8:** 1st eluate fraction **9:** 2nd eluate fraction **10:** Attempted overload of hLTB

expected, made this an unsuitable production method.

3.2.3 Protocol B: ^{15}N-^{13}C-hLTB production with long term expression

Long term expression

Since 5 hour production of hLTB at high cell density resulted in a lower yield than normal overnight production, the expression rate of hLTB over time was examined.

To investigate the production rate of hLTB, a culture prepared with normal glucose was allowed to express hLTB for 40 hours. 50 mL samples were collected and OD_{600nm} was measured at regular time points. The concentration was determined by binding hLTB on an affinity column, briskly eluting, and integrating the elution peaks (Section 2.2.2.2). The yield after 48 hour production was determined by allowing a 0.2 L culture to express hLTB for 48 hours and purifying as normal. The experiment was repeated with 58 hour expression since the protein concentration in the first experiment did not reach a plateau.

The plot in Figure 21 shows that the hLTB concentration in the growth medium increases with an almost constant rate for 40 hours after induction, even after the bacteria enter the stationary phase. This implies that hLTB production continues until the glucose has been consumed and all nutrients and metabolites are exhausted. However, the plot does not show the protein concentration inside the cells and it could be possible that the majority of hLTB was produced during the exponential growth phase, and that the observed increase of hLTB concentration in the medium is caused by slow secretion. Nevertheless, this shows that longer expression time increases the hLTB yield. It should be

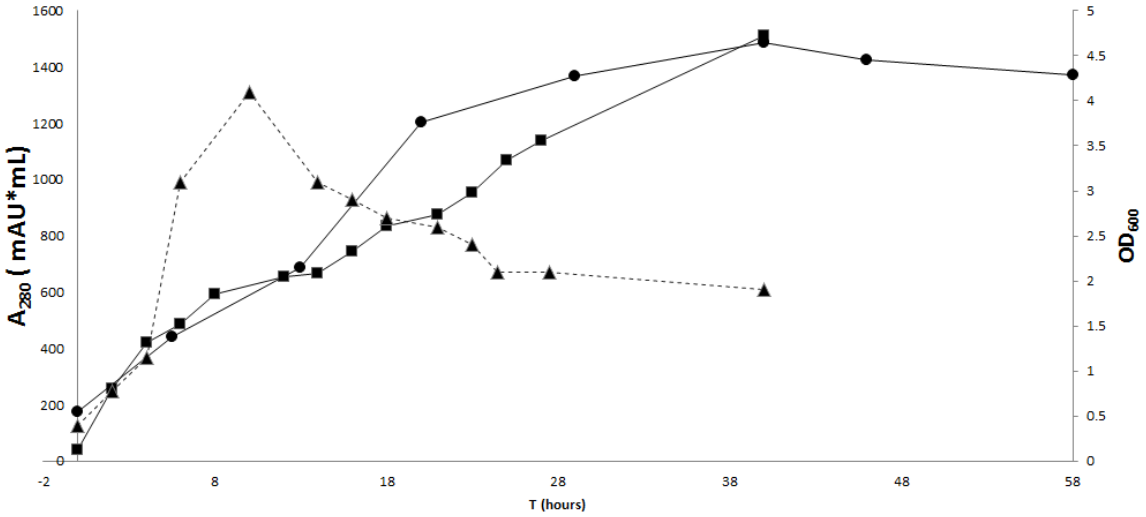


Figure 21: Protein concentration in supernatant and OD_{600nm} plotted versus time. ■ = hLTB concentration, first experiment, • = concentration of hLTB second experiment, ▲ = OD_{600nm}. The plot shows that the protein concentration in the supernatant increases until 40 hours have passed, even after the bacteria enters the stationary phase.

noted that these results are consistent with the findings of Lebens et al, who conducted a similar experiment for expression of CTB in *E. coli* [99]. The yield after 48 hours of expression was 1.8 mg hLTB per gram glucose used, ~3.5 times higher than protocol A and ~50% higher than the normal yield after overnight expression (~16 hours).

hLTB production with addition of glucose

Since hLTB is secreted from *Vibrio* sp. 60, overexpression will not cause an accumulation in the bacteria [95]. The limiting factor in hLTB production therefore appears to be exhaustion of nutrients. It was speculated that addition of glucose to a mature culture after 24 hours expression would increase the fraction of glucose used in production of hLTB, since the bacteria would then already be at a high concentration and use less glucose for growth. To investigate whether addition of more glucose during expression would result in an higher yield, a 200 mL culture was grown, and hLTB expressed for 48 hours. The initial glucose content was replenished after 24 hours. The final protein was purified and the yield quantified. The yield compared to production without glucose is presented in Table 12.

By adding the initial amount of glucose halfway through a 48 hour expression, the hLTB yield per g glucose was increased by two thirds (Table 12). Although more cultures should have been grown to confirm this result, it still suggests that a higher fraction of the glucose was used directly in hLTB production compared to 48 hour expression without addition of glucose.

Table 12: Comparison of hLTB yield after 48 hour expression with and without addition of glucose. Protein yield per g glucose was calculated with glucose used in preparation of the preculture

	w/o additional glucose	**With additional glucose**
mg hLTB per g glucose	1.65	2.28
mg hLTB per L culture	9.1	24

Several cultures during this project were grown with initial glucose concentrations varying from 0.4% to 2%. Although they are not directly comparable to the above results, none of them has a higher yield. A more systematic approach could have been chosen regarding initial glucose concentrations, but the final yield achieved in this section was considered sufficient.

Final ^{15}N-^{13}C-hLTB production

Following the results in section 3.2.3, the protocol for production of ^{15}N-^{13}C-labeled hLTB was extended to 72 hours to ensure complete exhaustion of nutrients. The initial glucose content was replenished after 24 hours. The final yield in a 200 mL culture was 5.2 mg ^{13}C-^{15}N-hLTB, or 2.48 mg ^{15}N-^{13}C-hLTB per g glucose. The ^{13}C-incorporation was determined to approximately 99.7% by MS (Appendix E).

3.3 D_2O adaption of *Vibrio* sp. 60

An TROSY-HNCO spectrum was successfully recorded using ^{15}N-^{13}C-labeled hLTB. However, no usable spectra correlating the backbone amides to $^{13}C_\alpha$ or $^{13}C_\beta$ could be obtained, presumably due to extreme peak broadening caused by direct dipole coupling to nearby protons (Described in section 1.3.1). It therefore became necessary to deuterate hLTB to decrease the effect of the dipole coupling.

Deuteration of hLTB is achieved by using the final protocol for production of ^{15}N-^{13}C-hLTB (Section 3.2.3) with 99% D_2O (v/v) and perdeuterated ^{13}C-glucose as carbon source, however, D_2O is toxic to organisms and bacteria must be adapted to high D_2O concentrations [100]. The protocol for D_2O adaption of *Vibrio* sp. 60, based on an article by Sosa-Peinado (2000) [101], is described in section 2.3. *Vibrio* sp. 60 was successfully adapted to 99% D_2O and the different growth steps are presented in Table 3. The OD_{600nm} was monitored in cultures with 30%, 60%, 90% and 99% D_2O and are presented in Figure 22.

D_2O alters hydrogen bond distances and reaction kinetics, which affects many cellular systems involved in metabolism, transport functions and biosynthesis of macromolecules [100]. This effect can be toxic to bacteria and growth in high D_2O-concentrations is char-

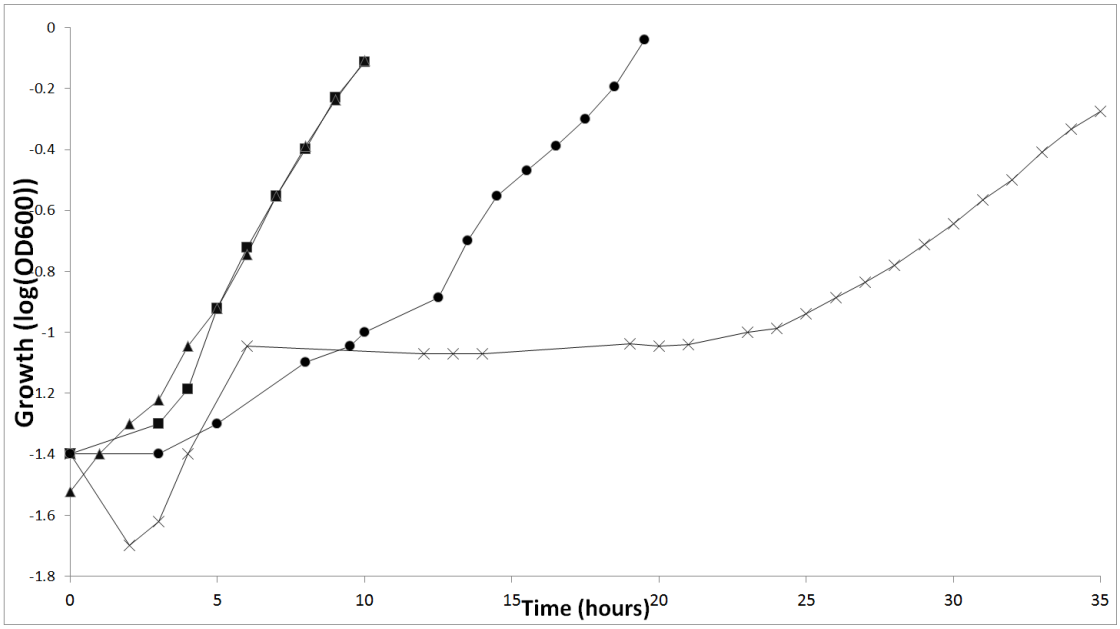


Figure 22: Semilogarithmic plot of *Vibrio* sp. 60 growth rates at different D_2O concentrations; ■ = 30%, ▲ = 60%, • = 90%, × = 99%. The figure illustrates the longer lag time and decreased growth rates as the D_2O concentration increase. The initial drop and quick recovery of OD_{600nm} in the 99% culture is probably caused by measurement error.

acterized by a long lag phase and reduced growth rate. The long lag phase is often a result of initial cell death, which causes a mutagenic selection of D_2O tolerant bacteria [102], and bacterial adaption to high D_2O concentrations is linked to a significant increase of the spontaneous mutation rate [103]. The growth rate usually returns to normal when the bacteria are adapted to D_2O.

Vibrio sp. 60 appears to be particularly fragile at D_2O concentrations above 85-90%. The culture died if it was pelleted between transfers, requiring that the new cultures were inoculated by direct transfer from the previous. Despite direct transfers, inoculation of cultures at >85% D_2O was followed by flocculation and an excessively long, 10-24 hour, lag phase (Figure 22) indicating initial cell death. Greater increments than 1% usually resulted in no growth after 48 hours. Furthermore, even after adaption to 99%, the growth rate did not improve. These problems appear to be greater for *Vibrio* sp. 60 than other bacteria [102, 103], and may be explained by an effect of D_2O that increases the osmotic pressure on bacteria [104]. This could affect *Vibrio* sp. 60 stronger due to the high salt content in the media (Appendix D).

3.4 Production of ^{15}N-^{13}C-^{2}H-hLTB

There are two general strategies for deuterating proteins [73]: random fractional deuteration and perdeuteration. The former is achieved by expressing bacteria in a D_2O/H_2O

fraction similar to the wanted incorporation degree. Perdeuteration, or complete, deuteration replaces ~99% of all protons with deuterons. Perdeuteration is achieved by growing bacteria in 99-100% D_2O with deuterated ^{13}C-glucose as carbon source and was the chosen strategy in this work to maximize the deuterium incorporation. Prior to ^{15}N-^{13}C-^{2}H-hLTB production, a 200 mL ^{15}N-^{2}H-labeled culture was produced to determine the yield.

^{15}N-^{13}C-^{2}H-hLTB was produced by growing a 200 mL and two 150 mL cultures in 99% D_2O with perdeuterated ^{13}C-glucose and $^{15}NH_4Cl$ as carbon and nitrogen source, respectively. The initial glucose content was replaced after 24 hours and the cultures were expressed for 72-84 hours. The total yield was ~5 mg, or 10 mg/L culture. This lower than normal yield was expected, as protein expression levels are lower in D_2O [73]. The ^{2}H incorporation was determined to 97.4% (Appendix E). This was slightly lower than expected since ^{15}N-^{13}C-^{2}H-hLTB was expressed in 99% D_2O. The lower ^{2}H-incorporation could be explained by incorporation of ^{1}H from the perdeuterated ^{13}C-glucose used to prepare the solution (97% ^{2}H), or because some biological reaction mechanisms may have a higher preference for H_2O than D_2O [100]. Nevertheless, the incorporation degree was high enough to acquire triple resonance spectra and did not cause any known problems.

3.4.1 Hydrogen-deuterium exchange

The backbone amides of perdeuterated hLTB are not observable by NMR before the amide deuterons are exchanged with protons (Section 1.3.1). Most amide deuteriums are spontaneously exchanged for hydrogen during the purification of perdeuterated hLTB [71]. However, a TROSY-HSQC spectrum of newly purified ^{15}N-^{2}H-hLTB lacked many peaks compared to a non-deuterated sample (Figure 23), indicating that the corresponding residues were buried in the protein and not solvent accessible [73]. Given enough time, buried backbone amides will exchange with the solvent due to spontaneous unfolding and refolding of the protein, however, for stable globular protein complexes, like hLTB, this process may be very slow [102]. To exchange the buried amide-deuterons, two approaches were attempted; heat treatment, and refolding of the protein.

Heat treatment

Proton exchange of buried residues is increased by heating the protein to near melting temperature. The increased thermal motion of the protein will then allow buried residues greater solvent contact [71]. The protocol is described in section 3.4.1.

A ^{15}N-^{2}H-labeled hLTB sample was heated to 60 °C for a total of 50 minutes to allow H-D-exchange. The subsequently collected TROSY-HSQC shows that approximately 9 peaks returned (Figure 23), however, it still missed many residues compared to a non deuterated TROSY-HSQC (Figure 19). To avoid gaps in the sequential assignment, it was important to exchange the maximum amount of amide deuterons. Furthermore, the weak intensity of some of the returned peaks indicate that prolonged heat treatment is

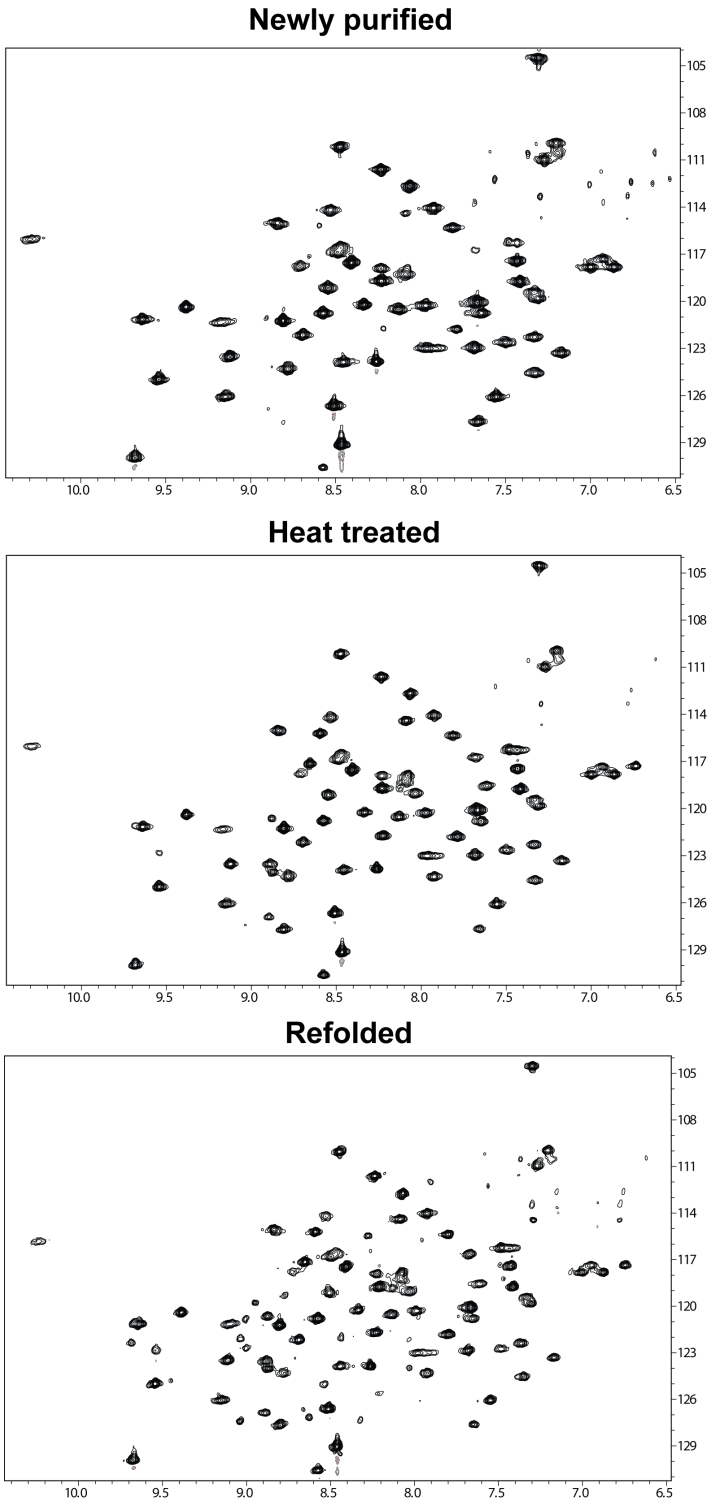


Figure 23: TROSY-HSQC spectra collected of newly purified, heat treated, and refolded ^{15}N-^{2}H-hLTB. The different spectra illustrate the return of amide peaks as a result of H-D-exchange. Approximately 9 peaks returned as a result of heat treatment and additional 13 peaks returned by refolding, compared to newly purified ^{2}H-^{15}N-hLTB.

necessary for complete exchange, thereby risking degradation of the protein. Subsequently, this method was not pursued further.

Denaturation and refolding of hLTB

In order to allow buried residues access to the solvent, a protein can be denatured and subsequently refolded. Solubilization and refolding of proteins is a problem commonly encountered when working with inclusion bodies [105]. Proteins can be denatured by adding chaotropic denaturants like GndHCl or urea to the solution. The refolding process is initiated by reducing the denaturant concentration by dilution or dialysis. During refolding, there is a competition between protein renaturation and aggregation that is affected by conditions like protein concentration, pH, temperature and the redox environment. Optimal refolding conditions are unique for every protein and must be determined experimentally to attain high refolding yields [105, 106].

hLTB was denatured in 6 M guanidinium chloride with mercaptoethanol to reduce the C9-C86 disulfide bond. Initial refolding was attempted on ^{15}N-^{2}H-labeled hLTB by dialysing away the guanidinine hydrochloride overnight, however, this resulted in aggregation and a low refolding yield. The aggregate was collected by centrifugation and resolubilized with guanidine hydrochloride for new refolding attempts. By an iterative process of dialysis, aggregation and resolubilization, the refolding protocol was gradually improved until the final protocol, presented in section 2.6.2.2, was developed. The main improvement of the yield was caused by three factors: the addition of a glutathione redox system to aid the formation of correct disulfide bonds, lower initial hLTB concentration to prevent aggregation of unfolded hLTB, and an additional overnight dialysis step at 0.5 M guanidine hydrochloride. A TROSY-HSQC spectrum of refolded ^{15}N-^{2}H-hLTB is presented in Figure 23. Although all peaks returned, some were weaker than compared to a ^{1}H-^{15}N-hLTB sample, suggesting incomplete H-D exchange, probably because of the initial ^{15}N-^{2}H-labeled hLTB sample was allowed a short exchange time. As a result, newly purified ^{15}N-^{13}C-^{2}H-hLTB was left denatured overnight prior to refolding.

The final protocol improved the yield of properly refolded hLTB from around 5% to 60%. The cumulative refolding yield was increased to 70-80% by collecting the aggregate, resolubilizing it and repeating the refolding process. It is unknown why the last 20-30% of the sample could not be refolded, but it could be caused by unfolded protein binding to the dialysis membrane or degradation of the protein during handling. Nevertheless, this method allowed complete H-D-exchange of all deuterons in hLTB, and the loss was acceptable.

3.5 Backbone assignment

The backbone amides of the pentameric hLTB was sequentially assigned in order to study the interaction between hLTB and LPS, Kdo and NEO by TROSY-HSQC titration experiments. To identify the residues involved in binding, it was necessary to perform a backbone

assignment where the different spin systems are sequentially linked and assigned to the protein sequence by comparing and analyzing the chemical shifts of the ^{13}C resonances in the protein backbone.

3.5.1 Data acquisition

Triple resonance spectra for backbone assignment are relatively insensitive [74], and a high sample concentration is required to collect usable spectra within a reasonable time. For hLTB, the maximum solubility allows approximately 4.8 mg in a normal 400 μL sample volume, however, only 3.6 mg useable hLTB was obtained necessitating the use of Shigemi tubes. These allow a smaller sample volume ($\geq$280 μL) without loss of field homogeneity. The final sample volume was 300 μL, collected at 35 °C. This temperature was chosen because it increased the resolution of the TROSY-HSQC spectrum, increased the solubility of hLTB and was closer to physiological temperature.

hLTB is a large complex of 58.5 kDa, and experienced severe peak broadening and signal loss in normal experiments due to short relaxation time [73]. It was therefore necessary to use TROSY experiments in order to acquire good quality spectra of the protein [75]. TROSY-HNCO and TROSY-HN(CA)CO spectra were initially acquired from the ^{15}N-^{13}C-labeled sample. However, the TROSY-HNCO spectrum displayed broad peaks (Figure 24), and and no peaks were visible in the TROSY-HN(CA)CO spectrum, suggesting that direct spin-spin coupling to nearby protons increased the transverse relaxation rate for ^{13}C, causing peak broadening and loss of signal [74]. This is especially significant for experiments that transfer the magnetization via $^{13}C_\alpha$ and $^{13}C_\beta$, which have directly bound protons [107]. Relaxation is a common problem in liquid state NMR of large proteins and a ^{15}N-^{13}C-^{2}H-labeled sample was produced to reduce the relaxation problems [74, 73]. The following spectra were initially acquired from ^{15}N-^{13}C-^{2}H-hLTB for backbone assignment: TROSY-HNCO [79], TROSY-HN(CA)CO [83], TROSY-HNCA [79], TROSY-HN(CO)CA [82], TROSY-HNCACB [82] and TROSY-HN(CO)CACB [81, 80, 77]. From here on, these spectra will be referred to without the TROSY-prefix. Deuteration of the protein was especially advantageous for the hLTB spectra, since it increased the overall sensitivity and resolution significantly [108] (Figure 24).

Deuteration of the protein cause an isotope effect that alters the ppm-values of the remaining nuclei [73]. The ppm values of the $^{1}H^{N}$ were generally reduced by ~0.25 ppm, and ^{13}C by ~2.8 ppm (Figure 24). However, these shifts were usually consistent and the spectrum was easily comparable with non-deuterated spectra by calibrating for the change in the different dimensions.

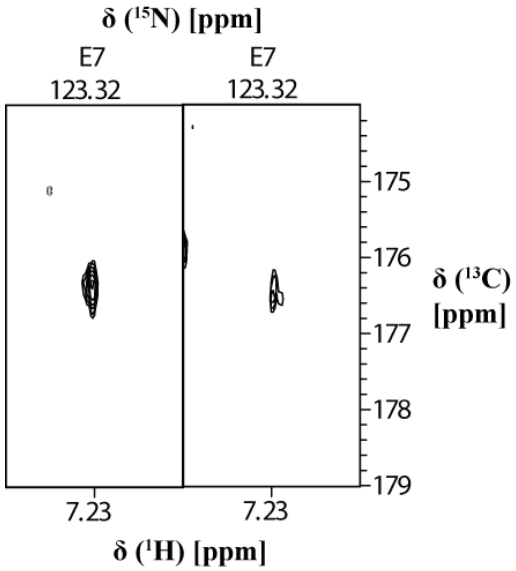


Figure 24: Comparison of two HNCO experiments collected on ^{15}N-^{13}C-^{2}H-hLTB (left) and ^{15}N-^{13}C-hLTB (right). The spectra illustrates the improved signal to noise in the ^{13}C dimension in deuterated hLTB compared to non-deuterated.

