N-Linked Glycopeptide Mimetics as Tools in Kinetic, Mechanistic and Structural Studies of Peptide *N*:Glycanase F

A thesis

presented in partial fulfilment

of the requirements for the degree

of

Doctor of Philosophy in Biochemistry

at

Massey University

by

Dirk Henning Lenz

Palmerston North

New Zealand

2003

Massey University
COLLEGE OF SCIENCES

Institute of Molecular BioSciences
Private Bag 11 222,
Palmerston North,
New Zealand
Telephone: 64 6 356 9099
Facsimile: 64 6 350 5688
Email: IMBS@massey.ac.nz

CERTIFICATE OF REGULATORY COMPLIANCE

This is to certify that the research carried out in the Doctoral Thesis entitled "N-linked Glycopeptide Mimetics as Tools in Kinetic, Mechanistic and Structural Studies of Peptide N:Glycanase F" in the Institute of Molecular BioSciences at Massey University, New Zealand, and at IRL, Lower Hutt (with the approval of the Head of Institute of IMBS):

(a) is the original work of the candidate, except as indicated by appropriate attribution in the text and/or in the acknowledgements;

(b) that the text, excluding appendices/annexes, does not exceed 100,000 words;

(c) all the ethical requirements applicable to this study have been complied with as required by Massey University, other organizations and/or committees which had a particular association with this study, and relevant legislation.

Please insert Ethical Authorisation code(s) here (if applicable): GMO 00/MU/22

Candidate's Name: Dirk H. Lenz

Signature: 

Date: 22 March 2004

Supervisor's Name: Gillian E. Norris

Signature: 

Date: 22 March 2004

Massey University
COLLEGE OF SCIENCES

Institute of Molecular BioSciences
Private Bag 11 222,
Palmerston North,
New Zealand
Telephone: 64 6 356 9099
Facsimile: 64 6 350 5688
Email: IMBS@massey.ac.nz

CANDIDATE'S DECLARATION

This is to state that the research carried out for my Doctoral thesis entitled "N-linked Glycopeptide Mimetics as Tools in Kinetic, Mechanistic and Structural Studies of Peptide N:Glycanase F" under the auspices of IMBS at IRL with the approval of the Head of Institute and in the Institute of Molecular BioSciences at Massey University, Palmerston North, New Zealand is all my own work. This is also to certify that thesis material has not been used for any other degree.

Candidate's Name: Dirk H. Lenz

Signature:

Date: 22 March 2004

Massey University
COLLEGE OF SCIENCES

Institute of Molecular BioSciences
Private Bag 11 222,
Palmerston North,
New Zealand
Telephone: 64 6 356 9099
Facsimile: 64 6 350 5688
Email: IMBS@massey.ac.nz

SUPERVISOR'S DECLARATION

This is to state that the research carried out for the Doctoral thesis entitled "N-linked Glycopeptide Mimetics as Tools in Kinetic, Mechanistic and Structural Studies of Peptide N:Glycanase F" was done by Mr Dirk Henning Lenz in the Institute of Molecular BioSciences at Massey University, Palmerston North, New Zealand and at IRL, Lower Hutt (with the approval of the Head of Institute, IMBS). This thesis material has not been used in part or in whole for any other qualification, and I confirm that the candidate has pursued the course of study in accordance with the requirements of the Massey University regulations.

Supervisor's Name: Gillian E. Norris

Signature: 

Date: 22 March 2004

ABSTRACT

PNGases (Peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidases (E.C. 3.5.1.52) cleave the carbohydrate chains from the asparagine side chains of glycoproteins. They are widely used to deglycosylate *N*-linked glycoproteins and glycopeptides for analytical purposes. PNGase F from *Flavobacterium meningosepticum* is the best characterised of this class of enzymes but little is known so far about the biological significance or the catalytic mechanism of these intriguing enzymes.

The substrate binding and cleavage mechanism of PNGase F has now been investigated.

The first part of this work describes the synthesis of various novel *N*-linked glycopeptide mimetics which were then used in kinetic investigations with PNGase F. To facilitate kinetic studies at low substrate concentrations, a discontinuous HPLC based assay using a fluorescently labelled ovalbumin glycopeptide had to be developed. These experiments led to a better understanding of the structural requirements for substrate binding which will aid the future development of potent PNGase F inhibitors.

In the second part of the thesis, a virtual *N*-linked glycopeptide from human lactoferrin was modelled into the active site region of PNGase F using molecular modelling techniques. This model has resulted in the proposal of a mechanism for catalysis that predicts an important role for Arginine 248, a residue that had previously not been considered part of the catalytic machinery. The model also provides a basis for explaining the substrate specificity of the enzyme. The mechanism is supported by kinetic studies with targeted PNGase F mutants. As a result of this study, new PNGase F mutants have been designed to test the current findings.

To my Family

To be consistent [...]: no penicillin, no lightning rods, no eyeglasses, no DDT, no radar and so on. We live technologically, with man as the master of nature, man as the engineer, and let anyone who raises his voice against it stop using bridges not built by nature. To be consistent, they would have to reject any kind of operation; that would mean people dying every time they had appendicitis. What an outlook! No electric-light bulbs, no engines, no atomic energy, no adding machines, no anesthetics-back to jungle!

From "Homo Faber", A Report by Max Frisch (1959)

ACKNOWLEDGEMENTS

This investigation would not have been possible without the valuable input from many people:

Natalie for her support and understanding during a truly difficult and exciting phase of my life. I wish you all the best for the final phase of your own PhD. Thumbs up.

My parents for their support of my scientific spirit.

Prof. Dr. Joachim Thiem for introducing me into the world of carbohydrates.

Gill Norris for truly passionate supervision as well as **her husband Robert** for a great lesson in Kiwi hospitality at the "Norris Hilton".

George Slim for his witty spirit, exquisite taste in music and enthusiasm for classic cars. I fully agree that a car (in its conceptual context) is a piece of art and not designed to simply get from A to B.

Carol Taylor for excellent supervision on the chemistry side.

Trevor Loo for being such a helpful scientific wizard at Massey's biochemistry lab.

Bryan Anderson for his great support with the 3D modelling work.

Richard Furneaux and his team for incorporating the GET-away German into their group.

Herbert Wong for excellent NMR service.

Posthumous **André Citroën, Pierre Boulanger, André Lefèbvre, Paul Mages and Flaminio Bertoni** the names of which are inseparably connected with a pinnacle of automobile engineering, "la déesse (DS)" (the goddess) [1].

The **DAAD (Deutscher Akademischer Austauschdienst)** for granting a PhD scholarship over the period of 3 years (DAAD Doktorandenstipendium im Rahmen des gemeinsamen Hochschulsonderprogramms III von Bund und Ländern).

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Boc	*t*-butyl oxycarbonyl
BTP	(1,3-bis[tris(hydroxymethyl)-methylamino]propane)
CNBr	cyanogen bromide
DBU	1,8-diazabicyclo[5.4.0]undec-7-ene
DCC	dicyclohexylcarbodiimide
DDQ	2,3-dichloro-5,6-dicyano-*p*-benzoquinone
DMF	dimethylformamide
DMSO	dimethylsulfoxide
DTPA	diethylenetriamine pentaacetate
EA	ethylacetate
eq.	equivalent
ER	endoplasmatic reticulum
ES-MS	Electrospray Mass Spectrometry
Eu	europium
FITC	fluoresceine isothiocyanate
Fmoc	9-fluorenyl methoxycarbonyl
FPLC	Fast Performance Liquid Chromatography
Fuc	fucose
Glc	glucose
GlcNAc	*N*-acetyl glucosamine
HBTU	2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HOBt	1-hydroxybenzotriazole
HPLC	High Pressure/Performance Liquid Chromatography
Man	mannose
NMR	Nuclear Magnetic Resonance
PE	petroleum ether
Phth	phthalimido
PMB	*p*-methoxybenzyl
PTSA	*p*-toluenesulphonic acid
r.t.	room temperature
RP-HPLC	Reverse Phase-HPLC

sat.	saturated
SDS	sodium dodecyl sulfate
SDS/PAGE	sodium dodecyl sulfate/Polyacrylamide Gel Electrophoresis
SPPS	Solid Phase Peptide Synthesis
TBDMS	*tert*-butyl dimethylsilyl
TFA	trifluoroacetic acid
TLC	Thin Layer Chromatography
TMSCN	trimethylsilyl cyanide
TMSOTf	trimethylsilyl trifluoromethanesulfonate
Tol	toluene

CHAPTER 1

Introduction

1.1 Peptide *N*:Glycanases

Glycopeptides and glycoproteins have been shown to play important roles in the mediation of physiological properties throughout the plant and animal kingdoms. Although many of the functions of carbohydrates such as their role in stabilising complex structures and in energy storage were known for many years, little was known about their function in physiological processes. This was mainly due to a lack of suitable analytical methods and the belief that carbohydrates linked to biological molecules were mere appendages and thus not important for function. Therefore glycobiology, the role of carbohydrates in biology, is a comparatively young discipline that has grown extensively in the last 20 years.

Protein glycosylation is now regarded as the most common post-translational modification of a protein and several functions have been attributed to these associated glycans *in vivo* [2]. While many of the glycosyl transferases and synthases, the enzymes which bring about glycosylation, have been functionally and biochemically characterised, little is known about the function of de-*N*-glycosylating enzymes. These enzymes, commonly known as PNGases or glycoamidases, have been widely used for structural studies of glycan moieties but hardly any attention has been paid to their *in vivo* functions [3].

PNGases (Peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidases, EC.3.5.1.52) represent a class of enzymes that cleave the glycosidic bond between the innermost *N*-acetyl-β-D-glucosamine and the asparagine sidechain of *N*-linked glycoproteins. Although PNGase activity has been identified in various organisms ranging from bacteria and plants through to mammals, their *in vivo* function has not yet been unequivocally established. PNGase F from *Flavobacterium meningosepticum*, a bacterial PNGase, is the best characterised of this class of enzymes but the exact catalytic mechanism it uses for hydrolysis remains unclear.

PNGases are expressed in very small quantities by their host organisms and hence are very difficult to purify to homogeneity [4]. In order to define the regions of their structures absolutely necessary for their function other members of this family of enzymes need to be characterised. For this purpose, efficient purification procedures

are essential as kinetic analyses and structural determinations require relatively large quantities of pure protein.

Affinity chromatography remains the method of choice for the purification of enzymes; therefore the synthesis of suitable inhibitors of PNGases is of growing interest.

The current concept for defining inhibitors divides them into two distinct classes: those that are substrate analogues and those that are transition state analogues. The synthesis of three different substrate analogues of PNGases has been reported in the literature, based on a chemoenzymatic approach using enzymes that are not commercially available [5-7]. Only one report describing the synthesis of a transition state analogue can be found [8] and to date, that compound has not yet been tested for inhibitory activity towards PNGases.

In summary, there is no method currently available that offers a facile synthesis of potent inhibitors of PNGases. Furthermore there is a lack of sufficient data that would allow some insight to be gained into the catalytic mechanism of this intriguing class of enzymes. Finally the limited number of PNGases which have been purified to homogeneity and structurally characterised is not sufficient to allow an understanding of the reasons for the functional differences observed in the different PNGase actvities discovered so far. For this reason further investigations into the design and synthesis of inhibitors of PNGases and their application in affinity chromatography and structural studies is required.

1.2 *In vivo* Functions of PNGases

It is generally acknowledged that the glycosylation of proteins and lipids is essential to deriving their biological functions. There are two common types of glycosylated proteins; *O*-linked and *N*-linked glycans. In *N*-linked glycans the carbohydrate chain is linked to the side chain of asparagine whereas in *O*-linked glycans the carbohydrate moiety is most commonly attached to the side chain of either a serine or threonine. In contrast to *O*-linked glycans, *N*-glycans have a distinctive core-region (Fig.1) and the reducing end of the sugar in the conserved *N*-glycosidic linkage region is in a β-configuration. However one exception to this generalisation has been reported: for a

glycopeptide called "Nephritogenoside" in which the glycan, consisting of three glucose residues, is α-*N*-glycosidically linked to the side chain of asparagine in a peptide comprising of 21 amino acids [9],[10].

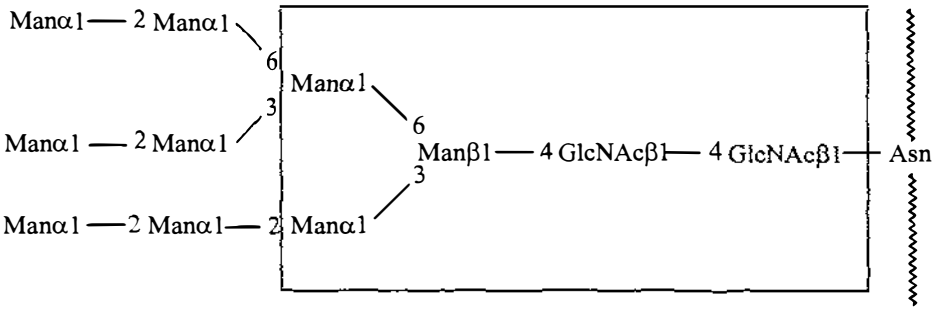
N-Glycans can be classified according to the carbohydrate structure at the non-reducing end. As shown in figure 1, the most common types are "high-mannose", "complex-" and "hybrid"- type glycans. High-mannose-type glycans contain only α-mannosyl residues whereas complex-type glycans contain no mannosyl residues other than those in the trimannosyl core. The hybrid-type *N*-Glycans have the characteristic features of both complex-type and high-mannose-type. The different types of *N*-glycans can vary in their branching pattern and bear additional substituents on the core as depicted by a dotted line.

Furthermore, both proteins and glycoproteins can be attached to a glycosylated phospholipid-structure, the so-called GPI-anchor (glycosyl phosphatidylinositol). The GPI-anchor is an alternative way to anchor a wide variety of proteins and glycoproteins onto the exterior surface of the eukaryotic plasma membrane.

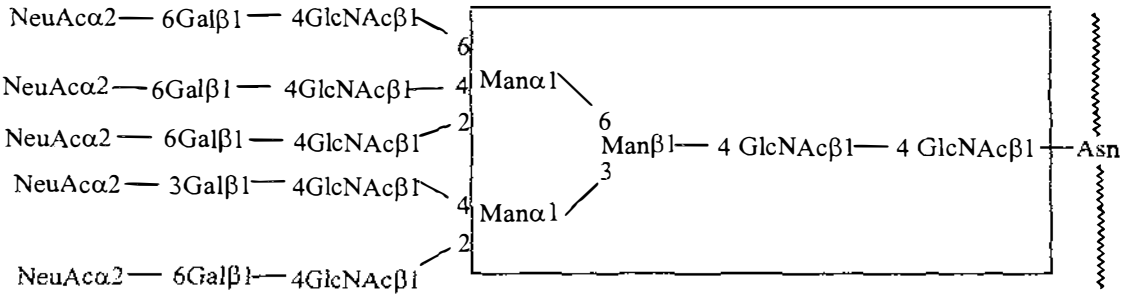
N-Glycosylation of nascent peptide chains occurs in the lumen of the RER (rough endoplasmatic reticulum) via an en-block transfer of a preformed dolichol pyrophosphoryl linked oligosaccharide precursor ($Glc_3Man_9GlcNAc_2$) onto the nascent polypeptide chain. This process is mediated by a specific oligosaccharyl transferase (OT) [11],[12].

The three glucosyl-residues on the non-reducing end of the oligosaccharide are believed to act as a signalling sequence for this process and the peptide sequence Asn-X-Ser/Thr (where X can be any amino-acid except proline or aspartic acid) is an absolute requirement [13 and references therein]. Subsequent to this oligosaccharide transfer, three glucose- residues and one specific mannose- residue are cleaved by ER-hydrolases before the properly folded glycoprotein is transferred to the Golgi-apparatus for further processing. In the Golgi-apparatus further mannosyl- residues are removed before elongation of the oligosaccharide chains can occur through the catalytic action of glycosyltransferases to form the different types of *N*-glycans.

I. High-mannose-type:



II. Complex-type:



III. Hybrid-type:

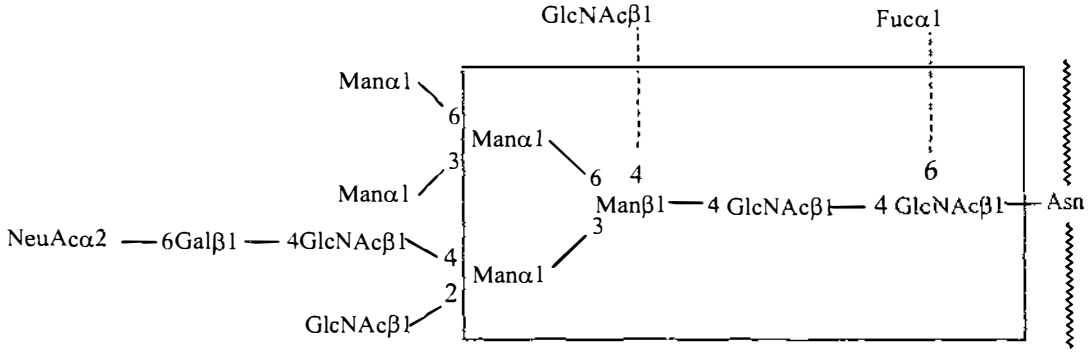


Figure 1: Common structures of *N*-linked glycans (box encloses common core).

N-Glycans play two principal roles. Firstly, they have a physical role involving the correct folding of the protein, solubility and stability towards degrading enzymes and secondly, they play an important role as determinants of molecular recognition [2],[11],[12],[14],[15],[16],[17].

It is well accepted that the oligosaccharide units are required for the correct folding of the proteins in the lumen of the ER as they largely influence the physiochemical properties of entire domains. Likewise, the oligosaccharide moieties of glycoproteins are important in determining the physical properties of tissues. An example of this function is seen in the adhesion of cells to the extracellular matrix where a variety of glycoconjugates have been shown to be directly involved in cell-cell adhesion. The heparan sulfate chains of proteoglycans for instance are designed to mediate such functions [2].

As protein-linked oligosaccharides are always located at the molecular surface they are optimally situated for serving as determinants for cellular recognition. There they can have an impact on immunogenic processes by masking potential antigenic sites or they can protect the cells from attack by microorganisms by masking receptor sites. An example of this is observed with the *O*-acetylation of the 9-position of a terminal sialic acid residue which prevents its recognition by the highly pathogenic influenza A viruse [2].

Oligosaccharides are also involved in the regulation of hormonal activity. Signal transfers mediated by cell surface receptors for growth factors (e.g. the insulin-receptor) were shown to be dependent on glycosylation of both the receptors and the growth factors [2].

Glycosylation also plays a role in the targeting of proteins within cells. A good example of this is the role of mannose-6-phosphate residues in the targeting of proteins to lysosomes. Another important example of the role of carbohydrates in molecular recognition events is their interaction with a class of carbohydrate binding proteins, called lectins. In animals, a class of lectins called the selectins mediate the targeting of leukocytes to sites of inflammation [2].

To date, many plant lectins (proteins that bind carbohydrate moieties with high affinity) have been characterised and are used for affinity chromatography of

glycoproteins. Their *in vivo* functions are not yet clear, but as an example they are known to play a role in the binding of nitrogen-fixing microorganisms to the root hairs of leguminous plants (*e.g.* soybeans, garden peas, white clover *etc.*) [2].

The deglycosylation of specific glycosylation sites in glycoproteins and glycopeptides has also recently been shown to play a physiologically important role [3],[18]. Berger *et al.* [18] define de-*N*-glycosylation as "the removal of a glycan from an *N*-glycosylprotein, by means of enzymes acting on the di-*N*-acetylchitobiosyl part of the invariant pentasaccharide inner core". Therefore their definition includes PNGases, as well as the corresponding endo-form called ENGases (endo-*N*-acetyl-β-D-glucosaminidase EC 3.2.1.96). This broad definition is justified by the fact that for both enzymes i.e. PNGases and ENGases, the resulting proteins will differ significantly in their physiochemical properties, resulting in a functional change of the molecule.

ENGases catalyse the cleavage of the GlcNAcβ1 $\rightarrow$ 4GlcNAc linkage of the pentasaccharide core, leaving behind an *N*-acetylglucosamine on the peptide or protein (figure 2). In contrast, PNGases cleave the linkage between the innermost GlcNAc and the asparagine side chain releasing the entire carbohydrate chain from the peptide/protein and in the process converting asparagine into aspartic acid [19].

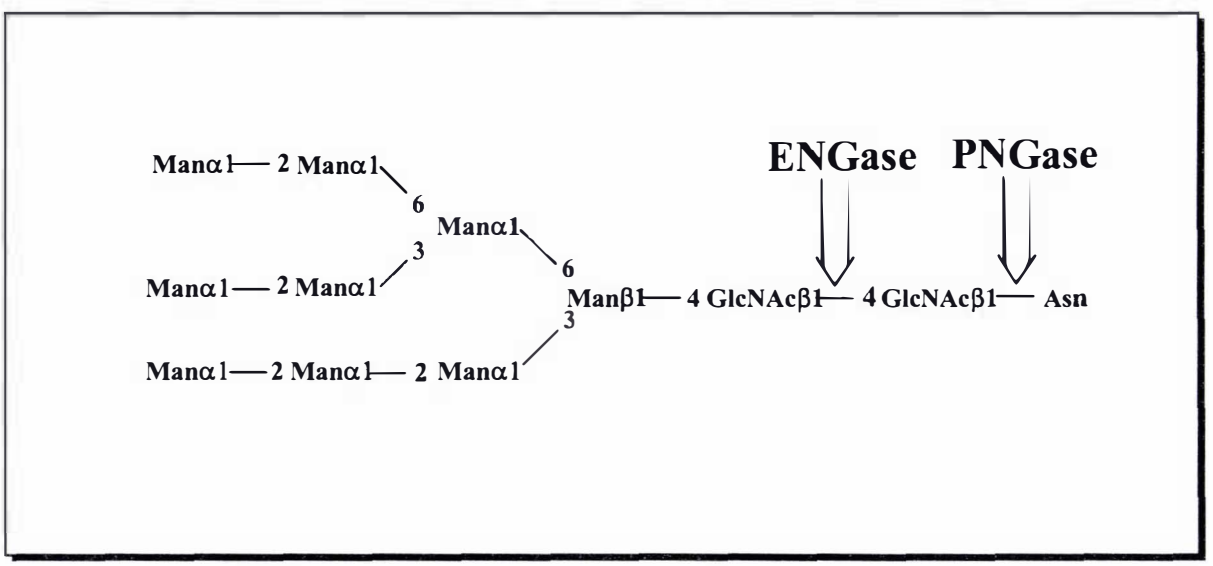


Figure 2: Catalytic action of glycoamidases.

Examples of the **proposed** functions of PNGases from different sources are as follows:

(a) Signalling in animals

Suzuki *et al.* [3] hypothesised that deglycosylating enzymes may play a general role in the mediation of physiological processes based on the finding of non-lysosomal PNGase activity in animal cells. Their search was initiated by the discovery of free glycan chains bearing an *N*,*N*'-diacetylchitobiosyl sequence at their reducing termini in the oocytes and early embryos of several fish species. They postulated two different functions existed for this animal PNGase:

(I) That de-*N*-glycosylation facilitated the dissociation of the vitellogenin-receptor complex after receptor mediated pinocytosis of vitellogenin into oocytes.

(II) That the de-*N*-glycosylation products of L-hyosophorin (Asp-Ala-Ala-Ser-**Asn**(CHO)-Gln-Thr-Val-Ser) played an as yet unknown bioactive role in the early embryogenesis of *Oryzia latipes* (Medaka fish). This was in spite of the fact that no evidence could be obtained to show that the resulting peptides or the glycans mediated any biological activity.

There is other circumstantial evidence that PNGase-catalysed reactions occur in animal cells based on differences in cDNA and Edman sequences of proteins. For example, hyaluronidase from the honeybee has an aspartic acid at residue 4 in the mature protein but the cDNA sequence indicates this residue is an asparagine. It is possible that a glycosylated proenzyme might be converted into the mature protein by the action of a cytosolic PNGase in the process becoming susceptible to protease activity [3 and references therein]. However, all these proposed functions of PNGase activity in animal cells still remain purely speculative and experimental evidence for the *in vivo* roles of the enzymes is lacking.

The change from Asn to Asp introduces an extra negative charge to the protein surface that is very likely to have an impact on its physiochemical properties such as

receptor binding or the susceptibility to proteases. Interestingly L-929, a PNGase from mouse fibroblast cells, shows an unusually high affinity for its substrate glycopeptides as well as for yeast mannan and a trimannose [20]. This enzyme may therefore be able to act in two capacities; as an enzyme and as a lectin. Parallels are drawn with carbohydrate-binding surface enzymes such as glycosyltransferases and glycosidases that are able to both bind and catalytically release potential substrates on opposing cell surfaces.

(b) Quality control

Suzuki *et al.* [21] proposed that the *in vivo* function of a PNGase from hen oviduct is in a "quality control system" for newly synthesised proteins. Their findings are based on the fact that hen ovalbumin exists primarily as a protein with a single *N*-linked glycan at Asn 292. However, a small proportion of the native protein has two *N*-glycan chains, one on Asn 292 and the other on Asn 311. They postulate that the di-glycosylated form is aberrantly synthesised and then selectively deglycosylated at Asn 311 by the catalytic action of an endogenous PNGase that they have isolated and named PNGase HO. Whether this site-specifically de-*N*-glycosylated ovalbumin, with Asn 311 converted to Asp 311, re-enters into the normal secretory pathway or is targeted to an intracellular degradative system remains unclear. Furthermore, PNGases have been shown to play a role in the processes that mediate the degradation of falsely folded and glycosylated glycoproteins [3].

In 1998 Suzuki *et al.* [22] reported the occurrence of a soluble PNGase activity in the yeast *Saccharomyces cerevisiae* using a highly sensitive PNGase assay based on a ^{14}C labelled asialofetuin glycopeptide. They speculated that this enzyme may also function in a quality control system for overglycosylated or misfolded proteins. Recently the same group reported the identification of the gene encoding the correspondent PNGase activity. Screening of the peptide sequence against various databases revealed that it is highly conserved in mammals, plants, insects, nematodes and fungi thus underlining the importance of this class of enzyme [23]. More recent research revealed that the *N*-terminal sequence of the soluble PNGase identified in mice, named m Png1p, showed homologies with various proteins found in the ubiquitin-related pathway [24]. These results were later interpretated by Suzuki *et al.*

[25] as a strong indication for the involvement of PNGase activity in the ER-associated degradation (ERAD) of misfolded glycoproteins. As many PNGases are cytosolic proteins it seems reasonable that they might be involved in the proteasomal degradation of glycopeptides/proteins. However, so far there is no scientific evidence for a link between the *in vivo* production of de-*N*-glycosylated proteins and the role of soluble PNGase (Png1p). Furthermore the *in vitro* function of this PNGase from yeast is limited to glycopeptides. Interestingly, in *S. cerevisiae* the deletion of the gene encoding for Png1p (PNG1) neither affected the viability of the cells nor their growth rate.

Weng et al. [26] reported the occurrence of a PNGase activity in the ER of rat liver. This PNGase activity might also be considered as part of a quality control mechanism operating in the biosynthesis of glycoproteins in the lumen of the ER.

(c) Immune system

Recently Selby *et al.* [27] observed potential PNGase activity in the cytoplasmatic processing of glycoproteins in lymphoblastoid cells from chimpanzee blood for the presentation by class I MHC (major histocompatibility complex) molecules. When Asn 234, which is glycosylated was replaced with Asp in a nonameric peptide fragment of the hepatitis C virus envelope glycoprotein E1, the concentration required to produce a response by cytotoxic lymphocytes (CTL) from a hepatitis C virus-infected chimpanzee was decreased at least 1000-fold. This experiment was based on the finding that CTL lines from chimpanzee liver were highly sensitive towards a nonameric epitope in the NS3 protein of HCV-1 (hepatitis C virus) that carried an Asp at position 2. This posed the question of whether envelope proteins or tumour antigens that are synthesised on membrane bound ribosomes (in contrast to cytoplasmic or nuclear proteins) undergo posttranslational cytoplasmic processing. The authors suggested that ER proteins may be translocated into the cytoplasm via a pore that is usually used for the extrusion of glycoproteins into the ER during cotranslational synthesis on membrane-bound ribosomes. PNGase-activity then deglycosylates the proteins changing Asn into Asp, thus generating a recognition site for specific CTL lines.

These findings support the results reported earlier on the posttranslational modification of tumour antigens [28]. A peptide recognised by melanoma-specific T cells was isolated from MHC class I molecules, sequenced using Edman methods and identified through homology searching. The peptide corresponded to residues 368-376 of tyrosinase, a melanoma gene product, except for the occurrence of Asp instead of Asn at position three. This change in the primary peptide sequence is believed to result from de-*N*-glycosylation by an unidentified cytosolic PNGase.

(d) Signalling in plants

The role of de-*N*-glycosylation enzymes in plant cells has been reviewed by Berger *et al.* [18].

Concanavalin A, a plant lectin, is not a glycoprotein although it is synthesised via an *N*-glycosylated precursor [29 and references therein]. De-*N*-glycosylation may therefore be necessary to its function as a carbohydrate binding protein. This theory is supported by the fact that both PNGase and ENGase activity have been detected in the seeds of *Canavalia ensiformis* (Jack bean).

Free oligosaccharides, especially those that include a fucosylated pentasaccharide core structure, are thought to play a role in the control of some physiological events [30]. For example free oligosaccharides have been shown to play a role in tomato fruit ripening [31],[32].

The recent discovery of PNGase actvity in the rice seeds of *Oryza sativa* [33] also suggests a physiological role for free glycans during the germination process. The activity of this PNGase, named PNGase Os, was monitored in the growing germ and in the coleoptile during germination. The results showed that the expression of PNGase activity was correlated with the growth activity. However, no scientific evidence has been reported for a role of the free glycans in triggering rice germination.

Wu *et al.* [34] reported that a glycosylation gradient on TTS (tobacco transmitting tissue) proteins may contribute to the guidance of pollen tubes from the stigma to the ovary during plant sexual reproduction in *Nicotiana tabacum*. The proteins are

deglycosylated and incorporated into the pollen tubes where they presumably serve as nutritients. The fact that the TTS proteins adhere to the pollen tube surface and tips is reminiscent of L-929 PNGase from mouse fibroblasts which is believed to have dual roles as both an enzyme and a carbohydrate binding protein [20]. Whether PNGase or ENGase activity is involved in the guidance of pollen tube growth is not clear. Interestingly, the substrate specificity of the plant PNGases so far characterised in *in vitro* experiments seems to be restricted to glycopeptides [18]. It is however possible that some as yet undiscovered plant PNGase is responsible for de-*N*-glycosylation of native proteins such as the lectin Concanavalin A [35].

(e) Nutritional functions in plants

Lhernould *et al.* [30] reported the increase of PNGase and ENGase activity under conditions of carbon starvation that resulted in the production of "unconjugated *N*-glycans" in cell suspensions of *Silene alba* (Evening lychnis, Caryophyllaceae). The finding, that the lower the initial sucrose supply the greater the increase in activity of both PNGase and ENGase, indicates that these enzymes are probably involved in nutritional processes in plant cells.

Berger *et al.* [29] monitored PNGase and ENGase activities during germination and postgerminative development in *Rhaphanus sativus* (radish). Both activities were prevalent (85-90 %) in cotyledons which serve as major sites of energy storage. Thus de-*N*-glycosylation is believed to be necessary to expose proteolytic sites for further degradation by proteases.

A recent report on the evidence of PNGase and ENGase activity in barley grains lends support to this theory [36]. PNGase and ENGase activity was monitored in barley grains during germination and postgerminal stages in the embryo and endosperm of barley grains. The results showed that more than 90 % of the PNGase activity was located in the endosperm, whereas the embryo had most of the ENGase activity. Again this indicates the involvement of PNGase activity in the mobilisation of storage proteins. As there is no hypothesis for the function of ENGases in the embryo or the radicle and the rootlets future research needs to focus on the physiological functions of both these de-*N*-glycosylation enzymes and their products.

(f) Nutritional functions in bacteria and fungi

To date only one PNGase from a bacterial (prokaryotic) source has been characterised- PNGase F from *Flavobacterium meningosepticum* (section 1.6). The role of this enzyme is yet unknown but again there are a number of proposals suggesting physiological functions. As this organism is primarily a soil bacterium one of the possible functions of this secreted PNGase is the degradation of foreign *N*-glycoproteins for nutritional purposes. The deglycosylated proteins would then become susceptible to the catalytic action of proteases. Parallels can be drawn with the PNGase At from the fungus *Aspergillus tubuginensis* [37] as this enzyme is secreted along with glycohydrolases and proteases indicating its possible role in nutrition.

In contrast, all other known PNGases (table 1 in section 1.3) are cytosolic or membrane bound as is the case with PNGase from the ER of rat liver [26]. This reflects the different physiological properties of prokaryotes (*Flavobacterium meningosepticum*) and simple eukaryotes (*Aspergillus tubuginenis*) as opposed to higher eukaryotic organisms. As "lower" organisms such as bacteria or fungi possess no digestive system but are dependent on exogenous energy sources, suitable mechanisms are required for the degradation of nutritional molecules before they can be taken up by the organism.

It is clear from the literature that PNGases from prokaryotes and eukaryotes differ in both structure and *in vivo* function. Those from the former are more likely to be secreted and have a function concerned with acquiring nutrients while those from eukaryotes are more likely to have more complex roles.The fact that recent research indicates a widespread distribution of PNGases throughout the plant and animal kingdom serves to underline their importance. The functions of PNGases will hopefully become clearer as more of these enzymes are characterised.

1.3 PNGases: Sources and Properties

After reviewing the proposed functions of PNGases from plants, animals and bacteria, their sources and specificities will be summarised in this section.

The following table gives an overview of the known PNGases and their sources:

Enzyme	Source	Reference
PNGase HO	Hen oviduct	Suzuki *et al.*, 1997 [21]
PNGase L-929	Mouse fibroblast	Suzuki *et al.*, 1993 [38] Suzuki *et al.*, 1994 [39],[20] Suzuki *et al.*, 1995 [40]
PNGase (ER)	ER of rat liver	Weng & Spiro, 1997 [26]
PNGase	Various mouse organs	Kitajima *et al.*, 1995 [41]
PNGase	various mammalian cell lines	Suzuki *et al.*, 1995 [42]
PNGase	*Oryzia latipes* (Medaka fish)	Seko *et al.*, 1991 [43],[44] Seko *et al.*, 1999 [45]
PNGase A	Almonds (*Prunus amygdalus*)	Taga *et al.*, 1984 [46] Altmann *et al.*, 1997 [47]
PNgase J	Jack bean (*Canavalia ensiformis*)	Sugijama *et al.*, 1983 [48] Sheldon *et al.*, 1998 [35]
PNGase Os	Rice seeds (*Oryza sativa*)	Chang *et al.*, 2000 [33]
PNGase GM	Soybean (*Glycine max*)	Kimura and Ohno, 1998 [49]
PNGase P	Pea (*Pisum sativum*)	Plummer *et al.*, 1987 [50]
PNGase	Barley (*Hordeum spec.*)	Vuylsteker *et al.*, 2000 [36]
PNGase R	Radish (*Raphanus sativus*)	Berger *et al.*, 1995 [29] Berger *et al.*, 1996 [51]
PNGase Se	White campion (*Silene alba*)	Lhernould *et al.*, 1995 [52]

PNGase (Png1P)	*Saccharomyces cerevisiae* (Yeast: Ascomycetes) (eukaryotic organisms)	Suzuki *et al.*, 1998 [22] Suzuki *et al.*, 2000 [23] Suzuki *et al.*, 2001 [24] Suzuki *et al.*, 2002 [25]
PNGase At	*Aspergillus tubuginensis* (Fungi)	Ftouhi-Paquin *et al.*, 1997 [53] Ftouhi-Paquin *et al.*, 1998 [37]
PNGase Sm	*Sphingobacterium multivorum*	Wang *et al.*, 1997 [5]
PNGase F	*F. meningosepticum* (Bacteria, Gram-)	Plummer *et al.*, 1984 [54] Kuhn *et al.*, 1995 [55]

Table 1: PNGases and their sources.

Recent research however [23] indicates that a far wider range of organisms express PNGase activity. Therefore this list can be expected to grow significantly in the near future.

PNGase F from the Gram-negative bacterium *Flavobacterium meningosepticum* and PNGase A from Almonds (*Prunus amygdalus*) are the most widely known of the PNGases. Both enzymes are commercially available and are widely used as tools in glycobiology. They will therefore be reviewed in more detail in section 1.5 and 1.6 respectively.

The catalytic activity of PNGases is highly dependent on the structure of the substrate. This includes the size and structure of the carbohydrate moiety as well as the primary structure and length of the peptide-chain. Most PNGases cannot act on the native glycoproteins and therefore the peptide length and sequence are critical factors in determining the substrate specificity. Additionally each enzyme shows pH- and temperature-optima. The effect of the oligosaccharide structure on catalytic rates has been thoroughly investigated by Altmann and co-workers [56] for PNGases A and F.

Data is also available for a number of other PNGases (see references for further examples).

As a general rule, catalytic rates have been shown to be dependent on the *N*-glycan-type. PNGase A for example can act on all sialic-acid containing complex-, high-mannose and hybrid-type glycans, but shows higher rates towards the complex substrates [57].

Sialylation at the non-reducing end of complex oligosaccharide-chains has been shown to be important for a number of PNGases. *S. cerevisiae* PNGase for instance shows a reduced activity on the sialylated fetuin glycopeptide I compared with asialoglycopeptide I [22]. This sensitivity to sialylation has also been reported for the L-929 PNGase [39], while PNGase A and F are seemingly not affected by the presence of sialic acids in substrates.