3.5.2 Assignment

The backbone assignment process is described in section 3.5.2. Following is a more detailed description of the assignment process of hLTB.

Initial peak picking and overlapping peaks

HN(CO)CACB and HNCACB may be enough to assign hLTB, however, hLTB is a large complex, and HNCA, HN(CO)CA, HN(CA)CO and HNCO were collected as well to aid the assignment. Following the acquisition of the spectra, the process of comparing spectra to identify the different ^{13}C'(i), ^{13}C$_\alpha$(i), ^{13}C$_\beta$i, ^{13}C'i-1, ^{13}C$_\alpha$i-1 and ^{13}C$_\beta$i-1 was initiated. We observed that the intensity of the peaks varied considerably from region to region in the different spectra. Careful comparison of all spectra was therefore necessary in order to assign many ^{13}C-peaks.

Figure 25 illustrates the quality of the different spectra for a typical spin system, belonging to T80. The HN(CO)CA spectrum, correlating the amides to ^{13}C$_\alpha$i-1 , usually provided strong peaks. The HNCA, correlating with both ^{13}C$_\alpha$i-1 and ^{13}C$_\alpha$ was considerably weaker, and the ^{13}C$_\alpha$i-1 is barely visible in the spectrum. HN(CO)CACB was usually strong, but disturbed by noise of unknown origin, possibly an artifact of the experiment. HNCACB was most prone to variation with some regions being well resolved, and others with so poor signal to noise that it was completely dependent on the other spectra in order to assign the spin system.

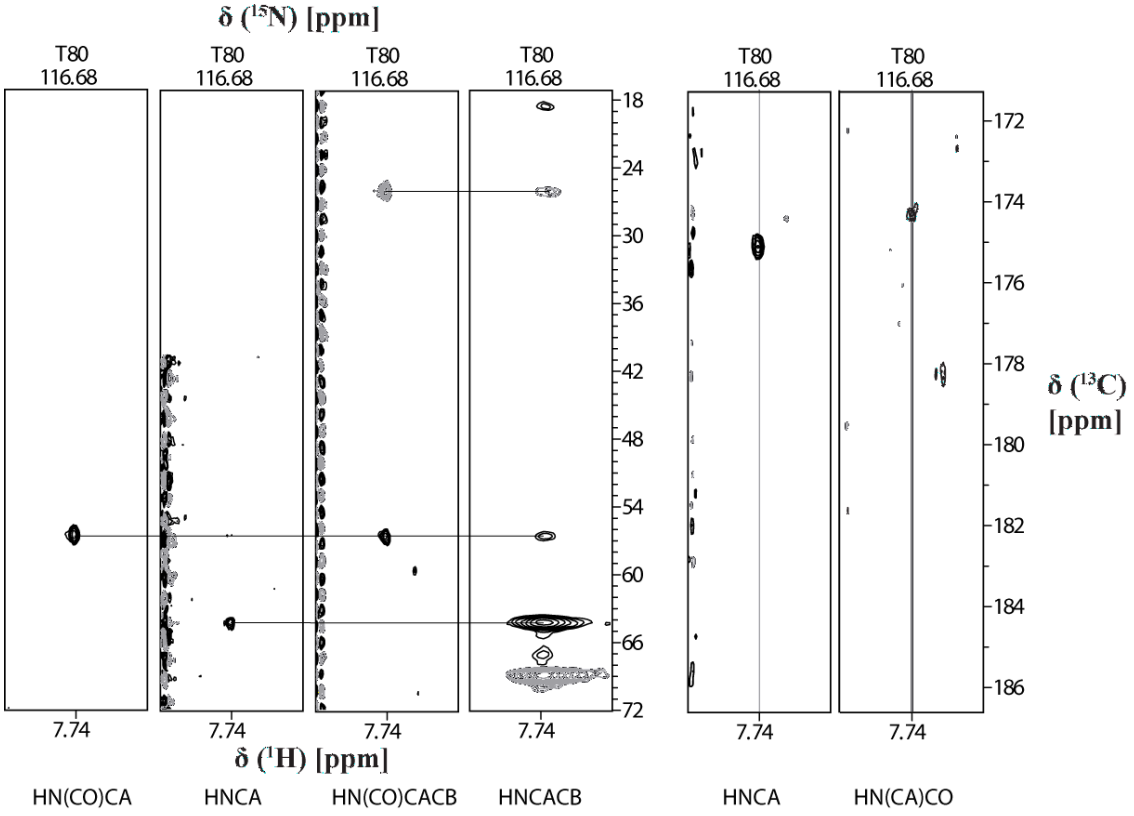


Figure 25: Strip view of the T80 spin system in different spectra, visualized in the ^{15}N-^{13}C plane. This figure illustrates the different quality in spectra for a typical spin system.

Overlapping peaks in the TROSY-HSQC spectrum causes the peaks in the ^{13}C dimension to appear directly above each other. This made it hard to differentiate the different spin systems since it is not known which ^{13}C peak that is correlating with which backbone amide. Furthermore, peaks in the ^{13}C dimension occasionally overlap, making the assignment process even harder. This problem is illustrated in Figure 27 for S30, which lies in a crowded region between 8.0-8.25 ^{1}H-ppm and 116-120 ^{15}N-ppm, containing 5 spin systems (S30, H57, L77, M68 and F42) in what appears to be 3 peaks (Figure 31). Several overlapping spin systems were resolved by carefully comparing the location of the ^{13}C peaks in the ^{15}N-^{13}C- and ^{1}H-^{13}C-planes to determine the exact ^{15}N-^{1}H-values of the amide correlated to each ^{13}C. Different ^{13}C-peaks with similar ^{15}N-^{1}H-values could then be grouped together to form a spin system.

The overlapping spin systems that were not distinguishable by measuring the ^{15}N-^{1}H alone, could be identified by tentatively assigning one ^{13}C peaks to an arbitrary spin system and match other spin systems with similar $^{13}C_{\alpha}$ values (Figure 26). The different candidates could then be assessed by comparing the other peaks the spin system. However, this method was not reliable before ~50% of the spin systems had been assigned due

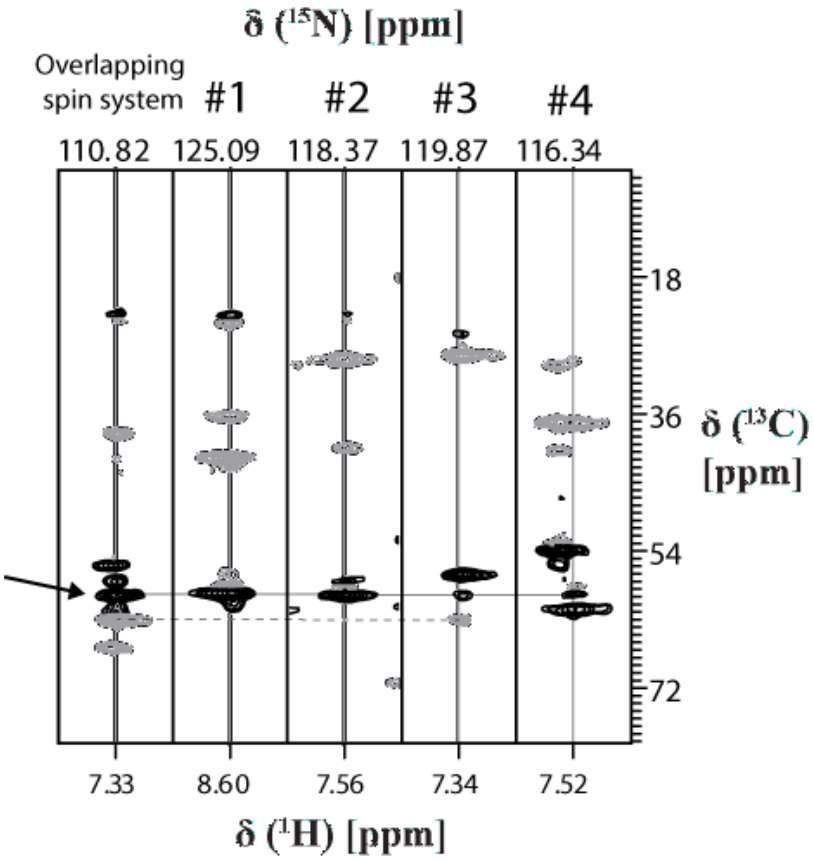


Figure 26: Strip view of an HNCACB spectrum in the ^{15}N-^{13}C plane displaying 5 undefined spin systems. One $^{13}C_{\alpha}$ in the overlapping spin systems has been picked (shown by arrow) and candidates with similar ^{13}C-ppm values are lined up. In this example, candidate #3 has a matching ^{13}C-ppm value for both $^{13}C_{\alpha}$ and $^{13}C_{\beta}$, indicating that it is a neighboring residue. This also indicates that the matching $^{13}C_{\beta}$ belongs to the same spin system as the original $^{13}C_{\alpha}$

to high number of possible candidates.

Linking and assigning the spin systems

When the majority of spin systems are picked, they can be linked together to form chains. The linking process is often time consuming since $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ in most residues show up within relatively narrow regions between 20-35 and 50-60 ^{13}C-ppm. With a large number of residues, every spin system can have several good linking candidates and care must be taken to prevent erroneous linking.

Chains of 3 or more spin systems were tentatively assigned to the sequence by identifying characteristic residues (Figure 10) and searching the sequence for matches. Figure 27 shows a chain of 6 spin systems assigned to S30-R35. Three different spins are characteristic, the S, A and G, indicating that the chain represents the residue sequence S/T-X-A-

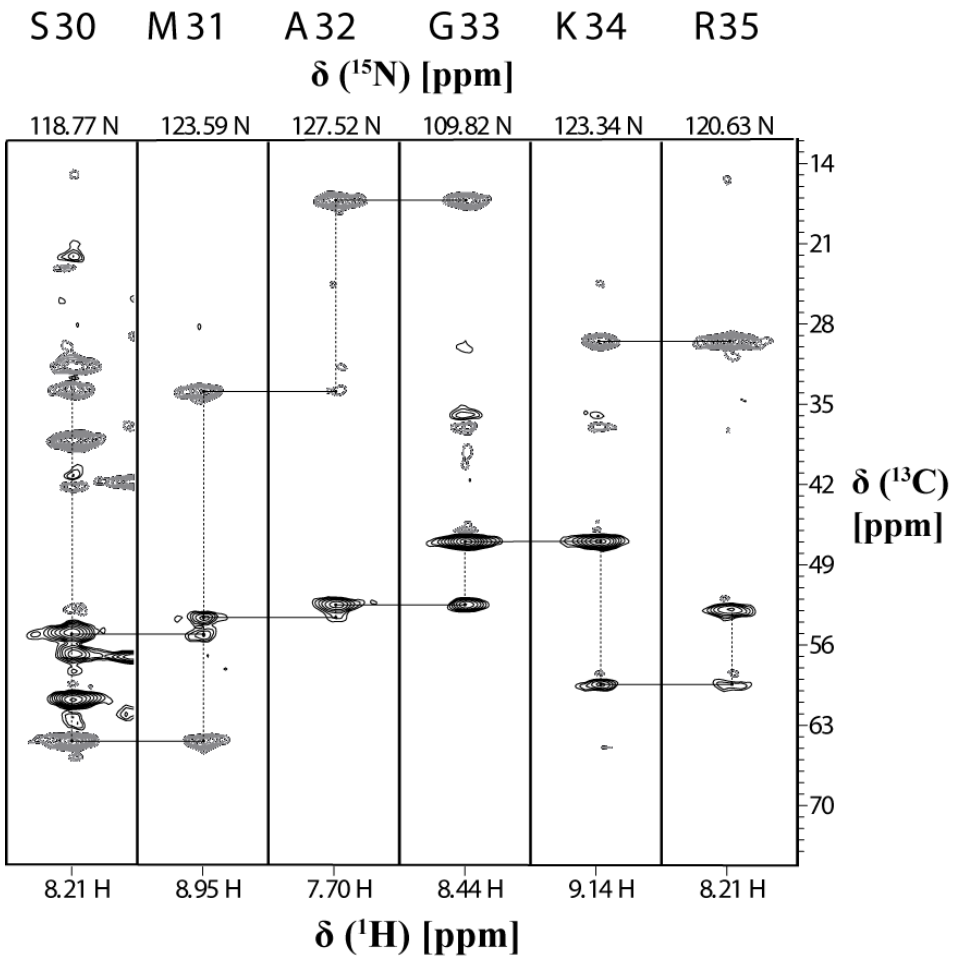


Figure 27: HNCACB. strip view view of residue 30-35 in the ^{15}N-^{13}C plane. The $^{13}C_{\alpha}$s (black) and $^{13}C_{\beta}$ (grey) are linked to the corresponding peaks in the next system by black lines. Dotted line indicate that they belong to the same spin system. The spin systems linked to S30, A32 and G33 display clear characteristics of T/S, A and G spin systems. G33 does not have a $^{13}C_{\beta}$, interrupting the linking of $^{13}C_{\beta}$. The high number of peaks in S30 is explained by overlap with H57.

G-X-X. This chain could therefore be assigned to residues S30-R35 since that is the only place in the sequence that fits. The remaining spin systems, although not very characteristic, were close to the average values (Figure 10). The link between K34 and R35 looks ambiguous in the HNCACB spectrum, because of the weak $^{13}C_{\alpha}$ and $^{13}C_{\beta}$. However, the corresponding HN(CA)CO spectrum confirms the link (Figure 28). This illustrates how the different spectra are used together to link and assign the spin systems.

At around 60% assignment, the remaining spin systems were characterized by poor signal to noise and the assignment therefore became a process of examining the noise level of the different spectra in an attempt to identify weak peaks. When a possible peak was identified it, was assigned to a spin system and compared with other spin systems. Possi-

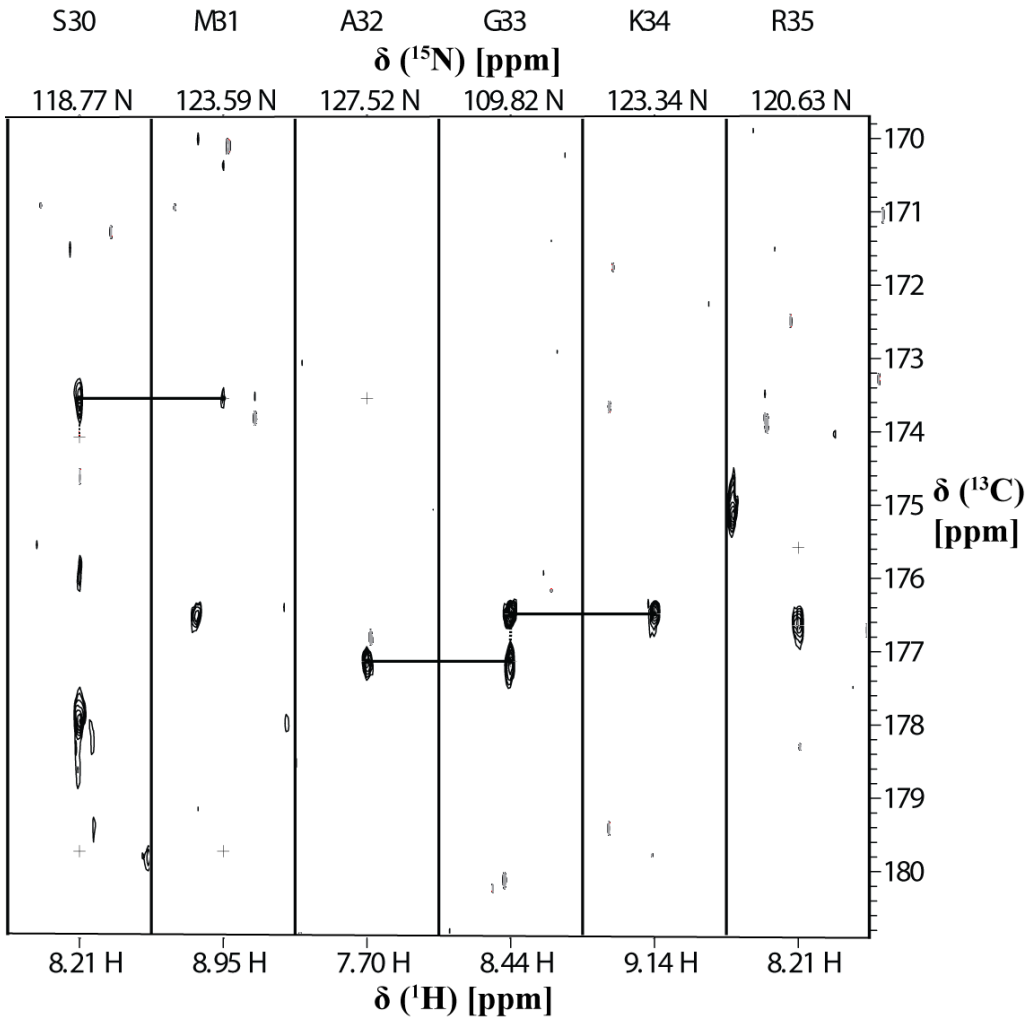


Figure 28: HNCACO. Strip view of residue 30-35 in the ^{15}N-^{13}C plane.

ble matches were linked and tentatively assigned to a part of the sequence and assessed in a iterative process leading to an eventual 74% assignment of the residues. The unassigned residues were mapped onto a crystal structure, revealing that the majority of the remaining residues were located in a large α-helix (Residues K62-T78). This indicates that the poor signal to noise in these residues could be caused by transverse relaxation and peak broadening by the rigid structure of the α-helix [73].

A 3D NOESY-[^{1}H-^{15}N-^{1}H]-ZQ-TROSY [109] (Called NOESY-HSQC from now on) was collected to aid further assignment (Figure 29). This experiment correlates the backbone amides with all neighboring protons by the NOE-effect, indicating which amides are sequential neighbors. This was, however, not unambiguous since it did not reveal which cross peak was the preceding or succeeding residue.

Figure 29 shows an ambiguous assignment in the HNCACB spectrum. The $^{13}C_{\alpha}$ is

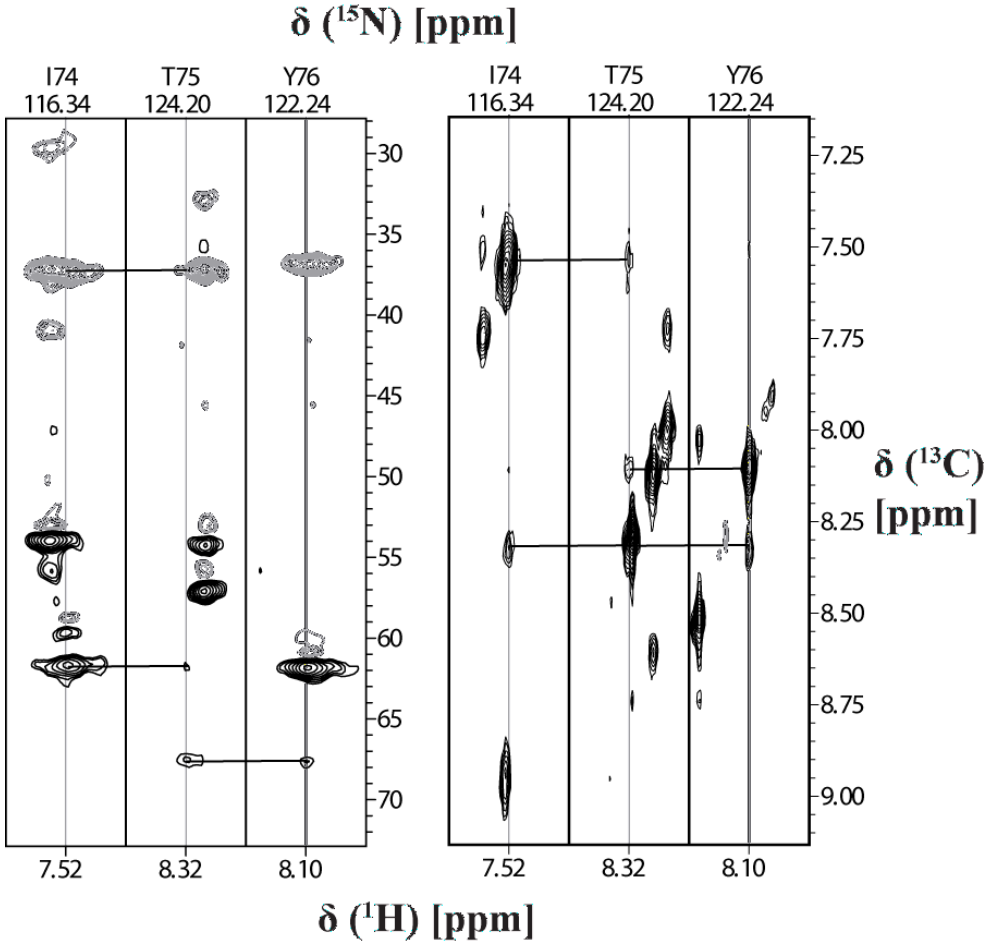


Figure 29: Comparison of HNCACB (left) and NOESY-HSQC (right) for residues I74-Y76. This shows how NOESY-HSQC can be used to confirm ambiguous assignments.

barely above the noise level and $^{13}C_\beta$ is on the shoulder of another peak in a nearby spin system. However, the assignment was confirmed by the NOESY-HSQC spectrum, which shows cross peaks from T75 to both neighbors and vice versa. I74 and Y76 also appear to have very weak cross peaks with each other, possibly because they are located in close proximity in the α-helix. The strong cross peaks due to close proximity of the amides made the NOESY-HSQC spectrum especially useful in the α-helix region.

The NOESY-TROSY spectrum also confirmed the assignment. For example, the M31 link to A32 in the HNCACB spectrum (Figure 27) was ambiguous due to an overlapping $^{13}C_\alpha$ and weak $^{13}C_\beta$ peak, and was not confirmed with HN(CA)CO due to poor signal to noise. However, the strong cross peaks between M31 and the neighboring residues S30 and A32 demonstrate that that they are in close proximity (Figure 30) and proved that the assignment is correct.

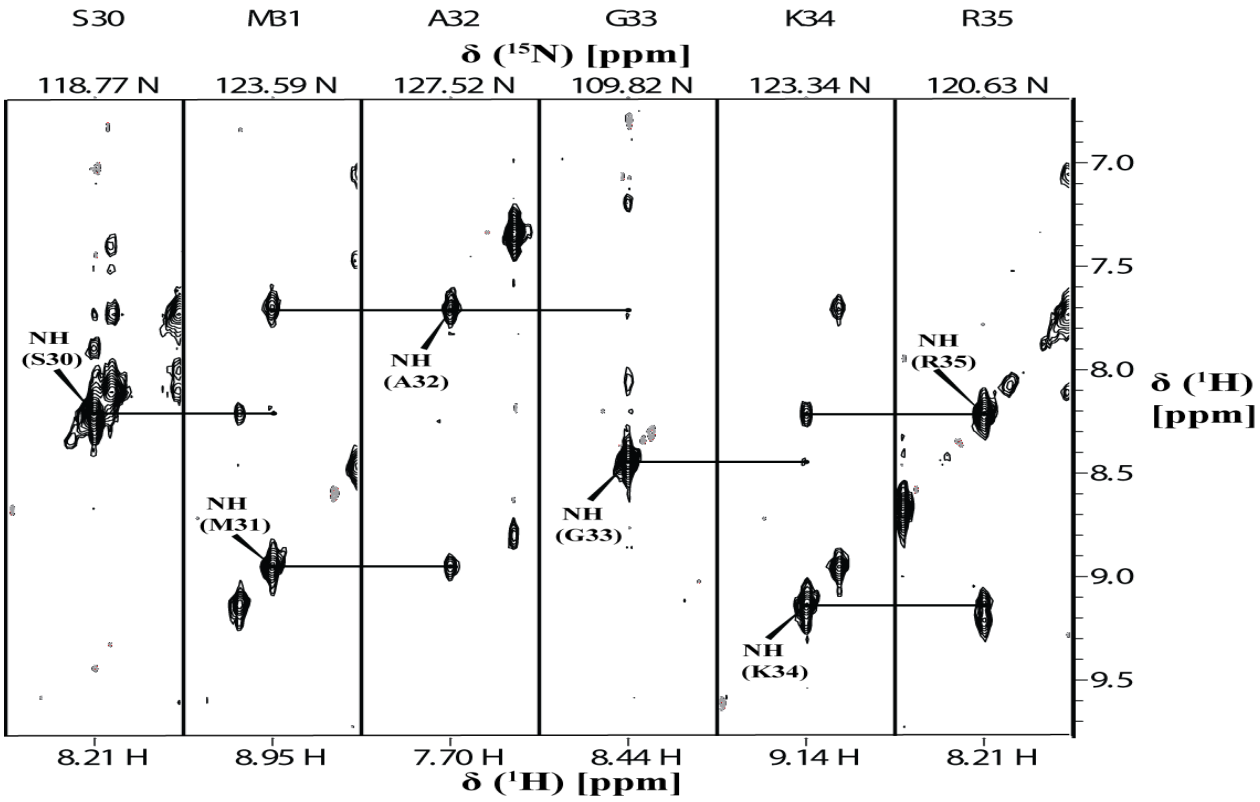


Figure 30: NOESY-HSQC. Strip view of residue 28-38.