PNGase At (*Aspergillus tubuginensis*) [53],[37] has a broad substrate specificity cleaving bi- and tri-antennary hybrid- as well as high-mannose-type glycans. In comparison to other PNGases it shows a relatively high thermostability retaining up to 70 % of its original activity during incubation at 62°C over 8h.

The question therefore arises as to whether significant differences exist in the substrate specificities of PNGases between plant-, animal- and bacterial sources. Differences have been reported for the minimal oligosaccharide requirements for enzymatic activity [40]. For L-929 PNGase from mouse fibroblast cells the trisaccharide core Manβ1$\rightarrow$ 4GlcNAcβ1$\rightarrow$ 4GlcNAcβ1$\rightarrow$ peptide has been shown to be the minimal structural requirement for enzymatic activity. For PNGase A a single GlcNAc is sufficient for catalytic action, whereas PNGase F requires at least an *N*,*N*'-diacetylchitobiosyl core to be present [56].

Fucosylation at the 3 and/or 6 position of the innermost GlcNAc of the pentasaccharide core has also been shown to be a crucial factor in substrate activity . L-929 PNGase [39] is inactive on such substrates whereas PNGase F [58] is only inactive on substrates which have a Fuc residue α 1$\rightarrow$3 linked to the innermost GlcNAc. PNGase GM (*Glycine max*) and PNGase R from radish can only cleave substrates bearing an α 1$\rightarrow$3 linked fucosyl residue on the innermost GlcNAc residue

[49],[29]. PNGase A on the other hand can process oligosaccharides with both types of substitution patterns [59].

As mentioned earlier, peptide length also has an effect on the catalytic activity. PNGase F for example can act on native glycoproteins whereas PNGase A only acts on glycopeptides. As a general rule PNGases will not cleave an oligosaccharide connected to Asp alone. Indeed, the length and sequence of the peptide has a significant influence on all PNGases so far characterised. This may be a factor in the site specificity observed for the deglycosylation of hen ovalbumin by PNGase HO. Asn 311 of this protein is selectively deglycosylated by PNGase HO, leaving Asn 292 glycosylated [21]. PNGase F, in contrast, deglycosylates both glycosylation sites. The molecular determinants for this site-specific reaction are not yet known.

Concerning the subcellular location, recent research indicates that most of the PNGase activity is located in the cytosol. PNGases can therefore generally be regarded as soluble enzymes [22]. The PNGase activity found in the ER of rat liver [26] could be regarded as an exception in this regard. As this enzyme has been assigned a function in the so called "quality control system" for newly synthesised glycoproteins, a comparison with PNGase HO might be reasonable because this enzyme is also believed to fulfil a similar function. Surprisingly, investigations have shown that more than 98 % of the PNGase HO activity is located in the cytosol with only a minor fraction being membrane associated [21].

In this context the soluble PNGases found in eukaryotic cells [25] and their proposed role in the proteasomal degradation of proteins has to be mentioned. Interestingly, recent research [60] has confirmed that these soluble PNGases appear to have significant homology to enzymes of the transglutaminase family, suggesting a common evolutionary lineage for PNGases and transglutaminases.

It is noteworthy that so far only two PNGases (PNGase F and yeast PNGase, see table 1) have been purified to homogeneity and completely characterised. This reflects the difficulty of isolation and purification as PNGases are only expressed in very small quantities. A large number of potential sources for PNGase activity have yet to be tested and will hopefully lead to a better understanding of this class of enzymes.

1.4 Assays for the Determination of PNGase Activity

Since PNGase A was first described by Takahashi in 1977 [61] many different analytical techniques for the determination of PNGase activity have been developed leading to the discovery of PNGase activity in a wide range of organisms.

To design suitable assays for PNGase activity several aspects have to be considered. Obviously, a substrate that meets the specific requirements for either one, a group or the whole class of enzymes is of prime consideration. Such a substrate may be prepared by the isolation or modification of a natural substrate or by synthesis using chemical or chemoenzymatic methods. Secondly, there has to be a suitable analytical method for the detection of the products of the enzymatic digest. Furthermore, a quantification of these products has to be possible in order to determine catalytic rates and to calculate the kinetic parameters such as V_{max} and K_M. This section will give a short overview of the current methods and focus on their practicability.

A standard procedure that is still being used by many laboratories is the so called discontinuous assay most commonly based on the HPLC (High Pressure Liquid Chromatography) separation and detection of the products derived from the enzymatic digest. When Plummer and co-workers [54] reported the discovery of PNGase F activity for the first time they used dansylated or dabsylated glycopeptides isolated from fetuin or hen-ovalbumin. Separation was achieved on a C18 RP (reversed phase) column and analyses were performed at 280 nm and 436 nm respectively. Bourgerie *et al.* [62] used a commercially available resorufin-labelled *N*-glycopeptide for their HPLC-based assay. A major disadvantage of this procedure was the heterogeneity of the substrate which required further purification before it could be used in this quantitative assay. A somewhat simpler and therefore less sensitive approach was reported by Norris and co-workers [63] who used an unlabelled glycopeptide (11mer) isolated from hen-ovalbumin. Separation and detection was again achieved by using HPLC, monitoring the difference between substrate and product peaks at 220 nm. The disadvantage of this procedure is the limited detection threshold. A common disadvantage of all these procedures is that they do not allow a continous monitoring of the enzymatic reaction as well as being time consuming due to the necessity of HPLC. A discontinuous quantitative approach using FPLC (Fast Performance Liquid Chromatography) has also been reported [64].

Another approach is the determination of ammonia released during the hydrolysis of the primary cleavage pruduct during the enzymatic hydrolysis [64-66]. The ammonia formed is allowed to react with *o*-phthalaldehyde and 2-mercaptoethanol to form a fluorophore. This method, however, lacks sensitivity [67]. Alternatively, the amount of released carbohydrate can be determined using the phenol-sulphuric method [64],[68]. This method also lacks the sensitivity required and is subject to interference from other compounds in the assay mixture.

Suzuki *et al*.[39] used a ^{14}C-labelled asialo glycopeptide for the determination of PNGase actvity in mouse-derived L-929 fibroblast cells. Analysis was achieved using either paper chromatography or paper electrophoresis. The disadvantage of this procedure is that the use of radionuclides requires special handling, storage and disposal protocols.

Tarentino *et al.* [69] reported the use of a triantennary pentaglycopeptide from fetuin which had been didansylated with ^{3}H dansyl chloride. The separation of the released peptide was achieved by paper chromatography and the quantitation by liquid scintillation spectrometry.

Fan and Lee [59] used high performance anion exchange chromatography (HPAEC) with pulsed amperometric detection (PAD) for the separation and quantification of the hydrolysis product. The disadvantage with this method is again that each analytical run takes between 20-30 minutes and requires a detector that is not commonly available in most laboratories.

A relatively modern and sophisticated concept is the use of bifluorescent-labelled or fluorescence-quenched glycopeptide substrates which allow the continuous measurement of enzymatic activity. The concept behind this method is that these labelled substrates bear two fluorescent groups on the two parts of the molecule that are later to be separated by the enzymatic activity. Furthermore the fluorescence of one of the groups has to be quenched by the second group. Lee *et al.* [70],[71] used a complex type glycopeptide in which dansyl groups (acceptor fluorophore) are attached to two terminal galactose residues and a naphthyl group (donor fluorophore) is attached to the peptide chain. The naphthyl fluorescence (λ_{em} = 335 nm) is quenched by the dansyl group which in turn emits fluorescence with a wavelength of

520 nm. When the bond between the carbohydrate moiety and the peptide moiety is cleaved by the action of PNGase, the energy transfer between the two fluorescent groups ceases to occur and the emission of the dansyl fluorescence diminishes or disappears. The advantages of this concept are that it allows for the continuous measurement of enzymatic activity with low concentrations of both substrate and enzyme and that no chromatographic step is required.

Vito *et al.* [72] have reported the synthesis of a bifluorescent substrate bearing the well characterised anthranilamide/nitrotyrosine donor/acceptor pair [73]. However this compound has not yet been tested for PNGases and does not meet the structural requirements for substrate recognition by these enzymes, although the authors claim this to be the case.

The latest concept for the determination of PNGase activity has been developed by Deras and co-workers [67] using an Eu^{3+} (europium) labelled high-mannose glycopeptide. In this method the removal of unreacted substrate and free carbohydrate is achieved by affinity chromatography with concanavalin A. The lanthanide ion (Eu^{3+}) then has to be dissociated from its chelator (diethylenetriaminepentaacetate = DTPA) using a so-called "enhancement-solution" to form a new highly fluorescent chelate. As the newly formed chelate has a long fluorescence decay period, analysis can be performed in plastic microtiter plates by time resolved fluorometry. This method stands out due to its low detection threshold (< 5 fmol) and the low amounts of protein (PNGase) and substrate required. However, being a discontinuous assay it is again relatively time consuming and subject to handling error. The binding to concanavalin A alone requires between 15-30 min. Also, competing chelators in the assay buffer, such as EDTA, have the potential to cause a significant loss of signal.

1.5 PNGase A from Almonds (*Prunus amygdalus*)

PNGase A was first described by Takahashi in 1977 [61]. The enzyme was partially purified from almond emulsin. Investigations showed that the enzyme cleaved the peptide-carbohydrate linkage in stem bromelain releasing an intact glycan chain and a peptide without any carbohydrate residues. Therefore this enzyme was defined as a new class of amidase [EC 3.5.1.52] in contrast to the hitherto known ENGase activity

[EC 3.2.1.96]. This activity was distinct from aspartyl-amidase activity [EC 3.5.1.26] where the enzymes cleave the asparagine-carbohydrate linkage but are inactive on any substrates bearing more than one amino acid. Risley and Van Etten [19] monitored the hydrolysis reaction of PNGase A using NMR techniques. They concluded that the reaction proceeds via a two step mechanism. In the first step the enzyme catalyses the hydrolysis of the glycopeptide converting asparagine to aspartic acid and releasing an intermediate oligosaccharide amine. In the second step the amine intermediate hydrolyses non-enzymatically to the final glycan. They therefore called the enzyme an amidohydrolase (amidase) indicating that the common name *N*-glycosidase was actually incorrect. Nevertheless the terminology *N*-glycosidase has "stuck".

In 1984 Taga and co-workers [46] reported the purification of PNGase A from defatted almond meal and from almond emulsin to homogeneity as judged by SDS/PAGE (sodium dodecyl sulfate/polyacrylamide gel electrophoresis). They reported the molecule to be a glycoprotein consisting of a single polypeptide chain of 66.8 kDa. This was revised in 1997 when Altmann and co-workers [47] reported the PNGase A molecule to be a heterodimer consisting of two subunits of 55 and 27 kDa that were not covalently linked by disulphide bridges.

These results were confirmed by SDS/PAGE and MALDI-MS (Matrix Assisted Laser Desorption Ionisation-Mass Spectrometry). This is reminiscent of the reported difference between the native and the recombinant form of the PNGase At (*Aspergillus tubuginensis*) [37],[53]. The native protein is synthesised as a single polypeptide of about 59 kDa that is subsequently cleaved into two non-identical subunits by protease activity. The recombinant enzyme on the other hand consists of a single polypeptide chain of 78 kDa. Interestingly, no difference in the specific activity for the two enzymes could be detected.

Investigations into the substrate specificity [56],[59] of PNGase A revealed the enzyme to have preferences for oligosaccharide structure and peptide length similar to those of PNGase F. As a general rule, PNGase A shows a higher tolerance towards peptide shortening than PNGase F. Also PNGase A is able to catalyse the cleavage of a single GlcNAc from the peptide. No preference for either complex or oligomannose glycopeptides could be shown by Altmann *et al.* [56] although Plummer and

Tarentino [74] reported that the enzyme showed a preference for complex-type glycopeptides.

Despite all similarities there are two striking differences between these two enzymes. Unlike PNGase F, the enzyme from almonds appears to be unable to cleave the glycan from glycoproteins even if the substrate has been denatured [56] which is a disadvantage for the use of the enzyme in the analysis of glycoproteins.

There is however contradictory information on this matter. Tarentino and Plummer [57] investigated the ability of PNGase A to remove oligosaccharide chains from intact glycoproteins. They reported that the oligosaccharides of Ribonuclease B and a FAB_{μ} fragment could not be removed from the native proteins. After denaturation of the glycoproteins by SDS (sodium dodecyl sulfate) and heating the enzyme was, however, able to act on these substrates.

Taga *et al.* [46] reported the complete removal of the glycan from denatured ovomucoid and ribonuclease B with PNGase A. Interestingly, in 1981 Takahashi and Nishibe [66] reported the isolation of three discrete species of glycoamidases from almond emulsin, one of which they reported to be active on native glycoproteins.

The second striking difference in comparison to PNGase F is that PNGase A can act on substrates bearing a fucosyl residue linked to the 3-OH of the innermost GlcNAc residue, a common structural feature of plant and insect glycoproteins.

So far the three-dimensional structure has not been reported for PNGase A, neither has the protein sequence nor the isolation or sequencing of its gene.

1.6 PNGase F from *Flavobacterium meningosepticum*

In 1984 Plummer and co-workers reported the detection of a Peptide:*N*-glycosidase F activity in preparations of ENGase F which until then had been used almost exclusively for the analysis of glycoproteins [54]. Even commercial preparations of ENGase F were shown to contain PNGase F activity. The two enzymes, ENGase F and PNGase F, were shown to exhibit different pH-optima. At pH 4 only ENGase activity could be detected whereas at pH 9.3 PNGase F was highly active and only a

slight level of ENGase activity was present. Attempts to separate the two different enzyme activities proved, however, to be labour intensive and difficult. Continued research on the purification of oligosaccharide-cleaving enzymes from cultural filtrates of *Flavobacterium meningosepticum* revealed that in addition to PNGase F and Endo F activity two new distinct endoglycosidases were expressed by the organism [75]. These were named Endo F_2 and Endo F_3.

In 1990 Tarentino and co-workers reported the molecular cloning and amino acid sequence of PNGase F [4]. In the same year Barsomian and co-workers succeeded in cloning the PNGase F gene and over-expressing the recombinant enzyme in *E. coli* cells [76]. It is noteworthy that in both approaches the enzyme was not secreted, as is the case in *Flavobacterium menongosepticum*, but both were produced intracellularly. The yield of soluble enzyme, however, was very low. In 2000 Loo and co-workers [77] succeeded in designing a promoter-based *E. coli* expression system in which the active enzyme is exported to the periplasm. Using this technique the active enzyme is expressed in relatively high yields (8 mg per litre of culture medium).

PNGase F has been crystallised and its three-dimensional structure determined independently by two groups [63], [78-80]. The PNGase F molecule was shown to consist of two tightly associated all β-domains that displayed a structural similarity to many viral coat proteins and animal lectins indicating an ability to bind carbohydrates. The active site of the enzyme was shown to be located in the top cleft of the molecule by Kuhn *et al.* in 1995 [55] by co-crystallising the enzyme with *N,N'*-diacetylchitobiose, which showed an inhibitory effect towards PNGase F. Furthermore site-directed mutagenesis experiments revealed potential amino acid residues involved in catalysis. However, the binding-mechanism for the peptide-bearing substrates has not yet been elucidated and the exact catalytic mechanism of the enzyme is still unknown.

The enzyme shows a very broad substrate specificity being able to act on high-mannose, complex (bi-, tri- and tetra-), hybrid and polysialic chain-containing glycopeptides and proteins [4 and references therein].

Fan and Lee [59] investigated the substrate requirements for PNGases F and A with synthesised glycopeptides bearing cellobiosyl, lactosyl and *N,N'*-diacetylchitobiosyl

cores. It was shown that neither of the enzymes could cleave cellobiose and lactose glycopeptides, indicating that the *N*-acetyl group is essential for the binding of the substrate. Substrates bearing only one GlcNAc residue attached to a protein chain were processed so slowly that a chitobiosyl core could be regarded as a minimal structural requirement for PNGase F.

Interestingly, the composition of the carbohydrate chain outside the conserved trisaccharide core structure was shown to have a very moderate influence on the hydrolysis rates. This was extensively investigated by Altmann and co-workers [56] who processed *N*-linked glycans with various exoglycosidases and assessed the influence of the carbohydrate shortening on the enzymatic activity of PNGase F and A. Down to the trisaccharide core, Manβ1-4-GlcNAcβ1-4-GlcNAcβ, carbohydrate shortening of the substrate did not show a significant influence on the enzymatic activity observed. Further shortening to an *N*,*N*'-di-acetylchitobiosyl core reduced the activity significantly. This effect was more pronounced for PNGase F than for PNGase A. The investigations by Fan and Lee [59], however, indicated that such compounds were still good substrates for both enzymes, PNGase A and F.

Fucosylation at the 3-position of the innermost GlcNAc, a structural feature of many plant and insect glycoproteins, abolishes the activity of PNGase F [58].

The length and the primary structure of the peptide chain, however, significantly affected the rate of hydrolysis observed. Structures consisting of a carbohydrate chain bound to a single Asn were not hydrolysed [81].

The most comprehensive investigation into this aspect was presented by Fan and Lee [59] who tested a wide variety of di-, tri- and pentapeptides *N*-glycosidically linked to *N*,*N*'-diacetylchitobiose as substrates for PNGase A and F. Their investigations revealed that PNGase F requires at least a tripeptide for significant activity and that glycotripeptides in which the glycoside chain is located at the *C*-terminus are poorer substrates than those in which the carbohydrate chain is located at the *N*-terminus. It was shown that a pentamer containing the glycosylation motif Asn-X-Ser/Thr was not required for activity and that the Asn must not necessarily be in peptide linkage on both sides [59]. The enzyme, however, showed a significantly reduced activity with such substrates.

Gosselin and co-workers [81] tested three different *N*-linked glycopeptide mimetics to assess the importance of structural features of the peptide backbone in the substrate recognition. It was concluded that PNGase F shows a stringent specificity for the conserved asparaginyl-carbohydrate linkage. Substrates in which the carbohydrate chain was either connected to the ε-amino group of lysine (via a succinyl bridge) or to the γ-amino group of glutamine were not processed by the enzyme. Acetylation of the proximal hydroxyl of serine in the peptide backbone of the substrate prevented processing by the enzyme. Thus the involvement of the proximal hydroxyl group in the catalytic mechanism seems likely although further experiments had to be carried out to confirm this hypothesis.

It has to be emphasised that unlike other PNGases, PNGase F shows an ability to act on native glycoproteins, albeit at often limited rates. In this context Chu [82] stresses the fact that since the enzyme is cleaving at the GlcNAc-Asn linkage the accessibility to the linkage is more critical than for example the ENGases. The high stability of PNGase F to SDS allows the deglycosylation of virtually any glycoprotein [83],[84]. Many different methods have been published on the use of PNGase F as a tool in glycomics [85-88].

1.7 Concepts of Inhibitors Towards PNGases

So far four concepts for the synthesis of suitable inhibitors towards PNGases have been reported in the literature; a high-mannose *C*-glycopeptide [5], a high-mannose glycopeptide analogue containing a glucose-asparagine linkage [6], complex glycopeptide derivatives *N*-linked to the side chain of glutamine [7], and a transition state *N*-glycosyl phosphonamidate [8] which has not yet been tested with the enzyme.

The first three concepts belong to the group of "so-called" substrate-analogue inhibitors. To a very large extent these compounds display structural features of the natural *N*-linked glycopeptide substrates to ensure recognition by the active site of the enzyme. However they all show a key structural difference in the conserved carbohydrate-peptide/protein linkage designed to impede hydrolysis by the enzyme and thus inhibit its activity. Due to their similarity with the natural *N*-linked glycans these compounds are also called *N*-linked glycopeptide mimetics. These inhibitors will be reviewed in section 1.7.1.

Transition state analogue inhibitors are stable molecules that are structurally designed to mimic the transition state of an enzymatic reaction. The *N*-glycosyl phosphonamidates will be reviewed in section 1.7.2.

1.7.1 *N*-Linked Glycopeptide Mimetics

The high-mannose *C*-glycopeptide published by Wang and co-workers in 1997 [5] was the first report describing the synthesis of an inhibitor for PNGases. The synthetic scheme employed in the synthesis of this compound used a chemoenzymatic approach.

In the first part of the synthesis the glucosamine-based building block **5** was synthesised bearing an anomeric β-configured aminomethyl-group suitable for the attachment to the side chain of aspartic acid in a pre-assembled peptide.

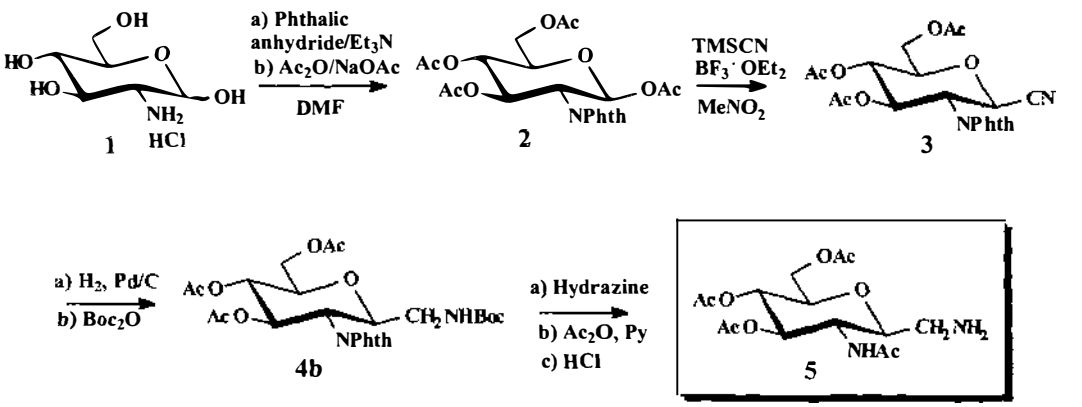


Figure 3: Synthesis of a *C*-glycoside building block for attachment to aspartic acid in pre-assembled peptides reported by Wang *et al.* [5].

After the attachment to the side chain of aspartic acid in a pre-assembled pentapeptide, this molecule was used as an acceptor substrate in the crucial enzymatic transglycosylation reaction to yield the desired high-mannose structure **8** (figure 4). The enzyme used for this reaction step was endoglycosidase A (Endo A), [89-91] which hydrolyses an intact glycan from a glycopeptide by cleaving between the two innermost GlcNAc residues of the invariant pentasaccharide core. This enzyme is not yet commercially available and was a gift from Takara Shuzo Co. in Japan to the authors.

In this synthesis, the donor substrate used was a high-mannose glycopeptide ($Man_9GlcNAc_2Asn$) that was isolated from soybean flour [92],[93]. During the transglycosylation step, the enzyme cleaves the glycosidic bond between the two innermost GlcNAc residues of the donor. However, instead of simply hydrolysing this glycosidic bond, it also posesses the capability to transfer the reducing end of the carbohydrate chain to the 4-position of the acceptor substrate **7**. The reported yield for the resulting high-mannose *C*-glycopeptide **8** was 26 %. However, the drawback of this synthesis is the extremely limited availability of endoglycosidase A.

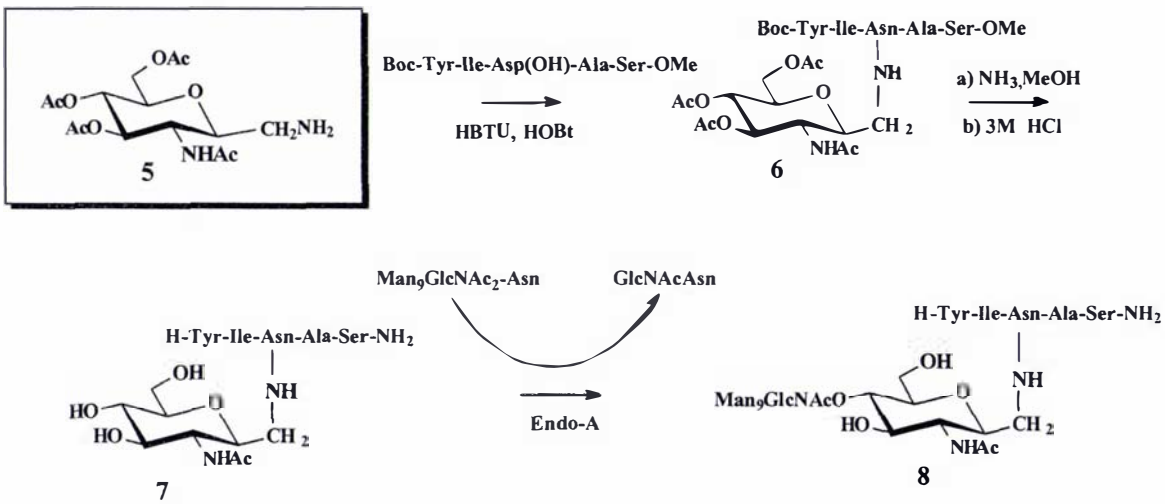


Figure 4: Chemoenzymatic synthesis of the high-mannose *C*-glycopeptide 8 by Wang *et al.* [5].

Such kinetically controlled enzymatic transglycosylations often proceed with low yields and limited regioselectivity. Nevertheless, they can permit a facile access to the desired product within one reaction step and without the extensive use of protecting group chemistry. Therefore, even if low yielding, they offer a considerable advantage over conventional chemical methods. The concept of enzymatic transglycosylation is well known [94].

Before conducting inhibition experiments using the high-mannose *C*-glycopeptide, the compound was tested as a substrate for the PNGases to be used in the inhibition trials (see below). The authors reported that none of the PNGases used was able to hydrolyse the *C*-glycopeptide.

This synthetic high-mannose *C*-glycopeptide was shown to have K_i- values of 1 and 43 μM for PNGase A and F respectively. The substrate used for these experiments was the corresponding *N*-glycopeptide. The K_i-values with PNGase Sm (*Sphingobacterium multivorum*) and PNGase HO (hen ovalbumin) were determined using a ^{14}C-labelled fetuin glycopeptide and calculated to be 152 and 157 μM respectively.

The authors concluded that their synthetic *C*-glycopeptide was an inhibitor of glycoamidases from various sources and could therefore be termed a broad-spectrum inhibitor for this class of enzymes.

A year later in 1998 a different research team, but again supervised by Yuan C. Lee [6], reported the synthesis of a high-mannose glycopeptide analogue containing a glucose-asparagine linkage (figure 5). Again the approach consisted of the straightforward synthesis of a peptide linked acceptor substrate which was then used in an enzymatic transglycosylation reaction employing endoglycosidase A. This approach was completely reminiscent of the synthesis of the high-mannose *C*-glycopeptide with the only difference being that the acceptor substrate used in the enzymatic step was a glucose *N*-linked to a pentapeptide.

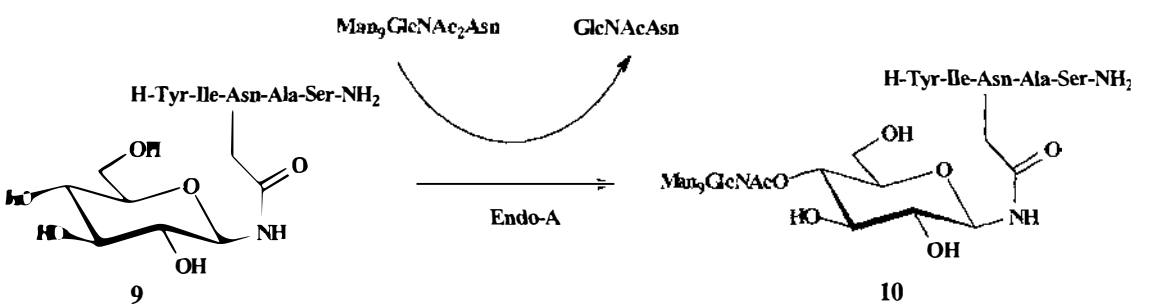


Figure 5: Enzymatic transfer of the donor Man$_9$GlcNAc to the glucosyl pentapeptide acceptor using endoglycosidase A by Deras *et al.* [6].

As with the high-mannose *C*-glycopeptide this glycopeptide mimetic was tested as a substrate for commercially available PNGase A and F. Neither of the enzymes were capable of hydrolysing this compound.

Inhibition experiments were performed with **10** (figure 5) testing its inhibitory activity in a digest of a glycopeptide isolated from bovine asialofetuin. The K_i values were determined to be 2 μM for PNGase A and 53 μM for PNGase F. The authors do not state why they tested the high-mannose glycopeptide mimetic against a substrate bearing complex glycans rather than the high-mannose glycan that had been used in the experiments with the *C*-glycopeptide.

The authors draw the conclusion that the acetamido group of the innermost sugar plays an important role in the recognition and cleavage of *N*-glycosidic bonds by PNGases. Furthermore, the inhibition constants are similar to those calculated for the inhibition of PNGase A and F using the *C*-glycopeptide. However, they do not take into consideration the fact that different substrates had been used in both inhibition trials.

A very interesting result reported by the authors is that the truncated glycopeptides $Man_9GlcNAc_2Asn$ and $Man_9GlcNAcGlcAsn$ as well as the free glycan from fetuin glycopeptide weakly inhibited PNGase F (Ki ~1 mM) but not PNGase A. Conversely, a shortened version of the glycopeptide mimetic having only a single glucose on the peptide backbone did not show any inhibitory effect. Thus the authors hypothesise an important role for the oligosaccharide chain in the inhibition of PNGases.

Haneda and Tanabe from the Noguchi Research Institute in Japan described the synthesis of a complex glycopeptide derivative *N*-linked to the side chain of glutamine and the testing of its inhibitory effect towards PNGase F in their Japanese patent of November 2000 [7].

As was the case for the two inhibitors previously described, this synthesis was also achieved using a chemoenzymatic approach employing an endoglycosidase. This time the authors prepared a glycopeptide as an acceptor substrate which had a single GlcNAc bound to the side chain of glutamine in the hexapeptide (H-Lys-Val-Ala-Gln(GlcNAc)-Lys-Thr-OH). Transglycosylation using endoglycosidase M from *Mucor hiemalis* [95],[96] and a disialo complex-type donor glycopeptide produced the glutamine linked glycopeptide mimetic **12** shown in figure 6.

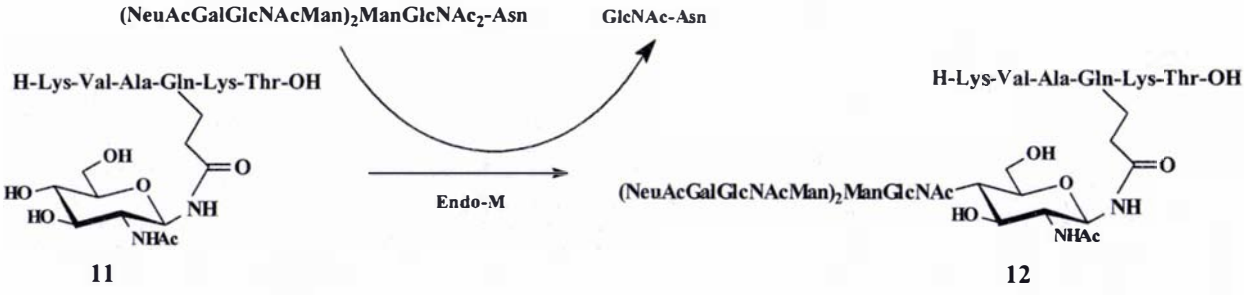


Figure 6: Synthesis of the glutamine bound disialo glycopeptide 12 using endoglycosidase M by Haneda and Tanabe [7].

This newly synthesised glycopeptide mimetic was tested as a substrate for both PNGases A and F. The authors reported that neither of the enzymes were capable of hydrolysing this glycopeptide.

Inhibition experiments conducted using the corresponding Asn-linked glycopeptide as a substrate showed that in a 1:1 mixture of the Asn and the Gln-linked glycopeptide the activity of PNGase F was reduced to 77 % compared to the control without any Gln-bound inhibitor present. No inhibition constants have been reported for this compound towards PNGase F nor has any kinetic data been reported with PNGase A.

The authors conclude that their Gln-linked glycomimetic is stable to the *in vivo* degradation by PNGases and therefore offers some potential in the design of physiologically stable glycopeptide based drugs. The potential for this class of mimetics to act as inhibitors of PNGases (PNGase F) may therefore be valuable in the future.

1.7.2 *N*-Glycosyl Phosphonamidates

The use of *N*-glycosyl phosphonamidates as potential transition state analogue inhibitors for PNGases was published by Ferro *et al.* in 1998 [8]. The main purpose of their work was to design haptens for generating catalytic antibodies with PNGase-like activity by binding *N*-glycosyl phosphonamidates onto suitable carrier proteins. Such structures might also find potential application as ligands in the affinity chromatography of glycoamidases.

Transition state analogue inhibitors are stable molecules that are structurally designed to mimic the transition state of an enzymatic reaction. There is scientific evidence that the active site of an enzyme is designed to optimally bind the transition state of the reaction catalysed. Therefore transition state mimetics may not only inhibit the enzymatic activity but may also bind to the enzyme with a higher degree of affinity than the natural substrate [97]. Figure 7 shows the proposed transition state during amide bond cleavage by glycoamidases. However, to date it is not known how the OH anion for the nucleophilic attack is generated [55] (see section 6.3 for a detailed discussion of the substrate binding and cleavage mechanism of PNGase F).

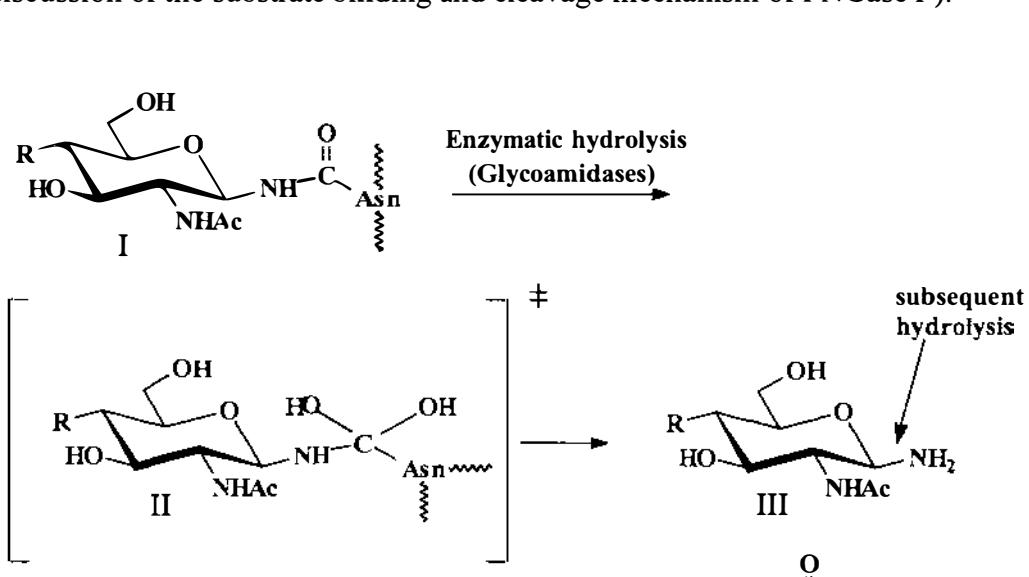


Figure 7: Proposed transition state during amide-bond cleavage (R = oligosaccharide).

The *N*-glycosyl phosphonamidates closely resemble the proposed transition state, mimicking geometry and charge distribution, but are designed to be stable towards cleavage by PNGases.

The key reaction step in the synthesis of *N*-glycosyl phosphonamidates, the coupling between a glycosyl amine and a phosphonochloridate, is illustrated in figure 8. The alkyl/aryl group (Bn in this example) might ultimately contain a functional group that will permit the attachment to a solid support e.g. for the design of an affinity resin.

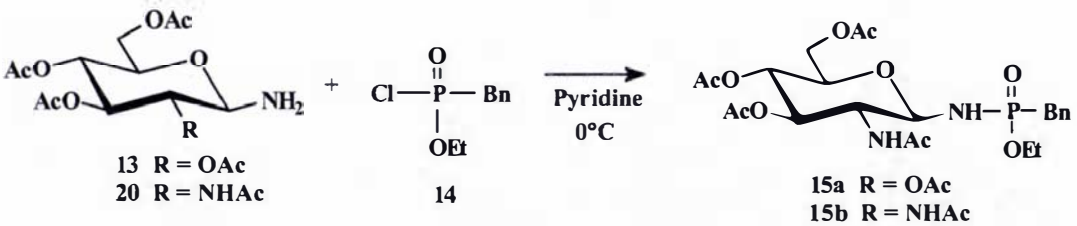


Figure 8: Synthesis of *N*-glycosyl phosphonamidates by Ferro *et al.* [8].

As an example the 1-amino-sugar **20** can be synthesised from the glucosamine hydrochloride **16** via its oxazoline **18** as shown in figure 9. Alternatively the deacetylated analogue of **20** can be synthesised directly from GlcNAc by dissolving the compound in a saturated ammonium hydrogen carbonate solution for 10 days at room temperature [59].

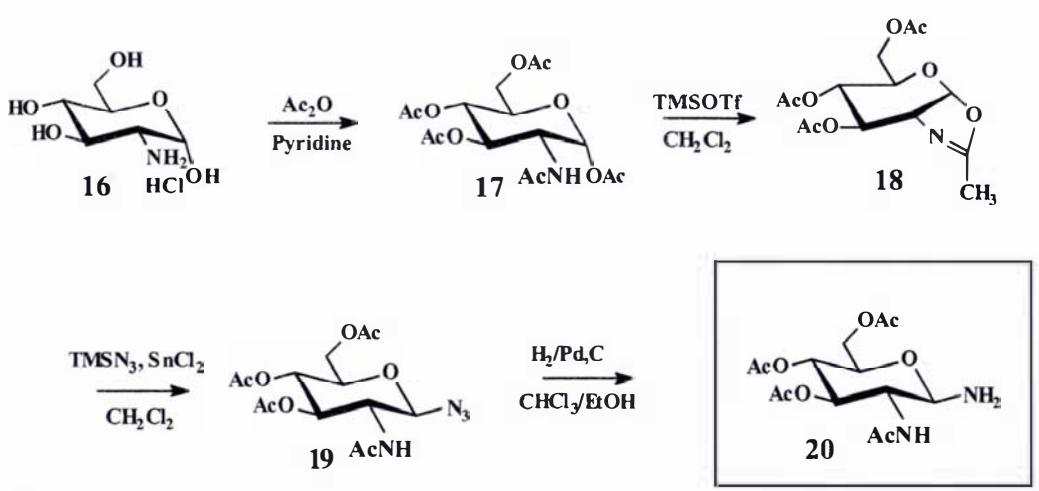
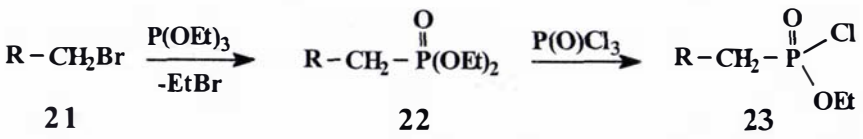


Figure 9: Synthesis of the glycosyl amine building block 20 [59].

Figure 10 outlines the general scheme for the synthesis of phosphonochloridates of the general structure donated by **23**. These compounds are accessible from the corresponding diethyl alkylphosphonates **22** using phosphorous oxychloride as reported by Morise *et al.* [98].



21 22 23

Figure 10: General pathway for the synthesis of phosphonochloridates (R = alkyl/aryl).

The diethyl esters donated by the general structure **22** can be synthesised from the corresponding alkyl bromides **21** as reported by Yamauchi *et al.* [99]. This reaction is commonly known as the *Arbuzov-Michaelis* reaction.

N-Glycosyl phosphonamidates bearing a phosphate ester group are referred to as "capped" phosphonamidates. Ferro *et al.* [8] successfully dealkylated these "capped" *N*-glycosyl phosphonamidates using TMSBr, a procedure published by McKenna *et al.* [100]. However, the authors indicated that the dealkylation of the amidates is not a prerequisite for the interaction of these compounds with bioactive peptides.

The interaction of *N*-glycosyl phosphonamidates with glycoamidases has not yet been reported. Therefore the inhibitory activity of these compounds towards PNGases remains purely speculative. A crucial factor for the success of these compounds will be determined by whether the transition state geometry is sufficient for recognition by the active site of the enzyme as the length of the carbohydrate and peptide moiety of the compounds presented by Ferro *et al.* clearly are not. Nevertheless, there is the potential that they may inhibit the enzyme and therefore will be able to be used for example as ligands for affinity chromatography.

While the use of phosphonate esters and amides as transition state analogues is well established their synthesis often remains difficult [101].

1.8 Aims of the Thesis

The aims of this work are to gain a deeper insight into the structural requirements for substrate recognition by PNGase F, develop a detailed understanding of the substrate binding and cleavage mechanism of this enzyme and synthesise potential inhibitors of PNGase F as possible candidates for use as affinity ligands.

For this purpose the development of pathways for the synthesis of new *N*-linked glycopeptide mimetics for use in kinetic and structural studies employing recombinant PNGase F will be required.

The synthesis of potential inhibitors will be based on the concept of substrate analogues, more specifically *C*-glycopeptides. In contrast to published methods the focus will be on the development of a straightforward chemical procedure which is not reliant on an enzyme that is not commercially available. Also the structures selected for these compounds will be guided by the minimal structural requirements for substrate recognition.

Furthermore, the synthesis of different *N*-linked glycopeptide mimetics which will show a distinct structural variation in comparison with the naturally occuring *N*-glycopeptide will be investigated.

The efficacy of all synthesised compounds as either substrates or inhibitors of PNGase F will be judged by kinetic studies. These studies are designed to provide a deeper insight into the structural requirements for substrate recognition and define possible candidates for the use as an affinity ligand.

Because of the limitations of the current HPLC based assay used in our laboratory for the determination of PNGase activity, it will be necessary to develop an assay using a fluorescently labelled substrate. The use of a fluorescent label is expected to give increased sensitivity and accuracy for the assay procedure.

Structural studies will also be carried out using the glycopeptide mimetics. One of the techniques employed will be the co-crystallisation of a potential ligand with recombinant PNGase F. If successful, analysis of the crystals will be carried out using X-ray crystallography and molecular replacement techniques. Furthermore molecular modelling studies using the resolved structure of PNGase F will be employed to develop a further understanding of the substrate binding and cleavage mechanism of this enzyme. These studies will be supported by the use of targeted enzyme mutants.

CHAPTER 2

Development of a New Assay for PNGases

2.1 Introduction

As outlined in section 1.4, many different methods have been developed for the determination of PNGase activity. Until recently the standard method used in our laboratory was a HPLC based assay using an unlabelled glycopeptide (11-mer) isolated from hen ovalbumin [63]. The isolation of the glycopeptide was straightforward and inexpensive using a cyanogen bromide (CNBr) digest of hen ovalbumin, followed by precipitation steps, gel filtration and preparative HPLC purification. Although the assay met the need of assessment of activity it lacked the sensitivity in the range of concentrations necessary for the accurate determination of K_M and V_{max} values. Therefore it was necessary to develop a substrate that could be detected at much lower concentrations. The addition of fluorescent labels to peptides to enhance detection has been widely used in a number of applications such as Edman sequencing [102]. There is a wide variety of commercially available fluorescent labels to chose from[1].

The fluorescent labelling of the glycopeptide from hen-ovalbumin was undertaken. There were three main requirements for the labelled substrate. Firstly, the labelled substrate had to be stable under the established assay conditions. Secondly, it had to have a range of sensitivity that would allow measurement of the product peak at low substrate concentrations (min. μM). Thirdly, the labelled substrate had to give a linear response with dilution. As the labelling sites (the *N*-terminal side and the side chain of Lysine) in the peptide moiety (EEKYNLTSVLHs) were at least two residues apart from the glycosylated asparagine, the labels were not expected to interfere with the affinity of the substrate for the active site of the recombinant PNGase F.

2.2 Labelling with Fluorescamine

Fluorescamine is intrinsically nonfluorescent[1] but reacts rapidly with primary aliphatic amines, including those in peptides and proteins, to yield a fluorescent derivative [103], (figure 11). The excitation wavelength used for fluorescamine labelled compounds is 390 nm and the emission wavelength is 480 nm.

---

[1] Molecular Probes, Inc.; Literature on " Reagents for Analysis of Low Molecular Weight Molecules" (http://www.probes.com/lit/bioprobes25/part12.html).

Excess reagent is rapidly converted to a nonfluorescent product by reaction with water meaning there is no additional chromatographic purification step required to remove excess reagent which would interfere in the detection process.

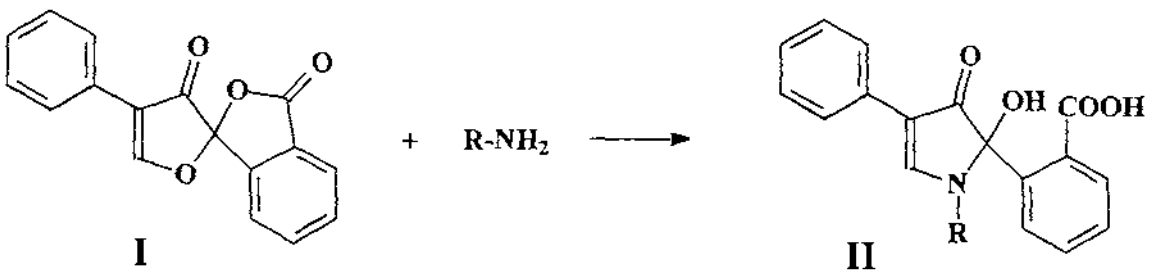


Figure 11: General scheme for the fluorescent labelling of primary amines using fluorescamine

The advantage in labelling with fluorescamine is its ease of use. The mixture of the enzymatic digest containing undigested hen ovalbumin glycopeptide and the deglycosylated peptide was simply labelled by adding an excess amount of fluorescamine in acetonitrile prior to HPLC analysis. This way the enzymatic digest could proceed without the interference of any label or side products of the labelling reaction present in the reaction vial. Furthermore no additional purification step was required.

The disadvantage in the use of fluorescamine is its instability under acidic conditions. Therefore the HPLC-procedure had to be modified. Eluant A (0.1 % TFA solution in doubly distilled water) was replaced with Tris-HCl pH 7.0 and eluant B was neat acetonitrile. The ionpairing agent trifluoroacetic acid (TFA) could not be used.

The separation was performed on a C_{18} analytical column (Jupiter, Phenomenex). Unfortunately the use of Tris-buffer resulted in a poor separation of the labelled substrate and product. This could be overcome by replacing eluant A (Tris-HCl pH 7.0) with a 10 mM sodium dihydrogenphosphate buffer pH 7.5.

Several attempts to establish a standard curve were undertaken but a linear response with dilution could not be attained. As this problem could not be overcome a different fluorescent label had to be employed.

2.3 Labelling with Fluoresceine Isothiocyanate

Fluoresceine derivatives which react with amines belong to a group of very common fluorescent reagents for covalently labelling proteins and peptides. Fluoresceine-5-isothiocyanate, also referred to as FITC "Isomer I", is an inexpensive reagent which is both easy to handle and commercially available.

The advantage of fluoresceine derivatives is their relatively high absorptivity, excellent fluorescence yield and their high solubility in water. Disadvantages are the occurrence of photobleaching and the pH- sensitivity of the fluorescence which is reduced below pH 7.

Photobleaching, however, will not cause a problem during the use of an FITC-labelled compound in a discontinuous assay. Such phenomena would have to be considered only in a continous assay where repeated excitation of fluorescence would reduce the fluorescence yield.

A disadvantage in comparison to the labelling with fluorescamine is the fact that the FITC-labelling involved an additional preparative HPLC step for the final purification of the labelled product.

The protocol for the FITC-labelling of the glycopeptide isolated from hen-ovalbumin was adapted from Hentz *et al.* [104] who used this method for the labelling of human insulin. Modifications were as outlined in Materials and Methods (section 8.6).

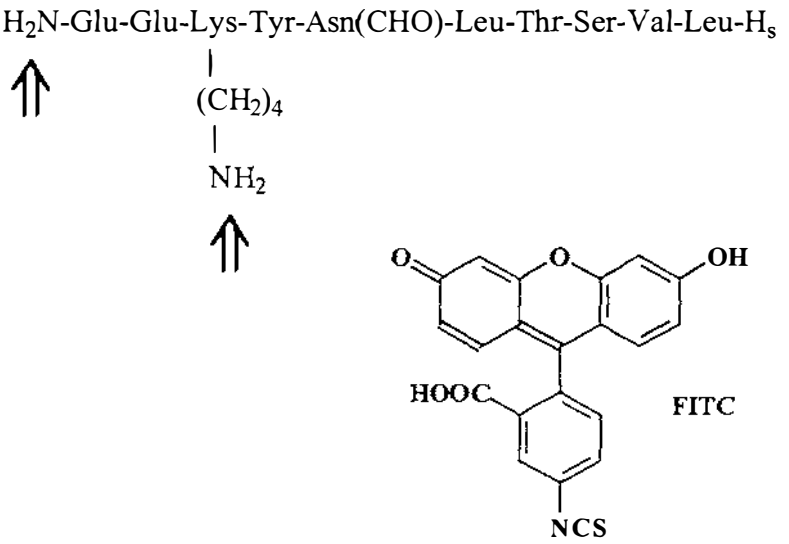


Figure 12: The ovalbumin glycopeptide (CHO indicates the carbohydrate chain on the glycosylated Asn) was labelled in the positions indicated by arrows using fluoresceine isothiocyanate (FITC). H_s = Homoserine lactone.

Figure 12 gives a schematic overview of the FITC-labelling of the glycopeptide from hen ovalbumin as well as the structure of fluorescein isothiocyanate. Note that the homoserine lactone of the *C*-terminus, which is derived from a methionine residue during the CNBr digest of the protein, was converted to the more stable open form by heating at reflux in BTP-buffer pH 8.5 ((1,3-bis[tris(hydroxymethyl)-methylamino]propane) after the labelling reaction.

The labelled glycopeptide was shown to be stable under the conditions of the standard HPLC-based discontinuous assay. The use of 0.08 % TFA as an ion pairing agent in the HPLC separation did not cause any stability problems with the labelled substrate. Critically, the FITC-labelled substrate gave a linear response with dilution. The sensitivity of the assay was increased significantly and the spectra were free of background noise, providing baseline separation for the substrate and the product of the enzymatic digest as shown in figure 13. The conditions for the assay are described in the general procedures (section 8.2).

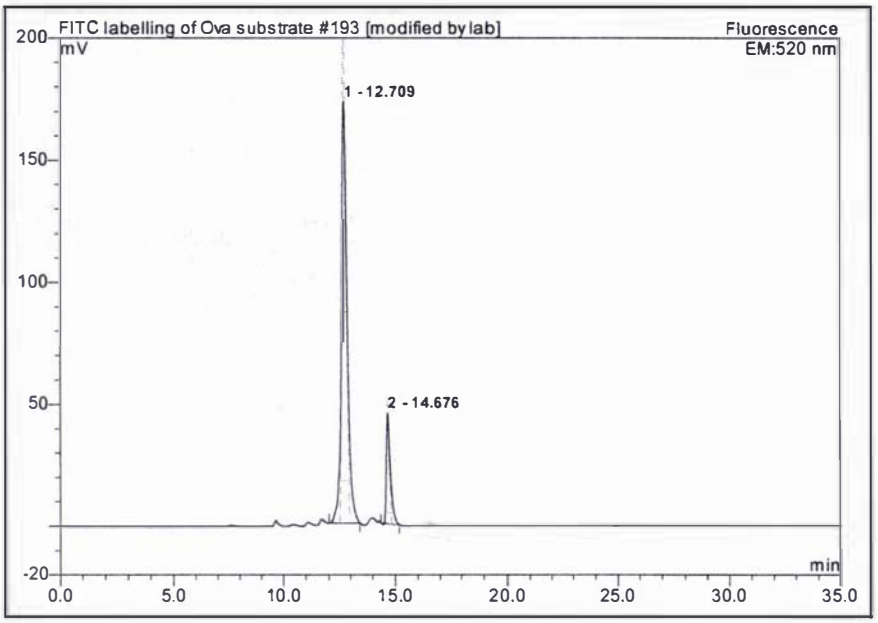


Figure 13: HPLC-chromatogram showing FITC-labelled Ova-substrate (peak 1) and the product of the enzymatic digest (peak 2) using recombinant PNGase F. Retention times are shown in minutes.

The excitation wavelength used was 495 nm and the emission wavelength 520 nm. It was shown that the labelled substrate was too unstable to be kept in a buffer solution in a fridge for longer than 2 days. In contrast the compound was shown to be stable in various buffers when stored below minus 18 °C in the dark.

Kinetic data for the PNGase F deglycosylation of this new fluorescently labelled glycopeptide has been determined. The results are presented in section 4.2. A comparison of K_M values using our recombinant PNGase F and the fluorescently labelled substrate with published data for similar substrates showed that the affinity of the substrate for the enzyme was not affected by the presence of the label.

CHAPTER 3

Synthesis of *N*-Linked Glycopeptides and Glycopeptide Mimetics

3.1 Introduction

A partial aim of this thesis is the development of pathways for the synthesis of new *N*-linked glycopeptide mimetics for the use in structural and kinetic studies with PNGase F. Section 1.7.1 gives an overview of PNGase inhibitors that have been published so far. All of these structures have been shown to inhibit the activity of PNGase F and to partially inhibit other PNGases. However, the one common drawback associated with all of these inhibitors is that in employing the existing strategy, each oligosaccharide moiety can only be added to a peptide by means of a crucial, low-yielding enzymatic transglycosylation reaction using enzymes that are not commercially available.

This work has therefore been directed towards the development of straightforward procedures based solely on chemical methods. Because of the difficult chemistry involved in the synthesis of complex oligosaccharides, the compounds synthesised during this investigation are simpler than those that have been published so far and their design is based on what is known about substrate recognition by PNGase F as outlined in section 1.6.

3.2 Synthesis of *C*-Glycopeptides

The high-mannose *C*-glycopeptide published by Wang et *al.* in 1997 [5] was the first report of the synthesis of a substrate analogue inhibitor of a PNGase and the assessment of its inhibitory properties using PNGases from different sources (section 1.7.1). The synthesis of this compound was lengthy and the chemoenzymatic step involved was low yielding. For this reason, the development of a straightforward procedure for the synthesis of *C*-glycopeptides was deemed to be both desirable and important.

When designing inhibitors for a specific enzyme, both the substrate binding and the mechanism of catalysis need to be considered. While the three dimensional structure of PNGase F is known the catalytic mechanism has not been unequivocally determined. The molecular determinants of substrate binding for PNGase F are not yet known and structural studies of PNGase F enzyme –substrate complexes have not been reported to date.

Initial modelling studies carried out in our laboratory, using the three dimensional structures of glycopeptides "virtually" cleaved from the structures of glycoproteins known to be substrates for PNGase F, showed that it is difficult to access the active site residues proposed by Kuhn *et al.* (1995) [55] who co-crystallised PNGAse F with *N,N'*-diacetylchitobiose.

Although our modelling studies led to the proposal of a possible substrate binding and cleavage mechanism during a later stage of this work (section 6.3), the current design of the *C*-glycopeptides rests on the results of kinetic studies reported in the literature, especially the work by Altmann *et al.* [56] and Fan and Lee [59].

The minimal substrate requirements for PNGase F in terms of the peptide and the carbohydrate moieties are a di-chitobiosyl core and a pentapeptide respectively (section 1.6). The pentapeptide Tyr-Ile-Asn-Ala-Ser, as well as a few tripeptides [59] have been shown to provide good recognition for PNGase F. Thus *C*-glycopeptides bearing a di-chitobiosyl core and a tri-or pentapeptide moiety have the potential to show inhibitory activity towards PNGases and could therefore be valuable tools in both structural and kinetic studies.

The development of a new pathway for the synthesis of such *C*-glycopeptides incorporating a di-chitobiosyl fragment and a tri-or pentapeptide moiety was regarded as being possible using a straightforward convergent approach.

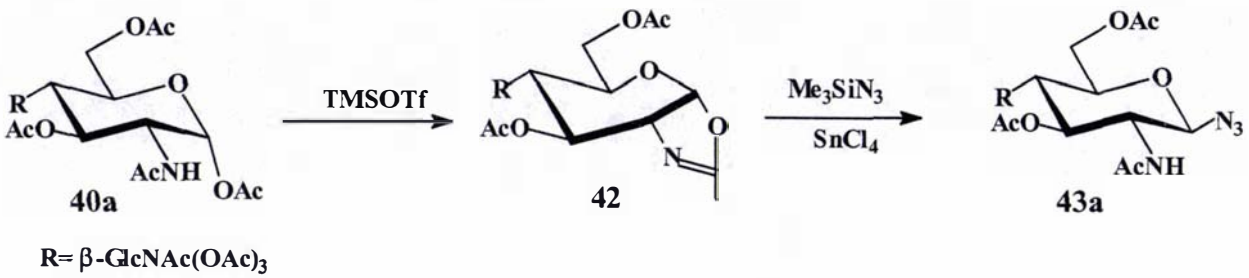
The synthetic approach was divided into three parts: (i) The preparation of the *C*-glycosyl building blocks, (ii) the preparation of the peptide building blocks and (iii) the final coupling reaction between the carbohydrate and the peptide building blocks to form the desired *C*-glycopeptides.

3.2.1 Preparation of the *C*-Glycosyl Building Blocks

The first synthetic target was a di-chitobiosyl *C*-glycosyl amine that could be attached to the side chain of aspartic or glutamic acid residues in various pre-assembled peptides.

The easiest approach seemed to be the functionalisation of chitobiose via its oxazoline. The oxazoline is accessible from the corresponding octaacetate. Oxazolines are susceptible to attack by nucleophiles. Therefore it should be possible

to introduce a β-configured functional group under Lewis acid catalysis as exemplified in scheme 1.



R= β-GlcNAc(OAc)$_3$

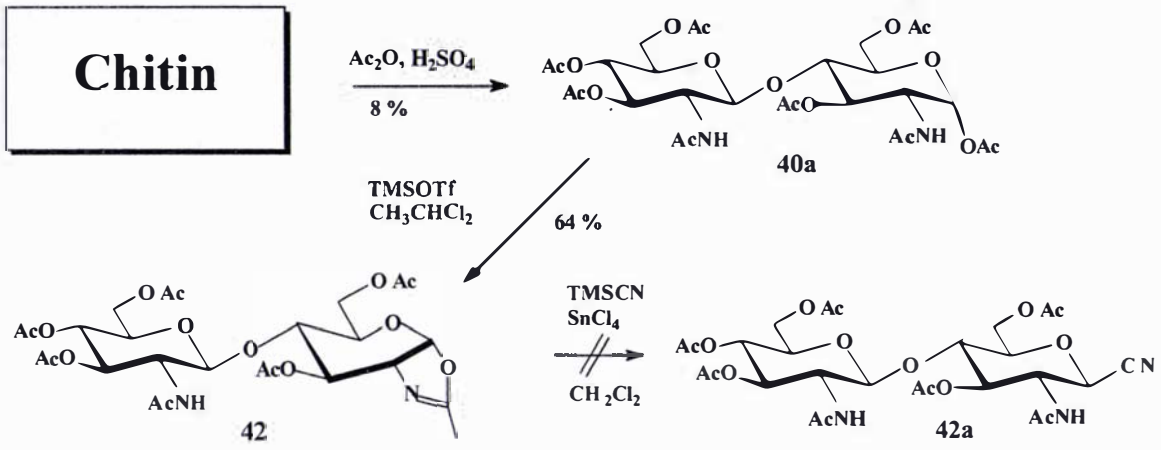
Scheme 1: Introduction of a β-configured azide into chitobiose as reported by Fan and Lee [59].

A search of the literature indicated that a nucleophilic opening of an oxazoline using azide displacement had been reported to proceed smoothly by several groups [105]. Based on this precedent, we hoped to be able to introduce a nitrile at the anomeric centre. The reason for chosing this functional group was that nitriles can be converted to the corresponding methylamines using catalytic hydrogenation giving access to the desired *C*-glycosidic building block.

For this purpose chitobiose-octaacetate **40a** was isolated from chitin (crab shells) using acetolysis conditions. The yields from this reaction are generally acknowledged to be below 10 % [106]. To increase the yields, higher initial temperatures (~ 60 °C) were used in order to produce a homogeneous solution in the shortest time possible (max. 4 h). Ultrasonication also proved to be beneficial. Once the chitin was partially degraded and solubilised, the reaction temperature was lowered (~35 °C) and the mixture stirred for 24 h. Many procedures use low initial temperatures that result in a slow initial degradation of the polymer. Consequently the concentration of soluble chitin fragments is very low throughout the acetolysis reaction and therefore these fragments are digested extensively to the monomers instead of the desired chitobiosyl fragments. The altered conditions employed allowed chitobiose-octaacetate to be isolated chromatographically pure in an 8 % yield as well as a minor amount of chitotriose.

The transformation into the corresponding oxazoline **42** was achieved in 64 % yield in dichloroethane using trimethylsilyl trifluoromethanesulfonate (TMSOTf) as catalyst.

Subsequent attempts to functionalise the molecule using trimethylsilyl cyanide (TMSCN), the most versatile source of cyanide [107] and a Lewis acid failed. A wide range of Lewis acids and reaction conditions using TMSCN were investigated but the desired product could not be isolated in any of the trial reactions (scheme 2).

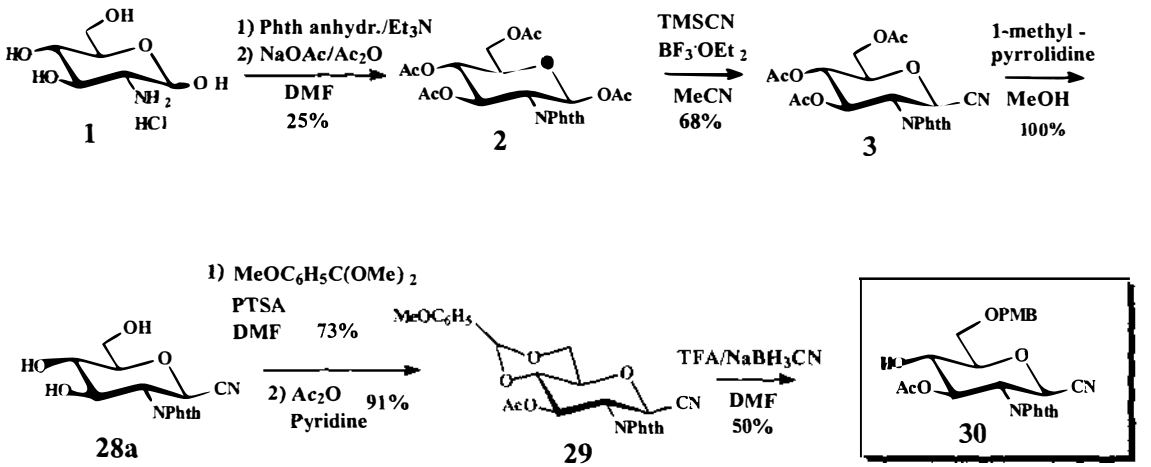


Scheme 2: Attempted functionalisation of chitobiose-derivatives using Lewis acids (e.g. tin tetrachloride).

Therefore the desired disaccharide had to be synthesised from its monosaccharide building blocks.

For this purpose the synthesis of a suitable acceptor molecule with a free hydroxyl group in the 4-position for the synthesis of the disaccharide was carried out. The acceptor **30** was synthesised from D-glucosamine hydrochloride **1** (scheme 3). Phthalimido-protection of the amine and acetylation of the hydroxyl groups was followed by reaction with TMSCN in the presence of boron trifluoride (BF_3), according to the procedure of Myers and Lee [107]. Deacetylation of **3** was achieved using *N*-methylpyrrolidine in methanol, as the *Zemplén*-procedure proved to be incompatible with the cyano function. This was followed by the introduction of a *p*-methoxybenzylidene acetal to block the 4- and 6-positions and subsequent acetylation of the 3-position. The regioselective opening of the acetal, following a procedure of Johansson and Samuelsson [108], using sodium cyanoborohydride as the reducing agent in DMF (dimethylformamide) under acidic conditions (trifluoroacetic acid) and in the presence of molecular sieves (3Å) gave the acceptor molecule **30**. The *p*-

methoxybenzyl (PMB) ether was chosen as a protecting group because this group can be cleaved easily under mild conditions.

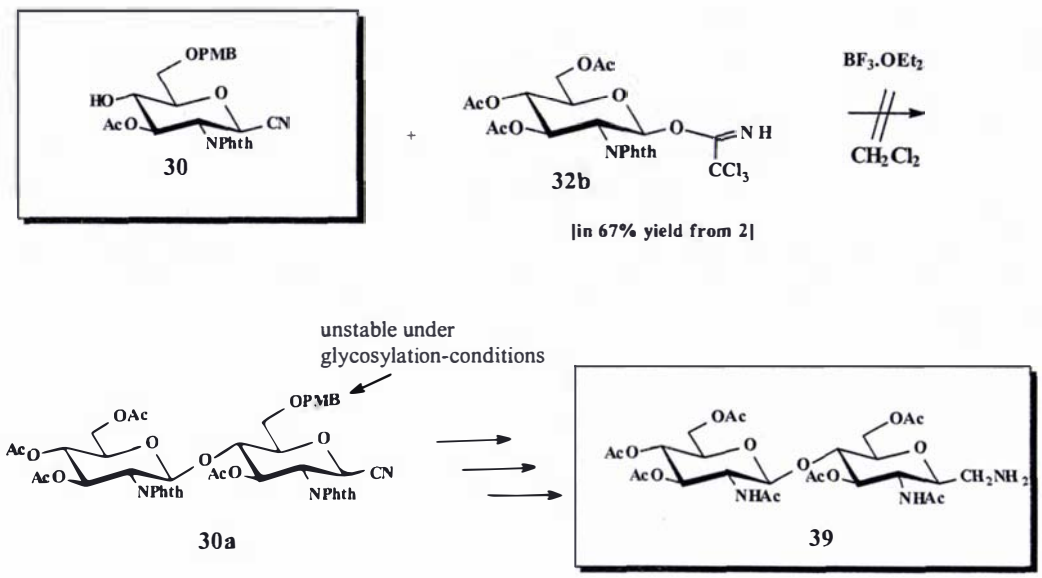


Scheme 3: Synthesis of the acceptor substrate 30.

Trichloroacetimidate **32b** was chosen as the glycosyl donor [109], (scheme 4). This molecule was synthesised from **2** in 67 % yield by selective deacetylation of the anomeric acetyl group using hydrazine acetate in DMF, followed by reaction with trichloroacetonitrile and DBU (1,8-diazabicyclo[5.4.0]undec-7-ene) in dichloromethane.

The desired glycosylation reaction between **30** and **32b** under BF_3-catalysis at 0 °C in dichloromethane was not successful due to the instability of the PMB group under the Lewis acid conditions employed (scheme 4).

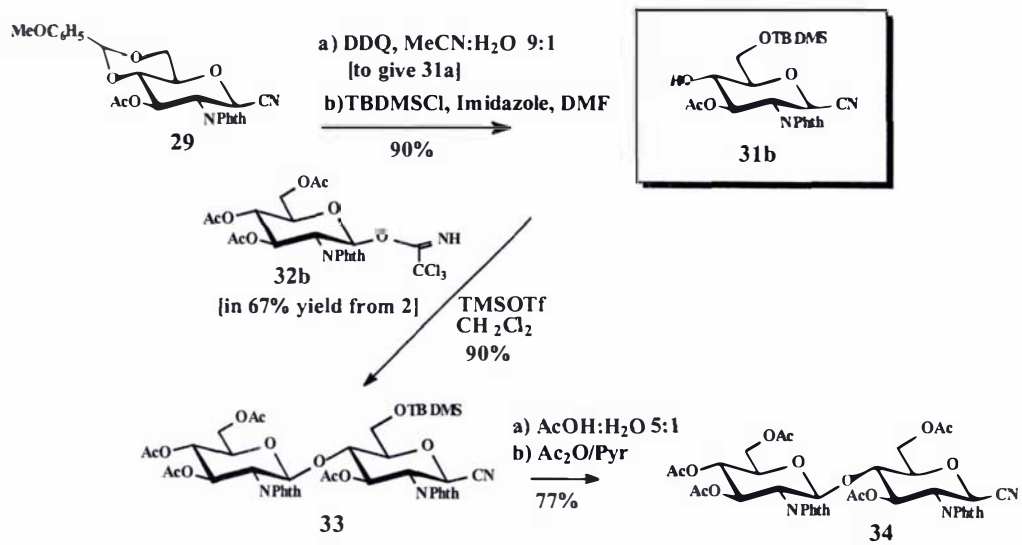
Lowering the temperature to – 40 °C did not improve this result. Following cleavage of the PMB group, the acceptor molecule was glycosylated at the 4- and 6-positions giving low yields of an undesired trisaccharide. Therefore the protecting group strategy had to be modified.



Scheme 4: Attempted glycosylation using BF_3 catalysis.

Instead of regioselectively opening the PMB acetal, it was cleaved completely (scheme 5). This was achieved quantitatively using DDQ (2,3-dichloro-5,6-dicyano-*p*-benzoquinone) in an acetonitrile-water system [110],[111]. Subsequent protection of the primary hydroxyl function as a TBDMS (*tert*-butyldimethylsilyl) ether resulted in the acceptor **31b** in 90 % yield (scheme 5).

The glycosylation reaction using this acceptor, with TMSOTf as catalyst, in dichloromethane at 0 °C gave the desired disaccharide **33** in 90 % yield (scheme 5).

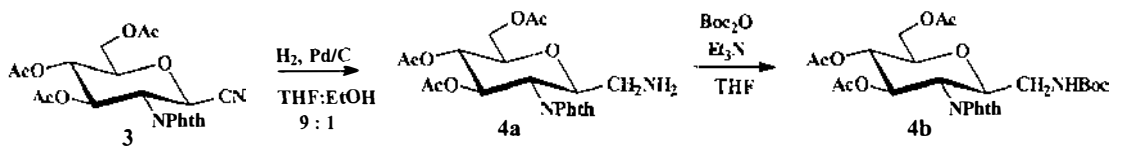


Scheme 5: Synthesis of the functionalised disaccharide building block 34.

The TBDMS group was then cleaved using aqueous acetic acid at 80 °C. The product was acetylated to give **34** in 77 % yield over the two steps.

This functionalised disaccharide was now available for further processing. The next crucial step was the transformation of the nitrile functional group into the corresponding methylamine group using catalytic hydrogenation. Furthermore, the phthalimido protecting group had to be transformed into an acetamide.

Towards this end, a similar scheme to that published by Wang *et al.* [5] for the analogous monosaccharide was used. They transformed the nitrile **3** to the Boc-protected methylamine **4b** in two steps (scheme 6). In the first step, the nitrile functional group was converted to the corresponding methylamine by hydrogenation in THF-EtOH using 10 % Pd on charcoal. In the second step, the free amino group was protected as a Boc-carbamate, an *N*-protecting group which ought to survive the de-phthalimidation and acetylation to give the 2-acetamido sugar.

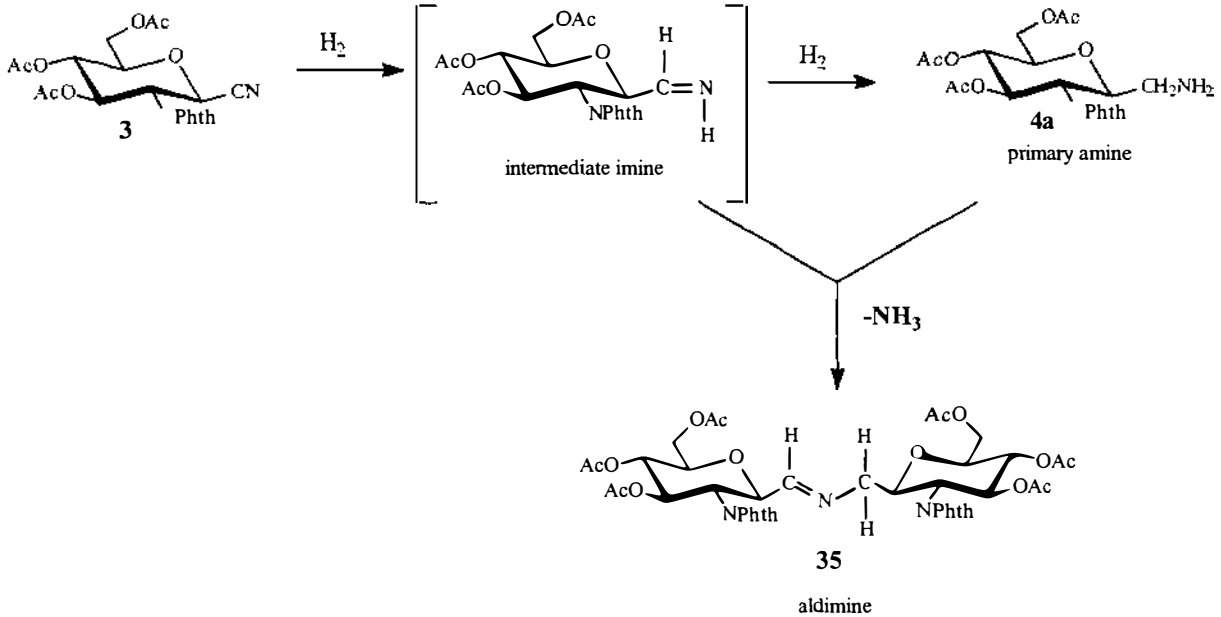


Scheme 6: Hydrogenation of the anomeric nitrile followed by Boc-protection as published by Wang *et al.* (1997) [5].

Employing these conditions used by Wang *et al.* in the pathway to convert nitrile **30a** into the disaccharide **39** (scheme 4) was not effective. To investigate the reasons for this failure, the corresponding monosaccharide was synthesised in order to reproduce the transformation on the exact substrate published by Wang *et al.* and prove its feasibility.

Using THF:EtOH 9:1 as the solvent system and catalytic amounts of 10 % Pd/C under a hydrogen atmosphere as shown in scheme 6, no primary amine could be isolated. However, the aldimine **35** (scheme 7) was obtained in almost 100 % yield as determined by NMR.

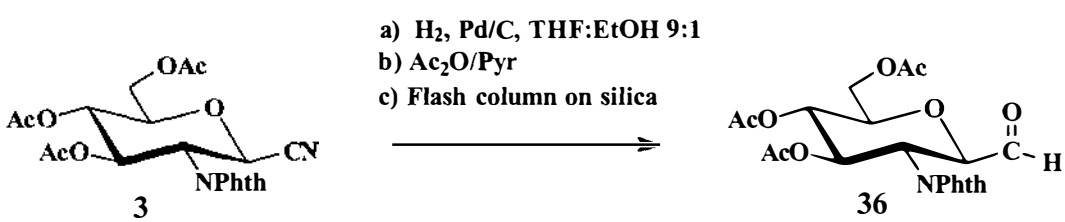
It has already been reported that the intermediate imine in the hydrogenation reaction can react with the desired amine, releasing ammonia to give an aldimine [112].