Summary of the assignment process

The final assignment of the TROSY-HSQC amide backbone peaks is presented in Figure 31.

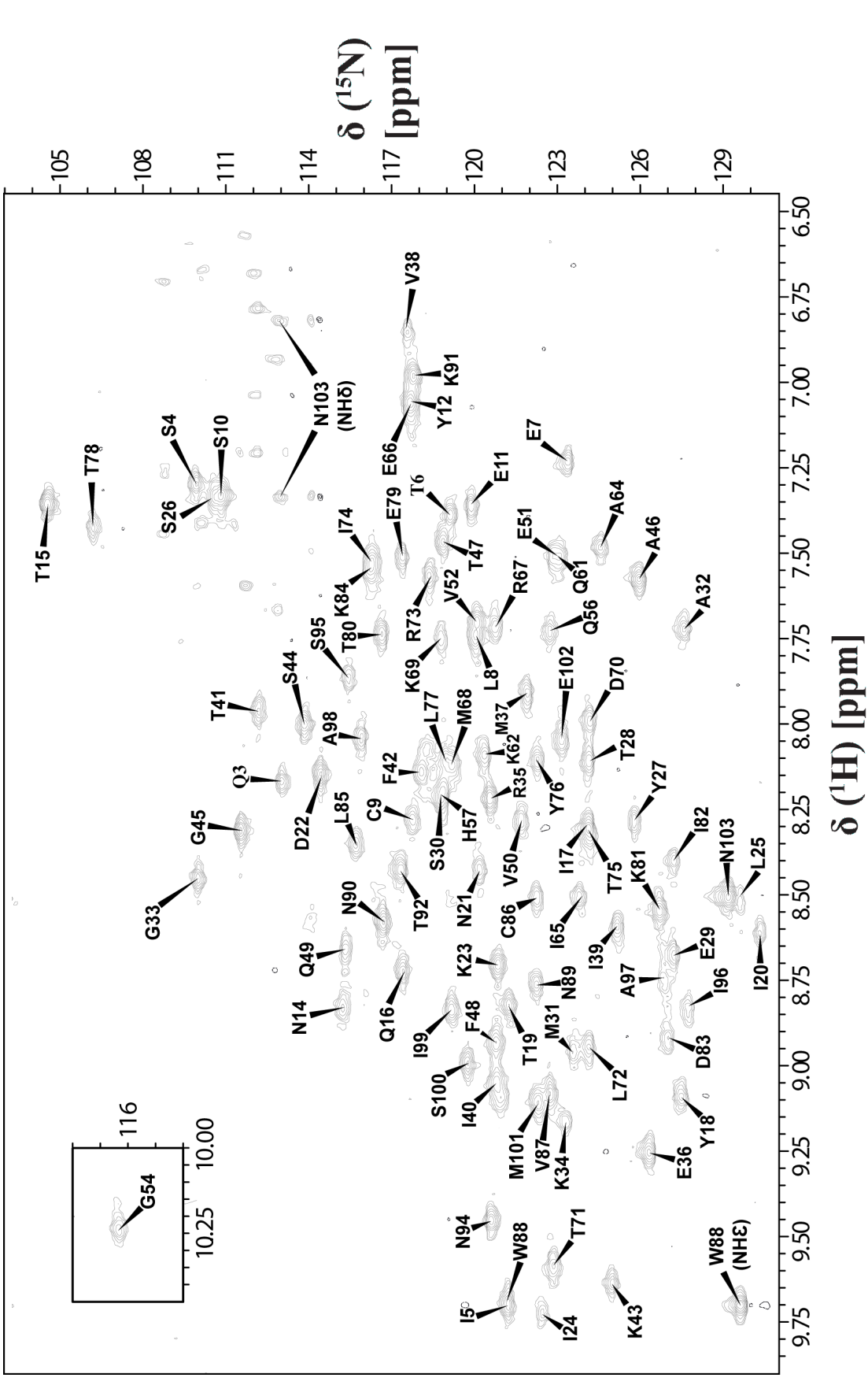


Figure 31: TROSY-HSQC spectrum showing the assignment of the peaks. Each peak represents an backbone amide or sidechain group. The insert contains G54, which is located far apart from the other peaks. The sidechains of W88 and N103 (appears as a doublet) were also assigned.

hLTB consists of 103 amino acids, which are potentially observable in all the experiments used in this work, except for the N-terminus (A1) and the 3 prolines (P2, P53 and P93), which are not observable in the TROSY-HSQC spectrum [73]. The N-terminus (A1) is not observable because of fast proton exchange between the amine group and solvent, which means that the proton signal can not be detected before it is exchanged with a water proton, giving it the same ppm value as water [71]. The three prolines are not observable due to lack of amide protons. However, their ^{13}C', $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ are observable from the succeeding residue in the triple resonance experiments and can thus be assigned, this would also be true for A1 if it was not succeeded by P2.

Of the 99 potentially observable backbone amide proton peaks in the TROSY-HSQC spectrum, 93 were identified and successfully assigned to the sequence (94%). The sidechain amide of N103 and indole group of W88 was also assigned. A possible peak is visible in the TROSY-HSQC at 8.55 ^{1}H-ppm, 113.5 ^{15}N-ppm (Figure 31). However, no correlations to any ^{13}C was visible. The 6 remaining residues, H13, S55, I58-S60 and K63 were not observable in the TROSY-HSQC spectrum and could not be assigned. However, some of their $^{13}C_{\alpha}$, $^{13}C_{\beta}$ and ^{13}C' were observable from the succeeding residue in the 3D spectra and could be assigned. Since they are all located on loops, it is possible that the absence of signal could be caused by intermediate conformational exchange, meaning that same residues experience a changing chemical environments during the experiment, leading to peak broadening and loss of signal [73].

Table 13: List of the different nuclei in the backbone and number how many of them was assigned.

Nuclei	**Expected**	**Assigned**	**%**
^{15}N	99	93	94
$^{1}H^{N}$	99	93	94
$^{13}C_{\alpha}$	102	98	96
$^{13}C_{\beta}$	99	95	96
^{13}C'	102	80	78
Total	501	459	92

The total number of assigned spins is presented in Table 13. Although 94% of the residues were identified, only ~92% of the total spins were assigned. This number was reduced because of the poor signal to noise in the HN(CA)CO and HNCO spectra.

In summary, all visible amide peaks in the TROSY-HSQC spectrum has been assigned, and can thus be used for the titration experiments of Kdo and hLTB. So far, hLTB ligand interactions have mostly been studied by binding assay studies, surface plasmon resonance X-ray crystallography. The completion of this assignment opens up for studying different

ligand interactions in solution with NMR, a complementary method that might give further insights into the hLTB interaction with these ligands.

3.5.3 Evaluation of the assignment by Talos+

The assignment of hLTB was evaluated by performing a secondary structure prediction and comparing the predicted structure with an earlier solved crystal structure.

The ^{13}C', $^{13}C_{\alpha}$, $^{13}C_{\beta}$, ^{15}NH and ^{15}N chemical shift values from the assignment were exported and analyzed by the TALOS+ web server [85]. The torsion angles of residue A1-Q3, G54, I58-S61 and N103 could not be calculated due to lack of assigned chemical shifts. The resulting secondary structure prediction was compared with the crystal structure of hLTB (PDB ID: 2O2L), presented in Figure 32.

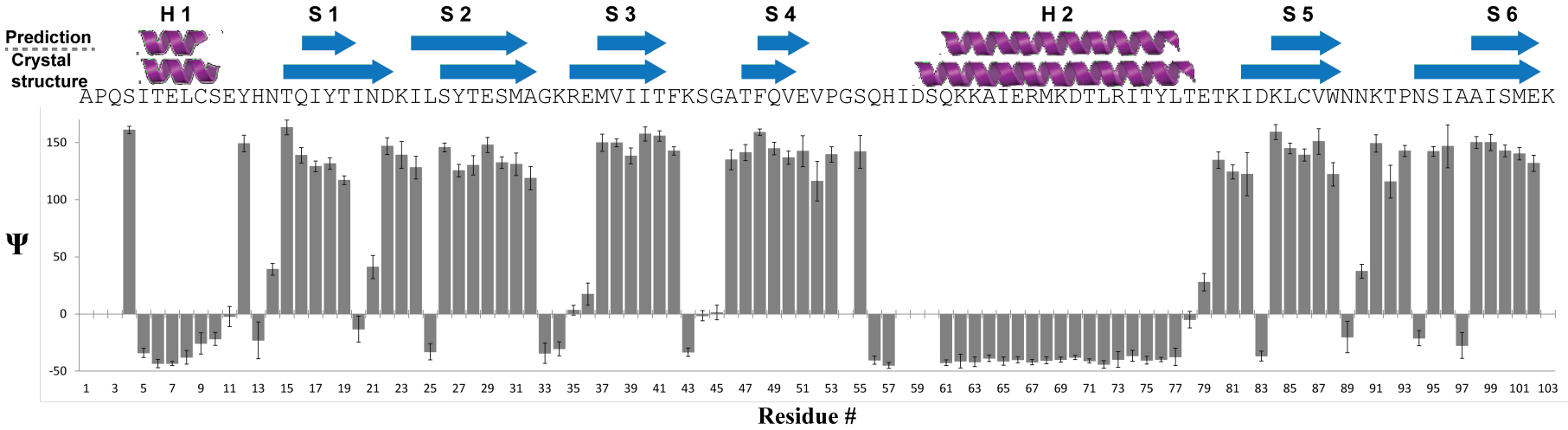


Figure 32: Secondary structure prediction of hLTB based on $^{13}C'$, $^{13}C_{\alpha}$, $^{13}C_{\beta}$, $^{1}H^{N}$, ^{15}N chemical shifts compared to the crystal secondary structure (PDB-ID:2O2L). ψ-angles with standard deviation bars are plotted below. Positive ψ-angles indicate β-sheet or β-sheet like structure.

The secondary structure prediction is in general agreement with the crystal structure, predicting two α-helices from residues I5 to C9 and Q61-L77 and six β-sheets between residues I17-T19, I24-M31, M37-T41, F48-E51, K84-W88 and A98-E102. The predicted number and location of β-sheets are fairly consistent, but generally shorter. This might be caused by the loops and bends that characterize the β-sheet regions between Y12-A46 and I82-E102. Such areas are generally of a more irregular structure, which might make a pronounced difference between the crystal- and solved hLTB structure. For example, A32-G33 are not predicted as β-sheet, however, by looking at the assignment (Figure 27, 28 and 30), they must be correct since those are the only clear AG-pair in the sequence. Furthermore, TALOS+ obtain a higher degree of ambiguous results when predicting these regions, and, since TALOS+ is programmed to minimize erroneous predictions, might decrease the amount of predicted secondary structures [85]. This notion is supported the ψ-torsion angles of residues neighboring β-sheets, which place them in the β-sheet region even though it was not predicted [72].

Overall, the predicted structure is consistent with the crystal structure. The results thus exclude the possibilities of large assignment errors. The assignment could have been further confirmed by collecting a ^{15}N-resolved ^{1}H-^{1}H NOESY spectrum to study the backbone relaxation properties [110], however, this was not done due to time limitations.

3.6 NEO titration

Neolactotetraose (NEO) is a secondary receptor that is presented on surface of intestinal epithelial cells, competing with the primary receptor of hLTB, GM1. The crystal structure of NEO bound to pLTB, published in 2011 [53], shows that the terminal N-galactose interacts with the primary binding site of pLTB. It is assumed that the binding is similar in hLTB, with the only difference being residue 13 (R in pLTB, H in hLTB). This difference may explain the different binding affinities to NEO between hLTB and pLTB with hLTB having a lower affinity than pLTB [111, 112]. The crystal structure also shows that NEO interact with residues that are connected to the secondary binding site, raising questions of cross-talk between the primary and secondary binding site. However the biological significance of this binding remains unknown.

A titration experiment was conducted two years ago by a group member to determine the binding affinity of NEO to hLTB (Figure 33). However, at that time, we did not have assignments for hLTB and were not able to determine the binding site. This presented itself as a good chance to partially confirm the assignment process, and further analyse the data by mapping the peak shifts onto the crystal structure of hLTB (PDB ID:2O2L). All peak shifts above $\Delta\delta = 0.15$ were chosen and mapped to the crystal structure.

The superimposed spectra of hLTB during titration shows that peaks move with vary-

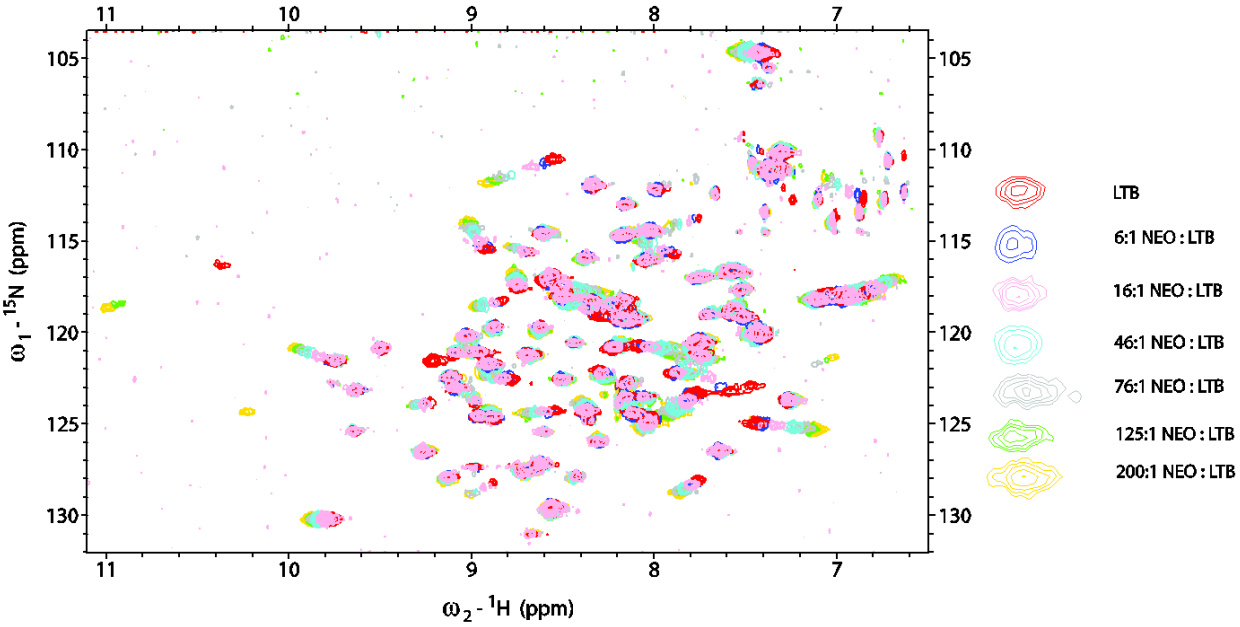


Figure 33: Superpositions of TROSY-HSQC spectra collected during titration of NEO to hLTB. (Prepared by Per Eugen Kristiansen)

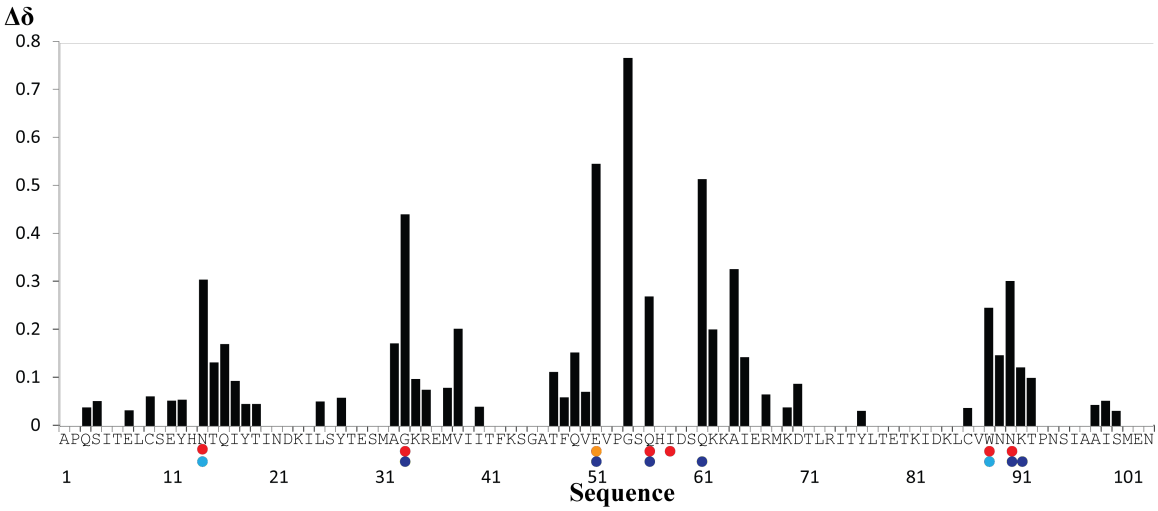


Figure 34: Bars showing the combined ^{15}N-^{1}H chemical shift change upon NEO-titration. to hLTB. Red dots indicate residues that participate in NEO binding [53]. Blue dots indicate residues that are involved in GM1-binding. Peaks that were not clearly shifting were not measured.

ing distance during the titration, indicating that the corresponding residues are differently affected upon NEO-binding (Figure 33). This could mean that the peaks that move the most are directly interacting with the ligand, while those that move less lie in the periphery of the interaction [72]. Some Aln/Gln peaks also shift, indicating that also the corresponding side chain is interacting as well.

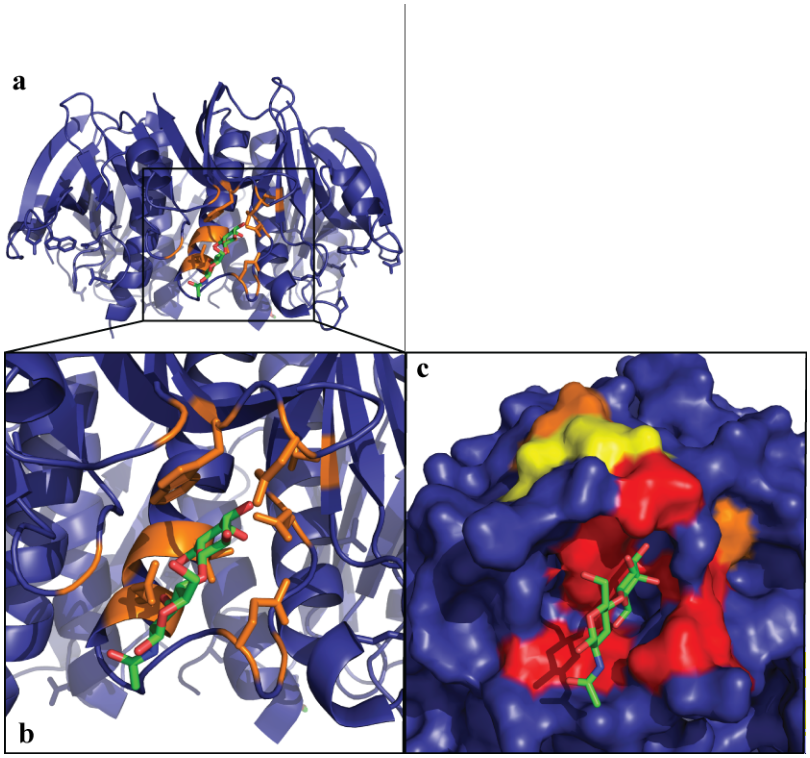


Figure 35: hLTB (PDB ID: 2O2L) superimposed with NEO from the pLTB structure (PDB ID: 2XRS). **a,b** Cartoon representation, orange indicate residues with combined ^{15}N-^{2}H-shifts $\Delta\delta$ above 0.15. **c** Surface representation heat map of interacting residues. Red = residues with $\Delta\delta$ above 0.175, Orange = residues with $\Delta\delta$ above 0.150, yellow = residues with $\Delta\delta$ above 0.100.

Residues N14, G33, E51, Q56, W88 and N90, are all located in the primary binding site and exhibit strong peak perturbations in the presence of NEO. (Figure 34 and 35). This confirms that NEO binding to hLTB is in the primary binding site, like pLTB.

G54, K62, A64 and I65 in the connecting helix also exhibit fairly strong perturbations, although they do not participate in NEO binding, suggesting that there is a conformational rearrangement. This supports the hypothesis of Holmner et al [53] that stabilization of the 51-59 loop upon NEO binding may lead to structural changes that lead to cross talk between the two binding sites, which might inhibit blood sugar binding. This is further supported by the observation that residues Q3, Q16, Y18, T47, F48, W89 and K92, which participate in blood sugar binding, also exhibit peak perturbation, although to a lesser degree. The NEO titration experiment therefore confirmed the binding location of NEO to hLTB, and supports the idea that there might be cross talk between the primary and secondary binding site.

Due to lack of assignments for the H13 backbone amide, we were not able to determine

to what extent this residue takes part in the binding. However, a well resolved peak with unknown origin shows up at 10.1 ^{1}H-ppm, 124 ^{15}N-ppm (Figure 33). This could be caused by interaction of the H13 side chain with NEO, which would change the pKa drastically, as a result of binding, and thus become visible in the TROSY-HSQC.

3.7 Initial LPS titration

LPS is presented on the surface of *E. coli* and anchors secreted hLTB to the bacterial membrane by binding to a site, distinct of the GM1 binding site [32]. Earlier binding studies by Mudrak et al. indicate that the putative LPS binding site overlaps with the blood sugar binding site [48]. However, there are no structural studies of the hLTB-LPS interaction to our knowledge. In order to study the interaction between LPS and hLTB, LPS-Ra (LPS without O-antigen) was titrated to hLTB (Table 6).

Upon addition of LPS to hLTB, the signal disappeared (Figure 36). This could be explained by LPS's ability to form aggregates [113, 114, 115]. When hLTB binds to the surface of these aggregates, it can no longer tumble freely, leading to peak broadening and loss of signal. This is supported by the observation that the LPS solution was slightly hazy prior to titration, indicating that it contained aggregates. The almost complete loss of signal before 1:1 hLTB-subunit:LPS ratio was reached (Figure 36), suggests a high to intermediate binding affinity in the μM to mM range. However, each B-pentamer can in theory bind 5 LPS molecules, complicating the interpretation of the experiment [48]. The binding of only one hLTB subunit to LPS would cause the signal of the whole B-pentamer to disappear, making the binding affinity hard to assess.

LPS binding to hLTB has been somewhat disputed. The Kuehn group first demonstrated that hLTB is secreted from the bacteria bound to the bacterial membrane in 2002 [32]. This finding was further supported by binding assays in 2003, showing that LPS of different lengths could inhibit the interaction between LT and *E. coli* cells [65]. Further binding studies in 2009 showed again that LT binds to the surface of *E. coli* cells [48] However, these findings have been challenged by Jansson et al [116], who performed competition studies with LPS immobilized on microtiter plates, demonstrating that neither LPS, LPS-Ra, LPS-Re or LipidA binds hLTB. The results in this thesis clearly support the findings of the Kuehn group. It is unknown what caused the negative results of Jansson et al, but it is speculated that the negative binding might be caused by improper orientation of LPS on the microtiter plates, preventing proper presentation of the LPS oligosaccharide to hLTB and, therefore binding [30].

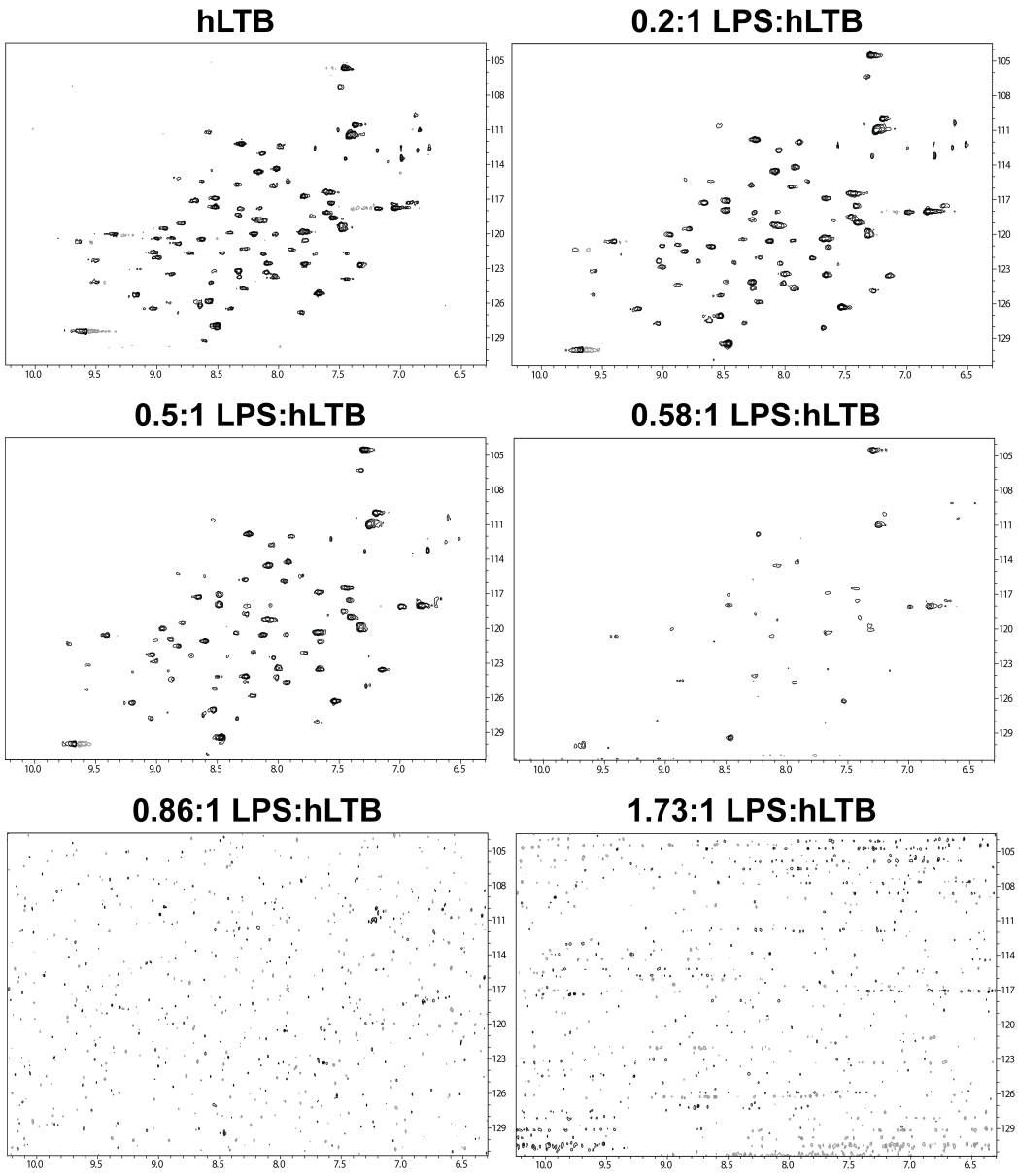


Figure 36: A selection of TROSY-HSQC spectra illustrating the loss of signal upon LPS titration to hLTB. The two last figures are rendered at a lower contour level than the rest.

3.8 Titration of LPS with dodecylphosphatecholine (DPC)

Light scattering studies show that LPS has a critical micelle concentration (CMC) of 14 μg/mL [115] Titration experiments cannot be conducted at such low concentrations because of low sensitivity in NMR. It has been shown that detergents disrupt aggregate structures and incorporate LPS into detergent micelles [117, 118, 119]. We therefore tried to incorporate LPS into micelle-forming detergents.

3.8.1 Choice of detergent

While choosing a detergent for the titration experiment, it was important that it was mild to hLTB, forms small micelles (< 20 kDa), and was available in perdeuterated form at a reasonable price. DPC had been extensively used in the group earlier, it is considered non denaturing, and is commonly used in liquid state NMR to study membrane systems [120]. It is also shown that 300 mM (11.7%) DPC is able to incorporate 1 mM LPS-Re into DPC micelles [119]. Upon addition of DPC to a 20 mg/mL LPS solution, it instantly turned clear, indicating that DPC breaks down the large aggregates and solubilizes LPS. DPC was therefore chosen as a detergent. SDS, deoxycholate and β-D-Maltopyranoside (DDM) was also assessed for this purpose, but ultimately found unsuitable.

3.8.2 Preparation of LPS-incorporated micelles

In order to form small structures of LPS it was attempted to incorporate LPS into DPC micelles. A protocol for incorporation of LPS into DPC was developed based on the Avanti Polars Lipids protocol for preparation of unilamelar vesicles [90], which can be found in section 2.7.

Chloroform was initially chosen as organic solvent, however, LPS turned out to be completely insoluble in this. In an attempt to increase the solubility of LPS, methanol was added to form a 1:1 chloroform:methanol solution without change. The experiment was repeated with methanol as organic solvent, but this did not improve solubility. MQ-H_2O was added to a 1:1 solution, resulting in apparent solubilization of some LPS. The experiment was continued and the finished product titrated to ^{15}N-hLTB. However, no peak movements or signal loss was observed, indicating that no LPS was present.

The hydrophilic nature of LPS display great heterogeneity that appear to depend on the presence of O-antigen and number of sugars in the core oligosaccharide. As discussed by Reaznia et al [121] and Raetz et al [113], smooth wild strain LPS that contain a full oligosaccharide core and O-antigen is extracted from bacteria by aqueous organic solvents as hot phenol-water, aqueous butanol and cold ethanol, while rough strains are more commonly extracted with chloroform and methanol. It could therefore be possible that LPS-Ra, that has a full core, but no O-antigen, may not be soluble in organic solvents due to the hydrophilic oligosaccharide, while Lipid A will cause it to self aggregate in aqueous solutions. Finding an ideal organic solution that completely dissolves LPS-Ra, as required by this method, was considered too time consuming. Hence, this method was not pursued further.

3.8.3 Titration of LPS solved in DPC to hLTB

Following the choice of DPC as detergent, two initial titration experiments were conducted by titrating LPS solubilized in 2% and 4% DPC to hLTB. However, these experiments resulted in signal loss without any observed peak movements, although at a higher concentration than in section 3.7. This indicates that LPS was only partially dissolved in DPC, and that larger LPS aggregates or LPS-DPC complexes may still be present. Mares et al [119] observed specific interaction between LPS-Re and a peptide, polymyxin, by dissolving LPS-Re in 300 mM (11.7%) DPC in a 40 mM acetate-buffer, pH 4.4 at 37 °C. Molecular dynamic studies combined with NMR-data implied that LPS-Re was incorporated into DPC micelles. We therefore decided to perform titrations at at DPC concentration as close to this as possible. Since hLTB starts precipitating above 8 % DPC, the titration experiment was conducted by adding LPS in a 10% DPC solution [1] to hLTB in a 6% DPC solution (Table 7). 1D NMR spectra were collected in addition to TROSY-HSQC to monitor the LPS concentration in solution. The temperature was increased to 35 °C in an attempt to aid DPC disruption of the LPS aggregates. Higher temperatures were preferable, but difficult to achieve at the time. The solution was sonicated at ~50 °C between every step in the titration in an attempt to disrupt the LPS aggregates.

The hLTB signal started deteriorating at 4.17:1 LPS/hLTB-subunit concentration (Figure 37), much later than in the initial titration experiments. The experiment showed that the mixture of DPC and LPS binds to a lesser extent than to LPS alone. However, we did not observe peak movements, even at high LPS concentrations.

To investigate the effect of higher DPC concentrations on LPS solubility, more DPC was added to a final concentration of 285 mM (10%). Although this precipitated hLTB, it was not expected to affect the DPC-LPS interaction. The 1D spectra (Figure 38) shows a broad signal between 3.4-4.1 ppm, believed to be overlapping signals from the LPS sugars. The signal increases as a response to higher LPS concentrations, but remains unchanged when more DPC is added after the titration experiment, indicating that more LPS is not solubilzed by adding more DPC after the titration experiment. Since no isolated peak belonging to LPS was identified, it was not possible to quantify the concentration of solubilized LPS in the solution.

Although increasing the DPC concentration after the titration did not affect further solubilization of LPS, the deviations in the experimental set up from Mares et al [119] may explain the different results. Buffer composition, pH and temperature are factors that could affect the equilibrium between aggregate LPS and LPS-DPC complex, pushing it towards the aggregate state and cause signal loss upon hLTB binding. The experimental factors could also affect the size of the LPS-DPC complex. The almost continuous sugar

[1] This was the only LPS solution available at the time

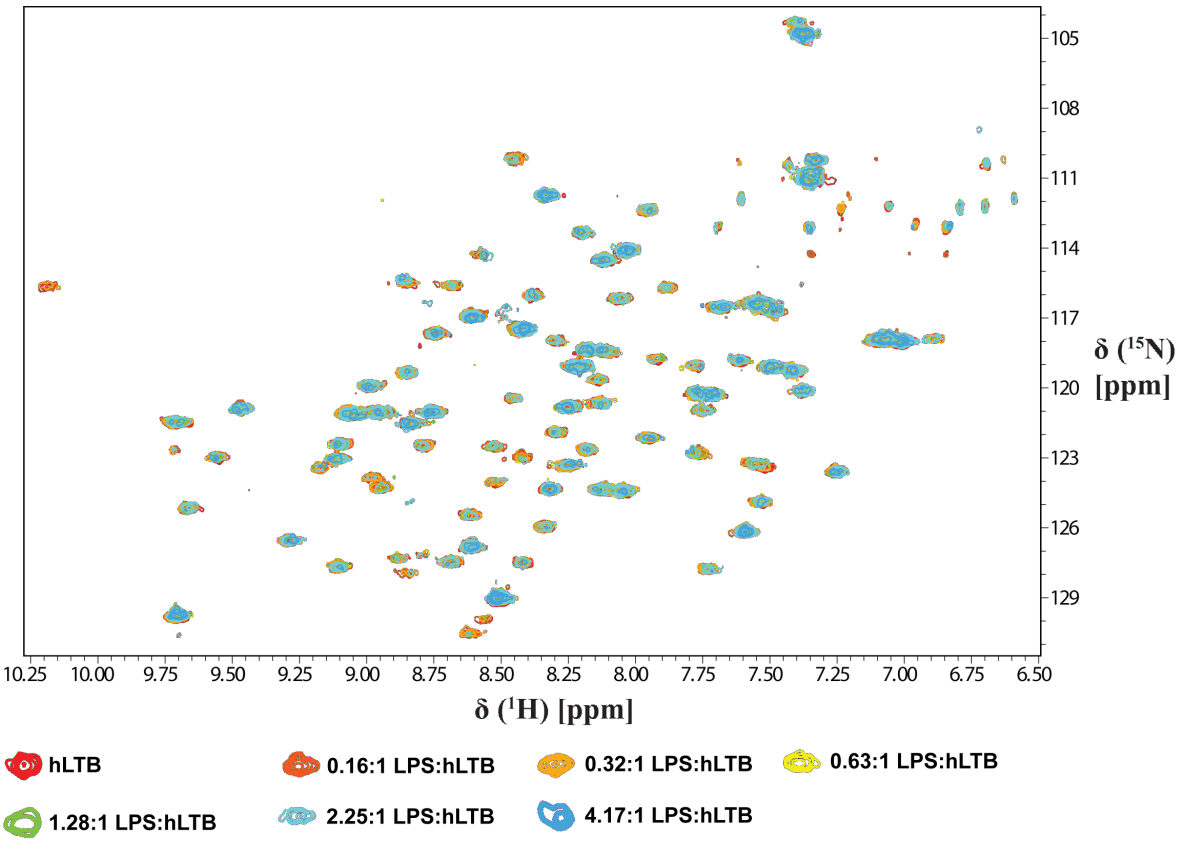


Figure 37: Overlay TROSY-HSQC spectra of LPS in 10 % DPC titrated to hLTB in 6.0 % DPC. The signal was significantly decreased indicating that hLTB still interacts with larger aggregates.

signal in the 1D spectra suggest that it is composed of broad peaks, indicating that the LPS-DPC complex has a large size. Although hLTB-binding to one of these complexes should be measurable by using TROSY experiments, it is possible that hLTB, which has 5 LPS binding sites, could bind to several complexes simultaneously [48] causing the total size to exceed the observable size, even with TROSY experiments. To investigate whether this is the case, the experiment should be repeated with perdeuterated hLTB [73]. Alternatively, the experiment could be repeated under the same conditions that were used by Mares et al, however, changing the conditions might perturb the shifts and complicate the interpretation of titration experiments. A proper 1D titration experiment with only LPS and DPC should also be conducted to determine the DPC level with maximum LPS-solubility.

Although the above hypotheses may explain why the signal disappears, they do not explain why the loss of signal is delayed in the presence of DPC. Mudrak et al [30], speculates that the presence of several side-by-side LPS molecules are important for hLTB binding, which could explain the lower affinity to LPS incorporated in DPC. However, it was also suspected that DPC may interferes with hLTB-LPS interaction, either by

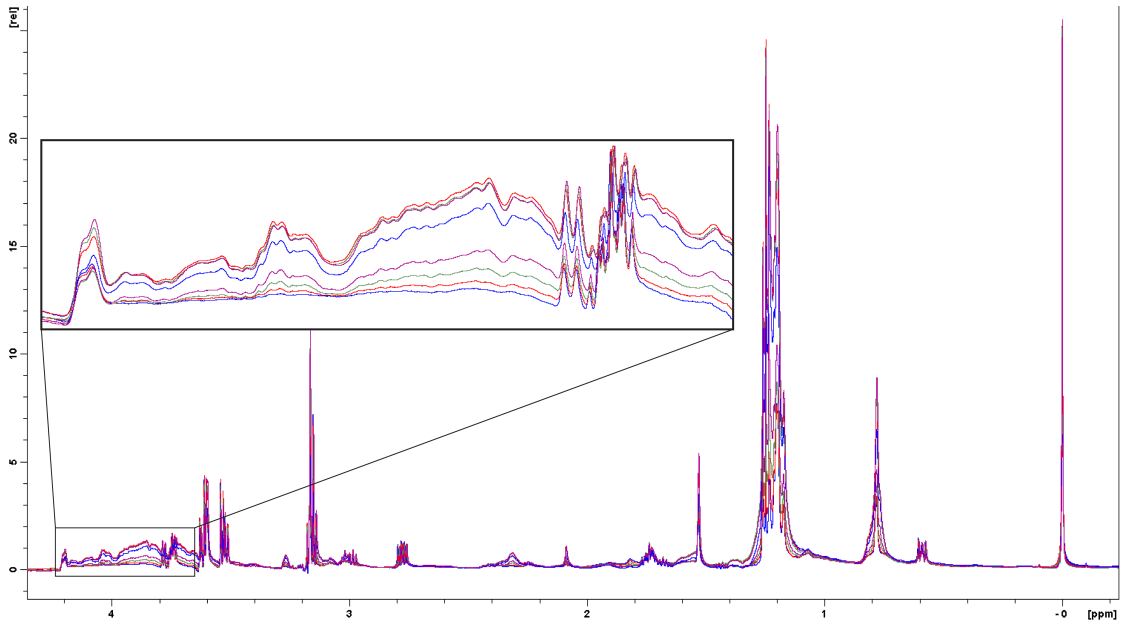


Figure 38: Overlayed 1D spectra of the LPS titration experiment in 6-10 % DPC. The putative LPS sugar peaks are located in the 3.6-4.2 ppm area and is shown in the insert. The LPS concentration in solution increases during titration. Further DPC addition does not further increase the LPS signal.

changing the conformation of the LPS sugars, or by interfering with hLTB.

3.8.4 Locating the DPC-hLTB interaction

DPC was titrated to hLTB (Table 8) and TROSY-HSQC spectra were collected to investigate the extent DPC interacts with hLTB. Some small peak movements were observed upon addition of DPC (Figure 39). The combined ^{1}H-^{15}N chemical shift changes ($\Delta\delta$) of the different residues was calculated [92] (Figure 40). A cut-off point of $\Delta\delta = 0.05$ was chosen and all residues above this value was mapped onto the crystal structure (PDB ID:2O2L). The binding affinity was higher than expected and the titration end point was reached too early, causing few datapoints.

The interaction between hLTB and DPC is localized to the apical interface between the subunits in close proximity to the blood sugar binding site (Figure 41). Residues G45, A46, T47 and F48, important residues in both blood group- and LPS binding [48, 57] lies in close proximity to T41, which is a part of the interaction between hLTB and DPC (Figure 41). This is especially important for A46 and T47, which has been shown to be critical for LPS binding [48]. Furthermore, Q16, Y18, G45, N88 and K90 are all affected by the interaction, although to a lesser degree (Figure 40). This could suggest that DPC binds to hLTB at a position that is close to the expected binding site of LPS, and thus might prohibit the binding of LPS. Although the binding affinity appears to be low, the

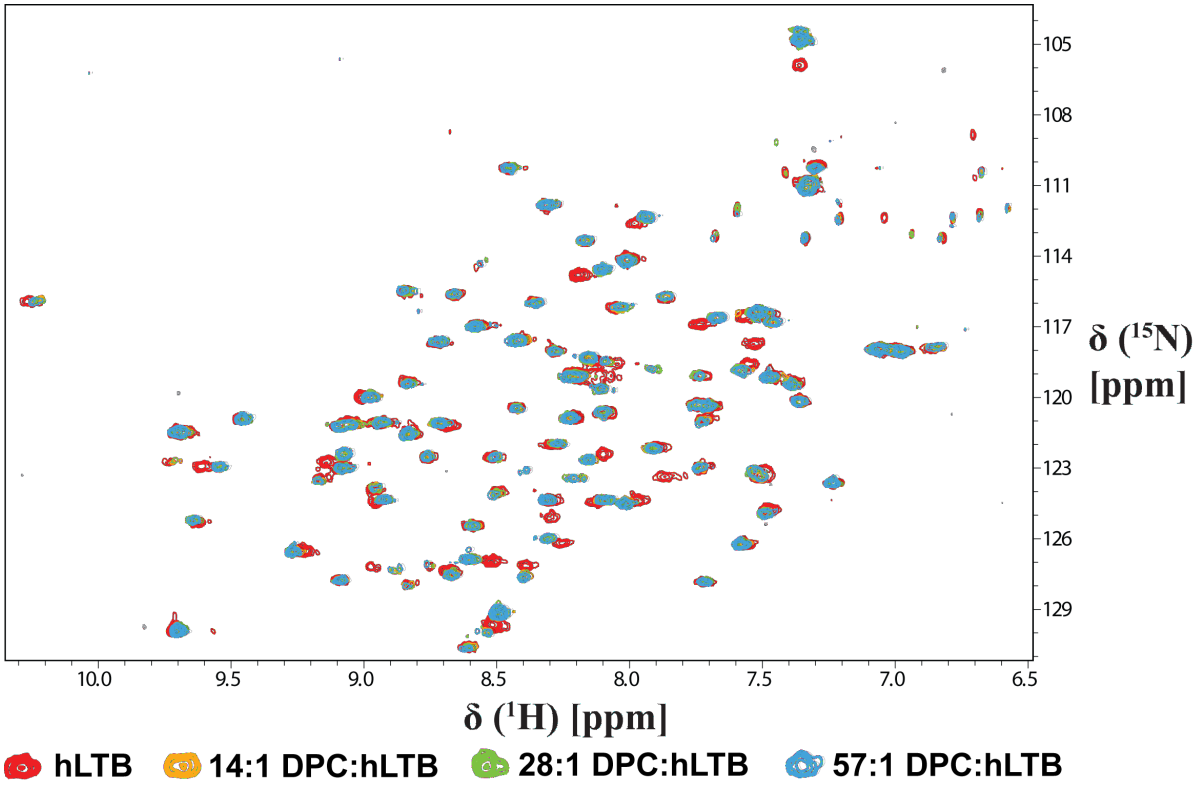


Figure 39: Overlay of TROSY-HSQC spectra of 0.1 mM hLTB collected at various DPC concentrations.

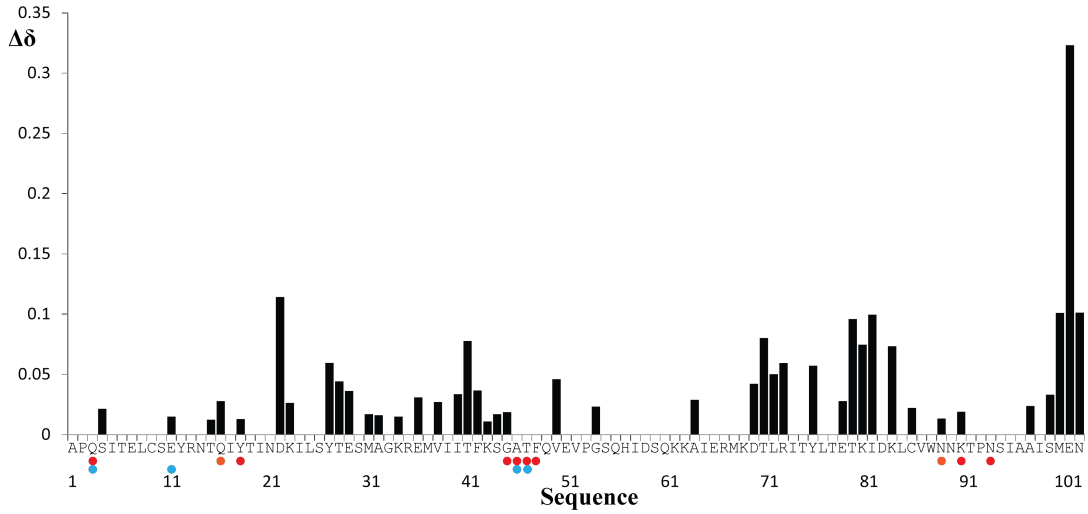


Figure 40: $\Delta\delta$ values of the different backbone amides of hLTB before and after addition of DPC. Residues involved in blood group antigen binding is marked in red [56], orange for water mediated contacts. Residues involved in LPS binding is marked blue [48]. Peaks with no clear shifts were not measured.

high concentration of DPC during the titration experiments could outcompete LPS binding. Since DPC does not prevent loss of signal before peak movements are observed and

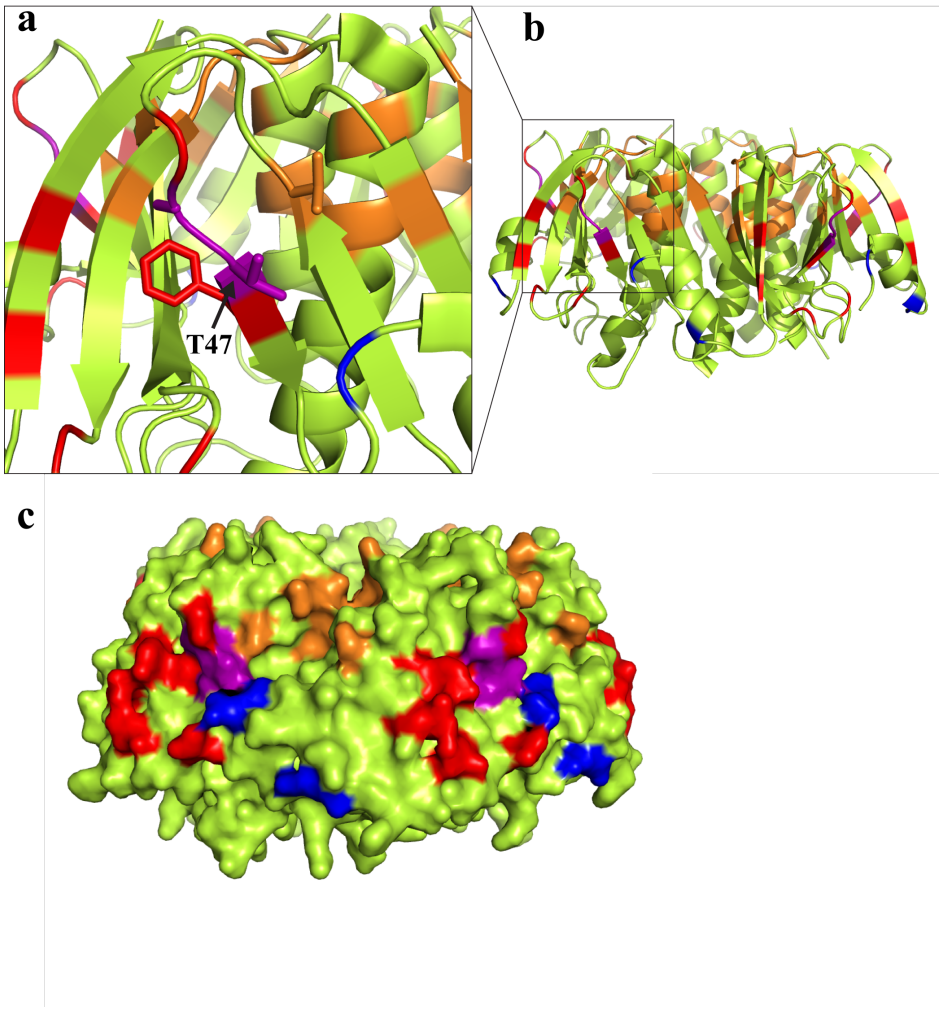


Figure 41: Peak movements of DPC titration mapped onto the crystal structure of hLTB (PDB-ID:2O2L). Orange = peaks affected by DPC binding. Blue = residues involved in LPS binding. Red = residues involved in blood sugar binding. Purple = residues involved in both blood sugar and LPS binding. The figure shows that the residues that are affected by DPC binding lies close to the secondary binding site.

might interfere with with LPS binding, further experiments with DPC were not pursued.