Scheme 7: Formation of the aldimine 35.

The aldimine may undergo reduction to the corresponding secondary amine but may also be stable as was observed in this case.

To further clarify these findings another experiment was conducted. If the hydrogenation of the monosaccharide bearing the anomeric nitrile functional group yielded a compound with the desired methylamino group, it should have been possible to protect the free amino function with a protecting group. The protecting group chosen was a simple acetamide, introduced using acetic anhydride in pyridine. The outcome of this three step approach consisting of hydrogenation, acetylation and column chromatography is shown in equation 1.



Equation 1: Formation of an aldehyde from the corresponding nitrile (see text).

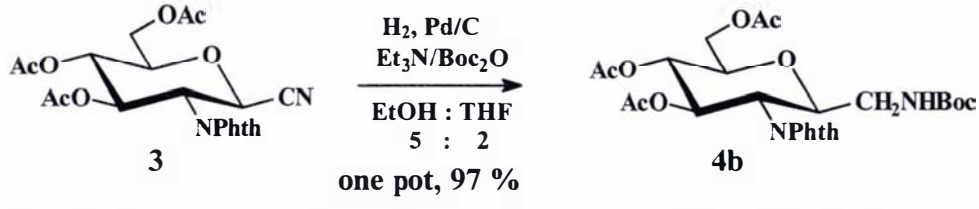
The isolation of the aldehyde **36** after column chromatography on silica was confirmed using NMR-techniques. Although this was a surprise it could be explained.

As the hydrogenation conditions led to the formation of the imine **35,** as shown in scheme 7, no structural change occurred during the attempted acetylation as there was no primary amino group to react with. This also explains why no change was observed in the TLC during this step. Subsequent column chromatography provided the conditions that led to the hydrolysis of the imine [112]. Residual water in the silica gel and the eluents as well as the acidic conditions present within the matrix of the column provided the required conditions for the formation of the aldehyde **36.**

No reasons could be identified for our inability to reproduce the hydrogenation of the nitrile as published by Wang *et al.* [5]. There may have been a variation in the quality of the reagents used.

As the hydrogenation of the anomeric nitrile was a crucial step in this synthetic pathway, the development of a convenient hydrogenation procedure was of particular importance. The approach taken to circumvent the formation of the aldimine was the *in situ* protection of the amine as a stable intermediate. Towards this end, an attempt was made to develop a one step procedure for the transformation of the glycopyranosyl cyanide to the corresponding *N*-(*t*-butoxycarbonyl)methylamine. The conditions established for this transformation to proceed smoothly in a single reaction step are shown in equation 2.

The β-configuration of the resulting glycopyranosyl methylamine **4b** was confirmed using nOe experiments which revealed that a nOe existed between H-1, H-3 and H-5.



Equation 2: One step transformation of glycopyranosylcyanides to the corresponding *N*-(*t*-butoxycarbonyl)methylamines [113].

Interestingly, the yield for the transformation is poor (52 %) when the reaction is performed under neutral conditions. In contrast, Saito *et al.* [114] reported the one-pot transformation of various 2-azido alcohols into 2-*N*-Boc-amino alcohols and preferred neutral conditions to prevent side reactions caused by base.

The general protocol for this reaction was also applicable to the disaccharide **34** as well as to two other compounds. Table 2 summarises these results. For a reprint of the resulting publication in *Tetrahedron Letters* [113] see Appendix II.

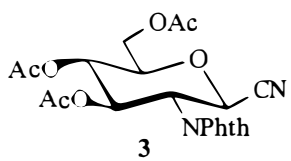
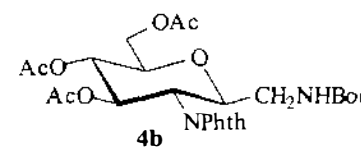
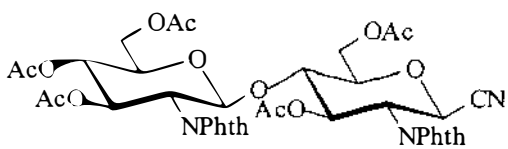
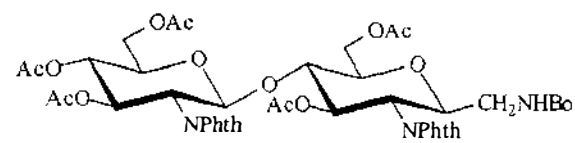
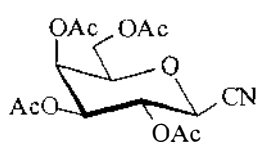
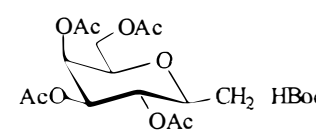
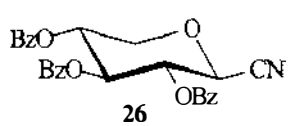
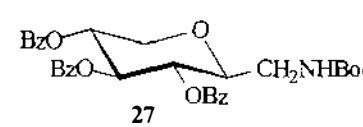
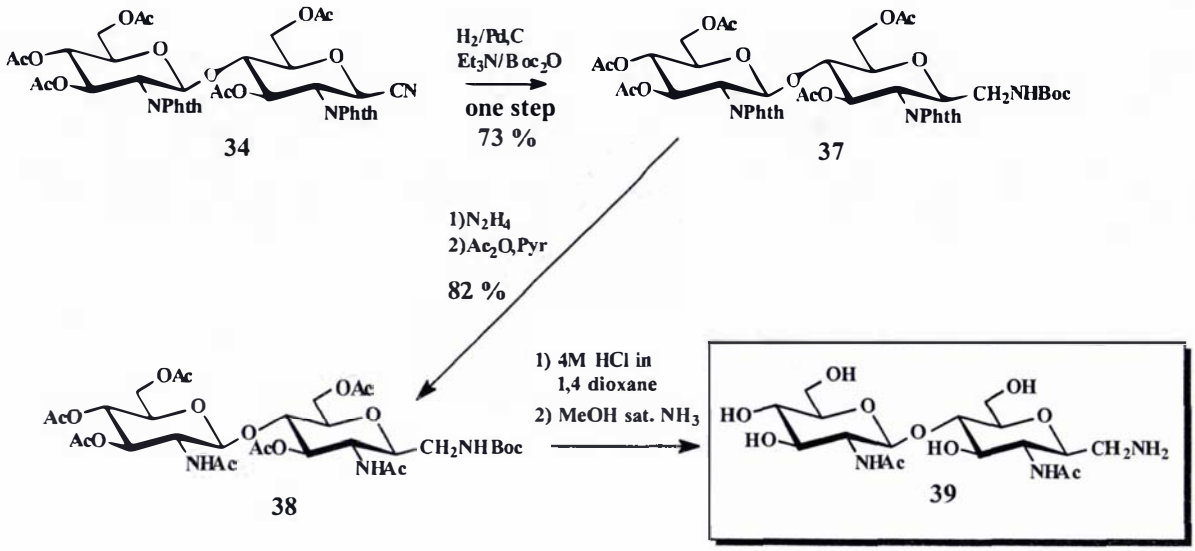
Glycopyranosylcyanide	**RCH_2NHBoc (Yield %)**
 **3**	 **4b** **(97 %)**
 **34**	 **37** **(82 %)**
 **24**	 **25** **(83 %)**
 **26**	 **27** **(51 %)**

Table 2: One-pot transformation of glycopyranosylcyanides to the corresponding Boc-carbamates.

After the successful completion of this crucial reaction step, the remaining steps in the formation of the *C*-glycosidic building block were accomplished without problems. De-phthalimidation was achieved by hydrazinolysis followed by acetylation to form the 2-acetamido sugar **38** (scheme 8). The Boc-protecting group was cleaved using 4M HCl in 1,4-dioxane as reported by Wang *et al.* [5]. The cleavage of the acetate esters was accomplished using a saturated methanolic ammonia solution to give **39** (scheme 8).



Scheme 8: Final reaction steps in the synthesis of the *C*-glycosyl amine 39.

This *C*-glycosyl amine was used without further purification in the coupling reaction with aspartic or glutamic acid in various pre-assembled peptides (section 3.2.3).

3.2.2 Preparation of the Peptide Building Blocks

Since its introduction by Bruce Merrifield in 1963, solid phase peptide synthesis (SPPS) has become the most widely used method for synthesising peptides [115].

Solid phase peptide synthesis using Fmoc (9-fluorenylmethoxycarbonyl) chemistry on an Applied Biosystems (Model 431 A) peptide synthesiser was used to synthesise two different pentapeptides. These compounds were synthesised on a Rink® amide resin [116]. Using this procedure the Fmoc-protected amino acid is coupled to the resin via its *C*-terminal carboxy function which is released from the resin as an amide after completion of the peptide assembly. The peptides were purified using preparative HPLC. Despite the strong advantages of SPPS compared to solution phase methods, the approach is relatively expensive due to the high cost of the Fmoc-protected amino acids.

Therefore the tripeptides required for the synthesis of the *C*-glycopeptides were synthesised using solution phase chemistry with Boc-protected amino acids and dicyclohexylcarbodiimide (DCC) activation. Compared to the solid phase synthesis this procedure was relatively time consuming but had the advantage that larger

amounts of material were obtained and purificaton was achieved by column chromatography (Materials and Methods).

The following list shows the peptides that were synthesised. The *N*-terminus of each tripeptide was substituted by a Boc-carbamate and the *C*-terminus esterified with a methyl group. The pentapeptides were acetylated at their *N*-termini and had an amide on their *C*-termini.

Peptide	(*N*-terminus) – Sequence - (*C*-terminus)
47b	Boc-**Asp**-Ala-Ser-OMe
49b	Boc-**Asp**-Ala-Thr-OMe
50	Ac-Tyr-Ile-**Asp**-Ala-Ser-NH_2
51	Ac-Tyr-Ile-**Glu**-Ala-Ser-NH_2

Table 3: List of peptides synthesised for the preparation of the *C*-glycopeptides.

3.2.3 Coupling of Carbohydrate and Peptide Fragments

The final phase in the assembly of the *C*-glycopeptides was the coupling of the carbohydrate and the peptide building blocks.

Two common approaches can be taken to synthesise *N*-linked glycopeptides and glycopeptide mimetics. The first involves a procedure which is commonly referred to as the building block approach, in which a glycosylated amino acid is incorporated into the growing peptide chain.

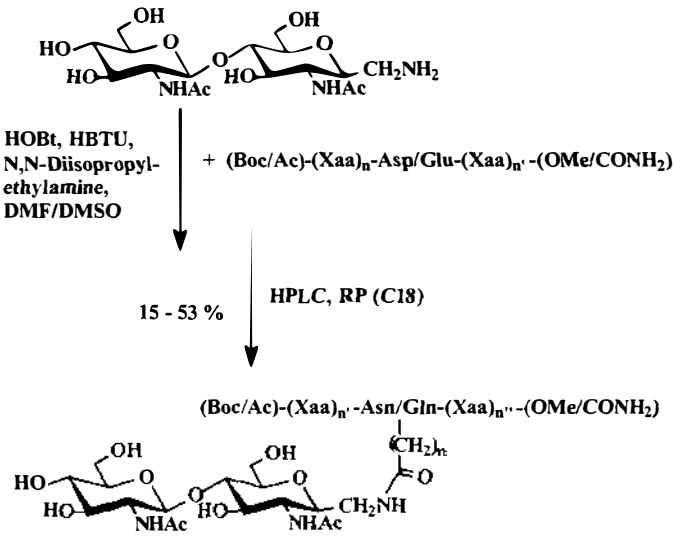
The second method, which has been adopted in this work, is the coupling of a glycosyl amine to the side chain of aspartic or glutamic acid in a pre-assembled peptide. This en-block approach has the advantage that the carbohydrate and peptide building blocks can be prepared separately and that the conditions used for the peptide synthesis do not interfere with the stability of the glycosidic bonds of the carbohydrate moiety, as is the case in the first approach (i.e. Boc chemistry is not useful for glycopeptide synthesis).

The coupling reaction represented a crucial step in this synthesis and conditions had to be carefully chosen in order to achieve the highest yields.

The most successful approaches for the coupling of glycosyl amines to aspartic acid containing peptides are those that have been reported by Anisfeld and Lansbury [117] and Wong and co-workers [118]. Two factors have been shown to be crucial to the success of these reactions: the choice of coupling reagent and the amount of base present in the reaction. Both have been thoroughly investigated by Anisfeld and Lansbury [117]. The use of HOBt (1-hydroxybenzotriazole) and HBTU (2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate in a DMSO (dimethylsulfoxide)/DMF solvent system using minimum amounts of a tertiary amine (diisopropylethylamine) have been reported to give satisfactory yields for the coupling. Wong and co-workers [118] who adopted these conditions for the coupling of an unprotected *N*,*N*-diacetylchitobiose glycosyl amine to a pentapeptide reported yields of approximately 85 % based on peak areas in HPLC-chromatograms.

In the current work a modification of these published conditions was used to couple the *C*-glycosyl amine **39** (section 3.2.1) to the side chain of aspartic or glutamic acid in pre-assembled peptides. The reactions were monitored using analytical RP (reverse phase)-HPLC (C_{18} column). Separation of the *C*-glycopeptides from any unreacted peptide and the coupling reagents present in the reaction mixture was achieved using preparative RP-HPLC.

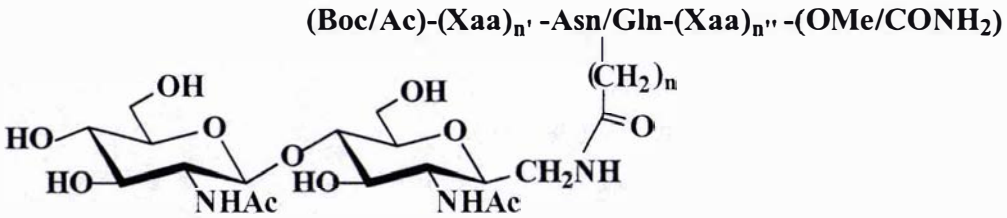
In order to drive the coupling reaction to completion a 1.5 molar excess of the *C*-glycosyl amine was used. Unfortunately, excess carbohydrate that did not react could not be recovered as it eluted in the DMF/DMSO fraction from the reverse phase column. Fortunately, unreacted peptide could be recovered from the column. Scheme 9 gives a schematic overview of the coupling reaction.



Scheme 9: Final coupling reaction in the synthesis of the *C*-glycopeptides after a modification of the methods by Wong *et al.* [118] and Anisfeld and Lansbury [117]. Xaa: unspecified amino acids

The reactions proceeded with average yields of 50 %. The coupling to the γ-carboxy group in glutamic acid gave a relatively poor yield of 15 %. This could be explained by the reduced electrophilic character of the γ-carbon atom due to the presence of the additional methylene group in comparison to aspartic acid. Yields have been calculated on the basis of the weight of isolated product. Calculation of the yields on the basis of the peak area of unreacted peptide using analytical HPLC suggested coupling rates of up to 90 %. One of the reasons for this discrepancy is almost certainly the loss of material that occurred during the HPLC-purification and subsequent lyophilisation.

Table 4 gives an overview of the yields of the coupling reaction between the di-chitobiosyl *C*-glycosyl amine and the side chain of aspartic or glutamic acid in the peptides shown.



Compound #	Peptide	% Yield
52	Boc-Asp-Ala-Ser-OMe	48
53	Boc-Asp-Ala-Thr-OMe	53
54	Ac-Tyr-Ile-Asp-Ala-Ser-NH_2	46
55	Ac-Tyr-Ile-Glu-Ala-Ser-NH_2	15

Table 4: Overall yields for the synthesis of the *C*-glycopeptides for which the general structure is shown above the table. The yields refer to the coupling between the di-chitobiosyl *C*-glycosyl amine 39 (section 3.2.1) and the respective peptide listed in the same row. Yields are calculated on purified material obtained after lyophilisation.

The *C*-glycopeptides produced were used in inhibition studies with recombinant PNGase F. The results of these trials will be discussed in section 4.4.1. For details on the synthesis of the *C*-glycopeptides see Materials and Methods.

3.3 Synthesis of α-Linked *N*-Glycopeptides and their β-Linked Analogues

In 1997 Fan and Lee [59] published a detailed study on substrate structure requirements of PNGases A and F, mainly using *N*,*N*-diacetylchitobiosyl based glycopeptides with varying peptide moieties. However, they only reported relative values to describe the activities of PNGase A and F for the different substrates. A K_M value was only reported in one case for PNGase A and a chitobiosyl based glycotripeptide.

In order to gain a better insight into the structural requirements for substrate binding of PNGase F the synthesis of two *N*,*N*-diacetylchitobiosyl based glycopeptides, one with a pentapeptide- and one with a tripeptide moiety, was envisaged. The use of these compounds in kinetic investigations for the determination of K_M and V_{max} with PNGase F is described in chapter 4.

In the first step, a β-configured *N*,*N*'-diacetylchitobiosylamine for the coupling to the side chain of aspartic acid in a pre-assembled tri-and a pentapeptide had to be synthesised.

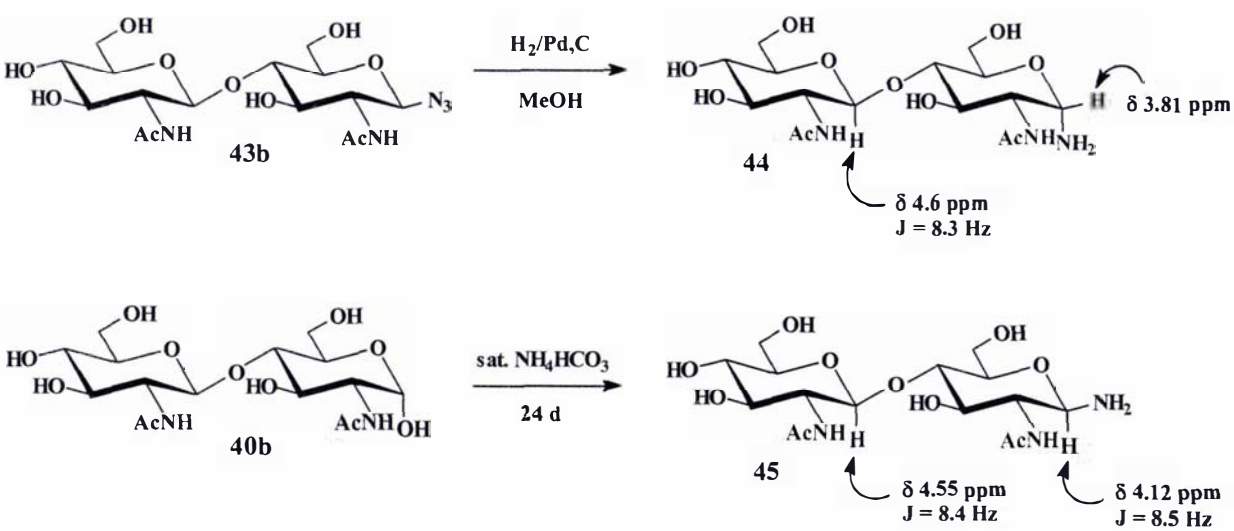
During the attempted synthesis of the β-configured *N*,*N*'-diacetylchitobiosylamine via hydrogenation of the deprotected β-di-chitobiosyl azide **43b** using 10 % Pd/C in methanol under an atmosphere of hydrogen, a single product (along with minor impurities) was obtained (scheme 10). TLC using ninhydrin stain indicated the presence of a primary amino function. Unfortunately the ^{1}H NMR spectrum (D_2O) did not allow for the determination of the anomeric configuration as only one doublet with a sufficient downfield shift and coupling constant characteristic for the anomeric proton of a β-linked glycopyranoside could be identified.

In order to prove the anomeric configuration of the glycosyl amine **44**, comparative NMR studies were carried out. For this purpose the desired β-di-chitobiosylamine was synthesised according to an alternative but time consuming procedure that is known to solely yield the β-amine [119],[120]. In this procedure the unprotected carbohydrate is dissolved in a saturated ammonium hydrogen carbonate solution for several weeks until completion of the transformation is shown by TLC. Excess ammonium hydrogen carbonate is then removed by repeated lyophilisation. Application of this method to α-di-chitobiose **40b** gave glycosyl amine **45** as a single product containing a primary amino group as judged by TLC and ninhydrin stain.

Comparison of the NMR-spectra of the glycosyl amines synthesised by these two different methods revealed a striking difference. As expected, the glycosyl amine **45** synthesised using a saturated ammonium hydrogen carbonate solution showed a β-configuration at C-1 (reducing end). In the ^{1}H NMR spectrum (D_2O) the anomeric proton on the reducing end gave rise to a doublet at δ 4.12 ppm with a coupling constant of 8.5 Hz. This is expected for β-configuration since H-1 and H-2 are in a

trans-diaxial relationship. The anomeric proton of the non-reducing GlcNAc gave rise to a doublet at δ 4.55 ppm with a coupling constant of 8.4 Hz, also consistent with a β-glycosidic linkage.

The ^{1}H NMR spectrum of the amine **44** resulting from hydrogenation of the β-azide did not contain the doublet at δ 4.12 ppm characteristic of the anomeric proton vicinal to the β-NH_2 group in compound **45**. Only a doublet at δ 4.6 ppm with a coupling constant of 8.3 Hz for the anomeric proton of the "internal" GlcNAc was identified in the spectrum of **44**. A broad signal at 3.81 ppm was tentatively assigned to H-1 at the reducing end of this compound. However, as it was impossible to extract a coupling constant the α-configuration could only be inferred by comparison to the ^{1}H NMR spectrum of glycosyl amine **45** (scheme 10).



Scheme 10: α- and β-configured di-chitobiosylamines. The chemical shifts and coupling constants of the anomeric protons in the products are shown for comparison.

Both glycosyl amines were coupled to the side chain of an aspartic acid residue in a tripeptide and a pentapeptide. The conditions used for amide bond formation were the same as those described in section 3.2.3. This gave rise to four different glycopeptides. Fortunately, the stereochemistry of the glycosidic linkage could now be determined in each case (figure 14 and 15).

A comparison of the ^{1}H-NMR spectra of the compounds **56** and **57** bearing a tripeptide moiety gave a clear indication of the differences in the anomeric

configuration (figure 14). Both compounds were soluble in deuterated water. Therefore the amide-proton was exchanged and the signal for the anomeric proton on the reducing end of the disaccharide could be resolved as a doublet. The anomeric configuration could then be determined from its coupling constant.

The ^{1}H NMR spectrum of the β-linked glycotripeptide **56** showed two doublets corresponding to the two anomeric protons: one doublet at δ 5.04 ppm with a coupling constant of 9.5 Hz corresponding to the reducing end and one doublet at δ 4.57 ppm with a coupling constant of 8.4 Hz corresponding to the "internal" anomeric proton.

The ^{1}H NMR spectrum of the α-linked glycotripeptide **57** showed an analogous signal at δ 4.60 ppm for the "internal" anomeric proton as in the case of the precursor glycosyl amine **44**. H-1 at the reducing end gave rise to a signal without much definition at δ 3.80 ppm.

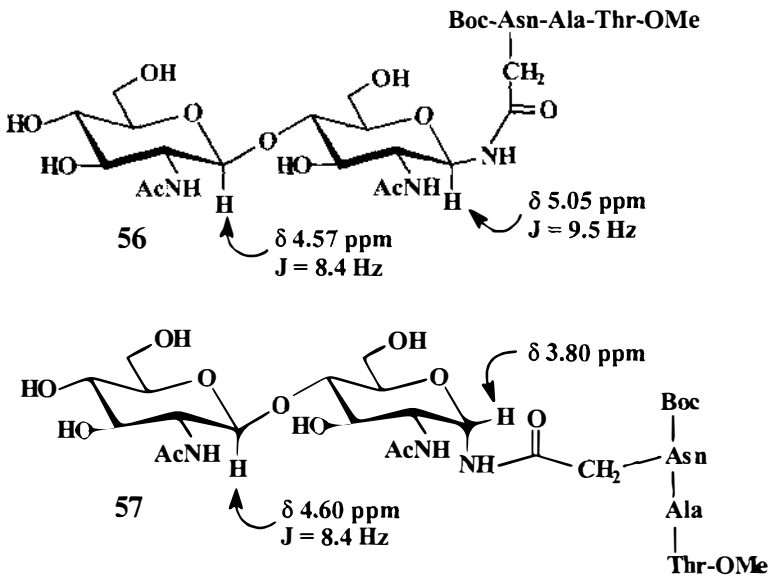
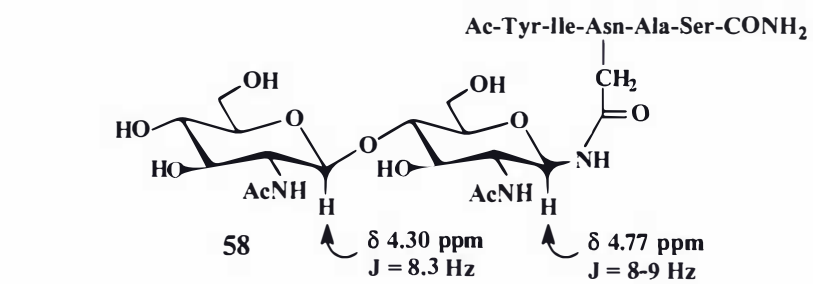


Figure 14: Chemical shifts and coupling constants of the anomeric protons measured in D_2O in the β- and α-linked glycotripeptides 56 and 57.

A comparison of the ^{1}H NMR spectra of the compounds **58** and **59** (figure 15) bearing a pentapeptide moiety proved more difficult. The compounds showed poor solubility in water so the NMR spectra had to be recorded in DMSO-d_6. As a consequence of this the amide proton in the *N*-glycosidic linkage does not exchange and the signal of the anomeric proton does not appear as a simple doublet. Furthermore, the spectra were more complex due to the appearance of proton signals from the hydroxyl groups which also do not exchange in deuterated DMSO.

For the β-linked glycopeptide **58** the anomeric protons were identified as two doublets. The "internal" proton appeared as a doublet at δ 4.30 ppm with a coupling constant of 8.3 Hz. The anomeric proton on the reducing end overlapped partially with a hydroxyl group but could clearly be identified with a chemical shift of δ 4.77 ppm using ^{1}H-^{1}H and ^{1}H-^{13}C-cosy experiments. The splitting pattern indicated a triplet for this anomeric proton signal and an approximate coupling constant of 8.5 Hz, indicating β-configuration.

As expected, the α-linked glycopeptide **59** did not give rise to the latter signal. A doublet with a chemical shift of δ 4.44 ppm and a coupling constant of 8.4 Hz could be identified for the "internal" anomeric proton. The signal for the anomeric proton at the reducing end was located within a multiplet at ~3.5 ppm, so unfortunately no coupling constants could be extracted.



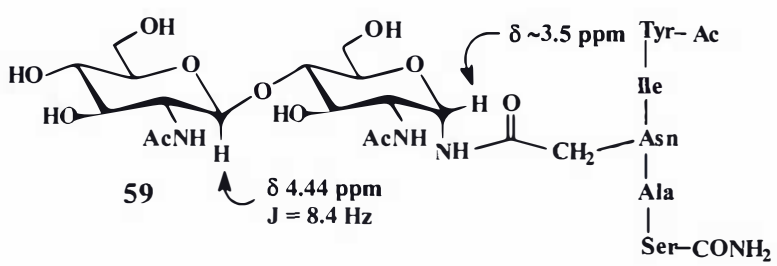


Figure 15: Chemical shifts and coupling constants of the anomeric protons measured in DMSO-d_6 in the β- and α-linked glycopentapeptides 58 and 59.

All four glycopeptides were fully characterised using ^{1}H and ^{13}C NMR spectroscopy. Direct comparison of the pairs of glycopeptides differing only in the configuration of the *N*-glycosidic linkage showed near complete agreement except for the differences already highlighted. Electrospray mass spectrometry was also used to determine the

overall mass of the compounds, which was shown to be the same for the pairs of anomers.

In conclusion, a method has been developed for the synthesis of α-linked *N*-glycopeptides. As in the synthesis of the *C*-glycopeptides, the glycosyl amine and the peptide building block have been prepared in separate steps followed by a coupling reaction to yield the glycopeptides. The yields for this step varied from 25-40 % for the purified material.

In the literature the synthesis of an α-linked *N*-glycopeptide had been reported for the total synthesis of "nephritogenoside" [121]. In this case the reduction of an α-D-glycosyl azide was achieved by hydrogenation over Pd/C. After coupling to Alloc-AspO^tBu an anomeric mixture of the resulting glycosyl asparagine was isolated.

Nakabayashi *et al.* [122] and Ogawa *et al.* [123] reported that α-D-glycosyl azides gave rise to anomeric mixtures upon hydrogenation/reduction over Lindlar catalyst whereas β-azides solely yielded the β-glycosyl amines.

It is noteworthy that many methods have been published for the stereospecific hydrogenation of β-glycopyranosyl azides in the synthesis of *N*-linked glycans [122], [124-126]. Nevertheless, it is a common fact that glycosyl amines are prone to anomerisation. The use of propane-1,3-dithiol [127],[128] in the presence of triethylamine seems to offer a general solution for the reduction of glycopyranosylazides to the corresponding glycosyl amines with retention of configuration.

Finally, the question arose whether the two α-linked *N*-glycopeptides could be useful as tools in the kinetic or structural studies with PNGase F.

In this context it has to be noted that generally in *N*-linked glycans the reducing end of the sugar in the conserved *N*-glycosidic linkage region has the β-configuration. However, as mentioned previously, one exception has been reported. "Nephritogenoside" is a glycopeptide in which the glycan consisting of three glucose residues is α-*N*-glycosidically linked to the side chain of asparagine in a peptide of 21 amino acids [9],[10].

It was therefore of great interest to investigate whether PNGase F would be capable of cleaving α-linked *N*-glycopeptides. The α-linked *N*-glycopeptides described have therefore been used in kinetic studies using recombinant PNGase F (section 4.4.2). The use of the β-linked *N*-glycopeptides in kinetic investigations with PNGase F is described in section 4.3.

3.4 Preparation of an *N*-Linked Glycopeptide with a Modified Core-Region

Most *N*-linked glycopeptides share a common pentasaccharide core structure consisting of three mannose (Man) and two *N*-acetylglucosamine (GlcNAc) residues as shown in figure 16. Additional sugar residues can be attached to the core, for example a fucose (Fuc) residue α (1→6) or α (1→3)-linked to the innermost GlcNAc is a common embellishment in many glycopeptides from plants and insects.

Detailed studies on substrate requirements for PNGases, as described in chapter one, have shown the importance of the core-region for substrate recognition. PNGase F is not capable of cleaving glycopeptides/proteins having a Fuc residue α (1→3)- linked to the innermost GlcNAc of the core region.

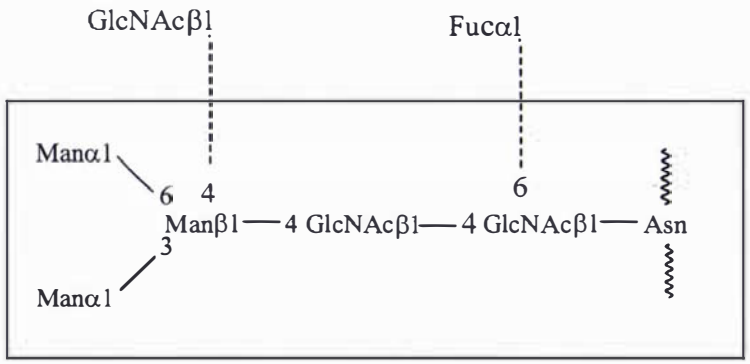


Figure 16: Common core structure of *N*-linked glycoproteins/peptides. Further sugar residues can be linked to the core as indicated by the dotted lines.

Of particular interest in this context is the importance of the β (1→4)-linked Man residue. Altmann and co-workers [56] have shown that the removal of this Man residue reduces the affinity of PNGase A for a fibrin glycopeptide by a factor of three. Unfortunately, corresponding data has not been reported for PNGase F. Although a di-chitobiosyl core has been shown to be the minimal structural requirement for

substrate recognition by PNGase F it can be assumed that the β (1→4)-linked Man residue also plays an important role. This is further supported by the fact that for L-929 PNGase from mouse fibroblast cells, the trisaccharide core Manβ1→4GlcNAcβ1→4GlcNAcβ1→peptide was shown to be the minimal structural requirement for enzymatic activity [40]. For this reason, the use of a glycopeptide substrate with a trisaccharide core in which the β (1→4)-linked Man residue has been replaced by a different sugar could provide valuable insights into the mechanism of substrate binding.

From a synthetic point of view the β-mannosidic linkage is a very difficult construct. The α-anomer is the favoured product under most glycosylation conditions due to the thermodynamic anomeric effect. The difficulty in forming a β-mannoside is reflected by the fact that the direct synthesis of such a linkage was not achieved until 1981 when Paulsen and Lockhoff reported the synthesis of a tri- and a tetraglycoside incorporating a β-mannosidic linkage [129]. Nowadays, more sophisticated methods are available [130].

Fan and Lee [59] showed that PNGase F is not capable of acting on glycopeptides bearing a lactosyl or cellobiosyl instead of a di-chitobiosyl core indicating that the *N*-acetyl group is essential for the binding of the substrate. However, the importance of the Man linked to the proximal GlcNAc has not been established. Kinetic studies using a substrate comprised of a pentapeptide linked to GlcNAcβ1→4GlcNAcβ1→4GlcNAc would give an indication of how tolerant the active site is to the replacement of Man in the core-region. *N*-acetylglucosamine is not only a C-2 epimer of mannose but the presence of the *N*-acetyl group imposes an additional structural feature to the active site with respect to binding. The synthesis of such a glycopeptide could be easily accomplished as the required chitotriose was obtained as a side product during the acetolysis of chitin as described earlier (section 3.2.1).

To this end the β-chitotriosylamine **46** was obtained by reacting chitotriose with a saturated ammonium hydrogen carbonate solution for twelve days. The ^{1}H NMR spectrum (D_2O) showed a doublet with a chemical shift of δ 4.11 ppm and a coupling constant of 8.6 Hz for the anomeric proton on the reducing end clearly indicating β-configuration (figure 17).

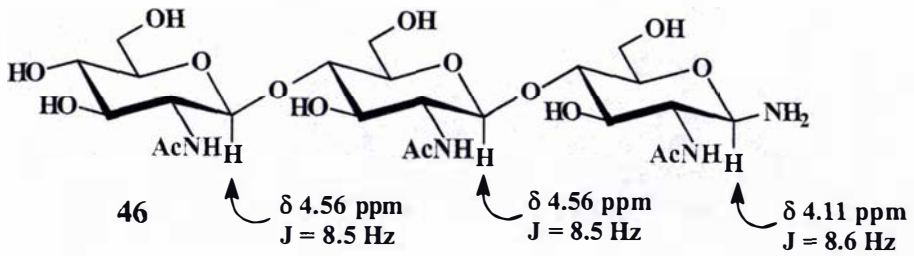


Figure 17: Chemical shifts and coupling constants of the anomeric protons in the shown chitotriosylamine building block (recorded in D_2O).

The glycosyl amine building block was then coupled to the side chain of aspartic acid in the pentapeptide Tyr-Ile-Asp-Ala-Ser using the procedure already described to produce the desired β-linked glycopeptide **60** in 41 % yield after HPLC-purification (figure 18).

Due to poor solubility in water the ^{1}H NMR spectrum had to be recorded in DMSO-d_6. The anomeric proton of the GlcNAc involved in the glycopeptide linkage was resolved as a triplet with a chemical shift of δ 4.76 ppm and a coupling constant of 9.3 Hz, indicating a β-configuration. Figure 18 shows the synthesised glycopeptide along with chemical shifts and coupling constants of the anomeric protons.

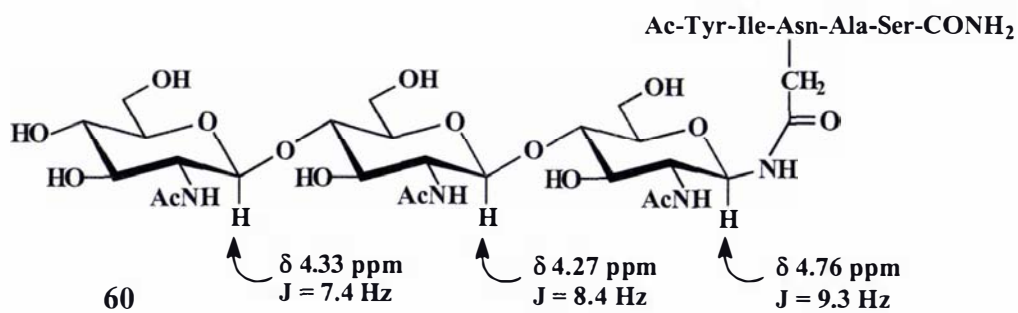


Figure 18: Structure of the glycopeptide mimetic 60 bearing a β (1-4)-linked GlcNAc instead of a mannosyl residue. The chemical shifts and coupling constants of the anomeric protons are derived from the ^{1}H NMR spectrum recorded in DMSO-d_6.

In summary, a straightforward, two-step procedure for the synthesis of a glycopeptide mimetic in which Man has been replaced by a GlcNAc residue has been developed starting from chitotriose .

This new glycopeptide mimetic has been used in kinetic studies using recombinant PNGase F (section 4.3).