3.9 Mild acid hydrolysis of LPS

Following the unsuccessful solubilization attempt of LPS with detergents, mild acid hydrolysis of LPS was attempted.

Oligosaccharides are easily hydrolyzed in acidic environments. The $\alpha(2 \rightarrow 6)$ bond connecting the core oligosaccharide to Lipid A has been shown to be particularly acid labile (See figure 6), allowing separation of the core oligosaccharide from Lipid A by mild acid hydrolysis [113]. Common protocols perform the hydrolysis at pH 1-4 at 100 °C for

1-2 hours in acetic acid, HCl of H_2SO_4 [122, 62, 123]. In this work, LPS dissolved in 2% DPC was hydrolyzed at 99 °C, pH 2.0 for 2 hours. DPC was added to keep LPS in solution during hydrolysis [117]. A sample of the hydrolyzed product was sent to MS for analysis (Figure 42) and the rest was titrated to a ^{15}N-hLTB sample.

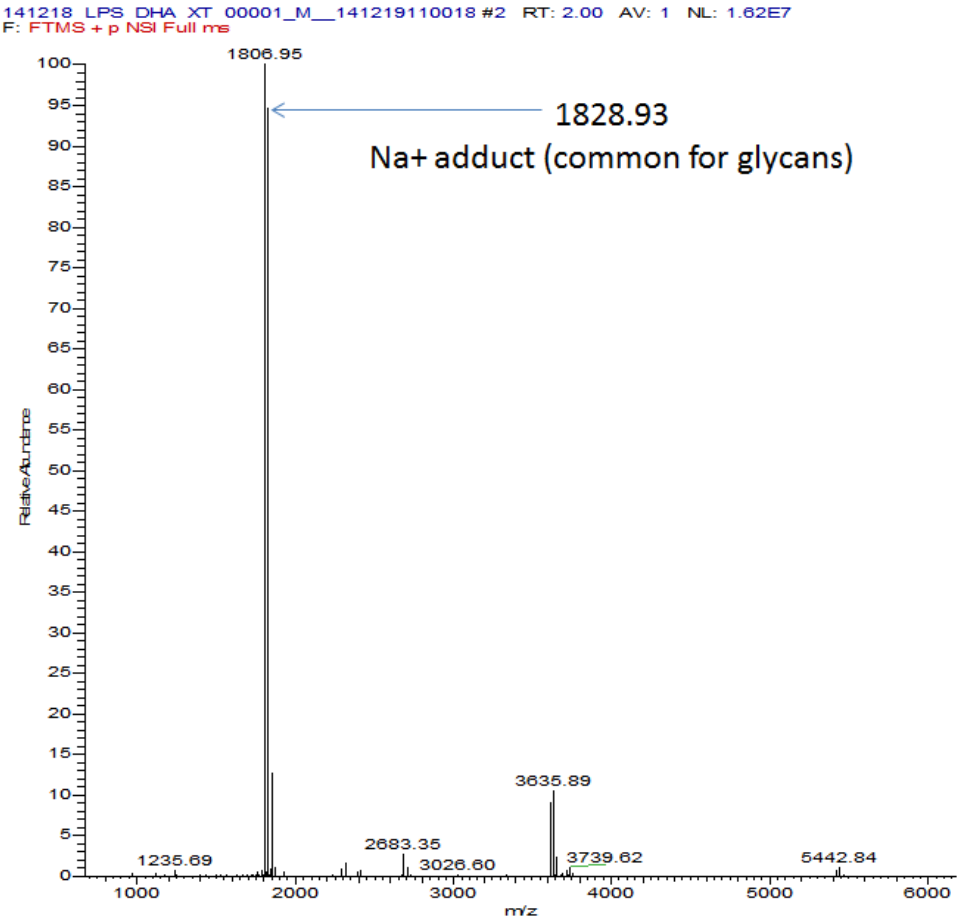


Figure 42: Top down MS spectrum of hydrolyzed LPS

The expected weight of hydrolyzed LPS-Ra was 1811.4 Da [62] (assuming hydrolysation of 2 Kdo sugars and PPEtn in *E. coli* K-12 LPS oligosaccharide[2]). The MS-results showed several molecules present in the sample with the most abundant at m/z 1806.95 and at m/z 1828.93. The 22 Da difference between the molecules can be attributed to a Na-adduct for glycans [124] (Figure 42). Neither of the masses was seen in a blank sample. However, the abundant mass at m/z 1806.95 did not correspond to the expected mass of 1812.4 Da [M+H$^+$]. There might be several explanations for the discrepancy between the theoretical and observed mass. The sample may contain an additional abundant molecule than LPS, or hydrolyzed LPS might have an other structure than the one reported in the litterature [62] and therefore different mass.

[2]The bonds between the different Kdo sugars were also expected to be acid labile since they are very similar to the bond connecting Kdo to Lipid A

The fact that the LPS was purchased from a supplier and the presence of the abundant peak in only the LPS sample, and not the blank, suggests that the hydrolysis of LPS was successful. However, the full spectrum showed that the sample contained multiple other peaks, indicating that small part of the LPS was hydrolyzed only at the expected bonds. To obtain higher amounts of pure, correctly hydrolyzed, core oligosaccharide, a larger amount of LPS could be hydrolyzed and the products separated by high performance liquid chromatography. However, we acquire LPS-Ra from a commercial source, and acquiring large amounts would not be economically viable.

The titration of the hydrolyzed core oligosaccharide to ^{15}N-hLTB showed no peak movement except for those earlier observed for DPC. This could be because of the apparently low hydrolyzed core oligosaccharide yield or interference in binding by DPC. Nevertheless, because of the supposed low hydrolysis yield and lack of peak movement, this approach was not pursued further.

3.10 Kdo binding experiments

Binding studies performed by Horstman et al [65] show that phosphorylation of the Kdo-sugar on LPS inhibits both LT and CT binding. This can be a determining factor in the different toxin delivery mechanisms of the two toxins since *Vibrio cholerae* LPS has a phosphorylated Kdo, while *E. coli* presents an unphosphorylated Kdo. Furthermore, the binding studies also show that LPS-Re (carrying only Kdo) and Kdo alone can bind to hLTB independently of the remaining core oligosaccharide [65], though, with much lower affinity. This indicates that Kdo participates in the binding between LPS and hLTB. To investigate the binding between Kdo and hLTB, Kdo was titrated to ^{15}N-hLTB and co-crystallized with hLTB.

3.10.1 Kdo titration experiments

To investigate the nature of Kdo binding, a 200 mM Kdo solution in NMR-buffer was titrated to 0.34 mM hLTB (Table 9).

The results in figure 43 show that there is no peak movement before Kdo:hLTB reaches ~100, indicating little to no binding. The combined ^{15}N-^{1}H chemical shift of residues that displayed any resemblance of peak movements were calculated (Figure 40). Residues with $\Delta\delta$ [92] values above 0.025 (K34, E51, R67, T92)) were chosen and mapped onto the crystal structure (PDB ID:2O2L) (Figure 44). Although three of the four residues (T92, E51 and K34) were located in vicinity of each other, they were located at a distance from the secondary binding site. This could support the notion by Mudrak et al [48] that free Kdo and LPS with a terminal Kdo-residue might bind to a separate binding pocket than the ordinary LPS-binding site. However, the peak shifts of the four chosen residues were barely above the background level and the generally small peak movements could also be

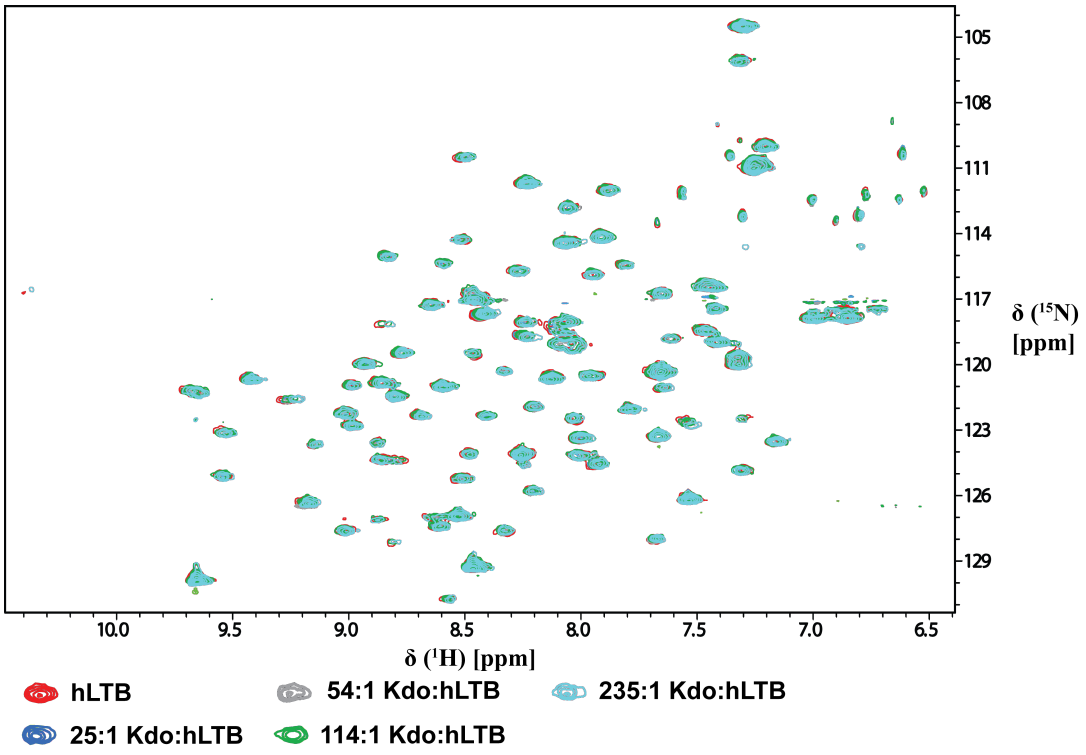


Figure 43: Overlay of TROSY-HSQC spectra collected during Kdo titration to hLTB.

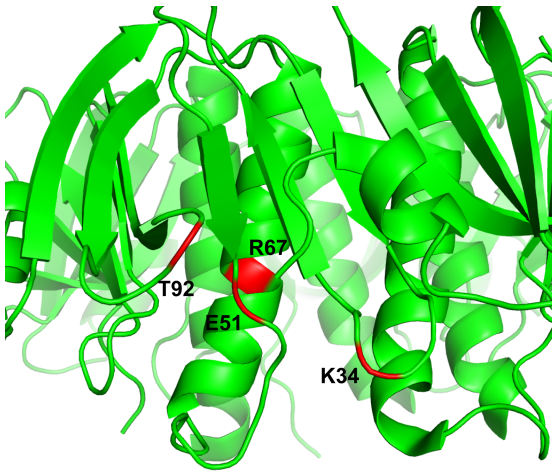


Figure 44: Residues with combined chemical shift ($\Delta\delta$) over 0.025 upon Kdo titration

a response to changes in the solution properties due to the large sugar concentration. No conclusions on the Kdo binding site can therefore be drawn from this experiment.

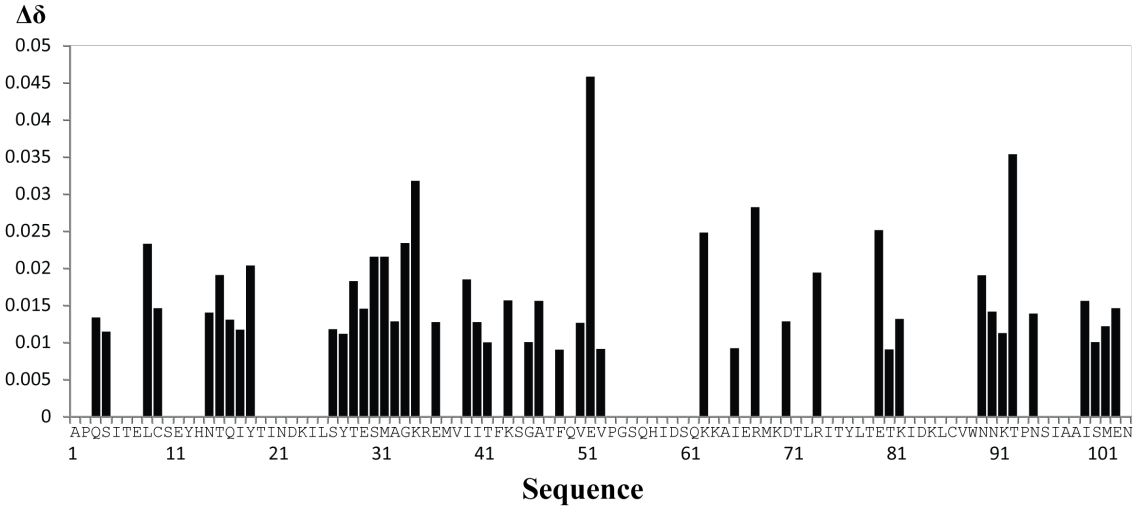


Figure 45: Combined ^{15}N-^{1}H ($\Delta\delta$) chemical shifts of Kdo titrated to hLTB.

3.10.2 Co-crystallization of Kdo with hLTB

To obtain a high resolution structure of Kdo bound to hLTB, several commercial crystallization screens were set up (Described Table 10, section 2.10.1).

Table 14: Commercial screen hits

Hit #	Screen	Well	Conditions
1	PGA	1-11	0.1 M Na-acetate, pH 5.0, 3% (w/v) PGA-LM, 30% (v/v) MPD
2	PGA	1-19	Na-acetate, pH 5.0, 3% (w/v) PGA-LM, 20% (v/v) MPD, 0.2 M K-cyanate
3	JCSG+	29	0.8 M NaH_2PO_4, 0.8 M KH_2PO_4, 0.1 M HEPES, pH 7.5

Single crystals were obtained under three conditions (Table 14) and screened further with regards to pH and precipitant concentration to increase the size and number of crystals. Crystals based on Hit #3 diffracted to ~6 Å (Figure 46a). This condition was further screened with different additives (Morpheus additive screen) and buffers at different pH-levels and precipitant concentrations. When the HEPES buffer was exchanged with TRIS-buffer, a new form of rod shaped crystals formed in the range of 0.45-0.80 M NaH_2PO_4 and KH_2PO_4, pH 6.5-7.5 (Figure 46b) at 5-8 mg/mL hLTB. These crystals generally diffracted to ~2.0-2.6 Å and were co-crystallized with 1:10 and 1:50 hLTB:Kdo ratios. Three datasets, one with 10x and two 50x Kdo concentration were collected and successfully solved. The most complete data set, from one of the 1:10 co-crystallization experiments, was further refined three times using Refmac5 (CCP4 software suite).

The final statistics are shown in table 15. A screenshot of the crystal structure is shown in figure 47.

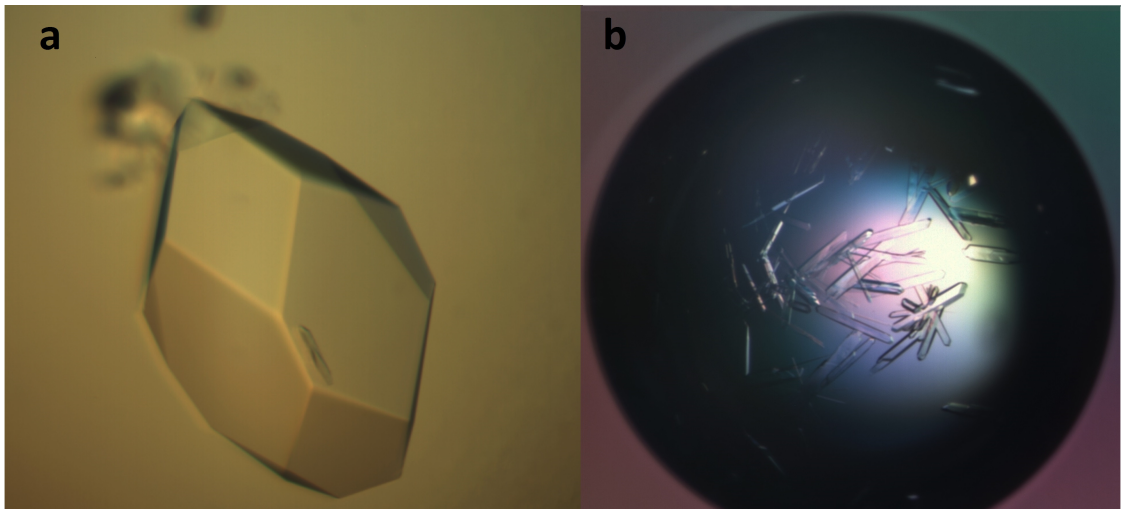


Figure 46: Typical crystals obtained at 0.45-0.8 M Na_2HPO_4, K_2HPO_4, pH 6.5-7.5 and 0.1 M HEPES (a), or 0.1 M TRIS (b). The pictures are not to scale.

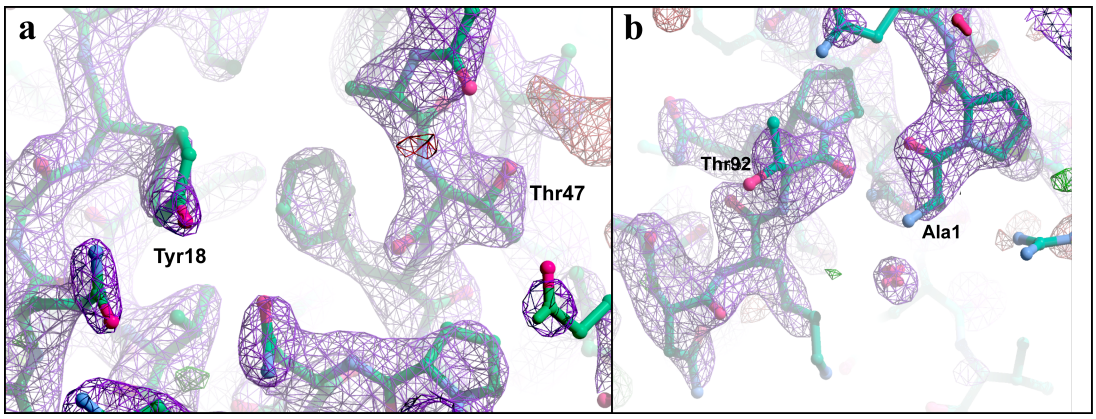


Figure 47: Raster3D screenshot from Coot [94] and solubilize it of the 1:10 hLTB:Kdo crystal structure contoured at 1.5σ. **a.** Secondary binding site **b.** Residues with peak shifts in the NMR-experiment. A water molecule has been modelled into the extra electron density in the background.

No extra electron densities indicating Kdo binding, was observed in, or near any of the secondary binding sites in the two pentamers in the asymmetric unit (Figure 47). The area around K34, E51 and T92 was also examined, without signs of extra electron density except for a small additional electron density that fit a water molecule.

There are several reasons for this apparent lack of binding. Since the NMR experiment in section 3.10 showed poor to no binding, it is possible that the binding affinity of Kdo to hLTB is too low to be observable at 1:10 concentration. An alternative explanation is that the cryoprotectant, propylene glycol (PG), outcompetes Kdo binding. In Heggelund et al [55] they used ethylene glycol (EG) as a regeneration solution in surface plasmon resonance to outcompete the binding of blood group sugars to the cholera toxin. Since

Table 15: Data collection and refinement statistics of hLTB co-crystallized with 10x Kdo concentration. Numbers in parentheses indicate the higher resolution shells [88]

Space group	$P4_12_12$
a (Å)	133.23
b (Å)	133.23
c (Å)	176.01
α (^{o})	90
β (^{o})	90
γ (^{o})	90
Resolution range (Å)	49.34-2.20 (2.24-2.20)
Completeness (%)	99.0 (92.1)
No. of reflections	476757 (23585)
No. of unique reflections	79924 (4167)
Multiplicity	6.0 (5.7)
Mean (I)/σ(I)	8.8 (0.6)
Mn(I) half-set correlation CC(1/2)	0.998 (0.265)
R_{merge}	0.111 (2.761)
Wilson B-factor (Å^2)	54.281
R_{factor}	0.250
R_{free}	0.290

$R = \frac{\sum_{hkl}||F_o|-|F_c||}{\sum_{hkl}|F_o|}$ where F_o and F_c are the observed and calculated structure factor amplitudes, respectively
$Rfree$ calculated from 5% of all unique reflections

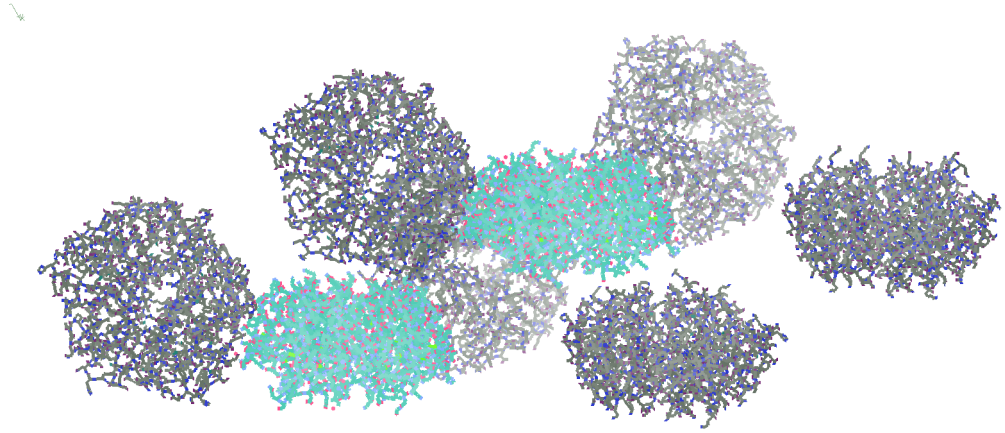


Figure 48: Screenshot from Coot showing the two pentamers (light blue) in the asymmetric unit along with neighboring pentamers related by symmetry (grey).