CHAPTER 4

Kinetic Investigations using PNGase F

4.1 Introduction

A partial aim of this study is the investigation of substrate analogues and potential inhibitors of PNGase F to gain a deeper insight into the structural requirements for substrate recognition by this enzyme and define possible candidates for use as affinity ligands. For this purpose the synthetic *N*-linked glycopeptides and glycopeptide mimetics described in chapter 3 were used in both kinetic and structural studies with recombinant PNGase F. The structural studies are presented and discussed seperately in chapter 6.

The work by Altmann and co-workers from 1995 [56] was the first detailed kinetic characterisation of PNGases and used some of the most well characterised glycans such as bovine fetuin (complex-type), hen ovalbumin (hybrid- and oligomannose type) and bovine fibrin (complex-type), as well as other glycopeptides. This work reported not only K_M values for the activity of PNGases A and F with many of these substrates but also described the minimal structural requirements for the substrates, the role of the carbohydrate and the peptide moiety, and the influence of chromophores introduced into the substrates on the activity observed.

Although the work by Altmann and co-workers as well as several other publications (discussed in section 1.6) report an *N*,*N*'-diacetylchitobiosyl residue to be the minimal structural requirement for the carbohydrate moiety no sufficient kinetic data, for example K_M, was published for such a substrate and PNGase F. This was of particular interest as the carbohydrate moiety of the potential *C*-glycopeptide inhibitors described in this thesis is based on this particular minimal structural requirement.

Kinetic investigations in conjunction with structural studies can provide important data that facilitates the design of inhibitors. The combination of both analyses can lead to the complete understanding of the catalytic mechanism of an enzyme.

In the following sections presenting the kinetic results K_M is generally referred to as a measure of affinity for a substrate in order to facilitate the discussion. However, the above assumption only applies if the contribution of k_2 to K_M is ignored which may not always be the case.

The data and experimental setup for the results presented in this chapter are shown in separate "Experimental Data Sheets" in Appendix I.

4.2 Fluorescently Labelled Ovalbumin-Glycopeptide

As outlined in section 1.4, the availability of a sensitive assay procedure is a prerequisite to investigate PNGase activity. Most assays reported in the literature are HPLC based discontinuous assays that measure either released carbohydrate or peptide moieties. The problem with such assays are a lack of both, sensitivity (due to the concentrations of substrate required) and reproducibility (arising from the inherent nature of the method). Sensitivity can be improved by labelling either the peptide or the glycan moiety with either a fluorophore or a radioactive moiety (section 1.4). For this work it was decided to label the peptide moiety of the glycopeptide (11mer) isolated from hen-ovalbumin, which was used as the "standard" substrate for PNGase F in our laboratory, with fluoresceine isothiocyanate as described in section 2.3. The labelled glycopeptide was stable under the conditions of the HPLC-based assay (section 8.2), gave a linear response, offered increased sensitivity and the spectra recorded were free of background noise providing baseline separation for the substrate and the hydrolysis product.

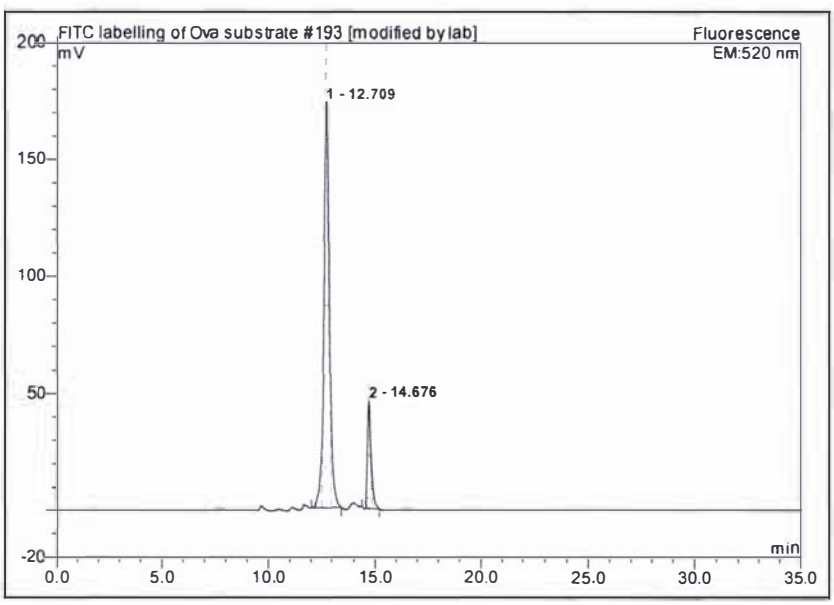


Figure 19: HPLC-chromatogram showing FITC-labelled Ova-substrate (peak 1) and the product of the enzymatic digest (peak 2). Experiments were conducted under conditions that gave approx. 10 % digest. Retention times are shown in minutes.

Figure 19 shows a typical example of a HPLC-chromatogram from the digestion of the labelled glycopeptide with recombinant PNGase F. The experiments were conducted under conditions that gave approximately 10 % hydrolysis or less in order to measure the initial velocities for each substrate concentration [131].

Before the labelled substrate was used, the characteristic kinetic values K_M and V_{max} were determined using recombinant PNGase F. This was important in order to evaluate the affinity of the enzyme for this new substrate and to compare the values obtained to the published data. The results are shown in figure 20 (for the data see **Appendix, Experimental Data Sheet #1**).

Parameter	Value	Std.Error
V_{max} ($\mu g*min^{-1}$)	6.85	0.30
K_M ($mg*mL^{-1}$)	0.13	0.02

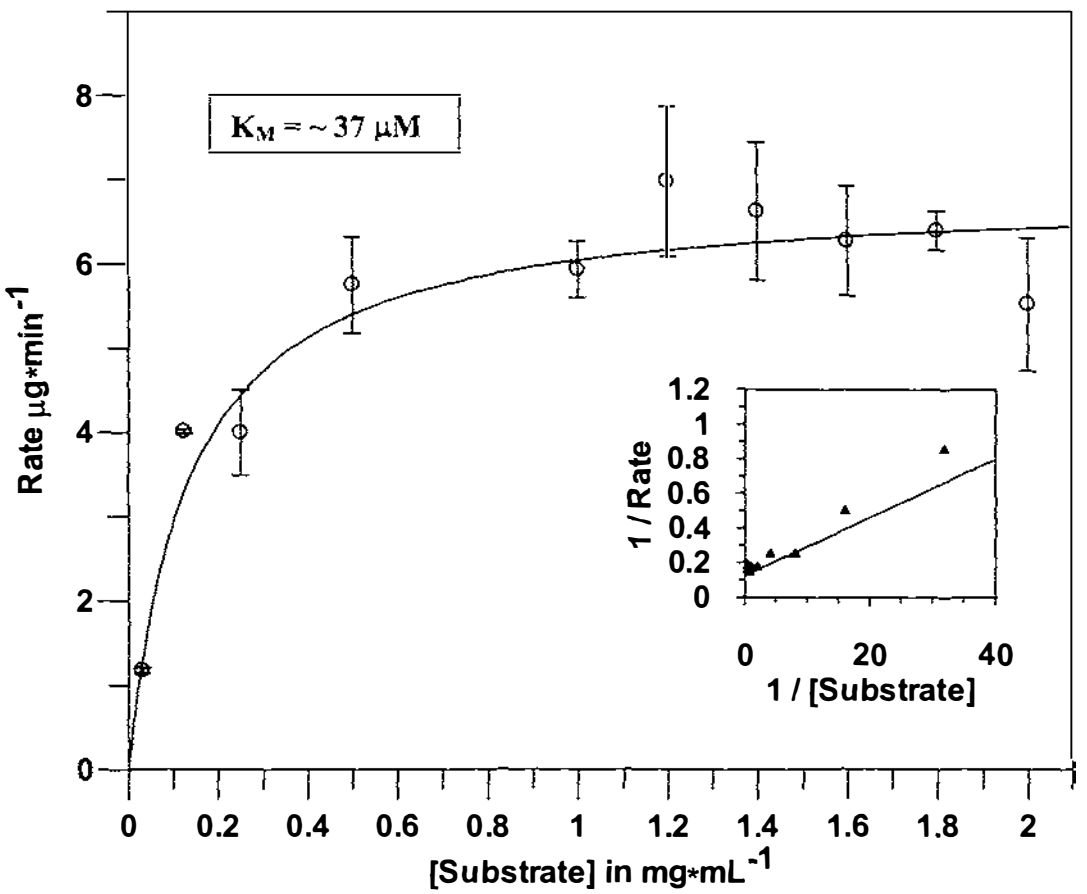


Figure 20: A plot of the rate of hydrolysis (y-axis) as a function of the substrate concentration (x-axis) using recombinant PNGase F and fluorescently labelled glycopeptide (section 2.3). The insert shows the Lineweaver-Burke plot of the data. K_M and V_{max} were calculated using GraFit Version 4.0 (Erithacus Software). Each data point represents the average of 2-3 independent experiments. Error bars correspond to the standard deviation for each determination.

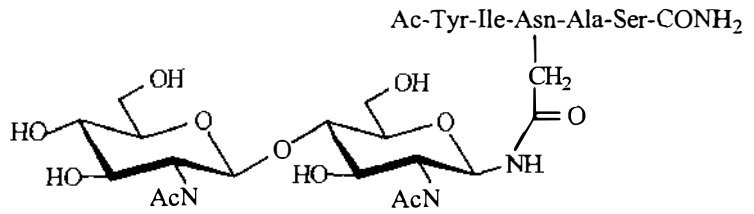
The K_M was calculated in units of mg/mL because the ovalbumin glycopeptide has been shown to be a mixture of different hybrid- and oligomannose glycans making the calculation of an exact mass impossible [132]. Using an approximate mass of 3.6 kD for the dilabelled substrate, the K_M is approximately 37 μM which is well in accordance with published results. Altmann *et al.* [56] reported a K_M of 31 μM for a resorufin labelled ovalbumin glycopeptide (10mer, obtained from pepsin digest) using commercial PNGase F (Boehringer-Mannheim). The authors reported that the presence and nature of a chromogenic label had a significant influence on PNGase activity and suggested that this could be due to conformational changes in the peptides induced by the chromophores.

As a summary, the kinetic investigations using the FITC-labelled glycopeptide showed that it was suitable to be used as a substrate for the inhibition studies undertaken as part of this work.

4.3 Synthetic *N*-Glycopeptides

During the course of this work several *N*-linked glycopeptides were synthesised as outlined in chapter 3. This section will describe the use of these compounds in kinetic studies with recombinant PNGase F and focus upon the interpretation of these results to further understand the requirements for substrate binding and recognition. The compounds were also used in inhibition studies with their *C*-glycopeptide relations as described in section 4.4.1.

Substrate recognition by PNGase F is dependent on both the glycan and peptide moiety of a glycopeptide or glycoprotein and for the sugar moiety it has been shown that an *N,N'*-diacetylchitobiosyl core is the minimal structural requirement [59],[56]. For the peptide moiety it has been shown that a pentamer containing the glycosylation motif Asn-X-Ser/Thr in which the Asn must be in peptide linkage on both sides is required for optimal activity [56],[59]. To date no sufficient kinetic data has been reported for a substrate fulfilling these minimal structural requirements using PNGase F. For this reason the glycopentapeptide **58** shown in figure 21 was synthesised and its kinetic data determined using recombinant PNGase F. This compound was also used as a substrate for inhibition experiments.



58

Figure 21: *N*-Linked glycopeptide 58 showing structural requirements for optimal activity using PNGase F according to published data (see text).

The peptide sequence of the above glycopeptide was chosen as it allowed a direct comparison with a published K_M value using native PNGase F and a high mannose glycopeptide with the same peptide sequence [5]. Catalytic rates were determined under conditions that gave approximately 10 % digest using a discontinuous HPLC-based assay (section 8.2). The results are shown in figure 22 (for the data see **Appendix, Experimental Data Sheet #2**).

A K_M of 2.24 mM shows that although the enzyme has only a rather moderate affinity for the compound depicted in figure 21, it will remove a dichitobiose from this glycopeptide. The relatively large error values were caused by the poor solubilty of this substrate. As a consequence of this there are sufficient data points at the high end of the concentration scale introducing a significant error into the non-linear fit. However, the derived values of V_{max} and K_M still give a good indication of the affinity of PNGase F for this compound.

A comparison of this K_M value to that of 80 μM for a high mannose glycopeptide with the same peptide structure [5] shows that the role of the carbohydrate moiety in the substrate recognition/binding by PNGase F has apparently been underestimated in the current literature, especially in the report by Fan and Lee [59]. The active site of the enzyme shows a 28 fold higher affinity for the high-mannose substrate in comparison to the substrate with a minimal *N,N'*-diacetylchitobiosyl core. It can therefore be concluded that the length of the carbohydrate chain does significantly affect the affinity of the enzyme for the substrate and that the role it plays in substrate binding and recognition needs to be established at a molecular level.

Parameter	Value	Std.Error
V_{max} (pmol*min^{-1})	28.8	5.7
K_M (mg*mL^{-1})	2.3	0.5

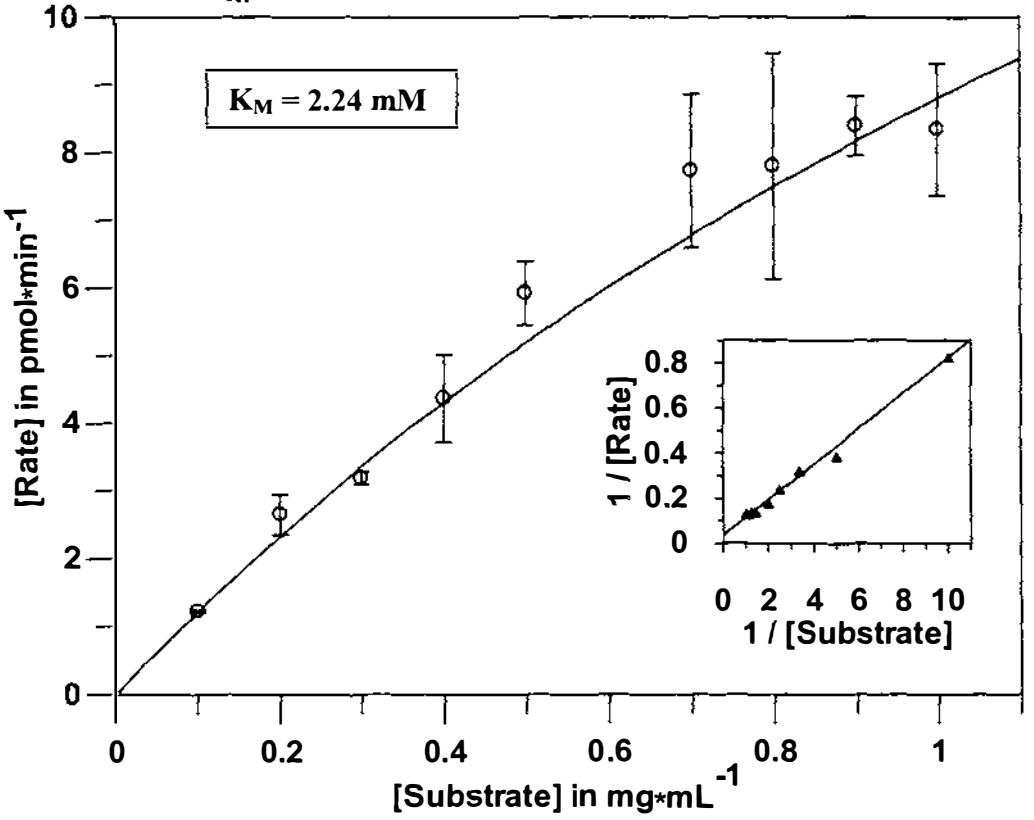
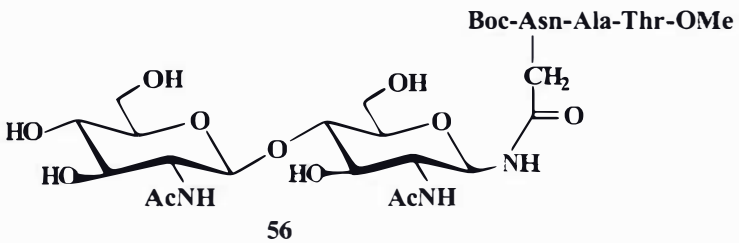


Figure 22: A plot of the rate of hydrolysis (y-axis) as a function of the substrate concentration (x-axis) using recombinant PNGase F and glycopeptide 58 shown in figure 21. The insert shows the Lineweaver-Burke plot of the data. K_M and V_{max} were calculated using GraFit Version 4.0 (Erithacus Software). Each data point represents the average of 2-5 independent experiments. Error bars correspond to the standard deviation for each determination.

The role of the peptide moiety was also of interest. The most comprehensive investigation into the role of the peptide was carried out by Fan and Lee [59] who tested a wide variety of di-, tri- and pentapeptides *N*-glycosidically linked to *N*,*N*'-diacetylchitobiose as substrates for PNGases A and F. Their investigations showed that PNGase F requires at least a tripeptide for any significant activity and that glycotripeptides in which the carbohydrate chain is located at the *C*-terminus of the peptide moiety are poorer substrates than those in which it is located at the *N*-terminus.

In light of the detailed investigations by Fan and Lee it was decided to synthesise glycotripeptide **56** with an *N,N'*-diacetylchitobiosyl core (figure 23) and compare the activity of PNGase F towards this substrate with that of the glycopentapeptide **58** shown in figure 21 and with the glycotripeptide of the same structure as **56** described by Fan and Lee [59].



56

Figure 23: *N*-Linked glycotripeptide 56 with the glycosylated Asn only in peptide linkage at the *C*-terminus.

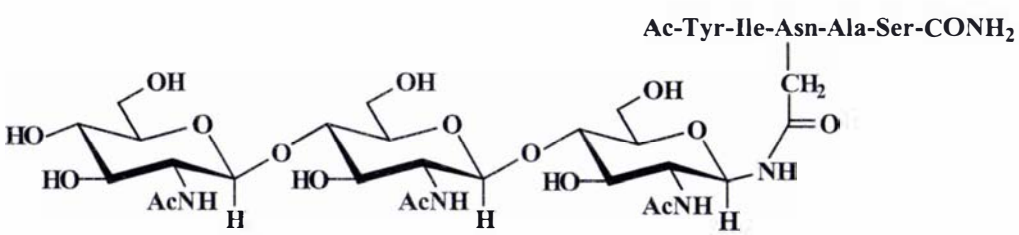
Kinetic investigations showed that the glycotripeptide **56** depicted in figure 23 is a very poor substrate for PNGase F. The activity was so low that the determination of K_M and V_{max} was impossible. This finding is in contradiction to the results of Fan and Lee who reported a much higher activity for PNGase F with this substrate compared with that of glycopentapeptide **58** shown in figure 21.

Fan and Lee compared the total percentage of hydrolysis within a fixed time frame using a constant amount of enzyme and substrate. According to this procedure PNGase F hydrolysed 98 % of the glycopentapeptide **58** and 83 % of the glycotripeptide **56**. Using the same experimental conditions 100 % hydrolysis of the glycopentapeptide **58** was observed and only 16 % hydrolysis of the glycotripeptide **56**. A large number of factors may have caused these differences ranging from the assay procedure to the enzyme used.

The important conclusion to be drawn from this result is that PNGase F shows a significantly reduced activity for a glycotripeptide where the glycosylated Asn occurs only with a peptide linkage on one side in comparison to a glycopentapeptide in which the Asn has peptide linkages on both sides.

The role of the structural elements of the carbohydrate moiety was also investigated. According to Altmann and co-workers [56] there is no precedence for any substantial influence of structural features of the glycan's non-reducing terminus. They conclude that both PNGase A and F show no preference for glycopeptides of either the complex or oligomannose type. The only structural feature that had a significant influence was fucosylation at the 3-position of the innermost GlcNAc, a structural feature of many plant and insect glycoproteins. This abolishes the activity of PNGase F [58].

Of particular interest in this context is the importance of the beta 1→4 linked Man residue, the third residue of the pentasaccharide core. For this reason glycopeptide **60** (figure 24) with a trisaccharide core in which the beta 1→4 linked Man residue is replaced by a GlcNAc residue was synthesised as described in section 3.4.



60

Figure 24: *N*-Linked glycopeptide 60 in which the beta 1→4 linked Man residue has been replaced by a GlcNAc (see text).

The glycopeptide **60** has the same peptide sequence as glycopeptide **58** (figure 21) and the high mannose glycopeptide reported by Fan and Lee [5], allowing the effect of changes in the carbohydrate moiety alone to be assessed. The kinetic investigation was carried out using the discontinuous assay described in section 8.2 and the results are shown in figure 25 (for the data see **Appendix, Experimental Data Sheet #3**). Again, similar solubility problems as described for glycopeptide **58** were encountered. However, when compared with glycopeptide **58** the solubility was slightly better due to the presence of the additional hydroxyl groups on the third GlcNAc residue.

Parameter	Value	Std.Error
V_{max} (pmol*min^{-1})	722.0	28.2
K_M (mg*mL^{-1})	0.43	0.04

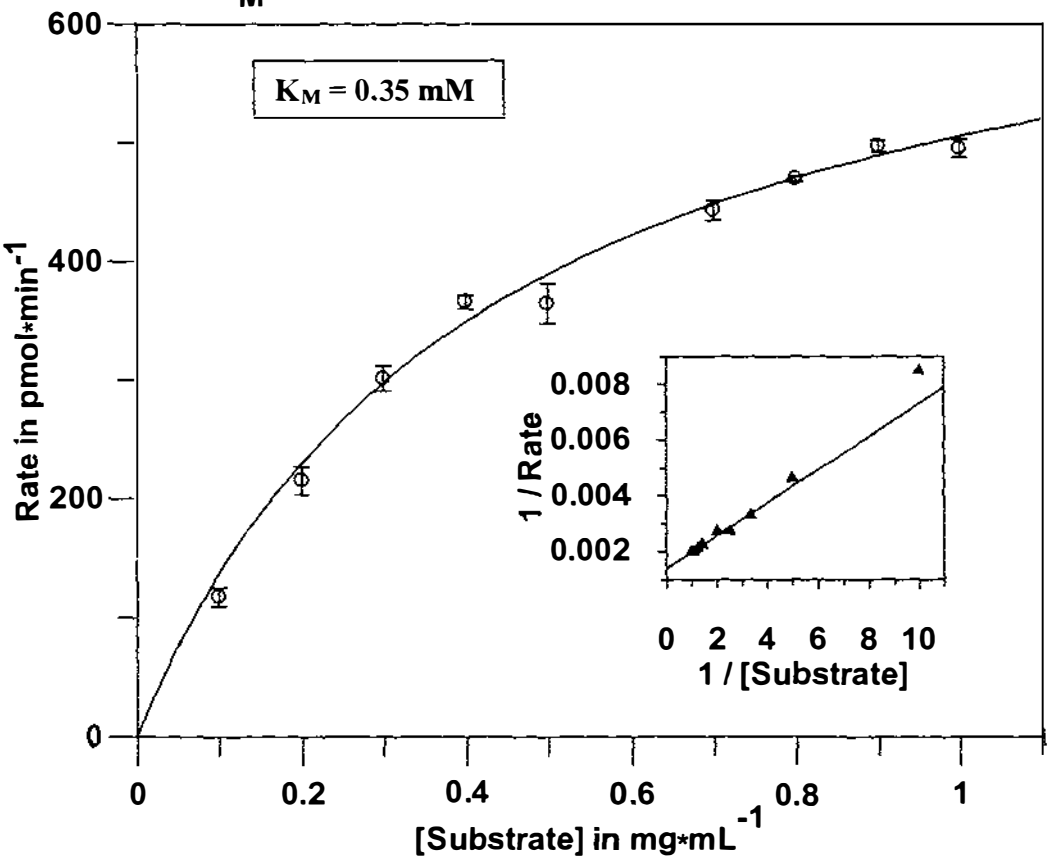


Figure 25: A plot of the rate of hydrolysis (y-axis) as a function of the substrate concentration (x-axis) using recombinant PNGase F and glycopeptide 60 shown in figure 24. The insert shows the Lineweaver-Burke plot of the data. K_M and V_{max} were calculated using GraFit Version 4.0 (Erithacus Software). Each data point represents the average of 2 independent experiments. Error bars correspond to the standard deviation for each determination.

A comparison of the K_M values obtained for the two substrates, 0.35 mM for the trichitobiose glycopeptide **60** and 2.3 mM for the diacetylchitobiosyl glycopeptide **58**, showed a ~ 6-fold increase for **60** relative to **58**, indicating a greater affinity of the recombinant PNGase F for the glycopeptide with the bigger glycan.

This is a very unexpected result as the beta mannosidic linkage in the core region of *N*-linked glycans is a very distinct and unique structural feature and it would not have been surprising if it had been important for substrate recognition. The fact that these

results show quite the opposite indicates that the length of the carbohydrate chain is more important than any other structural feature, which is in accordance with the results of Altmann and co-workers [56].

It is assumed that the reason for this observed increase in activity is an increase in the number of non-covalent interactions between the substrate and the enzyme. However, this can only be unequivocally established by the three dimensional structure of a PNGase molecule complexed with a glycopeptide mimetic/inhibitor.

4.4 Inhibition Trials with *N*-Linked Glycopeptide Mimetics

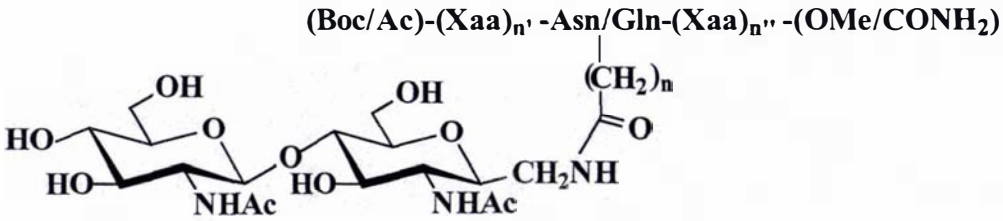
The synthesis of *N*-linked glycopeptide mimetics for the use in structural and kinetic studies has been described in chapter 3. This section will focus on the use of these compounds in inhibition trials with PNGase F.

As described in section 1.7, three different glycopeptide mimetics have been reported in the literature that successfully inhibit the activity of PNGase F and other PNGases. The drawback associated with all of these compounds is the lengthy and tedious procedure required for their synthesis. This study therefore aimed to develop a straightforward procedure for the synthesis of "simple" structures as potential inhibitors of PNGase F.

It was hoped that such compounds would be useful tools for gaining a deeper insight into the structural requirements for substrate recognition and the catalytic mechanism of this class of enzymes using techniques such as co-crystallisation and x-ray diffraction. Inhibitors could also find an application in the synthesis of affinity resins and therefore facilitate the purification of PNGases from a variety of sources.

4.4.1 *C*-Glycopeptides

The approach towards the straightforward synthesis of a variety of *C*-glycopeptides was based on the general concept of *C*-glycopeptide inhibition as published by Wang *et al.* in 1997 [5] and is described in section 3.2. Table 5 gives an overview of the different *C*-glycopeptides synthesised during the course of this study.



C-glycopeptide	(*N*-terminus) Sequence (*C*-terminus)
52	Boc-**Asn**-Ala-Ser-Ome
53	Boc-**Asn**-Ala-Thr-Ome
54	Ac-Tyr-Ile-**Asn**-Ala-Ser-NH_2
55	Ac-Tyr-Ile-**Gln**-Ala-Ser-NH_2

Table 5: The different *C*-glycopeptides synthesised for use in the kinetic and structural studies. The *C*-glycosyl moiety is either attached to the side chain of an asparagine or a glutamine in the tri- and pentapeptides.

The size of the carbohydrate moiety was based on the minimal structural requirements for substrate recognition as described by Fan and Lee [59], Altmann and co-workers [56], as well as other authors (sections 1.6 and 3.2). The amino acid sequence and length of the peptide moiety was solely based on the investigations by Fan and Lee [59] who assessed the activity of PNGase A and F on *N*,*N*'-diacetylchitobiosyl based glycopeptides in which only the peptide moiety was varied in order to assess the impact of the peptide on substrate recognition.

In this context it is necessary to briefly outline the rational basis for the *C*-glycopeptides chosen as potential inhibitors.

C-Glycopeptide **54** was chosen on the grounds that the corresponding glycopeptide was reported to display all of the structural features required for optimal activity with PNGase F. Therefore the *C*-glycopeptide could be expected to bind to the active site of the enzyme. Furthermore the high-mannose *C*-glycopeptide reported by Wang *et al.* [5] had an identical peptide sequence, but was not acetylated at the *N*-terminus. This allowed for the potential comparison of the inhibitory activity of these two compounds with respect to the role of the carbohydrate moiety.

The corresponding glutamine bound *C*-glycopeptide **55** was chosen because investigations showed that PNGase F was virtually inactive on the corresponding glycopeptide [59]. Furthermore Haneda and Tanabe [7] reported a glutamine bound glycopeptide that inhibited the activity of both, PNGase A and F (section 1.7.1). Compound **55** incorporates two structural features that could potentially inhibit PNGase activity. The first of these is the additional methylene group in the *C*-glycoside, the second is the additional methylene group in the side chain of glutamine in comparison to asparagine.

The *C*-glycotripeptides **52** and **53** were chosen because the corresponding glycopeptides have different activities towards PNGase A and PNGase F [59]. According to the data presented, an *N*,*N*'-diacetylchitobiosyl based substrate with the peptide sequence Asn-Ala-Thr was reported to be a good substrate for PNGase F whereas the enzyme showed a much lower activity when the peptide sequence was changed to Asn-Ala-Ser. In order to show whether threonine promoted tighter binding to the enzyme compared to serine, the corresponding *C*-glycopeptides were tested for their inhibitory effects.

Based on the current knowledge of the minimal structural requirements for substrate recognition by PNGase F and the reported inhibition of this enzyme by a *C*-glycopeptide (section 1.7.1), it was expected that an asparagine-bound *C*-glycopeptide of the general structure shown in table 5 would be recognised by the active site of PNGase F but not be hydrolysed. Such a compound would therefore have the potential to inhibit the hydrolysis of the corresponding *N*-linked glycopeptide.

Based on a report of the inhibitory effect of glutamine-bound glycopeptides by Haneda and Tanabe [7] similar results were expected for the glutamine-bound *C*-glycopeptides.

In an initial study the potential of the *C*-glycopeptides as inhibitors for PNGase F in the hydrolysis of the fluorescently labelled ovalbumin glycopeptide (FITC-ova) was investigated.

In a first set of experiments *C*-glycopeptides **54** and **55**, both bearing pentapeptide moieties, were tested as inhibitors. For this purpose the fluorescently labelled ovalbumin glycopeptide with a concentration of $2.63 * 10^{-4}$ mmol/mL and a molar

excess (~1.6 fold for **55** and ~2.5 fold for **54**) of one of the *C*-glycopeptides was incubated with recombinant PNGase F for a fixed time. The total rate of hydrolysis was determined and compared to the hydrolysis of the FITC-ova substrate in the absence of the *C*-glycopeptides. Only one set of data was collected as limited PNGase F was available and the intent of these experiments was only to identify compounds that needed more thorough testing. However, the initial results indicated that the two compounds showed no inhibitory effect whatsoever (for the data see **Appendix, Experimental Data Sheet #4)**.

A second set of experiments was carried out to indicate whether the lack of inhibition was the result of slow binding between the inhibitor and the enzyme. Therefore solutions of the enzyme were incubated with saturated solutions of the various *C*-glycopeptides for 24 h, then used to digest the fluorescently labelled ovalbumin glycopeptide (concentration: $1.01 * 10^{-4}$ mmol/mL). This time, all the *C*-glycopeptides shown in table 5 were tested. Although again only one set of data was collected, there was no indication of any inhibitory activity (for the data see **Appendix, Experimental Data Sheet #5**).

It was thus concluded that the *C*-glycopeptides had a much reduced affinity for the binding site of PNGase F compared to the ovalbumin-derived substrate and that this was probably due to the reduction in the size of the carbohydrate moiety of the *C*-glycopeptides. It could also have been possible, however, that the different arrangement of atoms at the scissile bond had an effect on the observed affinity. Therefore no further experiments were conducted using the fluorescently labelled ovalbumin glycopeptide as the substrate.

Although the synthesised *C*-glycopeptides failed to inhibit the activity of the labelled ovalbumin derived glycopeptide, it was possible that they might inhibit the hydrolysis of substrates with a lower affinity for the enzyme, substrates such as the corresponding *N*-linked glycopeptides. For this purpose the *N*-linked glycopeptides **56** and **58,** as shown in figure 26, were used as substrates. Their synthesis is outlined in section 3.3.

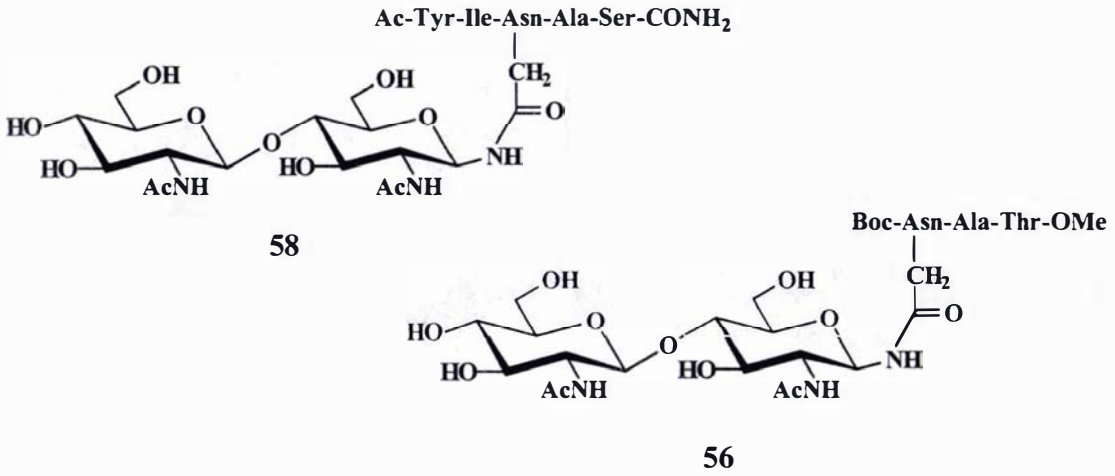


Figure 26: The corresponding *N*-linked glycopeptides 56 and 58 of the respective *C*-glycopeptides 53 and 54 as shown in table 5.

Before the two *N*-linked glycopeptides could be used as substrates in inhibition studies their ability to act as substrates had to be assessed. The results are described and discussed in section 4.3. Glycopentapeptide **58** was hydrolysed by PNGase F (K_M of 2.3 mM), albeit more slowly than the FITC-labelled ovalbumin glycopeptide (K_M of ~37 μM), while the glycotripeptide **56** was hardly processed at all. The latter result was inconsistent with the findings of Fan and Lee [59].

On the basis of these results further tests with the *C*-glycopeptides bearing a tripeptide moiety and in which the glycosylated asparagine or glutamine occurs with a peptide linkage on only one side, were abandoned. From the results of the kinetic investigations it was concluded that the active site of PNGase F has a very poor affinity for substrates with such structural features.

Similar results had been reported earlier by Tarentino and Plummer (1993) who indicated that the use of proteases in the preparation of glycopeptide substrates for PNGases can result in the formation of short glycopeptides resistant to hydrolysis by PNGase F [86].

Consequently, inhibition studies could only be carried out using the *N*-linked glycopentapeptide **58** as a substrate and the two *C*-glycopentapeptides **54** and **55,** as shown in table 5, as potential inhibitors.

In a first set of experiments the potential inhibitory activity of *C*-glycopentapeptide **54** was assessed using the corresponding *N*-linked glycopentapeptide **58** as a substrate. The potential substrate-inhibitor pair is shown below in figure 27.

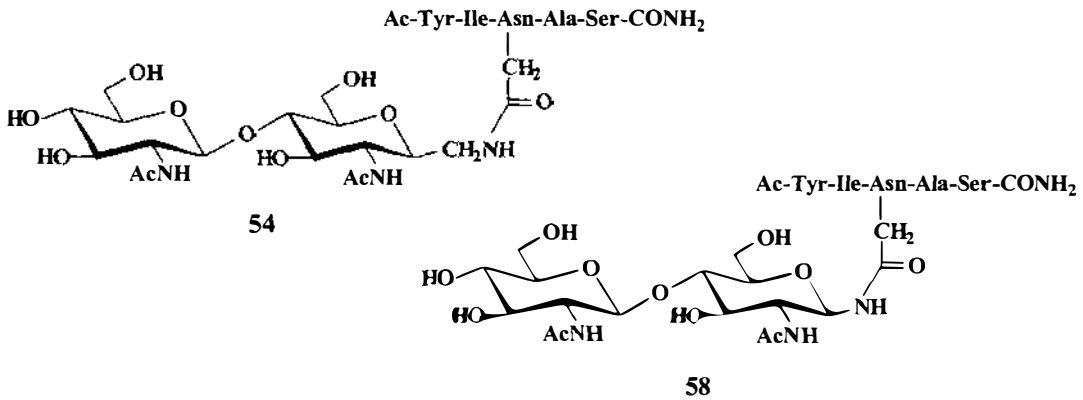


Figure 27: Asparagine-bound *C*-glycopeptide 54, a potential inhibitor of PNGase F and its naturally linked counterpart 58 which was used as a substrate in the inhibition study.