PG and EG is of similar structure, it is possible that this could also be the case for Kdo binding. Although Kdo was co-crystallized with both 10x and 50x Kdo concentration, this 10x data set was chosen since it had the highest resolution and completeness. Data sets with 50x Kdo concentrations were solved, but there were no positive electron densities that could fit Kdo after the first refinement and these lower resolution data sets were not pursued further. Overall, from the lack of binding in the NMR titration experiments and crystal structure, it can therefore be concluded that Kdo has low-to-none binding affinity to hLTB. We concluded that Kdo alone is not sufficient to study the interaction between hLTB and LPS. More effort should be made into solubilizing LPS or obtaining the hydrophilic core oligosaccharide for NMR and crystallization experiments.

4 Summary and future plans

ETEC and *Vibrio cholerae* cause diarrhea, and are endemic in the developing world. They both secrete structurally and functionally similar enterotoxins that share two binding sites, binding the primary receptor, GM1 and blood group antigens. hLTB has also been shown to bind to LPS, which may explain the different toxin delivery mechanisms and severities of the two diseases. Since the LPS binding site is believed to overlap with the secondary binding site, it could also explain why ETEC pathology does not appear to be blood group dependent

In this project, we wanted to localize and characterize the binding site of LPS to hLTB in order to achieve a greater understanding of these interactions. For this, an NMR backbone assignment of hLTB was completed as a preliminiary step towards titration experiments between with LPS and Kdo. Co-crystallization of Kdo with hLTB was also attempted.

4.1 Backbone assignment of hLTB

The production protocol of hLTB was improved and *Vibrio* sp. 60 was successfully adapted to grow in D_2O for expression of perdeuterated hLTB. This is the first time *Vibrio* has been adapted to D_2O to our knowledge. The improved expression protocol allows production of larger quantities of ^{15}N-labeled hLTB for more titration experiments.

Although the hLTB pentamer is a large complex, with a size of 58.7 kDa, 3D spectra for backbone assignment could be collected by using TROSY-experiments with deuterated hLTB. By combining several redundant experiments, all backbone amide peaks visible in the TROSY-HSQC spectrum were assigned. Thus, the assignment could be used to characterize hLTB-ligand interaction in the titration experiments. Only 6 residues residing in loops surrounding the primary binding site could not be identified due to lack of signal. Only 79% of the ^{13}C' could be assigned due to poor signal to noise. This lays the groundwork for future titration experiments and structural studies of ligand interactions with hLTB.

For future work, a new HNCO experiment with a higher number of scans should be recorded in order to assign the remaining ^{13}C' and decrease the number of ambiguities in the assignment. It has been shown that the loop around S55 and I58-S60 participates in GM1 binding to the primary binding site [53]. It is therefore possible that the loop is stabilized upon binding and that further assignment could be achieved by collecting new 3D-backbone spectra in the presence of GM1-pentasaccharide. However, the observable peaks may not represent the native state of the residues and it is uncertain if the new assignment could be used to study interaction with other ligands.

4.1.1 NEO titration

The analysis of the previously performed NEO titration experiment supports the hypothesis of Holmner et al [53] that there might be cross talk between the primary and secondary binding site upon NEO binding. This experiment demonstrates how crystallography and NMR titration experiments can give complementary information.

Future work should test the hypothesis of cross talk between the two binding sites by titrating ligands binding to the primary and secondary binding site in a competetive binding study.

4.2 LPS binding

The initial LPS titration experiment resulted in signal loss due to hLTB binding to LPS aggregates. To solve this problem, two approaches were attempted; detergent assisted solubilization and hydrolysis of the core oligsaccharide. DPC turned out to not be a very suitable detergent for this purpose, possibly because of formation of larger aggregates and possible interference with LPS-hLTB binding. Hydrolysis of the core oligosaccharide resulted in a glycan with similar weight as the expected core oligosaccharide, however, the yield was not quantified. The titration of the hydrolyzed oligosaccharide did not give any results, possibly due to DPC in the solution, or low hydrolysis yield. However, the signal loss clearly shows that LPS binds to hLTB.

Future work on studying the LPS-hLTB interaction by NMR should concentrate on downsizing the LPS aggregates. Following the finalization of the assignment process, more detergents and lipids can be quickly screened for suitability. Wang et al [125] have successfully studied LPS in aqeous solution by incorporating it into dihexanoylphosphatidylcholine (DHCP) micelles. Therefore, this lipid seems like a promising candidate. Alternatively, nanodiscs have lately proven to be a useful tool to study membrane mimetics and could also be tried [126]. Although hydrolyzation of LPS did not give any results, a protocol for extraction and purification of LPS from *E. coli* Ra-mutant can be developed for large scale hydrolysis to obtain pure core oligosaccharide. If this is achieved, the oligosaccharide could also be used in co-crystallization experiments.

4.3 Kdo binding

Kdo has been shown to bind to hLTB independently of the remaining core oligosaccharide on LPS. It was therefore titrated to hLTB and co-crystallized. The titration experiment shows that there is poor to no binding between Kdo and hLTB and the interaction could therefore not be studied by NMR. The crystal structure also did not show Kdo binding.

However the solved crystal structure is of higher resolution than the crystal structure currently available on PDB (PDB ID:2O2L). The structure should therefore be further

refined and uploaded to the database. The datasets collected from crystals co-crystallized with 1:50 hLTB:Kdo should also be further refined to reveal potential Kdo binding. A protocol for growing crystals overnight has been developed and crystals in the laboratory can be used to co-crystallize hLTB with even higher Kdo concentrations. These can also be used to try to co-crystallize or soak with other ligands. However, other cryoprotectants than PG should be investigated.

4.4 Conclusion

This project has laid the groundwork for studying the interaction between hLTB and different ligands by NMR, and progress has been made in studying the binding between LPS and hLTB by both crystallography and NMR. This work will hopefully contribute to the final elucidation of the LPS-hLTB interaction and function, giving a greater understanding of toxin delivery mechanisms and, perhaps, the lack of blood group dependence, important in ETEC pathogenesis.

References

[1] UNICEF/WHO, *Diarrhoea: Why children are still dying and what can be done*. World Health Organization, 2006.

[2] M. Kosek, C. Bern, and R. L. Guerrant, "The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000," *Bull World Health Organ*, vol. 81, no. 3, pp. 197–204, 2003.

[3] S. Huilan, L. G. Zhen, M. M. Mathan, M. M. Mathew, R. Olarte, J.and Espejo, K. M. U, M. A. Ghafoor, M. A. Khan, Z. Sami, and R. G. Sutton, "Etiology of Acute Diarrhea among Children in Developing Countries a Multicenter Study in 5 Countries," *B World Health Organ*, vol. 69, no. 5, pp. 549–555, 1991.

[4] F. Qadri, A. M. Svennerholm, A. S. Faruque, and R. B. Sack, "Enterotoxigenic *Escherichia coli* in developing countries: epidemiology, microbiology, clinical features, treatment, and prevention," *Clin Microbiol Rev*, vol. 18, no. 3, pp. 465–83, 2005.

[5] R. B. Sack, "The discovery of cholera - like enterotoxins produced by *Escherichia coli* causing secretory diarrhoea in humans," *Indian J Med Res*, vol. 133, pp. 171–80, 2011.

[6] H. Brüssow, H. Rahim, and W. Freire, "Epidemiological Analysis of Serologically Determined Rotavirus and Enterotoxigenic *Escherichia coli* Infections in Ecuadorian Children," *J Clin Microbiol*, vol. 30, no. 6, pp. 1585–7, 1992.

[7] F. Qadri, A. Saha, T. Ahmed, A. Al Tarique, Y. A. Begum, and A. M. Svennerholm, "Disease Burden Due to Enterotoxigenic *Escherichia coli* in the First 2 Years of Life in an Urban Community in Bangladesh," *Infect Immun*, vol. 75, no. 8, pp. 3961–3968, 2007.

[8] C. Wennerås and V. Erling, "Prevalence of enterotoxigenic *Escherichia coli*-associated diarrhoea and carrier state in the developing world," *J Health Popul Nutr*, vol. 22, no. 4, pp. 370–82, 2004.

[9] C. B. Dalton, E. D. Mintz, J. G. Wells, C. A. Bopp, and R. V. Tauxe, "Outbreaks of enterotoxigenic *Escherichia coli* infection in American adults: a clinical and epidemiologic profile," *Epidemiol Infect*, vol. 123, no. 1, pp. 9–16, 1999.

[10] E. J. Nelson, J. B. Harris, J. Morris, J. G., S. B. Calderwood, and A. Camilli, "Cholera transmission: the host, pathogen and bacteriophage dynamic," *Nat Rev Microbiol*, vol. 7, no. 10, pp. 693–702, 2009.

[11] C. C. Unger, S. S. Salam, M. S. Sarker, R. Black, A. Cravioto, and S. El Arifeen, "Treating diarrhoeal disease in children under five: the global picture," *Arch Dis Child*, vol. 99, no. 3, pp. 273–8, 2014.

[12] T. Ahmed, T. R. Bhuiyan, K. Zaman, D. Sinclair, and F. Qadri, "Vaccines for preventing enterotoxigenic *Escherichia coli* (ETEC) diarrhoea," *Cochrane Database Syst Rev*, vol. 7, p. CD009029, 2013.

[13] K. P. Allen, M. M. Randolph, and J. M. Fleckenstein, "Importance of heat-labile enterotoxin in colonization of the adult mouse small intestine by human enterotoxigenic *Escherichia coli* strains," *Infect Immun*, vol. 74, no. 2, pp. 869–875, 2006.

[14] T. Jonhson and L. K. Nolan, "Pathogenomics of the Virulence Plasmids of *Escherichia coli*," *Microbiol Immunol*, vol. 73, no. 4, pp. 750–774, 2009.

[15] R. A. Giannella and E. A. Mann, "*E. coli* heat-stable enterotoxin and guanylyl cyclase C: new functions and unsuspected actions.," *Trans Am Clin Climatol Assoc*, vol. 114, pp. 67–86, 2003.

[16] T. D. Connell and R. K. Holmes, "Characterization of hybrid toxins produced in *Escherichia coli* by assembly of A and B polypeptides from type I and type II heat-labile enterotoxins," *Infect Immun*, vol. 60, no. 4, pp. 1653–1661, 1992.

[17] Y. A. Begum, N. I. Baby, A. S. G. Faruque, N. Jahan, A. Cravioto, A. M. Svennerholm, and F. Qadri, "Shift in Phenotypic Characteristics of Enterotoxigenic *Escherichia coli* (ETEC) Isolated from Diarrheal Patients in Bangladesh," *Plos Neglect Trop D*, vol. 8, no. 7, 2014.

[18] W. Gaastra and A. M. Svennerholm, "Colonization factors of human enterotoxigenic *Escherichia coli* (ETEC)," *Trends Microbiol*, vol. 4, no. 11, pp. 444–452, 1996.

[19] M. K. Wolf, "Occurrence, Distribution, and Associations of O and H Serogroups, Colonization Factor Antigens, and Toxins of Enterotoxigenic *Escherichia coli*," *Clin Microbiol Rev*, vol. 10, no. 4, pp. 569–584, 1996.

[20] R. I. Walker, D. Steele, and T. Aguado, "Analysis of strategies to successfully vaccinate infants in developing countries against enterotoxigenic *E. coli* (ETEC) disease," *Vaccine*, vol. 25, no. 14, pp. 2545–2566, 2007.

[21] T. K. Sixma, S. E. Pronk, K. H. Kalk, E. S. Wartna, B. A. van Zanten, B. Witholt, and W. G. Hol, "Crystal structure of a cholera toxin-related heat-labile enterotoxin from *E. coli*," *Nature*, vol. 351, no. 6325, pp. 371–377, 1991.

[22] W. I. Lencer and B. Tsai, "The intracellular voyage of cholera toxin: going retro," *Trends Biochem Sci*, vol. 28, no. 12, pp. 639–645, 2003.

[23] E. K. Spicer and J. A. Noble, "*Escherichia coli* heat-labile enterotoxin. Nucleotide sequence of the A subunit gene," *J Biol Chem*, vol. 257, no. 10, pp. 5716–5721, 1982.

[24] H. Hofstra and B. Witholt, "Kinetics of Synthesis, Processing, and Membrane Transport of Heat-labile Enterotoxin, a Periplasmic Protein in *Escherichia coli*," *J Biol Chem*, vol. 259, no. 24, pp. 15182–15187, 1984.

[25] E. A. Merritt, S. E. Pronk, T. K. Sixma, K. H. Kalk, B. A. van Zanten, and W. G. Hol, "Structure of partially-activated *E. coli* heat-labile enterotoxin (LT) at 2.6 Å resolution," *FEBS Lett*, vol. 337, no. 1, pp. 88–92, 1994.

[26] T. Yamamoto, T. Tamura, M. Ryoji, A. Kaji, T. Yokota, and T. Takano, "Sequence analysis of the heat-labile enterotoxin subunit B gene originating in human enterotoxigenic *Escherichia coli*," *J Bacteriol*, vol. 152, no. 1, pp. 506–509, 1982.

[27] C. Cheesman, L. W. Ruddock, and R. B. Freedman, "The refolding and reassembly of *Escherichia coli* heat-labile enterotoxin B-subunit: analysis of reassembly-competent and reassembly-incompetent unfolded states," *Biochemistry*, vol. 43, no. 6, pp. 1609–1617, 2004.

[28] S. J. S. Hardy, J. Holmgren, S. Johansson, J. Sanchez, and T. R. Hirst, "Coordinated assembly of multisubunit proteins - Oligomerization of bacterial enterotoxins *in vivo* and *in vitro*," *Proc Natl Acad Sci U S A*, vol. 85, no. 19, pp. 7109–7113, 1988.

[29] S. J. Streatfield, M. Sandkvist, T. K. Sixma, M. Bagdasarian, W. G. Hol, and T. R. Hirst, "Intermolecular interactions between the A and B subunits of heat-labile enterotoxin from *Escherichia coli* promote holotoxin assembly and stability *in vivo*," *Proc Natl Acad Sci U S A*, vol. 89, no. 24, pp. 12140–12144, 1992.

[30] B. Mudrak and M. J. Kuehn, "Heat-Labile Enterotoxin: beyond G_{M1} Binding," *Toxins (Basel)*, vol. 2, no. 6, pp. 1445–70, 2010.

[31] M. Tauschek, R. J. Gorrell, R. A. Strugnell, and R. M. Robins-Browne, "Identification of a protein secretory pathway for the secretion of heat-labile enterotoxin by an enterotoxigenic strain of *Escherichia coli*," *Proc Natl Acad Sci U S A*, vol. 99, no. 10, pp. 7066–7071, 2002.

[32] A. L. Horstman and M. J. Kuehn, "Bacterial Surface Association of Heat-labile Enterotoxin through Lipopolysaccharide after Secretion via the General Secretory Pathway," *J Biol Chem*, vol. 277, no. 36, pp. 32538–32545, 2002.

[33] J. Holmgren, "Comparison of Tissue Receptors for Vibrio-Cholerae and *Escherichia coli* Enterotoxins by Means of Gangliosides and Natural Cholera Toxoid," *Infect Immun*, vol. 8, no. 6, pp. 851–859, 1973.

[34] J. P. Nataro and J. B. Kaper, "Diarrheagenic *Escherichia coli*," *Clin Microbiol Rev*, vol. 11, no. 1, pp. 142–201, 2011.

[35] J. Moss and S. Richardson, "Activation of Adenylate-Cyclase by Heat-Labile *Escherichia coli* Enterotoxin - Evidence for ADP Ribosyltransferase Activity Similar to That of Choleragen," *J Clin Invest*, vol. 62, no. 2, pp. 281–286, 1978.

[36] T. Yamamoto and T. Yokota, "Sequence of Heat-Labile Enterotoxin of *Escherichia coli* Pathogenic for Humans," *J Bacteriol*, vol. 155, no. 2, pp. 728–733, 1983.

[37] S. Schlör, S. Riedl, J. Blass, and J. Reidl, "Genetic Rearrangements of the Regions Adjacent to Genes Encoding Heat-Labile Enterotoxins (eltAB) of Enterotoxigenic *Escherichia coli* Strains," *Appl Environ Microbiol*, vol. 66, no. 1, pp. 352–358, 2000.

[38] T. Yamamoto, T. Tamura, T. Yokota, and T. Takano, "Overlapping Genes in the Heat-Labile Enterotoxin Operon Originating from *Escherichia coli* Human Strain," *Mol Gen Genet*, vol. 188, no. 2, pp. 359–359, 1982.

[39] T. Yamamoto, T. Nakazawa, T. Miyata, A. Kaji, and T. Yokota, "Evolution and Structure of two ADP-Ribosylation Enterotoxins, *Escherichia coli* Heat-Labile Toxin and Cholera Toxin," *FEBS Lett*, vol. 169, no. 2, pp. 241–246, 1984.

[40] J. D. Trachman and W. K. Maas, "Temperature Regulation of Heat-Labile Enterotoxin (LT) Synthesis in *Escherichia coli* is Mediated by an Interaction of H-NS Protein with the LT A-Subunit DNA," *J Bacteriol*, vol. 180, no. 14, pp. 3715–3718, 1998.

[41] J. Yang, M. Tauschek, R. Strugnell, and R. R. M., "The H-NS protein represses transcription of the eltAB operon, which encodes heat-labile enterotoxin in enterotoxigenic *Escherichia coli*, by binding to regions downstream of the promoter," *Microbiology*, vol. 151, no. 4, pp. 1199–1208, 2005.

[42] J. Teye and J. Trachman, "Effect of Acidic pH on Expression of Heat-Labile Enterotoxin (LT) by *Escherichia coli (E. coli)*," *Hostos Journal of StudentResearch*, vol. 3, no. 1, pp. 10–13, 2011.

[43] J. D. Trachman and M. Yasmin, "Thermo-osmoregulation of heat-labile enterotoxin expression by *Escherichia coli*," *Curr Microbiol*, vol. 49, no. 5, pp. 353–360, 2004.

[44] J. Yang, D. L. Baldi, M. Tauschek, R. A. Strugnell, and R. M. Robins-Browne, "Transcriptional regulation of the yghJ-pppA-yghG-gspCDEFGHIJKLM cluster, encoding the type II secretion pathway in enterotoxigenic *Escherichia coli*," *J Bacteriol*, vol. 189, no. 1, pp. 142–150, 2007.

[45] K. Takashi, I. Fujita, and K. Kobari, "Effects of short chain fatty acids on the production of heat-labile enterotoxin from enterotoxigenic *Escherichia coli*," *Jpn J Pharmacol*, vol. 50, no. 4, pp. 495–498, 1989.

[46] H. J. Binder, "Role of colonic short-chain fatty acid transport in diarrhea," *Annu Rev Physiol*, vol. 72, pp. 297–313, 2010.

[47] M. A. Lasaro, J. F. Rodrigues, C. Mathias-Santos, B. E. Guth, A. Regua-Mangia, Piantino, A. J. Ferreira, M. Takagi, J. Cabrera-Crespo, M. E. Sbrogio-Almeida, and L. C. de Souza Ferreira, "Production and release of heat-labile toxin by wild-type human-derived enterotoxigenic *Escherichia coli*," *FEMS Immunol Med Microbiol*, vol. 48, no. 1, pp. 123–131, 2006.

[48] B. Mudrak, D. L. Rodriguez, and M. J. Kuehn, "Residues of Heat-labile Enterotoxin Involved in Bacterial Cell Surface Binding," *J Bacteriol*, vol. 191, no. 9, pp. 2917–25, 2009.

[49] J. Holmgren, I. Lönnroth, J. Månsson, and L. Svennerholm, "Interaction of cholera toxin and membrane GM1 ganglioside of small intestine," *Proc Natl Acad Sci U S A*, vol. 72, no. 7, pp. 2520–2524, 1975.

[50] W. E. Minke, C. Roach, W. G. Hol, and C. L. Verlinde, "Structure-based exploration of the ganglioside GM1 binding sites of *Escherichia coli* heat-labile enterotoxin and cholera toxin for the discovery of receptor antagonists," *Biochemistry*, vol. 38, no. 18, pp. 5684–5692, 1999.

[51] E. A. Merritt, T. K. Sixma, K. H. Kalk, B. A. van Zanten, and W. G. Hol, "Galactose-binding site in *Escherichia coli* heat-labile enterotoxin (LT) and cholera toxin (CT)," *Mol Microbiol*, vol. 13, no. 4, pp. 745–753, 1994.

[52] W. B. Turnbull, B. L. Precious, and S. W. Homans, "Dissecting the cholera toxin-ganglioside GM1 interaction by isothermal titration calorimetry," *J Am Chem Soc*, vol. 126, no. 4, pp. 1047–1054, 2004.

[53] A. Holmner, A. Mackenzie, M. Ökvist, L. Jansson, M. Lebens, S. Teneberg, and U. Krengel, "Crystal Structures Exploring the Origins of the Broader Specificity of *Escherichia coli* Heat-labile Enterotoxin Compared to Cholera Toxin," *J Mol Biol*, vol. 406, no. 3, pp. 387–402, 2011.

[54] S. Teneberg, T. R. Hirst, J. Ångström, and K. A. Karlsson, "Comparison of the glycolipid-binding specificities of cholera toxin and porcine *Escherichia coli* heat-labile enterotoxin: identification of a receptor-active non-ganglioside glycolipid for the heat-labile toxin in infant rabbit small intestine," *Glycoconj J*, vol. 11, no. 6, pp. 533–540, 1994.

[55] J. E. Heggelund, E. Haugen, B. Lygren, Å. Mackenzie, Å. Holmner, F. Vasile, J. J. Reina, A. Bernardi, and U. Krengel, "Both El Tor and classical cholera toxin bind blood group determinants," *Biochem Bioph Res Co*, vol. 418, no. 4, pp. 731–735, 2012.

[56] A. Holmner, M. Lebens, S. Teneberg, J. Ångström, M. Ökvist, and U. Krengel, "Novel Binding Site Identified in a Hybrid between Cholera Toxin and Heat-labile Enterotoxin: 1.9 Å Crystal Structure Reveals the Details," *Structure*, vol. 12, no. 9, pp. 1655–67, 2004.

[57] Å. Holmner, G. Askarieh, M. Ökvist, and U. Krengel, "Blood Group Antigen Recognition by *Escherichia coli* Heat-labile enterotoxin," *J Mol Biol*, vol. 371, no. 3, pp. 754–64, 2007.

[58] S. Henry, R. Oriol, and B. Samuelsson, "Lewis Histo-Blood Broup System and Associated Secretory Phenotypes," *Vox Sang*, vol. 69, no. 3, pp. 166–182, 1995.

[59] S. M. Henry, "Molecular diversity in the biosynthesis of GI tract glycoconjugates. A blood group related chart of microorganism receptors," *TRANSFUS CLIN BIOL*, vol. 8, no. 3, pp. 226–230, 2001.

[60] A. Chaudhuri and C. R. DasAdhikary, "Possible role of blood-group secretory substances in the aetiology of cholera," *Trans R Soc Trop Med Hyg*, vol. 72, no. 6, pp. 664–665, 1978.

[61] A. S. Faruque, D. Mahalanabis, S. S. Hoque, and M. J. Albert, "The relationship between ABO blood groups and susceptibility to diarrhea due to *Vibrio cholerae* 0139," *Clin Infect Dis*, vol. 18, no. 5, pp. 827–828, 1994.

[62] C. R. Raetz and C. Whitfield, "Lipopolysaccharide Endotoxins," *Annu Rev Biochem*, vol. 71, pp. 635–700, 2002.

[63] C. R. Raetz, "Biochemistry of Endotoxins.," *Annu Rev Biochem*, vol. 59, pp. 129–170, 1990.