Experimental evidence showed that the *C*-glycopeptide did not inhibit the hydrolysis of the corresponding *N*-linked glycopeptide (for the data see **Appendix, Experimental Data Sheet #6**) . This result was surprising as the *C*-glycopeptide has all of the structural features that ensure recognition by the enzyme because the corresponding glycopeptide had been shown to be a moderate substrate for PNGase F. Therefore binding to the active site of the enzyme could have been expected and in the light of the findings by Wang *et al.* [5] inhibition of the *N*-glycopeptide by its corresponding *C*-glycopeptide analogue should have occurred.

The question therefore arises as to why the high-mannose *C*-glycopeptide inhibits the digest of its corresponding glycopeptide, whereas in contrast the shortened *C*-glycopeptide failed to inhibit the hydrolysis of its *N*-linked analogue.

As the peptide moiety of the high-mannose and the *N,N'*-diacetylchitobiosyl based *C*-glycopeptides as well as the corresponding glycopeptide substrates is identical, it is clear that the chain length of the carbohydrate moiety plays a major role in substrate binding, in a way that is not yet understood.

A comparison of the kinetic data for the two naturally linked glycopeptide substrates showed that PNGase F has a 28 fold higher affinity for the high-mannose substrate

than for the dichitobiosyl based analogue. If the presence of the additional methylene group in the *C*-glycopeptides is a structural feature that disrupts the interaction between the substrate and the enzyme, the effect on the shortened version of the substrate would be expected to be more drastic, as there are no other sugar residues to interact with the enzyme to stabilise the enzyme substrate complex. This hypothesis is also supported by the findings of the molecular modelling studies presented in section 6.3.

In fact it seems most likely that the observed effect of non-inhibition, in comparison to the high-mannose *C*-glycopeptide reported by Wang *et al.* [5], is based on the reduced affinity of PNGase F for *N*,*N*'-diacetylchitobiosyl based substrates. If the enzyme has a reasonable affinity for a substrate, as is the case for the high-mannose glycopeptide (K_M: 80 μM), the kinetic equilibrium of binding is shifted towards the formation of an enzyme-substrate complex. Bearing in mind that the substrate analogue *C*-glycopeptides are competitive inhibitors it can be argued that a sufficient affinity for the corresponding substrate is a prerequisite for inhibition of enzymatic activity to be observed. Although it is difficult to express the term "sufficient affinity" in absolute numbers, its implications are quite clear.

This concept can be illustrated by considering an enzymatic digest in which equimolar amounts of both, glycopeptide substrate and its substrate analogue *C*-glycopeptide inhibitor are present in the reaction medium. At a steady-state a constant amount of enzyme-substrate complex is present. Assuming that the enzyme shows a similar affinity for the substrate-analogue inhibitor and the corresponding substrate, an equal amount of an enzyme-substrate and an enzyme-inhibitor complex should be present at a steady state. As the substrate analogue inhibitor cannot be cleaved by the enzyme this will consequently result in a decrease in the rate of reaction as determined by the dissociation constant of the enzyme-inhibitor complex.

For a poor substrate, the kinetic equilibrium is shifted away from the enzyme-substrate complex to the free enzyme and the substrate. At a steady state there is consequently only a small amount of substrate associated with the active site of the enzyme. Furthermore the enzyme shows a low turnover rate for a poor substrate and the initial velocity is not necessarily proportional to the substrate concentration. This means that for a poor substrate, a plot of the initial velocity against the initial substrate

concentration at a constant total enzyme concentration will not generally result in a hyperbola, as is the case for the FITC-labelled ovalbumin glycopeptide and PNGase F as shown in figure 20 in section 4.2.

In this context Fan and Lee [59] have used the unscientific term "sluggishly" to compare the activity of PNGase F to that of PNGase A for an *N*,*N*'-diacetylchitobiosyl based substrate. Nevertheless, this expression is helpful in a metaphorical way to describe the processes involved.

When taken with the results obtained for the molecular modelling studies discussed in section 6.3, it can be concluded that the failure of the *C*-glycopeptide **54** (figure 27) to inhibit the activity of PNGase F is a result of the additional methylene group impeding the binding of this compound to the active site of PNGase F. This is not helped by the fact that the limited length of the carbohydrate moiety also dramatically affects the affinity of the enzyme for this potential inhibitor.

For this reason the testing of the inhibitory effect of the glutamine bound *C*-glycopentapeptide **55** (table 5) towards PNGase F was abandoned.

4.4.2 α-Linked *N*-Glycopeptides

The use of α-linked *N*-glycopeptides, the synthesis of which is described in section 3.3, in kinetic studies with PNGases has not been reported in the literature. For this reason it was desirable to investigate whether PNGase F would deglycosylate such compounds or whether they would inhibit PNGase F activity.

The natural occurrence of α-linked *N*-glycopeptides has only been confirmed in one case (section 3.3).

It could therefore be possible that, although PNGases should not be able to cleave such unusually linked glycopeptides, they may act as inhibitors because of the altered stereochemistry at the cleavage site.

As expected, PNGase F did not hydrolyse the two α-linked *N*-glycopeptides **57** and **59** shown in figure 28 (for the data see **Appendix, Experimental Data Sheet #7**) .

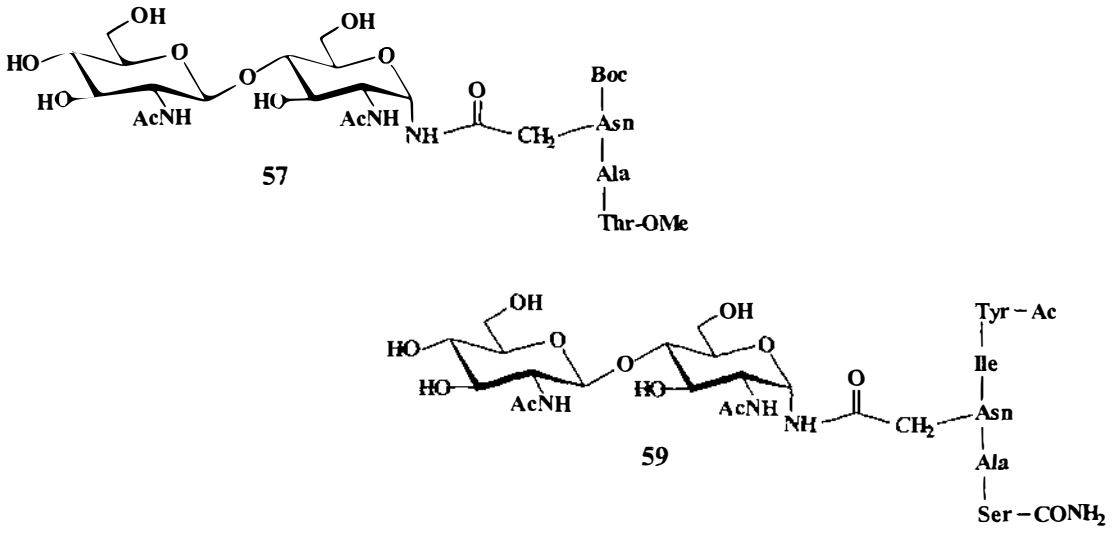


Figure 28: α-Linked *N*-glycopeptides 57 and 59.

Initially, similar reaction conditions were used as for previous digestions with the β-linked glycopeptides **56** and **58** (figure 26), but no activity was observed. The use of very concentrated enzyme preparations or prolonged reaction times did not change this result. As an example figure 29 shows overlayed HPLC- chromatograms for the digests of both, the β- and α-linked glycotripeptides **56** and **57**.

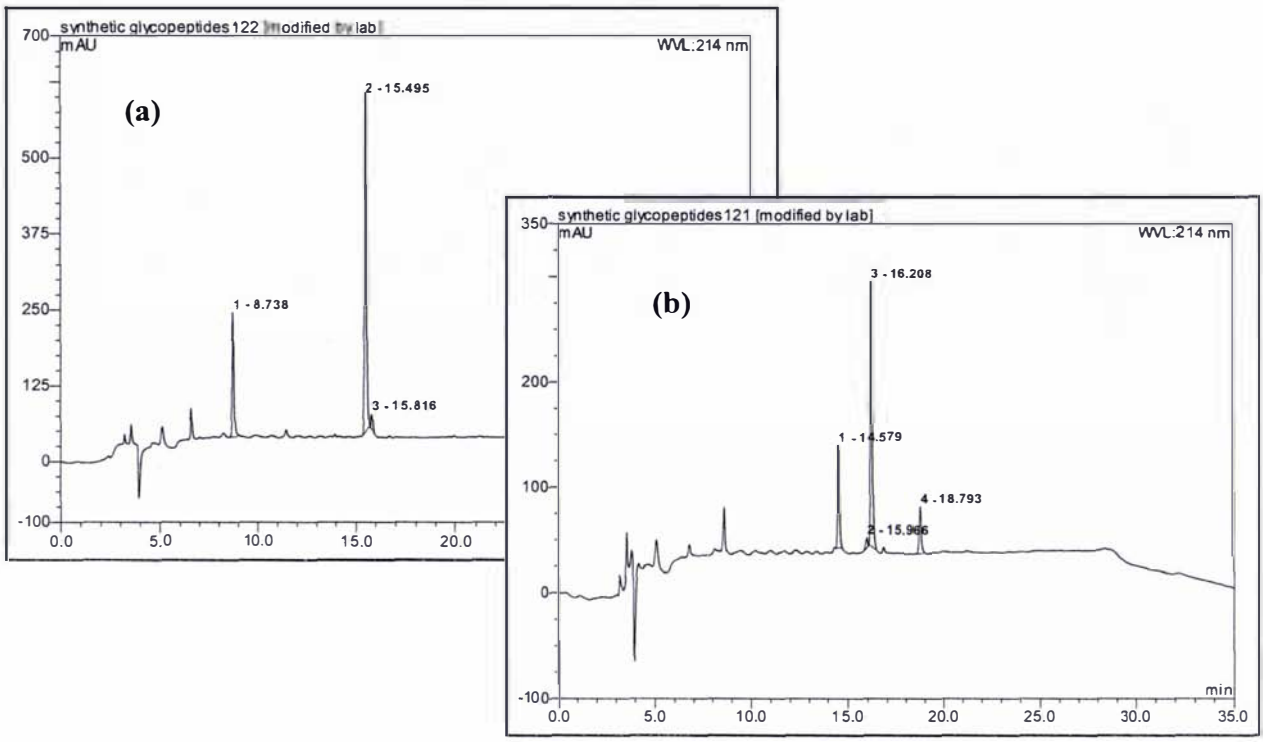


Figure 29: The HPLC-chromatogram (a) shows the attempted digest of the α-linked glycotripeptide 57 (peak 2). No peak corresponding to the cleavage product (tripeptide) can be identified. Chromatogram (b) shows the digest of the corresponding β-linked compound 56 (peak 3). Peak 4 corresponds to the hydrolysis product (tripeptide). Retention times are shown in minutes.

The HPLC-chromatogram for the digest of the α-linked glycotripeptide **57** (a) shows no peak at the retention time characteristic of the expected cleavage product, but only a signal with a retention time of 15.495 min (peak #2) corresponding to undigested **57**.

Using the same conditions for the corresponding β-linked glycotripeptide **56** (figure 23), as shown in chromatogram (b), resulted in the appearance of a peak that represented the hydrolysis product (tripeptide) with a retention time of 18.793 min (peak # 4). This result was confirmed by collecting the fractions corresponding to the signals and determining their mass using ES-MS.

A commercially available PNGase A preparation gave the same results.

These results led to the conclusion that α-linked *N*-glycopeptides cannot be cleaved by PNGase F or PNGase A.

The fact that the recombinant PNGase F and PNGase A could not hydrolyse α-linked glycopeptides suggests that they are either not recognised by the enzyme or that they might be able to bind to the active site but cannot be further processed. If the latter is true then the α-linked glycopeptides may be able to inhibit the enzyme.

To test this hypothesis, equimolar amounts of the α-linked glycopentapeptide **59** (figure 28) and the corresponding β-linked glycopeptide **58** (figure 26) were used in an inhibition experiment with recombinant PNGase F.

The results of the experiments showed that the α-linked glycopeptide **59** did not inhibit the activity of the enzyme (for the data see **Appendix, Experimental Data Sheet #8**). As an illustration the corresponding chromatograms of the HPLC-runs are shown in figure 30.

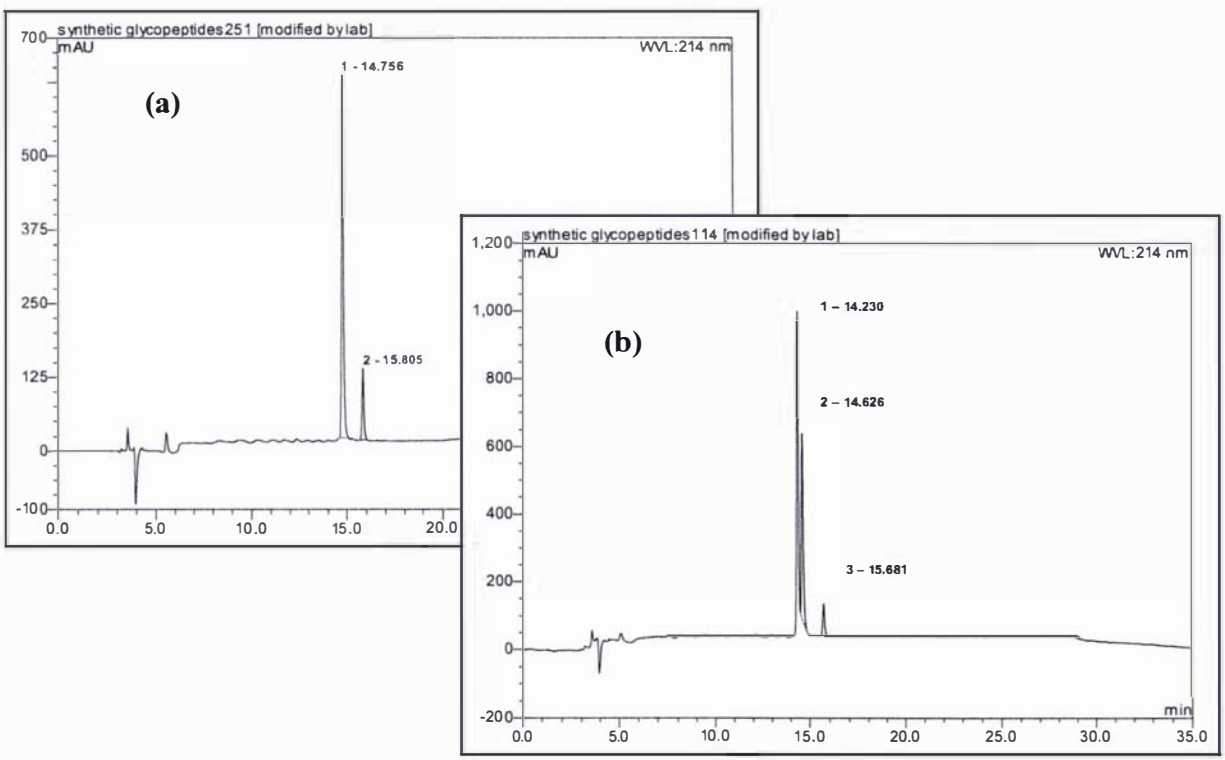


Figure 30: The HPLC-chromatogram (a) shows the digest of the β-linked glycopentapeptide 58 (figure 26). Peak 1 corresponds to the undigested substrate and peak 2 to the hydrolysis product (pentapeptide). Chromatogram (b) shows the digest of the same glycopentapeptide in the presence of the corresponding α-linked glycopentapeptide 59 (figure 28). Peak 1 corresponds to the α-linked glycopeptide, peak 2 to the substrate and peak 3 to the hydrolysis product (pentapeptide). The amounts of the α-and β-linked glycopeptides in the reaction mix were designed to be equimolar and the difference in the height of peak 1 and 2 accounts for an error introduced by the poor solubility of the compounds in aqueous buffers. Retention times are shown in minutes.

As a partial result this study has shown the importance of the carbohydrate moiety in substrate recognition by PNGase F.

To further examine the role of α-linked glycopeptides as potential inhibitors of PNGases, the capability of the enzyme to process an **α-linked high-mannose glycopeptide** needs to be tested. Depending on the results of the experiment, the effect of such a compound on the digest of the corresponding β-linked glycopeptide would confirm whether it acts as an inhibitor or not. Such an experiment would unequivocally establish whether the enzyme is able to bind and not process the α-linked glycopeptide, or whether the anomeric configuration of the glycoside is an important recognition factor.

As the only feasible pathway for the synthesis of such a compound would involve a chemoenzymatic approach, as described for the synthesis of the high-mannose *C*-glycopeptide reported by Wang *et al.* [5] using an enzyme not commercially available, this experiment could not be carried out during the course of this study.

4.5 Non-Specific Inhibition

The term "non-specific inhibition" as used in this thesis describes a process in which the enyzmatic activity is reduced by compounds that neither fit the category of substrate-analogue or transition-state inhibitors. The inhibitory activity of these compounds is termed "non-specific" because they are not structurally designed to target specific amino acids within the active site.

However, the term "non-specific inhibition" is not meant to be limited to the concept of competitive inhibition. It could as well belong to the concepts of "uncompetitive" (the inhibitor can only bind to the enzyme-substrate complex) and "non-competitive" inhibition (the inhibitor can bind to the enzyme regardless of whether a substrate molecule is bound) [131]. Consequently the compound could bind at a site different from the catalytic region inducing a conformational change and thus preventing substrate binding or cleavage. The emphasis is on the fact that the "unspecific inhibitor" is not specifically designed to target a certain binding site on the enzyme molecule.

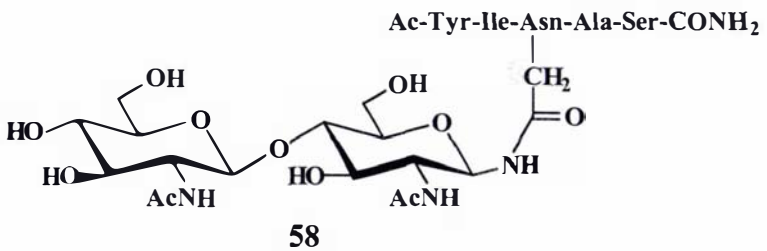
4.5.1 Chitobiose, Specific or Non-Specific Inhibition?

In 1995 Kuhn *et al.* [55] reported the co-crystallisation of PNGase F with *N,N'*-diacetylchitobiose claiming that their results revealed the oligosaccharide recognition residues of the active site. This was further backed by results obtained from site directed mutagenesis studies.

To begin with, the investigations conducted by Kuhn and co-workers revealed that *N,N'*-diacetylchitobiose inhibited the activity of PNGase F. The compound inhibited substrate hydrolysis by 36 % at a 12.5-fold molar excess relative to the substrate and by 50 % at a 100-fold excess. The substrate used in these experiments was a di-dansylated fetuin glycopeptide.

Experiments conducted in our laboratory using recombinant PNGase F confirmed that *N,N'*-diacetylchitobiose inhibited the activity of the enzyme when the fluorescently labelled ovalbumin glycopeptide (sections 2.3 and 4.2) was used as the substrate. As an example, a ~140-fold molar excess of chitobiose over the substrate concentration ($1.5 * 10^{-5}$ mmol/mL) reduced the enzymatic activity to 62 % when compared with the control runs (for the data see **Appendix, Experimental Data Sheet #9**).

The inhibitory activity of the dichitobiose was also tested using the *N,N'*-diacetylchitobiosyl based glycopentapeptide **58** shown in figure 31 as a substrate.



58

Figure 31: Substrate used for assessing the inhibitory activity of *N,N'*-diacetylchitobiose towards PNGase F.

In these experiments a 2.4-fold molar excess of chitobiose compared to the substrate concentration ($4.44 * 10^{-4}$ mmol/mL) reduced the enzymatic activity to 68 % compared with the control runs. Doubling the concentration of chitobiose to a 4.8-fold

molar excess reduced the activity to 58 % in comparison to the uninhibited reaction. A further increase by 50 % to a 7.2-fold molar excess over the substrate reduced the activity to 49 % compared to the runs in which there was no chitobiose present (for the data see **Appendix, Experimental Data Sheet #10**).

Two important conclusions can be drawn from these results: Firstly, the results confirm the findings by Kuhn and co-workers that a significant increase in the concentration of chitobiose results only in a relatively moderate increase in the level of inhibitory activity. Secondly, the results reveal a striking difference between the inhibition of the digest of the fluorescently labelled ovalbumin-glycopeptide and the *N,N'*-diacetylchitobiosyl based glycopentapeptide **58**. Whereas a ~140-fold molar excess of chitobiose was required to reduce the enzymatic activity to 62 % when the labelled ovalbumin-glycopeptide was used as a substrate, only a 2.4 molar excess was required to reduce the activity to approximately the same amount when the shortened glycopeptide was used as a substrate.

This is not surprising in light of the differences in the K_M values for the two substrates with PNGase F (sections 4.2 and 4.3) which indicates that the ovalbumin glycopeptide is a much better substrate than the shortened glycopentapeptide.

The three dimensional crystal structure obtained from the co-crystallisation of PNGase F with *N,N'*-diacetylchitobiose by Kuhn *et al.* [55] reveals that the diacetylchitobiose is held in a cleft in the enzyme through hydrogen bonds between the sugar hydroxyl groups and the side chain and main chain residues of Asp 60, Glu 118, Trp 120 and Arg 61. Hydrogen bonds to additional residues, Glu 206, Gly 190, Tyr 85 and Tyr 62 are mediated by bound water molecules.

One of these water molecules, water 346 according to the numbering by Kuhn *et al.*, forms a link between Asp 60, Glu 206, Arg 248, Tyr 85 and the substrate. It could therefore be possible that water 346 acts as a nucleophile, a common mechanism in enzymatic hydrolysis reactions. As water is a very weak nucleophile it needs to be activated through the interaction with an amino acid residue. Examples for this can be found in the catalytic mechanism of polymer degrading enzymes such as cellulases e.g. the cellobiohydrolases from the fungi *Clostridium thermocellum* [133] or *Trichoderma reesei* [134] in which aspartic acid residues have been proposed to assist in the nucleophilic attack of water on the substrate's glycosidic linkage. A further

example can be found in the mechanism of bacterial glycosylasparaginase in which a water molecule is activated for a nucleophilic attack by interaction with the primary amino group of a terminal threonine residue [135].

Despite the findings by Kuhn *et al.* the exact nature of the mechanism of action of PNGase F is, at this time, still unclear. A major drawback of their findings is the fact that the reducing end of the *N*,*N*'-diacetylchitobiosyl molecule in the obtained crystal structure is in an α-configuration whereas the corresponding GlcNAc residue in *N*-linked glycopeptide substrates is in a β-configuration.

These results by Kuhn *et al.*, however, fail to explain at least two significant observations:

1) This work has shown that α-linked glycopeptides cannot be processed by PNGase F. If the results by Kuhn *et al.* reveal the oligosaccharide binding sites, how can the reducing end of the *N*,*N*'-diacetylchitobiosyl molecule be in an α-configuration?

2) Results by Altmann *et. al.* [56] have shown that the length of the carbohydrate chain outside the *N*,*N*'-diacetylchitobiosyl core has a significant impact on the affinity for the active site of PNGase F, although this effect is less pronounced the longer the oligosaccharide chain. The reasons for this are still not clear.

The question therefore has to be asked: Is *N*,*N*'-diacetylchitobiose a specific or a non-specific inhibitor? If the compound was a non-specific inhibitor it may have bound to the PNGase F molecule in the active site in a conformation that does not reflect the true nature of substrate binding. In this context it is important to note that the ratio of *N*,*N*'-diacetylchitobiose to PNGase F in the crystallisation studies by Kuhn *et al.* was 30:1.

In order to answer this question, the model proposed by Kuhn *et al.* for the binding of chitobiose to the proposed active site region of the PNGase F molecule was compared to that of a "virtual" *N*-linked glycopeptide from human lactoferrin (residues 475-483) fitted into the active site region of PNGase F. If *N*,*N*'-diacetylchitobiose is a specific inhibitor, the innermost GlcNAc residues of the glycopeptide should show the same or at least similar interactions with the active site amino acid residues identified by Kuhn and co-workers.

For this purpose figure 33 shows an overlay of the lactoferrin glycopeptide modelled into the active site of the PNGase F molecule and the *N*,*N*'-diacetylchitobiose bound in the manner reported by Kuhn and co-workers.

The lactoferrin model used in this study (figure 32) was created by the virtual cleavage of a high resolution structure of the human lactoferrin glycopeptide [136]. It comprises nine amino acid residues (475-483) and a chain of two GlcNAc and one Man residue is *N*-glycosidically linked to Asn 478. The innermost GlcNAc is fucosylated on the 6-position.

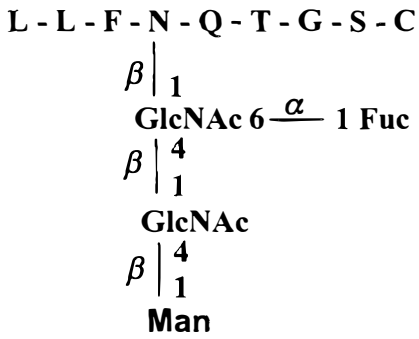


Figure 32: Amino acid sequence (one letter code) and carbohydrate moiety of the virtually shortened lactoferrin glycopeptide used in the modelling study.

The overlay shows an almost exact correspondence between the binding of the *N*,*N*'-diacetylchitobiose as shown by Kuhn *et al.* in their co-crystallisation studies and the two innermost GlcNAc residues of the lactoferrin glycopeptide fitted into the active site of the PNGase F molecule.

Assuming that the modelling of the lactoferrin glycopeptide into the active site of the PNGase F molecule is a true reflection of the enzyme's substrate binding mechanism, *N*,*N*'-diacetylchitobiose can be regarded as a specific inhibitor.

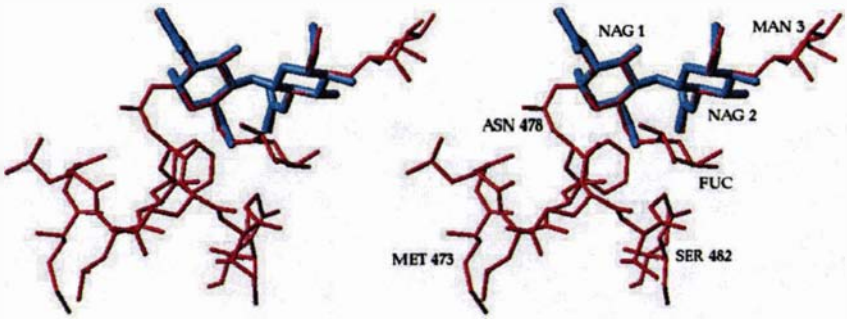


Figure 33: Stereoplot showing the overlay of the carbohydrate core of a lactoferrin glycopeptide (red, with Man β1-4 linked to NAG2 and Fuc α1-6 linked to NAG1) modelled into the active site region of the PNGase F molecule with *N,N'*-diacetylchitobiose (blue) bound in the way reported by Kuhn and co-workers from their co-crystallisation studies. Note that the glycopeptide is in a β-linkage whereas the dichitobiose is in an α-linkage. Amino acid residues of the enzyme's protein backbone are shown in red.

NAG = GlcNAc

4.5.2 PEG

Polyethylene glycol (PEG) is a polymer that is widely used as a precipitating agent in macromolecular crystallisation techniques [137]. The crystal structures of PNGase F reported by Kuhn and co-workers [78],[55] and by Norris *et al.* [63],[79] were obtained using this precipitant. Kuhn *et al.* used 25 % (w/v) PEG 4000 as a precipitating agent and Norris *et al.* were initially successful using 24 % (w/v) PEG, later changing the concentration of the precipitant to 20 %.

From previous work carried out in our laboratory (Loo,T.S., unpublished results) there was an indication that glycerol had an inhibitory effect on PNGase F. As glycerol and PEG are both structurally characterised by abundant hydroxyl functional groups it was necessary to establish whether this precipitating agent had any effect on the catalytic activity of the recombinant PNGase F.

Although there was no evidence of PEG associated with the original crystal structures [79] it was possible that the molecule was bound, but disordered, or that the waters observed in the structures were in fact associated with PEG molecules. If PEG does indeed inhibit the enzyme, this would have strong implications for any co-crystallisation study using this particular crystallisation protocol.

The substrate used for the inhibition studies was the *N,N'*-diacetylchitobiosyl based glycopentapeptide **58** (figure 31, section 4.5.1). This compound was chosen as a substrate because the *C*-glycopeptide **54** (table 5, section 4.4.1), which had been unsuccessfully used in a co-crystallisation study with PNGase F during the course of this work, had the same structure.

Inhibition studies using a substrate concentration of $4.44 * 10^{-4}$ mmol/mL showed a significant inhibitory activity of the PEG 4000. In the presence of 25 % (w/v) PEG the enzymatic activity was reduced to 14 % of that observed in the control runs which had no PEG present. The presence of 12.5 % (w/v) PEG reduced the activity to 36 % of that of the control runs (for the data see **Appendix, Experimental Data Sheet #11**).

Although the results show that the presence of 25 % (w/v) PEG, the same concentration of PEG used in the PNGase F co-crystallisation, inhibited the enzyme it is yet unclear whether the compound binds at the active site or whether it changes the conformation of the enzyme in a way that decreases the catalytic activity. On the basis of the chemical composition of polyethylene glycol [$HO(C_2H_4O)_nH$] with its abundant hydroxyl functional groups there is good reason to assume that it somehow binds to the active site, acting as a competitive inhibitor. Thus the presence of PEG in co-crystallisation studies with PNGase F was not desirable as it might have interfered with the binding of low affinity ligands.

4.5.3 Variation of PNGase F Activity in Different Buffer Systems

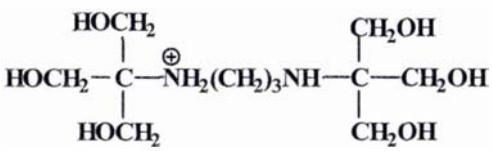
In 1995 Altmann and co-workers published a detailed report on substrate specificities of PNGases A and F [56]. In the discussion of their results they conclude that "Any specification of PNGase activity must therefore be accompanied by a thorough description of the substrate employed".

Likewise the biological buffer used in any investigation must be carefully chosen. Results obtained during the work for this thesis showed that the activity of the recombinant PNGase F was significantly reduced in BTP-buffer (1,3-bis[tris(hydroxymethyl)-methylamino]propane), a buffer system that was most commonly employed in this work. Not only were substrate solutions prepared in this buffer, but it was also used for the dilution of the recombinant PNGase F preprations.

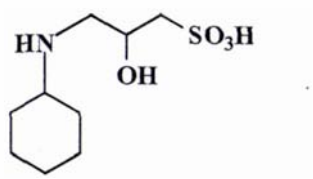
To unequivocally establish that BTP inhibited PNGase F activity, reactions were monitored in four different buffer systems, including BTP.

The buffers compared were BTP (1,3-bis[tris(hydroxymethyl)-methylamino]propane), CAPSO (3-(cyclohexylamino)-2-hydroxypropanesulphonic acid), EPPS (4-(2-hydroxyethyl)piperazine-1-propanesulfonic acid) and ammonium acetate. The chemical structures of these buffers are shown in figure 34.

Interestingly there were no significant differences observed between the PNGase F activities in CAPSO, EPPS and ammonium acetate-buffer all of them adjusted to pH 8.5. It was only in BTP-buffer pH 8.5 that the enzyme displayed a reduced level of activity.



BTP



CAPSO

EPPS

Ammonium acetate

Figure 34: Chemical structures of the different buffers used in the kinetic studies.

Using FITC-ova with a concentration of $1.5 * 10^{-5}$ mmol/mL as the substrate the enzyme showed only 80 % activity in 5 mM BTP in ammonium acetate (5 mM) compared to ammonium acetate buffer only (10 mM).

Using the *N,N'*-diacetylchitobiosyl based glycopentapeptide (figure 31, section 4.5.1) as the substrate with a concentration of $4.44 * 10^{-4}$ mmol/mL the enzyme showed only 68 % activity in 5 mM BTP in ammonium acetate (5 mM) compared to ammonium actetate buffer only (10 mM, for the data see **Appendix, Experimental Data Sheet #12**).

Comparing the chemical structures of the four different buffers in figure 34 shows one structural element that distinguishes the BTP-molecule from those of the other buffers; namely the clusters of primary hydroxyl groups [tris(hydroxymethyl)] on each end of the molecule. If the BTP-molecule was acting as a competitive inhibitor it is possible that the amino acid residues in the active site recognised its clustered primary hydroxyl groups because of their structural similarity with the primary hydroxyl groups of the carbohydrate moieties in natural glycopeptide substrates.

This hypothesis, however, can only be confirmed through structural studies. This problem will therefore be discussed further in section 6.3.

CHAPTER 5

Kinetic Investigations using Two PNGase F Mutants

5.1 Introduction

Site-directed mutagenesis is an important tool in the determination of catalytic amino acid residues within the active site of an enzyme. If the point mutation of a specific amino acid results in partial or complete loss of activity it can be assumed that it plays an important role in the catalysis of the specific reaction. However, a second possibility is that the mutation results in a disruption of the native folding of the enzyme thus impacting on the catalytic activity.

As briefly mentioned in chapter 4, Kuhn and co-workers [55] used site-directed mutagenesis studies in conjunction with crystallographic tools to investigate the carbohydrate binding mechanism of PNGase F.

Katiyar et al. [60] used site-directed mutagenesis studies to reveal the catalytic amino acid residues in the active site of yeast Peptide: *N*-Glycanase. The results of these studies in combination with crystallographic data allowed for the postulation of a very comprehensive mechanism for the catalysed de-*N*-glycosylation reaction. They identified a catalytic triad consisting of Cys 191, His 218 and Asp 235 as being essential for substrate cleavage. In this charge relay system the thiol group of the cysteine acts as the nucleophile to attack the carbonyl group of the *N*-glycosidic linkage. Histidine then picks up the liberated proton from the SH group, forming an imidazolium ion, which is in turn stabilised by the carboxylate ion of Asp 235. After liberation of the glycosyl amine which rapidly hydrolyses non-enzymatically to the reducing sugar, the His 218 assists in the nucleophilic attack of a water molecule to liberate the peptide residue from the active site resulting in the formation of free aspartic acid. The conclusion of the mutagenesis study was that loss of activity was either due to the absence of the charge relay or the disruption of the native folding of the enzyme.

However, these cytosolic PNGases appear to differ significantly from PNGases found in plant and bacteria in that they show different/unique carbohydrate binding properties, require an –SH group and a neutral pH for optimal activity [25]. Therefore these results do not contribute to a further understanding of the catalytic mechanism in PNGase F.

As discussed in section 4.5, the results by Kuhn and co-workers revealed that the active site was located in a deep cleft at the interface between the enzyme's two domains. This had also been proposed by Norris *et al.* [79] on the basis of their three-dimensional crystal structure of the PNGase F molecule. Norris and co-workers proposed that Asp 60 and Glu 206 might be important in the cleavage mechanism. These two amino acid residues are linked by a tightly bound water molecule which is also hydrogen bonded to Tyr 85 and Arg 248. Furthermore, His 193 appears to be able to interact with the substrate as it is located at the mouth of the active site cleft.

Although at that stage it was not clear whether a glycopeptide substrate could reach down far enough into this cleft, parallels were drawn to the mechanism of cellulases [133],[134] in which tryptophan residues assist in the substrate binding and aspartic acid residues assist as proton donors (direct acid catalysis) or in the nucleophilic attack of water.

Kuhn *et al.* [55] showed that the mutation of Asp 60 to Ala resulted in a total loss of activity, while the Glu 206 Ala mutant had less than 0.01 % of the activity of the wildtype, respectively. Other residues targeted in this study were Glu 118 and Tyr 85, both of which are in the active site region. Tyr 85 was chosen presumably because of the possibility of it acting as a proton donor and Glu 118 because of its position in the active site cleft. Changing Tyr 85 to Phe resulted in an activity of approximately 10 % and Glu 118 to Ala in an activity of 0.1 % compared to that of the wildtype.

Subsequent to these studies, His 193 was targeted for site directed mutagenesis on the basis of studies on the effect of various metal ions on the activity of PNGase F (Loo, T.S, unpublished results). It was observed that in the presence of metal ions (Co^{2+}, Fe^{3+}, Ag^{+}) the activity of the enzyme was markedly decreased. As the ability of histidine to form hydrogen bonds can be affected in the presence of these ions, this may have an effect on substrate binding.

A second residue, Arg 248, was also targeted because it had been shown to be hydrogen bonded to the water that was also in contact with Asp 60, a very important active site residue. Furthermore parallels were drawn to the role of arginine in the catalytic mechanism of phosphoglucose isomerase [138]. In this enzyme, which catalyses the reversible isomerisation of D-glucose 6-phosphate to D-fructose 6-phosphate, an arginine residue stabilises an intermediate *cis*-enediol(ate) during the

proton transfer between C1 and C2. A further possible similarity is that in this isomerase a water molecule acts as a nucleophile, which might also be the case in the cleavage mechanism of PNGase F.

A map of the active site region is shown in section 6.3 where a possible cleavage mechanism for *N*-linked glycopeptides by PNGase F is discussed. It is important to note that this model was developed from the results obtained from the kinetic studies with *N*-linked glycopeptides and glycopeptide mimetics presented in chapter 4 as well as the two mutant PNGases presented in this chapter.

The mutants were kindly donated by Trevor Loo (MSc).

5.2 H 193 A

In this mutant PNGase F which was labelled H 193 A, the histidine residue 193 was mutated to alanine in order to assess the importance of the imidazole ring for either substrate binding or catalytic activity.

The activity of this mutant was calculated in mmol of substrate digested per minute and the absolute amount of enzyme present (mmol * min^{-1} * mg^{-1}) using the fluorescently labelled ovalbumin glycopeptide described in sections 2.3 and 4.2 (FITC-ova) as the substrate. This allowed for a direct comparison of the catalytic activity of the wildtype PNGase F with the H 193 A and the R 248 A mutants.

The values of K_M and k_{kat} could not be determined during this work as, at the time, only very limited quantities of the mutant enzyme and FITC-labelled ovalbumin substrate were available.

The investigations showed that the H 193 A mutant showed approximately 15 % ($1.47 * 10^{-3}$ mmol * min^{-1} * mg^{-1}) of the wildtype activity ($9.67 * 10^{-3}$ mmol * min^{-1} * mg^{-1}), indicating that the imidazole of His 193 played a role in the catalytic process.

This result has to be seen in the context of the findings of Kuhn *et al.* [55] who reported that, compared to the wildtype enzyme, the D 60 N mutant had zero activity and the E 206 Q and E 118 Q mutants had 0.01 % and 0.1 % activity respectively. If His 193 was directly involved in the catalytic process a much more significant loss in activity could have been expected for this mutant. Furthermore, reference to the three

dimensional structure of PNGase F [79] shows that His 193 is too far away from the putative active site to be involved in the catalytic mechanism. It may however play a role in substrate binding.

When a lactoferrin glycopeptide was modelled into the active site, as described in section 6.3, there was very little freedom for movement of the substrate (figure 35). The modelled substrate enzyme complex shows that His 193 could play a role in substrate binding as the model allows for the possibility of a hydrogen bond to be formed between the imidazole NH and a carbonyl oxygen in the peptide moiety of the substrate. Therefore His 193 might assist in the binding and positioning of the susbtrate thus explaining the drop in activity when this amino acid is mutated to alanine.

It is unlikely that the mutation of His 193 to alanine would affect the proper folding of the protein as it is located on the outside of the structure, with the side chain pointing towards the opening of the cleft. However, this cannot be ruled out until a crystal structure of PNGase F co-crystallised with a non-cleavable substrate mimetic unambiguously reveals the molecular determinants of substrate binding.

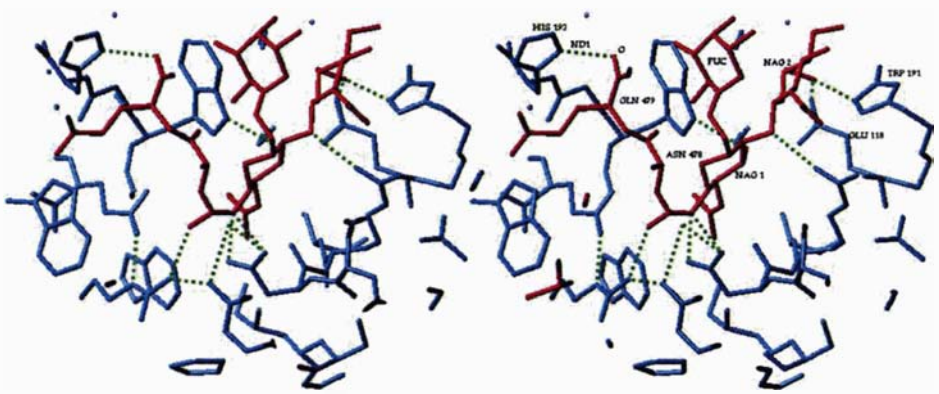


Figure 35: Stereoplot showing a lactoferrin glycopeptide (red) modelled into the active site region of the PNGase F molecule (blue). It shows the proposed hydrogen bonding between the imidazole NH in His 193 and a carbonyl oxygen in the substrate's peptide moiety, possibly stabilising the substrate in the active site. NAG 1 and NAG 2 denote the two innermost GlcNAc residues of the carbohydrate chain. The carbonyl carbon of the *N*-glycosidic linkage is shown in a tetrahedral conformation as expected for the transition state. Dotted lines indicate possible hydrogen bonds.

5.3 R 248 A

In this PNGase F, named R 248 A, Asp 248 was mutated to alanine.

As outlined in the introductory section of this chapter, the rational basis for this investigation was the suspected functional similarity of Arg 248 in the active site of PNGase F to that of Arg 272 in the catalytic mechanism of phosphoglucose isomerase [138], an enzyme that catalyses the reversible isomerisation of D-glucose 6-phosphate to D-fructose 6-phosphate.

Kinetic investigations using the R 248 A mutant revealed that this PNGase F only showed 0.098 % ($9.49 * 10^{-6}$ mmol * min^{-1} * mg^{-1}) of the wildtype activity ($9.67 * 10^{-3}$ mmol * min^{-1} * mg^{-1}). As was the case for the H 193 A mutant, it was not possible to determine the values of K_M and k_{kat} for this mutant as only very limited quantities of mutant enzyme and FITC-labelled ovalbumin substrate were available at the time.

Given the magnitude of the reduction observed in the catalytic activity compared to the wildtype PNGase F, it can be assumed that Arg 248 plays an important role in the catalytic mechanism of the enzyme. Arg 248, however, has a very large sidechain and replacing it with the considerably smaller alanine would leave a destabilising "hole" in the centre of the protein. It is therefore possible that the significantly decreased catalytic rate may simply be due to misfolding. Equally this may represent the importance of Arg 248 in the catalytic mechanism.

The molecular modelling studies presented in section 6.3 allow for a detailed discussion of a possible mechanism for the cleavage of *N*-linked glycopeptides by PNGase F.

In brief, a possible involvement of Arg 248 in the catalytic mechanism could be the hydrogen bonding between the NH_2^+ -group (pK_a 12.48) and the carbonyl oxygen of the *N*-glycosidic bond. This would help to stabilise the substrate in the active site, induce a positive charge on the carbonyl carbon and assist in the nucleophilic attack of an incoming nucleophile. This in turn could be carried out by a hydroxonium ion, possibly generated from water 422 which is fixed in position by hydrogen bonds to Tyr 85 and Asp 60 (see figure 36). The water molecule in this model could be considered analogous to the water 346 reported by Kuhn and co-workers [55].

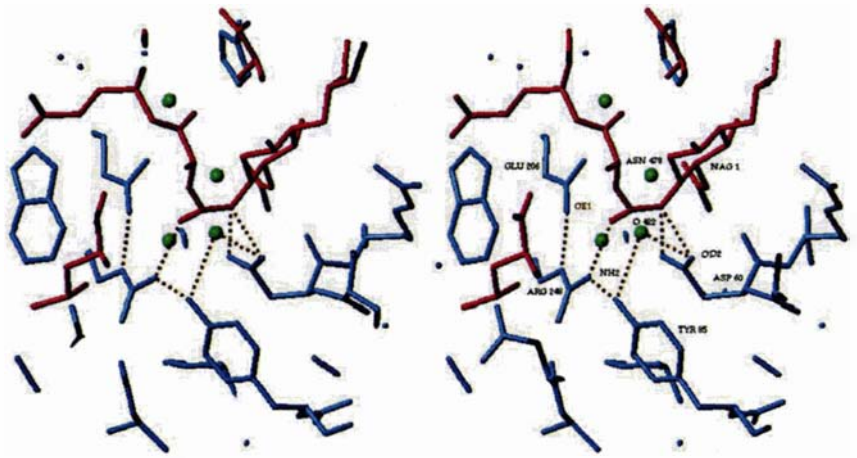


Figure 36: Stereoplot of a lactoferrin glycopeptide (red) modelled into the active site region of the PNGase F molecule (blue). The model shows that it is possible for the NH_2^+ of Arg 248 to donate a proton to the carbonyl oxygen of the *N*-glycosidic linkage, inducing a positive charge on the carbonyl carbon. This would facilitate the nucleophilic attack of a hydroxonium ion possibly generated by a proton transfer from the water 422 to Asp 60. The carbonyl carbon of the *N*-glycosidic linkage is shown in a tetrahedral conformation as expected for the transition state. Dotted lines indicate possible hydrogen bonds.

An important question that needs to be answered in this postulated mechanism is why the mutation of Arg 248 to Ala does not result in the complete loss of activity as reported for the D 60 N mutant by Kuhn *et al.* [55].

A possible explanation has to focus on the role of water 422 (figure 36) as a nucleophile that is situated in a position that would allow for a nucleophilic attack on the carbonyl carbon of the *N*-glycosidic bond.

The results by Kuhn *et al.* [55] suggest that there is an absolute requirement for D 60 in the catalytic mechanism. Their study supports our hypothesis that this residue is essential for the generation of a hydroxonium ion, a much more effective nucleophile than water, through its ability to abstract a proton from water 422. Therefore the mutation of aspartic acid to asparagine results in a total loss of activity.

In the R 248 A mutant, the hydroxonium ion for the nucleophilic attack can still be generated. However, with the lack of hydrogen bonding to the carbonyl oxygen of the *N*-glycosidic linkage, not only is the electrophilicity of the carbonyl carbon decreased but it is also not possible to stabilise the proposed tetrahedral transition state. The

nucleophilic attack of the hydroxonium ion could therefore be energetically as well as sterically unfavourable, yet not completely hindered. Therefore it is still possible for residual activity to be detected.

A very important experiment that could assist in the interpretation of the observed effect would be the generation of a 3D crystal structure of the R 248 A mutant. It would be very interesting to see whether the space occupied by the guanido group in the wildtype crystal structure is occupied by a water in the crystal of the mutant enzyme, possibly held in place by Tyr 85. The water would be much less efficient in hydrogen bonding and donating a positive charge to the carbonyl oxygen of the amide linkage. Consequently a mutant PNGase F in which both Arg 248 and Tyr 85 have been mutated to Ala should not exhibit any catalytic activity. In this mutant not only would the guanido group replacing water molecule but also water 422 (which is believed to be involved in the nucleophilic attack) not be stabilised.

CHAPTER 6

Molecular Modelling Studies

6.1 Introduction

Chapters 4 and 5 describe the kinetic investigations carried out during the course of this work. These investigations were aimed at a more detailed understanding of the structural requirements for substrate binding and the catalytic mechanism of PNGase F.

Likewise structural studies are a powerful technique for unravelling the complexities of catalytic mechanisms. Therefore molecular modelling studies, using the known structure of the PNGase F molecule, were used in this work.

As well as presenting new results, this chapter will also aim at clarifying the questions raised earlier from results of the kinetic studies for example the inhibitory effect of BTP as discussed in section 4.5.3.

6.2 Structural Investigations into Non-Specific Inhibition

Kinetic investigations using recombinant PNGase F in different buffer systems, as discussed in section 4.5.3, showed a significantly decreased activity of the enzyme in BTP buffer. Furthermore it was shown that PEG 4000, which is commonly used as a precipitating agent in co-crystallisation studies, also significantly reduced the activity of PNGase F (section 4.5.2).

The observed effects can be termed "non-specific inhibition" as neither the BTP molecule nor PEG 4000 show any similarities with natural glycopeptide substrates, its modified analogues or proposed transition states. If these molecules did bind to the active or binding site it would therefore be regarded as unspecific as it would not reflect the actual substrate binding mechanism of the enzyme.

Because a model of the BTP molecule was available, molecular modelling studies were conducted to explain the inhibitory effect of this buffer on the activity of PNGase F. For this purpose a model of the BTP molecule was fitted into the active site region of the PNGase F molecule in a way that allowed hydrogen bonding with active site amino acid residues.

Not surprisingly, the molecule could be fitted into the active site in various different ways and figure 37 shows one of these possibilities that would explain the observed inhibitory effect.

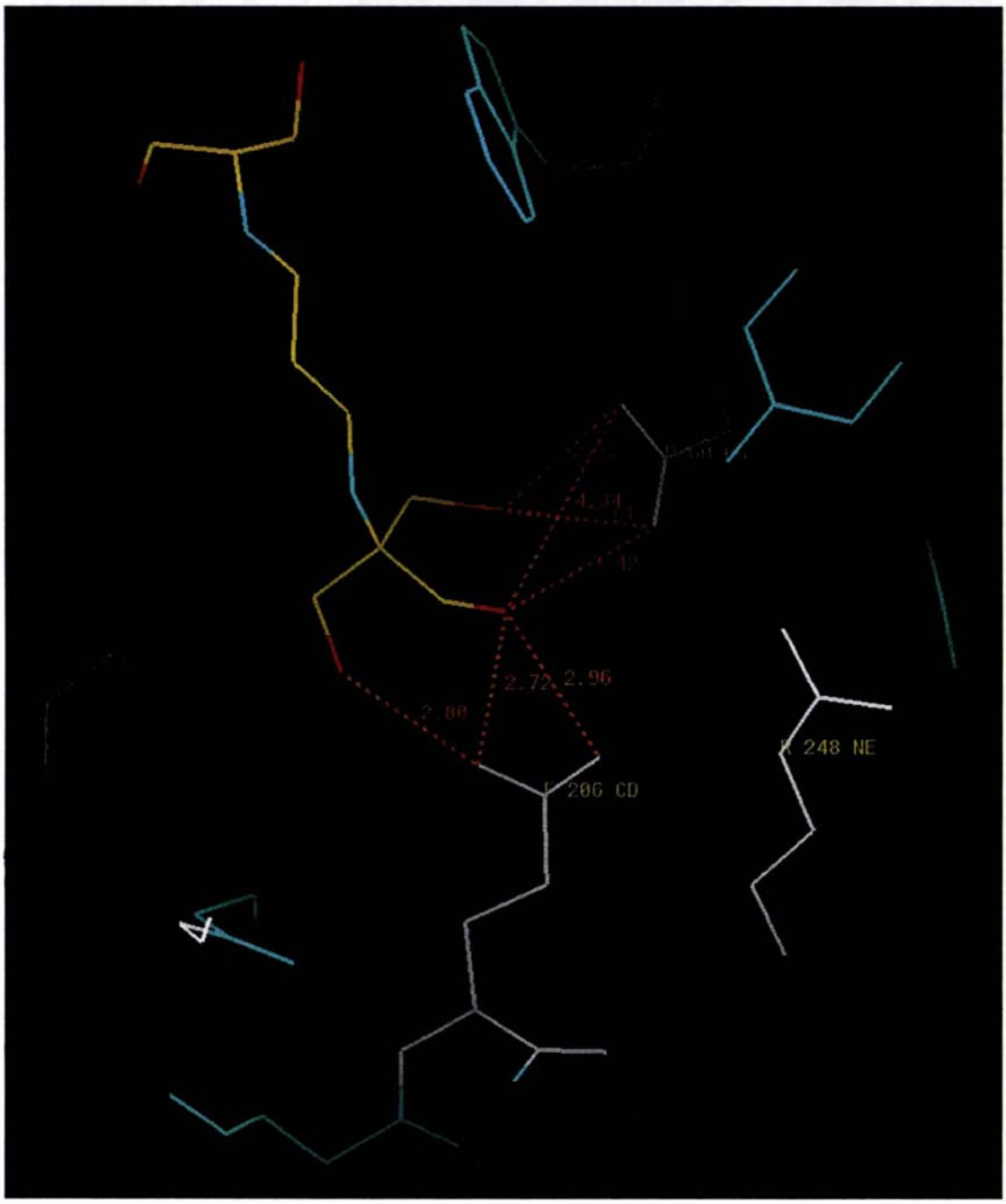


Figure 37: Possible interactions between the primary hydroxyl groups of the BTP molecule (mainly yellow) and two potentially important active site amino acid residues of the PNGase F molecule, Asp 60 and Glu 206. Distances are shown in Ångstrøms.

As was hypothesised in section 4.5.3, it seems most likely that the arrangement of the clustered hydroxyl groups in the BTP molecule allow for interactions with amino acid

residues in the active site that that would otherwise be involved in substrate binding. The inhibitory effect of BTP would therefore be competitive.

The model in figure 37 shows that it is possible for hydrogen bonds to be formed between the primary hydroxyl groups of the BTP molecule and OD of Asp 60 and OE of Glu 206. Mutagenesis studies carried out by Kuhn and co-workers [55] showed that Asp 60 is essential for activity and that the mutation of Glu 206 to Gln reduced the enzymatic activity to 0.01 % of the wildtype.

The model therefore provides an explanation for the observed inhibitory effect of BTP.

At this stage it is impossible to determine precisely why the BTP-molecule does compete with the binding of a glycopeptide substrate. However, it should be noted that the molecule is small and very abundant in the reaction mixture. In the inhibition study using FITC-ova as the substrate, a 300 fold excess of BTP was present relative to the amount of the glycopeptide. Using the shortened glycopeptide substrate **58**, a 10 fold molar excess of BTP was present. Furthermore the observed inhibitory effect was relatively weak. Therefore the results of this study should not be interpreted in any other way than that the use of BTP-buffer in kinetic investigations with PNGase F should be avoided. For future investigations it would be interesting to establish whether polyols with clustered hydroxyl groups are generally of any value in the inhibition of PNGases.

As there is no structure of PEG 4000 present in our structural database, it was not possible to use molecular modelling techniques to hypothesise about the possible interactions of this molecule with PNGase F.

However, based on the results obtained from the modelling studies with BTP, there is a precedence for this polymer to show interactions with active site amino acid residues. The multitude of hydroxyl groups present in a solution containing significant amounts of PEG (in the inhibition study an approx. 75 molar excess of PEG relative to the substrate was present when using 25 % (w/v) PEG) gives rise to a high possibility for the existence of interactions that might compete with the binding of potential ligands to the active site.

Therefore the only conclusion that can be drawn from these investigations is that the presence of PEG 4000 in co-crystallisation experiments with PNGase F is not desirable as it has the potential to interfere with the binding of potential ligands.

6.3 Structural Investigations into Substrate Binding

In 1995 Kuhn and co-workers were successful in revealing a part of the substrate-binding mechanism of PNGase F [55]. With the help of mutagenesis studies and the successful co-crystallisation of PNGase F with *N*,*N*'-diacetylchitobiose, potentially important active site amino acid residues were identified.

During the course of this work the co-crystallisation of PNGase F with the non-cleavable glycopeptide mimetic **54** (table 5, section 4.4.1) was studied in order to determine the binding mechanism for an entire, although shortened, glycopeptide. Unfortunately, this attempt was unsuccessful. Therefore molecular modelling was used to develop a comprehensive model for substrate binding and cleavage by PNGase F.

Having the three dimensional crystal structure of the PNGase F molecule [79],[80] available as well as the structure of a lactoferrin glycopeptide [136], attempts were made to fit this glycopeptide into the active site of the PNGase F molecule using molecular modelling techniques (SGI Graphics workstation; Software: Turbo Frodo (1996) Bio-Graphics, AFMB-CNRS, Marseille, France).

A result of this attempt is shown in figure 38. It is important to mention that the carbonyl carbon of the *N*-glycosidic linkage has been modelled in a tetrahedral conformation as would be expected for the transition state. This change in geometry made it easier to model both parts of the substrate, the peptide and the carbohydrate moieties, into the proposed active site.

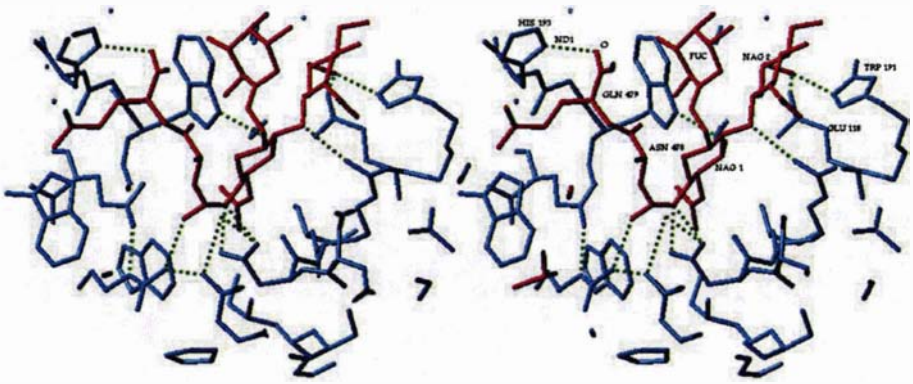


Figure 38: Stereoplot of a lactoferrin glycopeptide (red) modelled into the active site region of the PNGase F molecule (blue). NAG1 and NAG 2 donate the two innermost GlcNAc residues of the carbohydrate chain and a Fuc is α1-6 linked to NAG 1. The *N*-linkage region is shown at the bottom and the peptide moiety is pointing to the left. Dotted lines indicate possible hydrogen bonds.

Using this model, it has been possible to find an explanation for the results obtained from the site directed mutagenesis studies and to propose a mechanism for the hydrolysis of glycopeptide substrates by PNGase F.

With the substrate modelled into the active site of the PNGase F molecule as shown in figure 38, there exists the possibility of the formation of a hydrogen bond from the imidazole of His 193 to a carbonyl oxygen in the peptide chain. This would position the glycopeptide in a way that facilitates the entry of the innermost GlcNAc residues of the glycan into the cleft at the top of the enzyme. This hypothesis is supported by the results of the mutagenesis studies presented in section 5.2 which showed that the mutation of His 193 to Ala reduces the activity of PNGase F to approximately 15 % of the wildtype activity indicating that His 193 could play a significant role in the proper positioning and binding of the substrate.

A further amino acid residue that is believed to be involved in substrate binding is Glu 118 which displays a hydrogen bond to the 6-position of the second GlcNAc molecule. Again it can be postulated that this serves to stabilise and position the substrate in the active site. This is consistent with the results of the aforementioned co-crystallisation and site directed mutagenesis studies by Kuhn *et al.* [55]. In their crystal structure Glu 118 was found to form the same hydrogen bond and they showed that the mutation of this residue to Gln reduced the activity to a level of 0.1 % relative to the wildtype activity.

A possible cleavage mechanism can be postulated from the arrangement of the *N*-glycosidic linkage region of the substrate and the active site amino acid residues as shown in figure 39.

In this model, the NH_2^+ of Arg 248 could form a hydrogen bond to the carbonyl oxygen of the *N*-glycosidic linkage, inducing a positive charge on the asparagine carbonyl carbon. This would favour a nucleophilic attack from a hydroxonium ion, possibly generated by a proton transfer from the tightly bound water 422.

For this to occur it seems most likely for a proton to be transferred to the OD2 of Asp 60 which could subsequently be donated to the nitrogen of the glycosidic linkage, promoting its cleavage and the formation of the glycosyl amine as a primary cleavage intermediate. The model clearly shows that the Asp 60 is too far away from the carbonyl carbon to act as a nucleophile but that the water 422 is ideally placed to fulfill this function.

Furthermore, the model also accounts for the postulated involvement of Glu 206 in the catalytic mechanism. The possible formation of a hydrogen bond between the carboxylate of Glu 206 and the NH of the guanido group in Arg 248 seems to be part of a charge relay system [139]. A charge relay system describes a close hydrogen bonding network between active site amino acid residues in which charge densities can proceed in the opposite direction to a favoured equilibrium and therefore drive reactions that would otherwise be highly unfavoured. In this case it can be assumed that the hydrogen bonding between the guanido group of Arg 248 and the carbonyl oxygen of the glycosidic linkage could be stabilised by increasing the electron density on the guanido group, thus lowering its pK_a. Such a mechanism would explain why the mutation of Glu 206 to Gln reduced the catalytic activity to 0.01 % of the wildtype, but did not abolish it completely [55]. This would also be consistent with the results obtained from the mutation of Arg 248 to Ala which showed a similar level of reduction in activity. This indicates that both amino acid residues are equally important for the hydrogen bonding to the carbonyl oxygen of the substrate's *N*-glycosidic linkage, greatly facilitating the nucleophilic attack of the hydroxonium ion, a prerequisite for the cleavage reaction to proceed.

The postulated mechanism fits with the results of all of the mutagenesis studies discussed in this thesis. The only incongruity is the observed pH maximum of the

enzyme which is 8.5. It would seem unlikely that OD2 of Asp 60 would be able to accept a proton at this pH, given that the pKa for this residue is 3.90 in solution. However, the pKa of an ionizable group in a protein structure can be strongly influenced by its environment. If a carboxylic acid side-chain is in a hydrophobic environment, as is the case for the active site of the PNGase F molecule due to the wide abundance of the amino acid tryptophan, its pK_a will rise because the anionic form will be destabilised [139]. An example of this is seen for the catalytic mechanism of chymotrypsin, an enzyme with a pH optimum of approximately 8, in which an aspartate residue (pK_a 3.90) removes a proton from the imidazole side chain of a histidine residue (pK_a 6.04) [140].

The proposed mechanism also supports the role of Asp 60 in generating a hydroxonium ion for the nucleophilic attack, as water itself is a very weak nucleophile and would have to be activated in some way. Furthermore, without the donation of a proton to OD2 of Asp 60, the subsequent proton-transfer to the nitrogen of the *N*-glycosidic linkage would be disrupted.

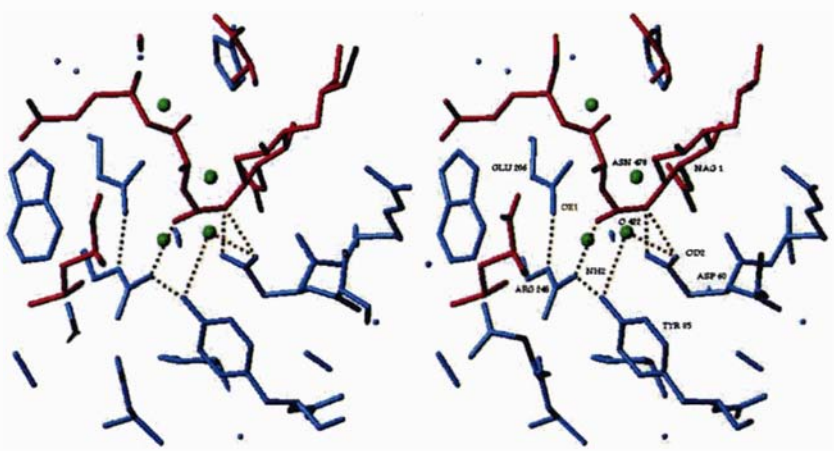


Figure 39: Stereoplot of a lactoferrin glycopeptide fitted into the proposed active site region of the PNGase F molecule. Dotted lines indicate possible hydrogen bonds to the amino acid residues believed to be involved in the catalytic mechanism.

In this context it is interesting to compare the proposed mechanism with that of glycosylasparaginase, an enzyme which carries out the same reaction but only recognises glycosylated asparagine. In glycosylasparaginase a threonine alkoxide is the nucleophile that attacks the carbonyl carbon of the glycosidic linkage with which

it forms a covalent bond. This serves three advantages: Firstly, the comparatively unreactive amide linkage is transformed into a much more reactive ester; secondly, an alkoxide is a much better nucleophile than a hydroxonium ion and thirdly, by forming a covalent bond to the substrate the cleavage reaction becomes intramolecular, which is entropically favoured over a bi-molecular reaction [139].

This mechanism which involves the formation of a covalent bond between the enzyme and its substrate, is expected to be favoured when it is not possible for the enzyme to form a sufficient number of non-covalent interactions to hold the substrate in place. This makes perfect sense for the glycosylasparaginase as its substrates do not contain a protein backbone as is the case for the natural substrates of PNGase F. A non-covalent interaction can therefore only be made to the carbohydrate moiety which might not be sufficient for proper binding.

For PNGase F, there appears to be the possibility for more interactions between the peptide as well as the carbohydrate moiety of the substrate and the enzyme which would help position the substrate. Hence it is possible for the catalytic mechanism to use a hydroxonium ion as a nucleophile. This would also explain the fact that *N*,*N*'-diacetylchitobiosyl based *N*-glycopeptides are relatively poor substrates for PNGase F as the extended carbohydrate chain would allow for a greater number of non-covalent interactions with the enzyme. This hypothesis would also be in accordance with the kinetic results reported in chapter 4, which showed that extending the carbohydrate chain of the chitobiosyl based glycopentapeptide by one GlcNAc unit significantly increased the affinity of the enzyme for such a substrate. It is, however, not yet fully understood to what extent the above hypothesis can be supported structurally.

Figure 40, which was derived from the molecular modelling studies, allows some speculation as to the importance of the β 1→4 linked Man residue in the trisaccharide core of *N*-linked glycopeptides for the activity of PNGase F. The model shows that the Man residue can increase the number of non-covalent interactions between the enzyme and the substrate by being sandwiched between two tryptophan residues in the active site pocket.

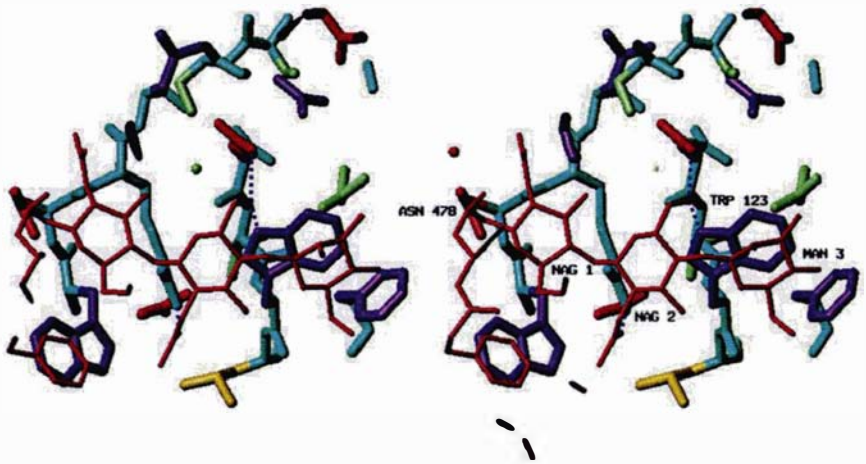


Figure 40: Stereoplot of a lactoferrin glycopeptide fitted into the proposed active site region of the PNGase F molecule. The model shows the possibility for Man 3 to be sandwiched between two Trp residues thus increasing the number of non-covalent interactions between the substrate and the enzyme.

In this context it is also noteworthy that Altmann *et al.* [56] showed that the extension of the carbohydrate chain beyond the trisaccharide core of glycopeptide substrates only has a very moderate influence on the catalytic activity of PNGase F.

The results of our modelling studies are consistent with these observations as there appears to be no obvious way that extra carbohydrate residues outside the invariant trisaccharide core of *N*-linked glycopeptides can interact with the protein thus increasing the number of non-covalent interactions.

This is further exemplified by figure 41 which shows a space-filled model of the active site cleft of PNGase F with the shortened lactoferrin glycopeptide fitted into the groove.

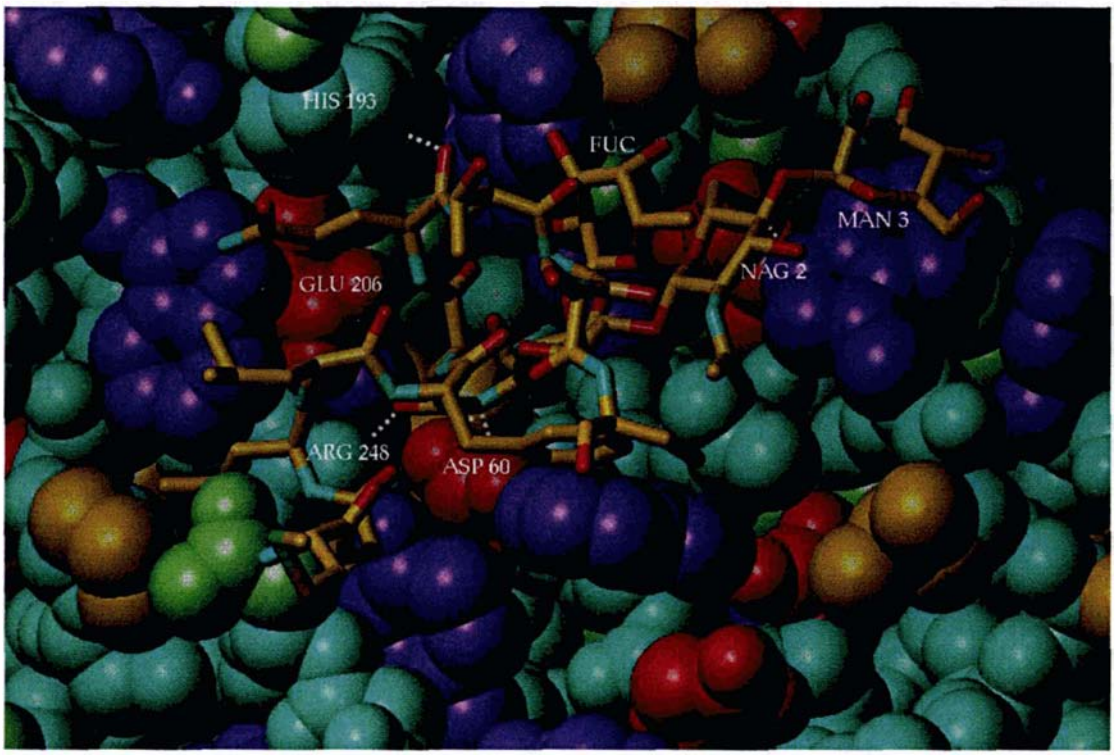


Figure 41: A model of the human lactoferrin glycopeptide fitted into the active site of the PNGase F molecule. The wide abundance of Trp, the aromatic side chain residues of which are shown in purple, clearly shows the hydrophobic environment within the active site region. Of the substrate's carbohydrate chain, the two innermost GlcNAc residues (only NAG 2 labelled), a Man β1-4 linked to NAG 2 and a Fuc α1-6 linked to the innermost GlcNAc can be seen. Amino acid residues believed to be involved in susbtrate binding and cleavage have been labelled with possible hydrogen bonds indicated by dotted lines.

In summary, the molecular modelling studies presented in this section allow for the postulation of a model that supports a possible substrate binding and cleavage mechanism for PNGase F. The proposed mechanism is in accordance with the results of the kinetic and mutagenesis studies presented in this thesis as well as the results of related studies published in the literature.

The challenge for the future is therefore to design experiments which will provide valuable data against which this model of substrate binding and the cleavage mechanism can be judged. This will need to involve further mutagenesis and kinetic studies as well as the co-crystallisation of PNGase F with a non-cleavable glycopeptide mimetic. Whilst the latter would unequivocally confirm or dismiss the proposed mechanism, the former could provide valuable data to evaluate the current ideas.

A further target for the mutagenesis studies has already been identified; a mutant PNGase F in which R248 and Y85 are both mutated to alanine. According to the current model this mutant is expected to show no activity as it cannot protonate the carbonyl oxygen of the *N*-glycosidic linkage and furthermore the bound water molecule 422, which is believed to be involved in the generation of the hydroxonium ion as the nucleophile, cannot be held in position.

The current results have also been linked to the kinetic investigations presented in chapter 4. The modelling study provides a way to explain the results of non-inhibition using the *C*-glycopeptide mimetics and it has also provided valuable data towards the understanding of the non-specific inhibition effects observed.

The effect of non-specific inhibition seems to offer a great potential for the further development of inhibitors for PNGase F. If polyols with clustered hydroxyl groups are indeed capable of competetively inhibiting the activity of PNGase F it would be interesting to see whether the attachment of such compounds to a peptide backbone would increase their affinity for the active site and result in a whole new series of inhibitors.

Such compounds could be very valuable tools in further structural investigations as well as in the design of affinity resins for the one-step purification of PNGases.

CHAPTER 7

Summary

The current work can be divided into four different sections: the synthesis of *N*-linked glycopeptide mimetics, their use in kinetic investigations with recombinant PNGase F, kinetic investigations using targeted PNGase F mutants, and structural studies involving co-crystallisation and molecular modelling techniques.

The aims of this work were to gain a deeper insight into the structural requirements for substrate recognition by PNGase F, develop a detailed understanding of the substrate binding and cleavage mechanism of this enzyme and synthesise potential inhibitors of PNGase F as possible candidates for use as affinity ligands.

A straightforward method for the synthesis of *N*-linked glycopeptide mimetics, more specifically *C*-glycopeptides, as potential inhibitors of PNGase F has been developed. The identification of these targets was based on the current knowledge of the minimal structural requirements for substrate recognition by PNGase F. The synthesis was successfully accomplished in a convergent approach by the separate assembly of *C*-glycosidic and peptide building blocks which were then joined together in a final coupling step.

Furthermore, the synthesis of naturally linked *N*-glycopeptides and glycopeptide mimetics with, for example, an altered anomeric configuration or a variation in the common carbohydrate core, was also achieved. The syntheses described in this thesis are based on simple, robust procedures which produce good yields of material in a minimal number of reaction steps.

To facilitate kinetic studies at low substrate concentrations a discontinuous HPLC based assay using a fluorescently labelled ovalbumin glycopeptide was developed. For the first time K_M and V_{max} values for this new substrate with recombinant PNGase F were determined.

Furthermore, K_M and V_{max} for an *N*,*N*'-diacetylchitobiosyl based glycopentapeptide as well as for a glycopeptide mimetic, showing a variation in the common carbohydrate core, with PNGase F were determined for the first time. It was shown that outside the substrate's *N*,*N*'-diacetylchitobiosyl core the length rather than the constitution of the carbohydrate chain is the crucial factor determining the display of enzymatic activity. In addition, it was shown that PNGase F cannot cleave α-linked *N*-glycopeptides.

It was further established that *C*-glycopeptides based on the minimal structural requirements for substrate recognition are not inhibitors of PNGase F and have therefore no value as affinity ligands.

The inhibition of PNGase F by compounds that do not resemble natural substrates or analogues thereof, termed non-specific inhibition, was also investigated as part of this work. This resulted in the identification of compounds that should not be used in studies with PNGase F, for example BTP buffer and PEG.