[64] A. L. Horstman and M. J. Kuehn, "Enterotoxigenic *Escherichia coli* Secretes Active Heat-labile Enterotoxin via Outer Membrane Vesicles," *J Biol Chem*, vol. 275, no. 17, pp. 12489–12496, 2009.

[65] A. L. Horstman, S. J. Bauman, and M. J. Kuehn, "Lipopolysaccharide 3-Deoxy-D-manno-octulosonic Acid (kdo) Core Determines Bacterial Association of Secreted Toxins," *J Biol Chem*, vol. 279, no. 9, pp. 8070–5, 2003.

[66] N. C. Kesty, K. M. Mason, M. Reedy, S. E. Miller, and M. J. Kuehn, "Enterotoxigenic *Escherichia coli* vesicles target toxin delivery into mammalian cells," *EMBO J*, vol. 23, no. 23, pp. 4538–4549, 2004.

[67] M. J. Kuehn and N. C. Kesty, "Bacterial outer membrane vesicles and the host-pathogen interaction," *Genes Dev*, vol. 19, no. 22, pp. 2645–55, 2005.

[68] A. M. Johnson, R. S. Kaushik, D. H. Francis, J. M. Fleckenstein, and P. R. Hardwidge, "Heat-labile Enterotoxin Promotes *Escherichia coli* Adherence to Intestinal Epithelial Cells," *Infect Immun*, vol. 191, no. 1, pp. 178–186, 2009.

[69] M. H. Levitt, *spin dynamics : Basics of Nuclear Magnetic Resonance*. John Wiley Sons, 2nd ed., 2008.

[70] P. W. Atkins and J. De Paula, *Physical Chemistry*. Oxford University Press, 10th ed., 2014.

[71] K. Wüthrich, *NMR of Proteins and Nucleic Acids*. Wiley, 1986.

[72] D. G. Reid, *Protein NMR Techniques*. Humana Press, 1997.

[73] J. Cavanagh, *Protein NMR Spectroscopy : Principles and Practice*. Academic Press, 2nd ed., 2007.

[74] T. Yamazaki, W. Lee, C. H. Arrowsmith, D. R. Muhandiram, and L. E. Kay, "A Suite of Triple Resonance NMR Experiments for the Backbone Assignment of ^{15}N, ^{13}C, ^{2}H Labeled Proteins with High Sensitivity," *J Am Chemical Soc*, vol. 116, no. 26, pp. 11655–11666, 1994.

[75] K. V. Pervushin, R. Riek, G. Wider, and K. Wüthrich, "Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution," *Proc Natl Acad Sci U S A*, vol. 94, no. 23, pp. 12366–12371, 1997.

[76] G. Bodenhausen and D. J. Ruben, "Natural abundance nitrogen-15 nmr by enhanced heteronuclear spectroscopy," *Chem Phys Lett*, vol. 69, no. 1, pp. 185–189, 1980.

[77] M. Salzmann, G. Wider, K. Pervushin, H. Senn, and K. Wüthrich, "TROSY-type Triple-Resonance Experiments for Sequential NMR Assignments of Large Proteins," *J Am Chemical Soc*, vol. 121, no. 5, pp. 844–848, 1999.

[78] K. V. Pervushin, G. Wider, and K. Wüthrich, "Single Transition-to-single Transition Polarization Transfer (ST2-PT) in [^{15}N,^{1}H]-TROSY," *J Biomol NMR*, vol. 12, no. 2, pp. 345–348, 1998.

[79] L. E. Kay, M. Ikura, R. Tschudin, and A. Bax, "Three-Dimensional Triple-Resonance NMR-Spectroscopy of Isotopically Enriched Proteins," *J Magn Reson*, vol. 89, no. 3, pp. 496–514, 1990.

[80] M. Ikura, L. E. Kay, and A. Bax, "A Novel Approach for Sequential Assignment of ^{1}H, ^{13}C, and ^{15}N Spectra of Proteins: Heteronuclear Triple-Resonance Three-Dimensional NMR Spectroscopy. Application to Calmodulin," *Biochemistry*, vol. 29, no. 19, pp. 4659–4667, 1990.

[81] S. Grzesiek and A. Bax, "Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR," *Journal of the American Chemical Society*, vol. 114, no. 16, pp. 6291–6293, 1992.

[82] S. Grzesiek and A. Bax, "An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins," *J Magn Res*, vol. 99, no. 1, pp. 201 – 207, 1992.

[83] R. T. Clubb, V. Thanabal, , and G. Wagner, "A Constant-Time 3-Dimensional Triple-Resonance Pulse Scheme to Correlate Intraresidue $^{1}H^{N}$, ^{15}N, and ^{13}C' Chemical-Shifts in ^{15}N-^{13}C-Labeled Proteins," *J Magn Res*, vol. 97, no. 1, pp. 213–217, 1992.

[84] D. S. Wishart and B. D. Sykes, "The ^{13}C Chemical-Shift Index: A Simple Method for the Identification of Protein Secondary Structure Using ^{13}C Chemical-shift Data," *Journal of biomolecular NMR*, vol. 4, no. 2, pp. 171–180, 1994.

[85] F. Shen, Y.and Delaglio, G. Cornilescu, and A. Bax, "TALOS+: a hybrid method for predicting protein backbone torsion angles from nmr chemical shifts," *Infect Immun*, vol. 44, no. 4, pp. 213–223, 2009.

[86] A. McPherson, *Introduction to Macromolecular Crystallography*. Wiley-Blackwell, 2nd ed., 2009.

[87] A. McPherson, *Crystallization of Biological Macromolecules*. Cold Spring Harbour Laboratory Press, 1999.

[88] D. Blow, *Outline of Crystallography for Biologists*. Oxford: Oxford University Press, 2002.

[89] B. Rupp, *Biomolecular Crystallography : Principles, Practice, and Application to Structural Biology*, vol. Garland Science. 2010.

[90] "Avanti Polar Lipids preparation of liposomes." `http://www.avantilipids.com/index.php?option=com_content&view=article&id=1384&Itemid=372`. Accessed: 2013-12-15.

[91] R. L. J. Keller, *Optimizing the process of nuclear magnetic resonance spectrum analysis and computer aided resonance assignment*. PhD thesis, 2004.

[92] F. Schumann, H. Riepl, T. Maurer, W. Gronwald, K.-P. Neidig, and H. Kalbitzer, "Combined chemical shift changes and amino acid specific chemical shift mapping of protein protein interactions," *J Biomol NMR*, vol. 39, no. 4, pp. 275–289, 2007.

[93] E. P., "Scaling and assessment of data quality," *Acta Crystallogr D Biol Crystallogr*, vol. 62, no. 1, pp. 72–82, 2006.

[94] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. Read, J., A. Vagin, and K. S. Wilson, "Overview of the CCP4 suite and current developments," *Acta Crystallogr D Biol Crystallogr*, vol. 67, no. 4, pp. 235–242, 2011.

[95] R. Leece and T. R. Hirst, "Expression of the B-subunit of *Escherichia coli* heat-labile enterotoxin in a marine *Vibrio* and in a mutant that is pleiotropically defective in the secretion of extracellular proteins," *J Gen Microbiol*, vol. 129, no. 6, pp. 533–540, 1992.

[96] S. Trygsland, "Understanding the cell surface interactions of enterotoxigenic *Eschericia coli*," Master's thesis, University of Oslo, 2010.

[97] M. R. Green and J. Sambrook, *Molecular cloning: a laboratory manual*, vol. 3. Cold Spring Harbor Laboratory Press, 4th ed., 2012.

[98] J. Marley, M. Lu, and C. Bracken, "A method for efficient isotopic labeling of recombinant proteins," *J Biomol NMR*, vol. 20, no. 1, pp. 71–5, 2001.

[99] M. Lebens, S. Johansson, J. Osek, M. Lindblad, and J. Holmgren, "Large-Scale Production of *Vibrio cholerae* Toxin-B Subunit for Use in Oral Vaccines," *Bio-Technology*, vol. 11, no. 13, pp. 1574–1578, 1993.

[100] O. Mosin and I. Ignatov, "Phenomenon of Biological Adaption to Deuterium Oxide," *Journal of Medicine, Physiology and Biophysics*, vol. 5, 2014.

[101] A. Sosa-Peinado, D. Mustafi, and M. W. Makinen, "Overexpression and Biosynthetic Deuterium Enrichment of TEM-1 β-Lactamase for Structural Characterization by Magnetic Resonance Methods," *Protein Expres Purif*, vol. 19, no. 2, pp. 235–245, 2000.

[102] S. W. Englander and N. R. Kallenbach, "Hydrogen exchange and structural dynamics of proteins and nucleic acids," *Q rev Biophys*, vol. 16, no. 4, pp. 521–655, 1984.

[103] M. Lu, Y. Zhang, and H. Zhang, "Isotope effects on cell growth and sporulation, and spore heat resistance, survival and spontaneous mutation of bacillus cereus by deuterium oxide culture," *Afr J Microbiol Res*, vol. 7, no. 8, pp. 604–611, 2013.

[104] P. Andjus and Vuĉelié, "D_2o-induced cell excitation," *J. Membrane Biol.*, vol. 115, no. 2, pp. 123–127, 1990.

[105] A. P. Middleberg, "Preparative Protein Refolding," *Trends Biotechnol*, vol. 20, no. 10, pp. 437–443, 2002.

[106] H. Lilie, E. Schwarz, and R. Rudolph, "Advances in refolding of proteins produced in *e. coli*," *Curr Opin Biotech*, vol. 9, no. 5, pp. 497–501, 1998.

[107] D. T. Browne, G. L. Kenyon, E. L. Packer, H. Sternlicht, and D. M. Wilson, "Studies of Macromolecular Structure by ^{1}3c Nuclear Magnetic Resonance. II. A Specific Labeling Approach to the Study of Histidine Residues in Proteins," *J Am Chem Soc*, vol. 95, no. 4, pp. 1316–1323, 1973.

[108] R. A. Venters, C. C. Huang, n. Farmer, B. T., R. Trolard, L. D. Spicer, and C. A. Fierke, "High-level ^{2}h/^{1}3c/^{1}5n labeling of proteins for nmr studies," *J Biomol NMR*, vol. 5, no. 4, pp. 339–344, 1973.

[109] K. V. Pervushin, G. Wider, R. Riek, and K. Wüthrich, "The 3D NOESY-[^{1}H,^{15}N,^{1}H]-ZQ-TROSY NMR experiment with diagonal peak suppression," *P Natl Acad Sci USA*, vol. 96, no. 17, pp. 9607–9612, 1999.

[110] S. Morin, "A practical guide to protein dynamics from ^{15}N spin relaxation in solution," *Prog Nucl Magn Reson Spectrosc*, vol. 59, no. 3, pp. 245–262, 2011.

[111] S. Teneberg, A. Berntsson, and J. Ångström, "Common Architecture of the Primary Galactose Binding Sites of Erythrina Corallodendron Lectin and Heat-labile Enterotoxin From *Escherichia coli* in Relation to the Binding of Branched Neolactohexaosylceramide," *J Biochem*, vol. 128, no. 3, pp. 481–491, 2000.

[112] K. A. Karlsson, S. Teneberg, J. Ångström, A. Kjellberg, T. R. Hirst, J. Berström, and H. Miller-Podraza, "Unexpected Carbohydrate Cross-binding by *Escherichia coli* Heat-labile Enterotoxin. Recognition of Human and Rabbit Target Cell Glycoconjugates in Comparison With Cholera Toxin," *J Biol Chem*, vol. 4, no. 11, pp. 1919–1928, 1997.

[113] C. R. Raetz, "Biochemistry of endotoxins," *Annu Rev Biochem*, vol. 59, pp. 129–70, 1990.

[114] Y. A. Knirel and M. A. Valvano, *Bacterial Lipopolysaccharides : Structure, Chemical Synthesis, Biogenesis, and Interaction with Host Cells*. Springer, 2011.

[115] N. C. Santos, A. C. Silva, M. A. Castanho, J. Martins-Silva, and C. Saldanha, "Evaluation of Lipopolysaccharide Aggregation by Light Scattering Spectroscopy," *Chembiochem*, 2003.

[116] L. Jansson, J. Ångström, M. Lebens, and S. Teneberg, "No direct binding of the heat-labile enterotoxin of *Escherichia coli* to *E. coli* lipopolysaccharides," *Glycoconj J*, vol. 27, no. 1, pp. 171–9, 2010.

[117] M. Caroff, A. Tacken, and L. Szabó, "Detergent-accelerated hydrolysis of bacterial endotoxins and determination of the anomeric configuration of the glycosyl phosphate present in the "isolated lipid a" fragment of the *Bordetella pertussis* endotoxin," *Carbohydr Res*, vol. 175, no. 2, pp. 273–82, 1988.

[118] J. Kubiak, J. Brewer, S. Hansen, and L. A. Bagatolli, "Lipid Lateral Organization on Giant Unilamellar Vesicles Containing Lipopolysaccharides," *Biophys J*, vol. 100, no. 4, pp. 978–86, 2011.

[119] J. Mares, S. Kumaran, M. Gobbo, and O. Zerbe, "Interactions of Lipopolysaccharide and Polymyxin Studied by NMR Spectroscopy," *J Biol Chem*, vol. 284, no. 17, pp. 11498–506, 2009.

[120] D. A. Kallick, M. R. Tessmer, C. R. Watts, and C. Y. Li, "The use of dodecylphosphocholine micelles in solution NMR," *J Magn Reson Ser B*, vol. 109, no. 1, pp. 60–65, 1995.

[121] S. Rezania, N. Amirmozaffari, B. Tabarrei, M. Jeddi-Tehrani, O. Zarei, R. Alizadeh, F. Masjedian, and A. H. Zarnani, "Extraction, purification and characterization of lipopolysaccharide from *escherichia coli* and *salmonella typhi*," *Avicenna Journal of Medical Biotechnology*, vol. 3, no. 1, pp. 3–9, 2011.

[122] T. Hofstad and B. Aadnegard, "Oligosaccharides obtained by partial hydrolysis of lipopolysaccharide from Fusobacterium nucleatum-F1," *J Periodontal Res*, vol. 21, no. 2, pp. 163–168, 1986.

[123] N. K. Birkeland and T. Hofstad, "Oligosaccharides Obtained by Partial Hydrolysis of Lipopolysaccharides from *Leptotrichia-Buccalis*," *Scandinavian Journal of Dental Research*, vol. 93, no. 5, pp. 432–435, 1985.

[124] "Mass Spectrometry of Glycans." `http://www.sigmaaldrich.com/technical-documents/articles/biology/glycobiology/mass-spectrometry-of-glycans.html`. Accessed: 2015-01-05.

[125] W. Wang, H. J. Sass, U. Zähringer, and S. Grzesiek, "Structure and Dynamics of ^{13}C,^{15}N-labeled Lipopolysaccharides in a Membrane Mimetic," *Angew Chem Int Ed Engl*, vol. 47, no. 51, pp. 9870–9874, 2008.

[126] T. H. Bayburt and S. G. Sligar, "Membrane protein assembly into Nanodiscs," *FEBS Lett*, vol. 584, no. 9, pp. 1721–1727, 2010.

[127] S. Panke, A. Meyer, C. M. Huber, B. Witholt, and M. G. Wubbolts, "An Alkane-Responsive Expression System for the Production of Fine Chemicals," *Appl Environ Microbiol*, vol. 65, pp. 2324–32, 1995.

Appendices

A hLTB sequence

hLTB: APQSITELCSEYHNTQIYTINDKILSYTESMAGKREMVIITFKSGATFQVEVPGSQHIDS QKKAIERMKDTL-
RITYLTETKIDKLCVWNNKTPNSIAAISMEN

B Chemicals and materials

Chemical and commercial buffers	**Vendor**
2-keto-3deoxyoctonate (Kdo) ammonium salt	Sigma
Acetic Acid	Merck
Ampicillin	AppliChem
Ammonium chloride ($^{14}NH_4Cl$)	Sigma
Ammonium chloride ($^{15}NH_4Cl$, 99% ^{15}N)	Cortecnet
BME vitamin mix (100x)	Sigma
β-mercaptoethanol	Sigma
Carbenicillin	AppliChem
Coomasie Brilliant Blue G250	Amersham Biosciences
Calsium chloride ($CaCl_2$)	Merck
Chloroform	VWR
D-(+)-Galactose	Fluka
D-Glucose	Fluka
D-Glucose ($^{13}C_6$, 99% ^{13}C)	Cortecnet
D-Glucose ($^{13}C_6$-$^{2}H_{12}$, 99% ^{13}C, 97% ^{2}H)	Cortecnet
Deuterium oxide (D_2O)	euriso-top
Disodium hydrogen phosphate (Na_2HPO_4)	G-Biosciences
Dipotassium hydrogen phosphate (Na_2HPO_4)	G-Biosciences
Dodecylphosphocholine (DPC)	Avanti
Dodecylphosphocholine (d_{38}-DPC)	Isotec
Ethanol (Absolut prima)	Arcus
Guanidine hydrochloride	Fluka
Hydrogen-Chloride (HCl)	Merck
Isopropyl-β-D-thiogalactopyranoside	Sigma
L-glutathione oxidized	BioChemica
L-glutathione reduced	BioChemica
Lipopolysaccharide from *E. coli* EH100 Ra mutant (LPS-Ra)	Sigma
Magnesium sulfate ($MgSO_4$)	VWR
Methanol	VWR
Milli-Q H_2O	Millipore
MES SDS running buffer, 20x	Invitrogen
NuPAGE	Life technologies
NuPAGE Loading Buffer (4x)	Life technologies
Propylene glycol	Ampresco
Potassium chloride (KCl)	Fluka
Potassium dihydrogen phosphate (NaH_2PO_4)	G-biosciences
SeeBlue Plus standard, molecular weight marker	Life technologies
Sodium azide (NaN_3	Merck
Sodium carbonate (Na_2CO_3)	Sigma
Sodium chloride (NaCl)	Prolabo
Sodium dihydrogen phosphate (NaH_2PO_4)	G-biosciences
Sodium hydroxide (NaOH)	Kebo Lab
Thiamine	Sigma
Tris-hydrogen chloride	Chalbiochem

Materials	**Vendor**
5mm, symmetrical microtube, matched D_2O, NMR-tube	Shigemi
5mm, THIN WALL, 7", 600 MHz NMR-tube	Wilmad
Crystallization Plate, 96 MRC	SWISSCI
Cuvettes	Sarstedt
Crystal Clear Sealing Tape	Henkel
Eppendorf tubes, 1.5 mL	Eppendorf
Litho loops	Molecular dimensions
NuPage Bis-Tris 4-12% gel	Life Technologies
PCR cuvette	Eppendorf
Silanized glass cover slides	Hampton research
Snake skin dialysis bag 3500 MWCO	Thermo scientific
TPP tissue culture plates, 24 well	Sigma
Vivaspin 20 Centrifugal concentrators	Sartorius
Vivaspin 5 Centrifugal concentrators	Sartorius

Crystallization Screens	**Vendor**
JCSG+ suite	Molecular Dimensions
Morpheus	Molecular Dimensions
PGA	Molecular Dimensions

Kits	**Vendor**
Morpheus additive kit	Molecular Dimensions

C Equipment

Instruments and lab equipment	Vendor
ÄKTA purifier-900	GE Healthcare
Avance II 600 MHz NMR-instrument	Bruker
Avanti Centrifuge J-26 XP	Beckman Coulter
Bench pH-meter, 3510	Jenway
Biofuge	Fresco Heareus
BX41 microscope	Olympus
Centrifuge 5810 R	Eppendorf
CP224s Analytical Balance	Sartorius
Electrophoresis power supply-EPS 601	GE Healthcare
IKA-WERK VF2 vortex mixer	Janke & Kunkel
NanoPhotometer	IMPLEN
Oryx4 robot	Douglas Instruments
PCB precision balance	Kern
Peltier Thermal Cycler (PTC)-200	MJ Research
Spectrafuge Mini Centrifuge	Labnet International Inc.
SteREO Discovery V12	Zeiss
TC-3000 PCR Thermal Cycler	Techne
Tuttnauer 3870-ML, autoclave	Tuttnauer
Ultrasonic cleaner	VWR

D Buffers and media

Tris storage buffer, 1000 mL	
Tris-HCl pH 7.5	20 mM
10 mL 1 M stock	
NaCl	200 mM
20 mL 5 M stock	
Add MQ-H_2O to 1000 mL, adjust to pH 7.5 by addition of HCl or NaOH	

PBS-buffer, 1000 mL		
NaCl	137 mM	8 g
KCl	2.7 mM	0.2 g
Na_2HPO_4	10.1 mM	1.44 g
NaH_2PO_4	2 mM	0.78 g
Add MQ-H_2O to 1000 mL, adjust to pH 7.2 by addition of HCl		

NMR buffer, 1000 mL		
NaCl	50 mM	2.9 g
Na_2HPO_4	13 mM	1.8 g
NaH_2PO_4	2 mM	0.78 g
Sodium azide	0.02%	0.2 g
Add MQ-H_2O to 1000 mL, adjust to pH to 6.5 by addition of HCl		

Galactose elution buffer, 100 mL		
D-(+)-Galactose	300 mM	5.4 g

Add 1 × PBS up to 100 mL, filter

M9 minimal medium, 1000 mL	
10X M9 salts (extra salty)	100 mL
BME Vitamin mix	10 mL
US* Trace elements	1 mL
Antibiotics (100 mg/mL)	1 mL
Glucose (50%)	8 mL
NH_4Cl	1 g
$MgSO_4$ (1 M)	2 mL
Thiamine (5 g/L)	1 mL
$CaCl_2$ (1 M)	0.1 mL
NaN_3 (20%)	2 mL

Sterilize all solutions by autoclaving or filtering before it is added to the buffer.
Ampicillin or carbenicillin was used as antibiotic depending on the expression time.

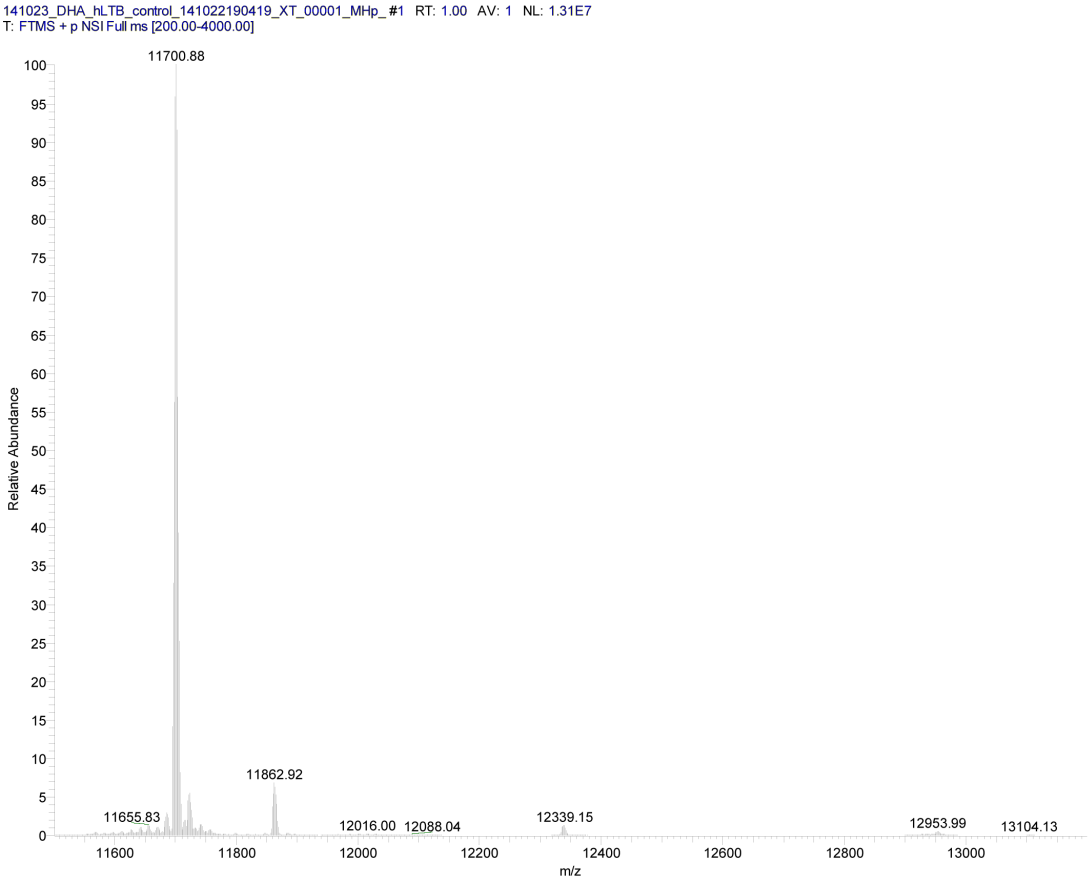
10X M9 salts (extra salty), 1000 mL	
Na_2HPO_4	60 g
KH_2PO_4	30 g
NaCl	150 g

Add MQ-H_2O to 1000 mL; adjust to pH 7.2; filter or autoclave.

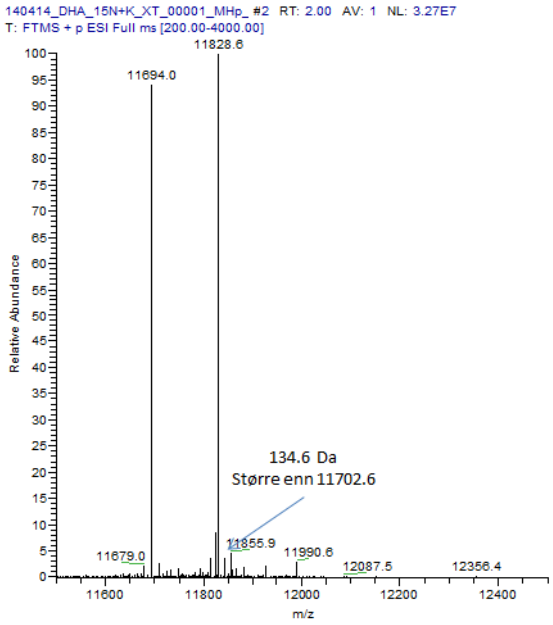
US* Trace elements, 1000 mL [127]	
$MnCl_2 \cdot 4H_2O$	1.50 g
$ZnSO_4$	1.05 g
H_3BO_3	0.30 g
$Na_2MoO_4 \cdot 2H_2O$	0.25 g
$CuCl_2 \cdot 2H_2O$	0.15 g
$FeSO_4 \cdot 7H_2O$	4.87 g
$CaCl_2 \cdot 2H_2O$	4.12 g

Coomasie staining solution	
Coomasie blue	0.05%
Acetic acid	10%
Ethanol	20 %

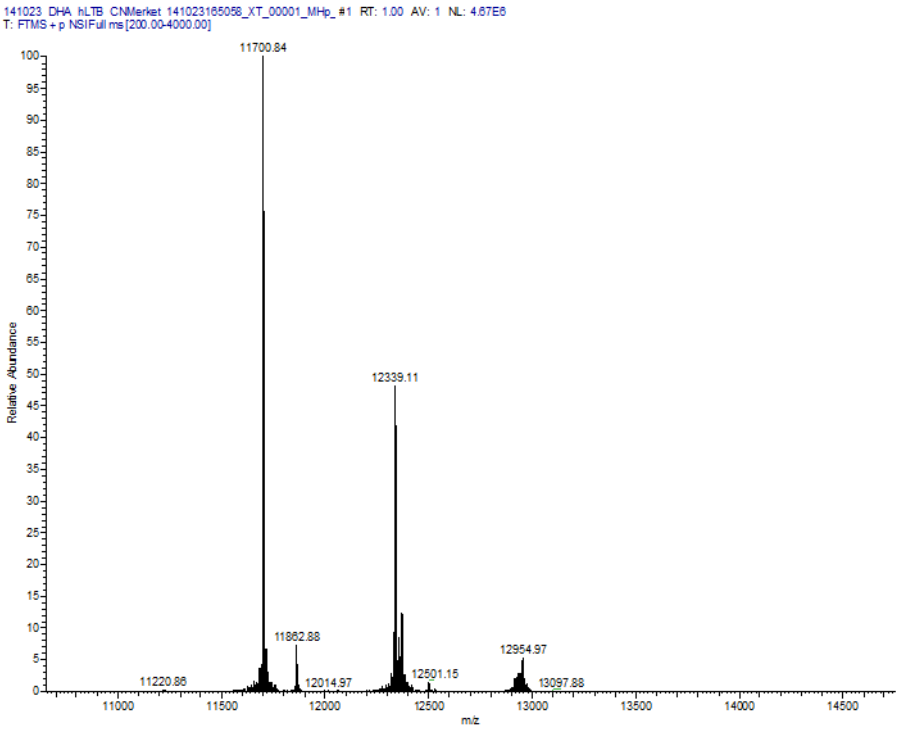
E MS results of hLTB



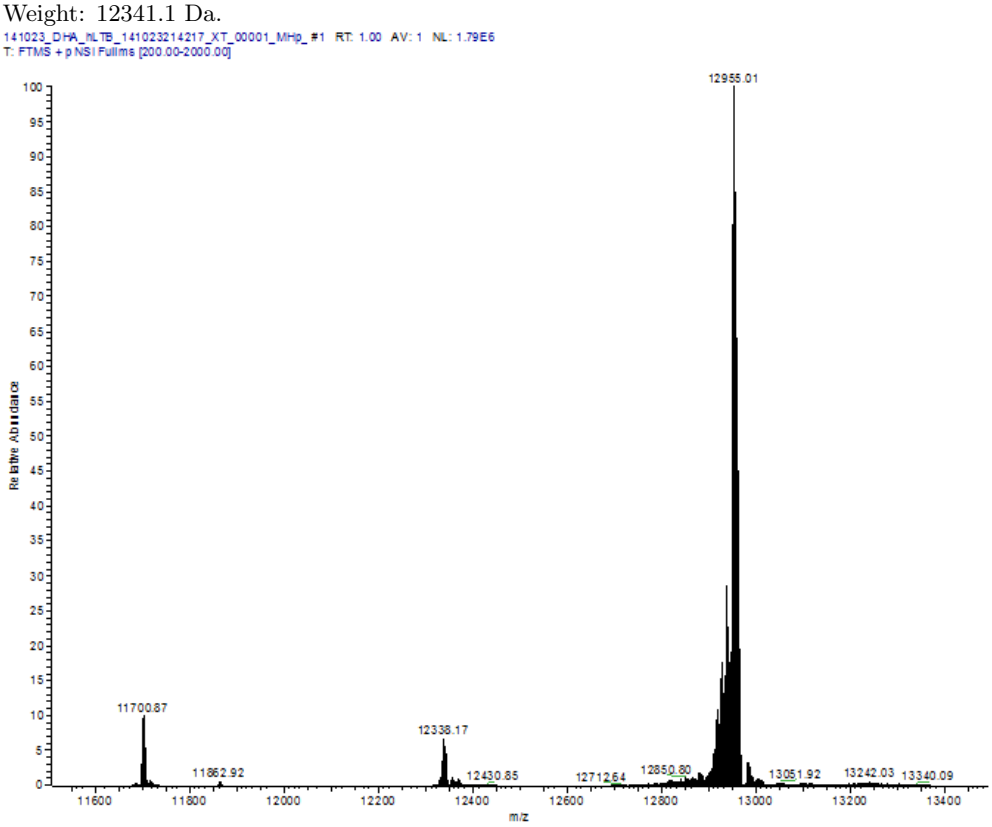
Over: Direct injection Top-down MS-spectrum of unlabeled hLTB. Weight: 11702.88 Da. (Theoretical weight 11702.3 Da)



Over: Direct injection Top-down MS-spectrum of ^{15}N-labeled hLTB and normal hLTB as controll. Weight: 11837.2 Da.



Over: Direct injection Top-down MS-spectrum of ^{15}N-^{13}C-labeled hLTB and normal hLTB as control. Weight: 12341.1 Da.

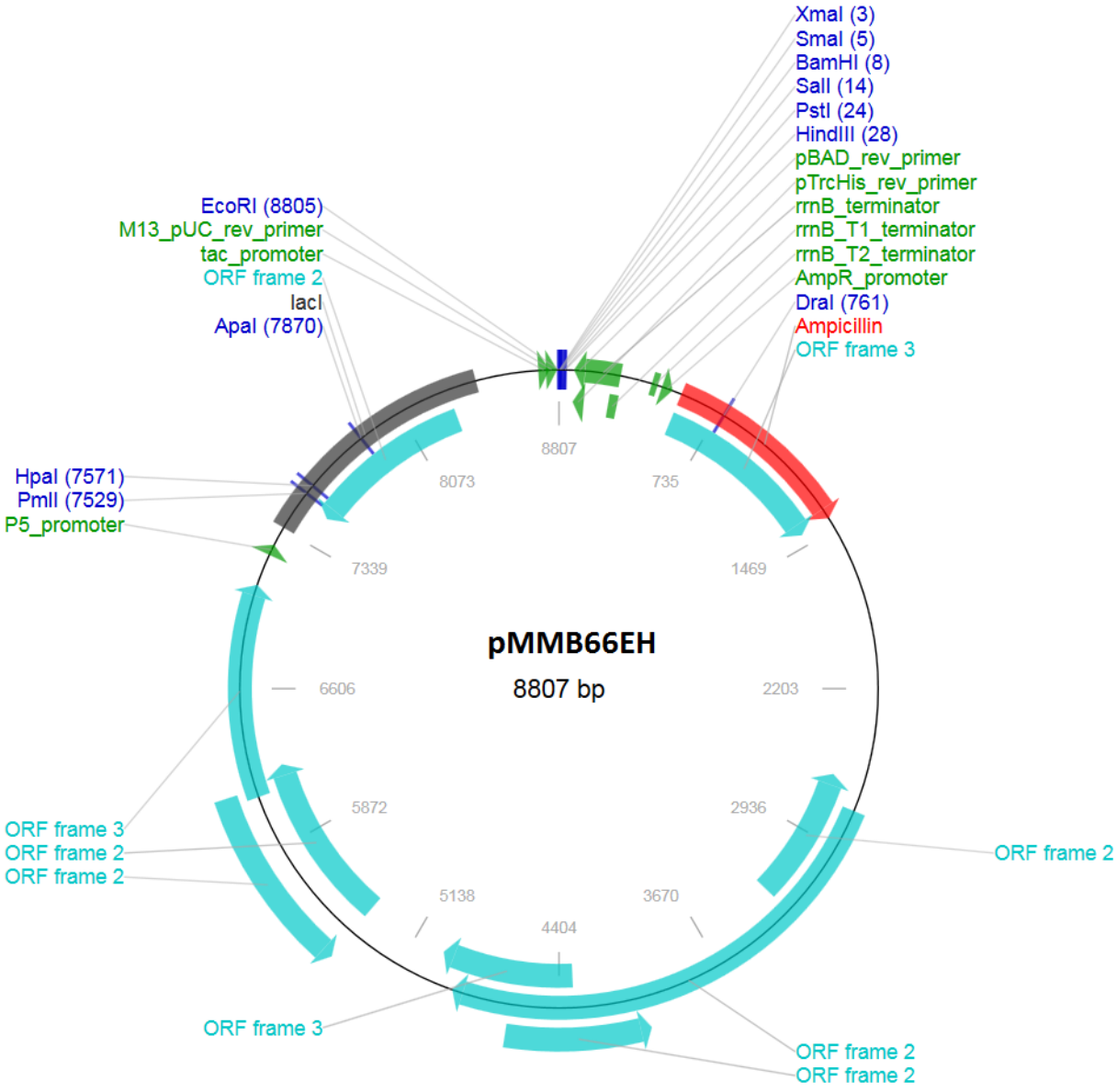


Over: Direct injection Top-down MS-spectrum of refolded ^{15}N-^{13}C-^{2}H-labeled hLTB. Weight 12957.01 Da. Theoretical weight calculated to 12334.03 Da by adding number of ^{15}N, ^{13}C and ^{2}H in the protein minus exchangeable ^{2}H.

The degree of incorporation was calculated by dividing the observed weight difference between two

degrees of incorporation, e.g. ^{15}N-hLTB and ^{15}N-^{13}C-hLTB, with the theoretical difference, assuming that incorporation of the different isotopes was independent of each other.

F Plasmid pMMB66EH



G Chemical shift values

ResID	ResName	Atomname	Shift	ResID	ResName	Atomname	Shift
3	Q	CB	29.336	11	E	H	7.343
3	Q	N	112.981	11	E	CB	28.356
3	Q	C	174.601	11	E	N	119.867
3	Q	CA	54.354	11	E	C	175.222
3	Q	H	8.154	11	E	CA	57.043
4	S	CB	66.261	12	Y	C	174.794
4	S	C	174.571	12	Y	N	117.661
4	S	H	7.279	12	Y	CB	40.739
4	S	N	109.867	12	Y	CA	57.878
4	S	CA	55.924	12	Y	H	7.049
5	I	H	9.672	14	N	N	115.275
5	I	C	175.463	14	N	CA	54.16
5	I	CA	63.952	14	N	CB	36.67
5	I	N	121.2	14	N	H	8.812
5	I	CB	36.608	15	T	C	175.721
6	T	CB	66.807	15	T	CA	57.961
6	T	CA	65.874	15	T	H	7.339
6	T	H	7.386	15	T	N	104.678
6	T	N	119.057	15	T	CB	72.007
7	E	H	7.226	16	Q	CB	32.679
7	E	CB	29.325	16	Q	H	8.727
7	E	N	123.324	16	Q	C	173.706
7	E	C	178.948	16	Q	N	117.382
7	E	CA	58.513	16	Q	CA	54.135
8	L	N	120.014	17	I	CB	37.024
8	L	C	178.074	17	I	C	176.562
8	L	H	7.741	17	I	H	8.29
8	L	CB	41.979	17	I	CA	57.003
8	L	CA	58.206	17	I	N	124.068
9	C	H	8.267	18	Y	CA	57.003
9	C	CB	38.551	18	Y	H	9.09
9	C	CA	57.957	18	Y	C	175.524
9	C	N	117.75	18	Y	N	127.389
9	C	C	176.569	18	Y	CB	40.812
10	S	CB	62.936	19	T	C	180.561
10	S	H	7.325	19	T	H	8.813
10	S	N	110.818	19	T	N	121.196
10	S	CA	59.666	19	T	CA	62.781

ResID	ResName	Atomname	Shift	ResID	ResName	Atomname	Shift
19	T	CB	67.644	27	Y	N	125.779
20	I	C	175.916	27	Y	C	172.92
20	I	CB	36.026	28	T	C	171.359
20	I	N	130.318	28	T	CB	70.051
20	I	CA	59.189	28	T	H	8.115
20	I	H	8.599	28	T	CA	60.594
21	N	CB	36.697	28	T	N	124.08
21	N	N	120.1	29	E	C	174.072
21	N	H	8.419	29	E	H	8.685
21	N	C	173.881	29	E	CB	33.512
21	N	CA	53.499	29	E	CA	54.161
22	D	CB	43.014	29	E	N	127.182
22	D	C	175.275	30	S	C	179.72
22	D	H	8.153	30	S	N	118.773
22	D	N	114.455	30	S	CB	64.266
22	D	CA	52.729	30	S	H	8.208
23	K	N	120.841	30	S	CA	54.832
23	K	CB	32.305	31	M	CB	33.761
23	K	C	174.962	31	M	C	173.543
23	K	H	8.678	31	M	H	8.954
23	K	CA	53.937	31	M	N	123.586
24	I	C	176.542	31	M	CA	53.475
24	I	H	9.726	32	A	CB	17.044
24	I	CB	38.775	32	A	N	127.523
24	I	CA	62.928	32	A	C	177.131
24	I	N	122.39	32	A	H	7.703
25	L	N	129.588	32	A	CA	52.353
25	L	CB	41.731	33	G	C	176.481
25	L	C	174.84	33	G	CA	46.752
25	L	CA	56.61	33	G	H	8.439
25	L	H	8.496	33	G	N	109.823
26	S	N	110.523	34	K	N	123.344
26	S	CB	66.612	34	K	CB	29.294
26	S	CA	55.758	34	K	H	9.141
26	S	H	7.337	34	K	CA	59.335
26	S	C	170.554	35	R	CA	52.785
27	Y	CA	56.2	35	R	N	120.63
27	Y	H	8.272	35	R	CB	29.294
27	Y	CB	41.287	35	R	C	176.637

ResID	ResName	Atomname	Shift	ResID	ResName	Atomname	Shift
35	R	H	8.214	43	K	C	176.625
36	E	H	9.22	43	K	H	9.62
36	E	CB	24.179	44	S	CA	58.804
36	E	N	126.201	44	S	H	8.002
36	E	C	174.278	44	S	C	175.186
36	E	CA	56.497	44	S	N	113.748
37	M	CB	35.435	44	S	CB	63.161
37	M	C	172.686	45	G	H	8.301
37	M	CA	54.305	45	G	N	111.528
37	M	H	7.898	45	G	C	173.917
37	M	N	121.801	45	G	CA	43.849
38	V	C	174.047	46	A	N	125.958
38	V	CA	59.392	46	A	CB	18.235
38	V	H	6.856	46	A	C	175.59
38	V	CB	36.391	46	A	H	7.573
38	V	N	117.58	46	A	CA	52.541
39	I	CA	59.392	47	T	CB	71.26
39	I	CB	41.724	47	T	H	7.453
39	I	C	176.215	47	T	CA	61.073
39	I	H	8.605	47	T	N	118.823
39	I	N	125.091	47	T	C	171.482
40	I	C	173.913	48	F	CB	43.239
40	I	CB	41.347	48	F	C	173.379
40	I	CA	57.948	48	F	CA	55.575
40	I	H	9.054	48	F	N	120.73
40	I	N	120.791	48	F	H	8.939
41	T	CA	59.116	49	Q	CB	31.242
41	T	N	112.228	49	Q	C	175.32
41	T	CB	72.055	49	Q	N	115.299
41	T	C	172.964	49	Q	H	8.656
41	T	H	7.97	49	Q	CA	52.621
42	F	N	118.058	50	V	H	8.3
42	F	C	177.183	50	V	CA	62.372
42	F	CB	42.406	50	V	N	121.678
42	F	H	8.15	50	V	CB	31.242
42	F	CA	57.893	51	E	H	7.497
43	K	N	124.962	51	E	CA	55.644
43	K	CB	31.012	51	E	N	122.728
43	K	CA	58.804	51	E	CB	30.948

ResID	ResName	Atomname	Shift	ResID	ResName	Atomname	Shift
51	E	C	175.951	66	E	N	117.595
52	V	CB	31.806	66	E	H	7.057
52	V	C	175.005	66	E	CB	27.91
52	V	CA	59.541	67	R	H	7.732
52	V	H	7.698	67	R	CA	58.659
52	V	N	120.03	67	R	N	120.744
54	G	N	115.23	67	R	CB	28.84
54	G	H	10.154	68	M	H	8.105
54	G	CA	44.765	68	M	CA	56.453
54	G	C	175.301	68	M	N	119.102
56	Q	CA	56.517	68	M	C	179.711
56	Q	H	7.741	68	M	CB	29.166
56	Q	N	122.613	69	K	H	7.737
56	Q	CB	31.616	69	K	CB	29.471
57	H	CB	37.947	69	K	N	118.65
57	H	H	8.205	69	K	CA	61.591
57	H	CA	60.666	70	D	H	7.996
57	H	C	177.931	70	D	N	124.128
57	H	N	118.797	70	D	CB	38.452
61	Q	CA	58.203	70	D	CA	57.448
61	Q	H	7.518	71	T	CB	65.865
61	Q	CB	29.701	71	T	CA	66.937
61	Q	N	122.915	71	T	H	9.59
62	K	C	175.671	71	T	N	122.829
62	K	H	8.065	72	L	N	124.047
62	K	CB	30.618	72	L	C	177.394
62	K	N	120.386	72	L	CA	57.877
62	K	CA	60.617	72	L	CB	40.369
64	A	C	179.748	72	L	H	8.942
64	A	CB	17.834	73	R	CA	59.715
64	A	H	7.47	73	R	N	118.372
64	A	CA	54.366	73	R	CB	28.816
64	A	N	124.413	73	R	H	7.561
65	I	N	123.699	74	I	C	178.226
65	I	CB	37.797	74	I	H	7.525
65	I	CA	65.08	74	I	CA	61.676
65	I	C	177.885	74	I	N	116.336
65	I	H	8.467	74	I	CB	37.195
66	E	C	178.555	75	T	N	124.205

ResID	ResName	Atomname	Shift	ResID	ResName	Atomname	Shift
75	T	CA	67.489	83	D	N	126.943
75	T	H	8.321	83	D	C	174.494
76	Y	CB	36.827	84	K	CA	53.876
76	Y	N	122.241	84	K	C	178.18
76	Y	CA	61.875	84	K	N	116.289
76	Y	C	179.301	84	K	CB	37.054
76	Y	H	8.101	84	K	H	7.552
77	L	CA	56.902	85	L	CB	45.337
77	L	H	8.116	85	L	H	8.347
77	L	C	178.235	85	L	C	175.457
77	L	CB	41.673	85	L	CA	52.707
77	L	N	118.919	85	L	N	115.589
78	T	H	7.405	86	C	CA	54.299
78	T	CB	70.309	86	C	H	8.517
78	T	N	106.041	86	C	CB	38.491
78	T	CA	61.224	86	C	N	122.153
79	E	N	117.515	87	V	CB	33.488
79	E	CB	26.155	87	V	CA	57.501
79	E	H	7.509	87	V	H	9.084
79	E	C	175.238	87	V	N	122.845
79	E	CA	56.558	87	V	C	174.472
80	T	CB	68.805	88	W	CA	58.674
80	T	C	174.281	88	W	CB	27.816
80	T	CA	64.144	88	W	H	9.701
80	T	H	7.744	88	W	N	121.13
80	T	N	116.678	89	N	CA	51.793
81	K	CB	31.851	89	N	C	175.596
81	K	C	176.167	89	N	CB	37.818
81	K	N	126.638	89	N	H	8.754
81	K	CA	57.668	89	N	N	122.155
81	K	H	8.52	90	N	N	116.684
82	I	CB	38.198	90	N	CB	35.901
82	I	H	8.383	90	N	C	173.865
82	I	C	174.807	90	N	H	8.596
82	I	N	127.199	90	N	CA	52.155
82	I	CA	60.273	91	K	CB	35.183
83	D	CB	40.924	91	K	C	174.553
83	D	CA	55.72	91	K	CA	53.496
83	D	H	8.932	91	K	H	6.971

ResID	ResName	Atomname	Shift	ResID	ResName	Atomname	Shift
91	K	N	117.878	100	S	CB	67.011
92	T	N	117.261	100	S	N	119.676
92	T	C	173.024	101	M	N	122.261
92	T	CB	70.957	101	M	C	172.231
92	T	H	8.42	101	M	CA	56.004
92	T	CA	58.92	101	M	CB	35.935
94	N	CA	54.556	101	M	H	9.111
94	N	CB	38.853	102	E	H	7.996
94	N	H	9.419	102	E	CB	32.173
94	N	C	174.301	102	E	CA	55.828
94	N	N	120.499	102	E	C	175.129
95	S	C	174.829	102	E	N	123.065
95	S	H	7.863	103	N	C	179.835
95	S	CA	54.76	103	N	H	8.503
95	S	N	115.46	103	N	N	129.138
95	S	CB	64.861	103	N	CA	54.549
96	I	N	127.508	103	N	CB	39.571
96	I	C	172.855				
96	I	CA	64.554				
96	I	H	8.839				
96	I	CB	39.012				
97	A	H	8.733				
97	A	C	176.943				
97	A	N	126.777				
97	A	CB	17.705				
97	A	CA	51.987				
98	A	CB	20.661				
98	A	CA	51.063				
98	A	C	173.749				
98	A	N	115.811				
98	A	H	8.031				
99	I	C	171.978				
99	I	CA	58.672				
99	I	N	119.235				
99	I	CB	42.428				
99	I	H	8.836				
100	S	H	8.994				
100	S	CA	56.004				
100	S	C	173.944				