The kinetic studies employing targeted PNGase F mutants revealed two new active site amino acid residues that are very important for the catalytic activity of PNGase F, namely His 193 and Arg 248. These results in conjunction with molecular modelling studies using the three dimensional structure of PNGase F and that of a virtual lactoferrin glycopeptide resulted in the proposal of a comprehensive substrate binding and cleavage mechanism for this enzyme. The results also allowed for an explanation of the observations made during the kinetic studies as well as significant results reported in the literature.

The co-crystallisation studies using a *C*-glycopeptide mimetic and recombinant PNGase F did not reveal any potential inhibitor bound to the active site region.

CHAPTER 8

Materials and Methods

8.1 General Methods, Materials and Suppliers

Thin Layer Chromatography (TLC):

Silica gel 60 F_{254} plates from Merck (Darmstadt, Germany) were used for thin layer chromatography. Detection was achieved by using UV-light and/or by dipping the plate into a solution containing 88 % ethanol, 10 % water and 2 % conc. sulphuric acid followed by heating with a heat gun. Compounds containing a free amine could be developed to a purple colour by dipping the plate into a solution of 200 mg of ninhydrin in 94 mL *n*-butanol, 4 mL water and 2 mL glacial acetic acid followed by heating.

Flash Chromatography

Silica Gel 60, 0.04-0.06 mm (230-400 mesh ASTM) from Scharlau Chemie S.A. (Spain) was used for flash chromatography.

High Performance Liquid Chromatography (HPLC):

The purification of peptides and glycopeptides was carried out on a Philips PU 4100 liquid chromatograph connected to a PU 4110 UV/VIS detector using a Phenomenex Jupiter C_{18} column (5μ, 250 x 10 mm). Peaks were monitored at λ 214 nm.
Kinetic investigations were carried out on a Dionex Summit HPLC system connected to a UVD-340S Photo Array Detector and a RF 2000 Fluorescence detector using a Phenomenex Jupiter C_{18} column (5μ, 250 x 4.60 mm). Peaks were monitored at λ 214 nm on the UV detector. The excitation wavelength for the fluorescence detector was 495 nm and the emission wavelength was 520 nm.

Nuclear Magnetic Resonance (NMR)- Spectroscopy:

All ^{1}H-NMR and ^{13}C NMR spectra as well as 2D experiments were recorded on a Bruker Avance 300 spectrometer. The proton frequency is 300.13 MHz and that for ^{13}C is 75.47 MHz . The Avance 300 is equipped with a Quad Nuclei Probe and the sample temperature is normally 30 °C. If necessary ^{1}H-^{1}H-COSY, ^{1}H-^{13}C-HMQC and DEPT- experiments were performed to achieve the full assignment of signals. Nuclear Overhauser effect spectroscopy (nOe) was used to determine the stereochemistry of the *C*-glycopeptides.

Electrospray Mass Spectrometry (ES-MS)

Electrospray mass spectrometry was carried out on a Perkin-Elmer API 300 ES-IMS.

Solid Phase Peptide Synthesiser:

Solid phase peptide synthesis was carried out on an Applied Biosystems Peptide Synthesiser Model 431 A using Rink® resins and 9-fluorenylmethoxycarbonyl (Fmoc)-chemistry.

Solvents

All solvents used for chemical reactions were obtained from Aldrich® in Sure/Seal ™ bottles. Solvents for recrystallisations and work-up procedures were reagent grade from various suppliers. Solvents used for column chromatography were distilled in our laboratory prior to usage. HPLC-grade acetonitrile was used for HPLC-chromatography. Water used was distilled and filtered by a Milli Q® system.

Reagents

All chemicals used were reagent grade obtained from various suppliers (mainly Aldrich®, Acros Organics and Lancaster).

Microanalysis

Microanalytical investigations were carried out by the Campbell Microanalytical Laboratory in the Department of Chemistry at the University of Otago, New Zealand/Aotearoa.

8.2 General Procedures

Solid Phase Peptide Synthesis

For the solid phase peptide synthesis 0.42 g of Rink® resin (loading 0.6 mmol/g) was reacted with 1 mmol of each amino acid using Fmoc-chemistry. The peptide chains were assembled from the *C*- to the *N*-terminus. The peptides were cleaved from the resin with a solution of 9.5 mL TFA and 0.5 mL water. After filtration and washing of the resin with chloroform, the solvent was evaporated using a rotary evaporator, the residue taken up in water and the solution washed twice with cold ether. After lyophilisation the peptides were purified on a Phenomenex Jupiter C_{18} column (5µ, 250 x 10 mm) at a flow rate of 3 mL /min. A linear gradient was employed over 1h, with initial conditions 100 % water (0.1 %TFA) and final conditions 20 % water (0.1 % TFA) and 80 % acetonitrile (0.08 % TFA). Peaks were monitored at λ 214 nm. The peptides were dissolved and injected as solutions in DMF.

Solution Phase Peptide Synthesis

Solution phase peptide synthesis was carried out using *N*-*t*-butoxycarbonyl (Boc)-chemistry with dicyclohexyl carbodiimide/1-hydroxybenzotriazole activation in DMF. The general protocol was as follows: the carboxyl-protected amino acid component (1.0 equiv.) and the Boc-protected carboxyl component (1.5 equiv.) were coupled in DMF at 0 °C using HOBt (1.5 equiv.) and DCC (1.6 equiv.). After 12 h of stirring at room temperature the DCU (dicyclohexylurea) was filtered off and the solvent evaporated. The residue was taken up in chloroform and washed with sat. (saturated) sodium hydrogen carbonate solution, brine and water. After drying over sodium sulfate, the residue was subjected to column chromatography using silica gel.

Boc-Group Cleavage

For the solution phase peptide synthesis Boc-groups were cleaved using 4 M hydrochloric acid in 1,4-dioxane at 0 °C following a procedure previously described by Wang *et al.* [5].

Synthesis of β-Glycopyranosylamines

β-Configured glycopyranosylamines were synthesised following the procedure of Likhosherstov et al. [119].

The deacetylated glycoside was dissolved in sat. ammonium hydrogen carbonate solution and the solution left standing at r.t. (room temperature) for 24 d. If the reaction flask is sealed, the reaction does not go to completion. A practical method was to close the flask with a rubber septum into which was then inserted a Luer-lock needle. The progress of the reaction was monitored by TLC ($CHCl_3$:MeOH, 1:1, ninhydrin stain) and upon completion (very minor traces of starting material might still be detectable) the solution was repeatedly lyophilised to remove the ammonium hydrogen carbonate. The resulting compound was used without further purification.

Synthesis of Glycopeptide Mimetics

The coupling reaction between the free amino-group of the glycoside building block and the carboxy group of the side chain in aspartic or glutamic acid in pre-assembled peptides was carried out using a method based on that of Anisfeld and Lansbury [117] and S.Y.C. Wong *et al.* [118]: 1.5 eq. (equivalent) of the glycoside and 1eq. of the peptide component were coupled in DMSO/DMF at r.t. using 3 eq. HOBt, 1 eq. 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) and 1 eq. *N,N*-diisopropylethylamine. The reaction was monitored by HPLC (for conditions see below). After completion (typically 4 h) the glycopeptide was isolated from the reaction mixture by RP HPLC (Phenomenex Jupiter C18, 5μ, 250 x 10 mm) at a flow rate of 3 $mL*min^{-1}$. The following method was used to elute the product (solvent A:100 % water, 0.1 % TFA; solvent B: MeCN, 0.08 % TFA): 1) isocratic flow of 100 % A for 5 minutes, 2) linear gradient: 100 % A → 65 % A and 35 % B over 30 min.

Method for the HPLC-based Discontinuous Assay using FITC-Ova

Kinetic investigations using FITC-ova as the substrate were carried out on a Dionex Summit HPLC system connected to a RF 2000 Fluorescence detector using a Phenomenex Jupiter C_{18} column (5μ, 250 x 4.60 mm). The excitation wavelength was 495 nm and the emission wavelength was 520 nm. The following linear gradient was

used (solvent A:100 % water, 0.1 % TFA; solvent B: MeCN, 0.08 % TFA): 1) 80 % A and 20 % B → 60 % A and 40 % B over 15 min. 2) 60 % A and 40 % B → 30 % A and 70 % B over 10 min. 3) 30 % A and 70 % B → 80 % A and 20 % B over 10 min. The flowrate was 1 $mL*min^{-1}$.

Method for the HPLC-based Discontinuous Assay using Synthetic *N*-Glycopeptides

Kinetic investigations using synthetic *N*-glycopeptides as the substrate were carried out on a Dionex Summit HPLC system connected to a UVD-340S Photo Array Detector using a Phenomenex Jupiter C_{18} column (5μ, 250 x 4.60 mm). Peaks were monitored at λ 214 nm. The following linear gradient was used (solvent A:100 % water, 0.1 % TFA; solvent B: MeCN, 0.08 % TFA): 1) isocratic flow of 100 % A for 2 min., 2) 100 % A → 60 % A and 40 % B over 20 min., 3) 60 % A and 40 % B → 100 % A over 8 min. The flowrate was 1 $mL*min^{-1}$.

8.3 Synthesis of the Glycoside Building Blocks

In general all reactions were conducted under an inert atmosphere using dry argon gas.

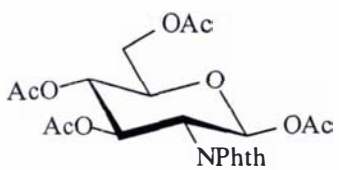
The *C*-glycosidic compounds have not been named using the systematic heptitol nomenclature. This decision was made as the current naming retains the relevant analogies with the "naturally" linked glycosides.

1,3,4,6-Tetra-*O*-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranoside (2)

This well known compound was synthesised according to the procedure of Dasgupta and Garegg [141] from D-glucosamine hydrochloride **1**.

A dispersion of D-glucosamine hydrochloride (10.89 g, 50.1 mmol) in DMF (70 mL) containing triethylamine (17.2 mL, 123.4 mmol) and phthalic anhydride (11.25 g, 75.9 mmol) was stirred at r.t. for 20 min. followed by heating to 70 °C for 15 min. Anhydrous sodium acetate (~ 10 g) and acetic anhydride (100 mL) were quickly added to the flask and the mixture stirred at 100 °C for 5 h. The dark brown solution was then allowed to cool to r.t. and poured into 1 L of ice water. The product was extracted with dichloromethane (500 mL) and the organic phase washed with water and sodium hydrogen carbonate solution. After drying over sodium sulfate the product was subjected to column chromatography on silica eluting with Tol:EA 4:1 (v/v). The product was recrystallised from ethanol and the solids washed with ice-cold ethanol to remove the last traces of brown colour, giving the title compound **2** (4.71 g, 20 %) as a white powder.

(2) $C_{22}H_{23}NO_{11}$
R_f (Tol:EA 4:1): 0.24
Yield: 4.71 g, 20 %
white powder



^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 7.83-7.77 (m, 2H, phthaloyl), 7.71-7.66 (m, 2H, phthaloyl), 6.45 (d, 1H, $^3J_{1,2}$ = 8.9 Hz, H-1), 5.82 (dd, 1H, $^3J_{3,4}$ = 9.1 Hz, H-3), 5.14 (t, 1H, $^3J_{4,5}$ = 10.1 Hz, H-4), 4.40 (dd, 1H, $^3J_{2,3}$ = 10.6 Hz, H-2), 4.30 (dd, 1H, $^2J_{6a,6b}$ = 12.5 Hz, $^3J_{6a,5}$ = 4.5 Hz, H-6a), 4.10 (dd, 1H, $^2J_{6b,6a}$ = 12.5 Hz, H-6b), 3.96 (m, 1H, H-5), 2.04, 1.97, 1.93, 1.79 (4s, 3 H each, 4 x $COCH_3$).

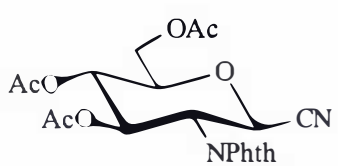
^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 169.61, 168.98, 168.44, 167.59, 166.36 (6C, 4 x $\underline{C}OCH_3$, 2 x CO phthaloyl), 133.49, 130.26, 122.79 (6C, phthaloyl), 88.81 (1C, C-1), 71.67 (1C, C-3), 69.54 (1C, C-4), 67.37 (1C, C-5), 60.57 (1C, C-6), 52.54 (1C, C-2), 19.72, 19.58, 19.37 (4C, 4 x $COCH_3$).

3,4,6-Tri-*O*-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl cyanide (3)

The title compound was synthesised using the procedure of Myers and Lee [107] from the well known compound *1,3,4,6-Tetra-O-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranoside* **2** [141].

To a stirred solution of **2** (4.66 g ,9.76 mmol) and TMSCN (2.45 mL, 19.52 mmol) in dry acetonitrile (40 mL) was added $BF_3{\cdot}OEt_2$ (48 % BF_3, 2.56 mL, 9.76 mmol) and stirring continued at r.t. for 2 h. The reaction was monitored by TLC ($CHCl_3$:MeOH 100:1). The starting material and product have very similar R_f values but stain differently using ethanolic sulphuric acid. After completion the solvent was evaporated, the residue taken up in chloroform (100 mL) and washed with chilled water (3x 25 mL). After drying over sodium sulfate the dark brown syrup was subjected to column chromatography on silica eluting with $CHCl_3$:MeOH 100:1 (v/v) to give **3** (2.74 g, 63 %) as a white foam.

(3) $C_{21}H_{20}N_2O_9$
R_f ($CHCl_3$:MeOH 100:1): 0.63
Yield: 2.74 g, 63 %
white foam, powder



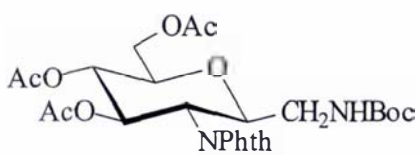
^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 7.93-7.87 (m, 2H, phthaloyl), 7.84-7.76 (m, 2H, phthaloyl), 5.75 (dd, 1H, $^3J_{3,4}$ = 9.3 Hz, H-3), 5.38 (d, 1H, $^3J_{1,2}$ = 10.84 Hz, H-1), 5.19 (t, 1H, $^3J_{4,5}$ = 10.00 Hz, H-4), 4.67 (t, 1H, $^3J_{2,3}$ = 10.6 Hz, H-2), 4.30 (dd, 1H, $^2J_{6a,6b}$ = 12.6 Hz, $^3J_{6a,5}$ = 4.8 Hz, H-6a), 4.19 (dd, 1H, $^2J_{6b,6a}$ = 12.6 Hz, $^3J_{6b,5}$ = 2.2 Hz, H-6b), 3.89 (m, 1H, H-5), 2.13, 2.04, 1.87 (3s, 3H each, 3 x $COCH_3$).

^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 170.88, 170.33, 169.54 (3C, 3 x $\underline{C}OCH_3$), 135.21, 131.44, 124.45 (6C, phthaloyl), 114.55 (1C, CN), 70.77 (1C, C-3), 68.25 (1C, C-4), 64.81 (1C, C-1), 61.93 (1C, C-6), 52.84 (1C, C-2), 21.05, 20.87, 20.66 (3C, 3 x $CO\underline{C}H_3$).

N-tert-Butoxycarbonyl-(3,4,6-tri-_O_-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)methylamine (4b)

Triethylamine (4.37 mL, 10 eq.), di-*tert*-butyl-dicarbonate (2.06 g, 3 eq.) and 10 % Pd/C (700 mg) were added sequentially to a solution of glycopyranosylcyanide **3** [107], (1.4 g, 1 eq.) in EtOH:THF (5:3, 48 mL). The mixture was stirred under an atmosphere of hydrogen at 30 °C for 6 h. The reaction mixture was filtered, the filtrate evaporated and the residue taken up in dichloromethane. The organic phase was washed with sat. sodium hydrogen carbonate solution, brine and water, dried over sodium sulfate and then subjected to column chromatography using PE:EA 1:1 (v/v) as the eluent to give **4b** (1.29g,75 %) as a white foam.

(4b) $C_{26}H_{32}N_2O_{11}$
R_f (PE:EA 1:1): 0.39
Yield: 1.29 g, 75 %
white foam/powder



^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 7.83-7.71 (m, 4H, phthaloyl), 5.80 (t, 1H, $^3J_{3,2}$ = 9.7 Hz, $^3J_{3,4}$ = 9.8 Hz, H-3), 5.15 (t, 1 H, $^3J_{4,3}$ = 9.8 Hz, $^3J_{4,5}$ = 9.5 Hz, H-4), 4.84 (m, 1H, NH), 4.51 (m, 1H, H-1), 4.34 (t, 1H, $^3J_{2,3}$ = 9.7 Hz, H-2), 4.30 (dd, 1H, $^2J_{6a,6b}$ = 12.4 Hz, $^3J_{6a,5}$ = 4.8 Hz, H-6a), 4.14 (dd, 1H, $^2J_{6b,6a}$ = 12.4 Hz, $^3J_{6b,5}$ = 2.3 Hz, H-6b),

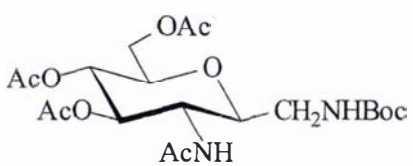
3.84 (ddd, 1H, $^{3}J_{5,6a}$ = 4.8 Hz, $^{3}J_{5,6b}$ = 2.3 Hz, $^{3}J_{5,4}$ = 9.5 Hz, H-5), 3.33-3.30 (m, 2H, CH_2N), 2.12, 2.03, 1.83 (3s, 3H each, $COCH_3$), 1.39 (s, 9H, Boc).
^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 171.05, 170.56, 169.87 (3C, 3 x $\underline{C}OCH_3$), 155.84 (1C, CO, Boc), 134.66, 131.76, 124.00 (6C, phthaloyl), 80.07 (1C, $\underline{C}(CH_3)_3$, Boc), 76.34 (1C, C-5), 74.91 (1C, C-1), 72.09 (1C, C-3), 69.42 (1C, C-4), 62.67 (1C, C-6), 51.75 (1C, C-2), 41.48 (1C, CH_2N), 28.65 (3C, $C(\underline{C}H_3)_3$, Boc), 21.14, 20.99, 20.77 (3C, 3 x $CO\underline{C}H_3$).

Microanalysis calculated for $C_{26}H_{32}N_2O_{11}$ (0.25 H_2O): C, 56.46; H, 5.92; N, 5.07. Found: C, 56.41; H, 6.11; N, 4.85.

N-tert-Butoxycarbonyl-(2-acetamido-3,4,6-tri-_O_-acetyl-2-deoxy-β-D-glucopyranosyl)methylamine (4c)

According to the method of Wang *et al.* [5] hydrazine-monohydrate (2.8 mL, 58 mmol) was added to a solution of **4b** (1.22 g, 2.23 mmol) in ethanol (20 mL). The mixture was refluxed at 80 °C for 2 h during which time a white precipitate developed. After evaporation of the solvent and azeotroping with toluene the residue was dissolved in pyridine (30 mL) and acetic anhydride (20 mL). The mixture was stirred at r.t. for 12 h. The volume of the reaction mixture was reduced by rotary evaporation (1/3) and then the mixture was poured into ice water (40 mL) with vigorous stirring. Stirring was continued for 1 h before the solvent was removed by rotary evaporation. Chloroform (100 mL) was added to the residue to give a suspension which was washed with water to remove the white flaky material. The organic phase was washed successively with 0.1 M hydrochloric acid, sat. sodium hydrogen carbonate solution and water, dried over sodium sulfate and the subjected to column chromatography on silica eluting with $CHCl_3$:EtOH 35:1 (v/v) to give the title compound **4c** (915 mg, 89 %) as a white foam.

(4c) $C_{20}H_{32}N_2O_{10}$
R_f ($CHCl_3$: EtOH 35:1): 0.07
Yield: 915 mg, 89 %
white foam/powder



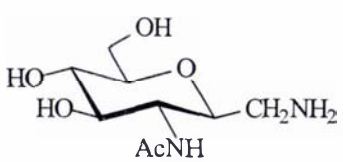
^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 6.28 (d, 1 H, $^3J_{NH,2}$ = 8.9 Hz, NHAc), 5.18-5.02 (m, 3 H, H-3, H-4, NHBoc), 4.24 (dd, 1 H, $^2J_{6a,6b}$ = 12.3 Hz, $^3J_{6a,5}$ = 4.6 Hz, H-6a), 4.11 (dd , 1H, $^2J_{6b,6a}$ = 12.3 Hz, $^3J_{6b,5}$ = 2.2 Hz, H-6b), 3.98 (q, 1H, $^3J_{2,3 \text{ and } 2,1}$ = 9.5-10 Hz, H-2), 3.68-3.47 (m, 3H, H-1, H-5, ½ CH_2N), 3.09-3.01 (m, 1H, ½ CH_2N), 2.08, 2.03, 2.02 and 1.95 (4s, 3H each, 4 x $COCH_3$), 1.44 (s, 9H, Boc).

^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 172.06, 171.58, 171.02, 169.69 (4C, 4 x $COCH_3$), 156.29 (1C, CO, Boc), 79.91 (1C, C-1), 75.94 (1C, C-5), 74.34 (1C, C-3), 68.98 (1C, C-4), 62.65 (1C, C-6), 51.98 (1C, C-2), 42.05 (1C, CH_2), 28.70 (3C, $C(CH_3)_3$, Boc), 23.44, 21.03, 21.01, 20.91 (4C, 4 x $COCH_3$).

(2-Acetamido-2-deoxy-β-D-glucopyranosyl)methylamine (4d)

Hydrochloric acid in 1,4-dioxane (4 M, 7 mL) was added to a solution of compound **4c** (831 mg, 1.81 mmol) in dichloromethane (20 mL). The mixture was stirred at r.t. for 30 min. and the solvent evaporated by rotary evaporation. The residue was azeotroped with toluene, taken up in methanol (30 mL) and the mixture saturated with ammonia at 0 °C. The flask was stoppered and stirring continued at r.t. for 24 h. After rotary evaporation of the solvent, a light yellow syrup (366 mg, 87 %) was obtained.

(4d) $C_9H_{18}N_2O_5$
Yield: 366 mg, 87 %
positive ninhydrin stain
light yellow syrup



^{1}H NMR (300 MHz, MeOH-d_4): δ ppm: 3.82 (dd, 1H, $^2J_{6a,6b}$ = 11.7 Hz, $^3J_{6a,5}$ = 1.8 Hz, H-6a), 3.59 (dd, 1H, $^2J_{6b,6a}$ = 11.7 Hz, $^3J_{6b,5}$ = 5.6 Hz, H-6b), 3.47-3.38 (m, 3 H, H-2, H-3, H-1), 3.04 (dd, 1H, $^2J_{CH,CH'}$ = 13.4 Hz, $^3J_{CH,1}$ = 2.1 Hz, CH), 2.85 (dd, 1H, $^2J_{CH',CH}$ = 13.4 Hz, $^3J_{CH',1}$ = 6.6 Hz), 1.92 (s, 3H, $COCH_3$).

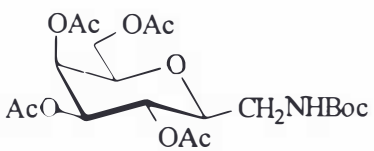
^{13}C NMR (300 MHz, MeOH-d_4): δ ppm: 175.16 (1C, $\underline{C}OCH_3$), 82.06 (1C, C-5), 76.66 (1C, C-3), 76.24 (1C, C-1), 72.40 (1C, C-4), 63.13 (1C, C-6), 54.69 (1C, C-2), 42.92 (1C, CH_2), 23.05 (1C, $CO\underline{C}H_3$).

ES-MS: m/*z* for $C_9H_{18}N_2O_5$ $(M+H)^+$: 235.0; calculated: 235.3.

N-tert-Butoxycarbonyl-(2,3,4,6-tetra-_O_-acetyl-β-D-galactopyranosyl) methylamine (25)

Triethylamine (270 µL, 10 eq.), di-*tert*-butyl-dicarbonate (132 mg, 3 eq.) and 10 % Pd/C (80 mg) were added sequentially to a solution of glycopyranosylcyanide **24** [107], (72 mg, 0.2 mmol, 1 eq.) in EtOH:THF 5:3 (8 mL). The mixture was stirred under an atmosphere of hydrogen at 30 °C for 6 h (att: starting material and product show the same R_f in PE:EA 1:1 as well as in $CHCl_3$:EtOH 35:1). The reaction mixture was filtered, the filtrate evaporated and the residue taken up in dichloromethane. The organic phase was washed with sat. sodium hydrogen carbonate solution, brine and water, dried over sodium sulfate and then subjected to column chromatography using PE:EA 1:1 (v/v) as the mobile phase to give **25** (77 mg , 83 %) as a white foam.

(25) $C_{20}H_{31}NO_{11}$
R_f (PE:EA 1:1): 0.61
Yield: 72 mg, 83 %
white foam/powder



^1H NMR (300 MHz, $CDCl_3$): δ ppm: 5.35 (dd, 1H, $^3J_{4,5}$ = 1.0 Hz, $^3J_{4,3}$ = 3.2 Hz, H-4), 5.05 (m, 1H, H-2), 4.96 (dd, 1H, $^3J_{3,4}$ = 3.2 Hz, $^3J_{3,2}$ = 10.1 Hz H-3), 4.84 (m, 1H, NH), 4.08 (dd, 1H, $^2J_{6a,\,6b}$ = 11.1 Hz, $^3J_{6a,5}$ = 6.7 Hz, H-6a), 3.99 (dd, 1H, $^2J_{6b,6a}$ = 11.1 Hz, $^3J_{6b,5}$ = 6.6 Hz, H-6b), 3.82 (dt, 1H, $^3J_{5,6a}$ = 6.7 Hz , $^3J_{5,6b}$ = 6.6 Hz, $^3J_{5,4}$ = 1.0 Hz , H-5), 3.48-3.38 (m, 2H, H-1, ½ CH_2), 3.18-3.07 (m, 1H, ½ CH_2), 2.09, 1.99, 1.98, 1.90 (4s, 3H each, $COCH_3$), 1.38 (s, 9H, Boc).

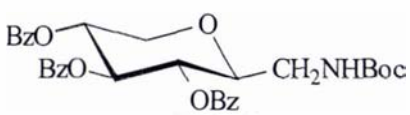
^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 169.36, 169.17, 169.12, 168.78 (5C, 5 x $\underline{C}OCH_3$), 154.64 (1C,CO Boc), 76.10 (1C, C-1 under $CDCl_3$ multiplett), 72.19 (1C, C-5), 70.98 (1C, C-3), 66.64 (1C, C-4), 65.90 (1C, C-2), 60.46 (1C, C-6), 40.35 (1C, CH_2), 27.37 (3C, $C(\underline{C}H_3)_3$, Boc), 19.65, 19.57 (4C, 4 x $CO\underline{C}H_3$).

Microanalysis calculated for $C_{20}H_{31}NO_{11}$: C, 52.06; H, 6.77; N, 3.04. Found: C, 51.91; H, 6.99; N, 3.00.

N-tert-Butoxycarbonyl-(2,3,4-tri-O-benzoyl-β-D-xylopyranosyl)methylamine (27)

Triethylamine (0.53 mL, 10 eq.), di-*tert*-butyl-dicarbonate (250 mg, 3 eq.) and 10 % Pd/C (100 mg) were added sequentially to a solution of glycopyranosylcyanide **26** (180 mg, 0.38 mmol, 1 eq.) in EtOH (8 mL). The mixture was stirred under an atmosphere of hydrogen at 30 °C for 6 h. The reaction mixture was filtered, the filtrate evaporated and the residue taken up in dichloromethane. The organic phase was washed with sat. sodium hydrogen carbonate solution, brine and water, dried over sodium sulfate and then subjected to column chromatography using PE:EA 1:1 (v/v) as the mobile phase to give **27** (111 mg, 51 %) as a colourless syrup.

(27) $C_{32}H_{33}NO_9$
R_f (PE:EA 2:1): 0.48
Yield: 111 mg, 51 %
colourless syrup



^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 7.98-7.85 (m, 6H, benzoyl), 7.55-7.26 (m, 9H, benzoyl), 5.87 (t, 1H, ^3J = 9.6 Hz, H-3), 5.42-5.34 (m, 2H, H-4, H-2), 5.02 (m, 1H, NH), 4.41 (q, 1H, $^3J_{5a,4}$ = 5.6 Hz, H-5a), 3.75-3.70 (m, 1H, H-1), 3.65-3.61 (m, 1H, ½ CH_2), 3.54 (t, 1H, $^2J_{5b,5a}$ = 10.8 Hz, H-5b), 3.22-3.14 (m, 1H, ½ CH_2), 1.43 (s, 9H, Boc).

^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 133.78, 133.55, 130.25, 130.19, 130.07, 129.42, 128.82, 128.68 (15 C, benzoyl), 77.39 (1C under $CDCl_3$- multiplett, C-1), 74.17 (1C,

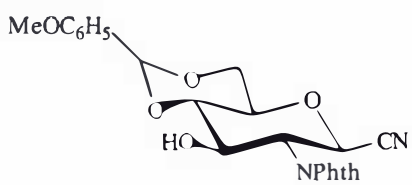
C-3), 70.64 (2C, C-2, C-4), 67.47 (1C, C-5), 42.75 (2C, CH_2), 28.72 (3C, $C(\underline{C}H_3)_3$, Boc).

Microanalysis calculated for $C_{32}H_{33}NO_9$: C, 66.77; H, 5.78; N, 2.43. Found: C, 66.82; H, 6.02; N, 2.70.

4,6-*O*-*p*-Methoxy-benzylidene-2-deoxy-2-phthalimido-β-D-glucopyranosyl cyanide (28b)

Glycopyranosylcyanide **3** [107], (2.74 g, 6.17 mmol) was treated with a solution of *N*-methylpyrrolidine (2.5 % in methanol, 100 mL) at r.t. The mixture was stirred for 12 h to give the corresponding deacetylated glycopyranosylcyanide **28a** after rotary evaporation of the solvent [R_f(EA:MeOH 9:1): 0.65]. The crude triol was obtained as a creamy white residue which was used without further purification in the synthesis of **28b**. To **28a** (1.96 g , 6.17 mmol) in dimethylformamide (20 mL) was added 4-methoxy-benzaldehyde dimethylacetal (1.57 mL, 1.5 eq.) and the solution acidified until the pH was below 3 using *p*-toluenesulfonic acid. The resulting mixture was stirred under reduced pressure at 50 °C for 2 h using a rotary evaporator, then neutralised with triethylamine. The solvent was removed by rotary evaporation and the residue taken up in ethyl acetate (50 mL). After washing sequentially with sat. sodium hydrogen carbonate solution, brine and chilled water the residue was dried over sodium sulfate and then subjected to column chromatography on silica using PE:EA 1:1 (v/v) as the eluent to give **28b** (1.97 g, 73 %) as a white foam. The ^{1}H-NMR and ^{13}C-NMR still showed minor impurities but the compound was sufficiently clean for the synthesis of intermediate **29**.

(28b) $C_{23}H_{20}N_2O_7$
R_f (PE:EA 1:1): 0.57
Yield: 1.97 g, 73 %
white foam/powder

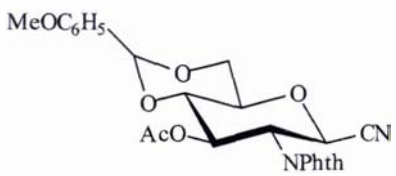


^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 7.84-7.81 (m, 2H, phthaloyl), 7.71-7.68 (m, 2H, phthaloyl), 7.38 (d, 2H, $^3J_{CH2\ ortho,\ CH2\ meta}$ = 8.7 Hz, 2 x CH, benzylidene), 6.85 (d, 2H, $^3J_{CH2\ meta,CH2\ ortho}$ = 8.7 Hz, 2 x CH, benzylidene), 5.49 (s, 1H, $MeOC_6H_5C\underline{H}$, benzylidene), 5.10 (d, 1H, $^3J_{1,2}$ = 10.5 Hz, H-1), 4.55 (m, 2H, H-2, H-3), 4.31 (dd, 1H, $^2J_{6a,6b}$ =10.3 Hz, $^3J_{6a,5}$ = 5.9 Hz, H-6a), 3.76 (s, 3H, OMe), 3.72- 3.69 (m, 1H, H-6b), 3.60-3.54 (m, 2H, H-5, H-4).

^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 167.93 (2C, 2 x CO, phthaloyl), 134.98, 131.65, 124.22 (6C, phthaloyl), 128.02, 114.07 (4C, 4 x CH, benzylidene), 102.20 (1C, $MeOC_6H_5\underline{C}H$), 81.63 (1C, C-5), 71.71 (1C, C-4), 68.70 (1C, C-3), 68.28 (1C, C-6), 65.31 (1C, C-1), 55.64 (1C, OMe), 54.71 (1C, C-2).

3-*O*-Acetyl-4,6-*O*-*p*-methoxy-benzylidene-2-deoxy-2-phthalimido-β-D-glucopyranosyl cyanide (29)

Compound **28b** (1.97 g, 4.5 mmol) was dissolved in pyridine (50 mL) and acetic anhydride (50 mL) and stirred at r.t. for 12 h to give crude *3-O-acetyl-4,6-O-p-methoxy-benzylidene-2-deoxy-2-phthalimido-β-D-glucopyranosyl cyanide* **29** (2.0 g, 92 %) which, after azeotroping with toluene, was used in the synthesis of **31a**. The compound could be purified by chromatography on silica using PE:EA 1:1 (v/v) as the eluent. However, this was not necessary for the next reaction step. It should be noted that the title compound is very insoluble in both petroleum ether:ethyl acetate mixtures as well as in chloroform, a fact that renders the column purification rather tedious.



(29) $C_{25}H_{22}N_2O_8$
R_f (PE:EA 1:1): 0.62
Yield: 92 % (crude)
brownish syrup/foam (white powder after column purification)

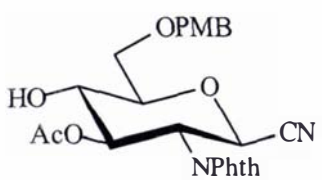
^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 7.90-7.85 (m, 2H, phthaloyl), 7.78-7.73 (m, 2H, phthaloyl), 7.37 (d, 2H, $^3J_{CH\ ortho,\ CH\ meta}$ = 8.7 Hz, 2 x CH, benzylidene), 6.88 (d, 2H, $^3J_{CH\ meta,\ CH\ ortho}$ = 8.7 Hz, 2 x CH, benzylidene), 5.82 (t, 1H, $^3J_{3,4}$ = 9.4 Hz, H-3), 5.50 (s, 1H, $MeOC_6H_5C\underline{H}$, benzylidene), 5.43 (d, 1H, $^3J_{1,2}$ = 10.9 Hz, H-1), 4.65 (t, 1H, $^3J_{2,3}$ = 10.4 Hz, H-2), 4.42 (m, 1H, H-6a), 3.81-3.73 (m, 6H, H-4, H-5, H-6b, OMe), 1.89 (s, 3H, $COCH_3$).

^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 170.44 (1C, $\underline{C}OCH_3$), 167.05 (2 C, 2 x CO, phthaloyl), 135.17, 129.35, 124.40 (6C, phthaloyl), 127.96, 114.06 (4C, 4 x CH, benzylidene), 114.91 (1C, CN), 102.20 (1C, $OMeC_6H_5\underline{C}H$), 78.83 (1C, C-5), 71.81 (1C, C-4), 69.72 (1C, C-3), 68.36 (1C, C-6), 65.13 (1C, C-1), 55.89 (1C, OMe), 53.33 (1C, C-2), 20.79 (1C, $CO\underline{C}H_3$).

3-*O*-Acetyl-6-*O*-*p*-methoxybenzyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl cyanide (30)

To a solution of **29** (320 mg, 0.66 mmol) in dimethylformamide (10 mL) over molecular sieves (3 Å) at 0 °C was added neat sodium cyanoborohydride (420 mg, 10eq.) followed by the slow addition of TFA (1 mL, 20 eq.) and the mixture was warmed to r.t. The reaction was monitored using TLC (PE:EA, 1:1). After completion, sat. sodium hydrogen carbonate solution was added, the molecular sieves filtered off and the product extracted with dichloromethane. After drying over sodium sulfate the solvent was removed *in vacuo* and the residue subjected to column chromatography on silica eluting with PE:EA 1:1 (v/v). Title compound **30** was obtained as a colourless syrup (160 mg, 50 %).

(30) $C_{25}H_{24}N_2O_8$
R_f (PE:EA 1:1): 0.38
Yield: 160 mg, 50 %
colourless syrup



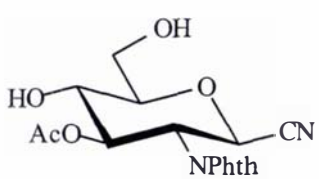
^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 7.90-7.83 (m, 2H, phthaloyl), 7.77-7.69 (m, 2H, phthaloyl), 7.29 (d, 2H, $^3J_{CHortho,CH\ meta}$ = 8.6 Hz, 2 x CH, benzyl), 6.90 (d, 2H, $^3J_{CH\ meta,CH\ ortho}$ = 8.6 Hz, 2 x CH, benzyl), 5.61 (dd, 1H, $^3J_{3,2}$ = 10.3 Hz, $^3J_{3,4}$ = 8.6 Hz, H-3), 4.60-4.50 (m, 3H, CH_2 benzyl, H-2), 3.85-3.71 (m, 7H, OMe, H-4, H-5, H-6a, H-6b), 1.91 (s, 3H, $COCH_3$).

^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 171.29 (1C, $\underline{C}OCH_3$), 135.09, 131.51, 124.33 (6C, phthaloyl), 129.87 (2C, benzyl), 115.14 (1C, CN), 114.33 (2C, benzyl), 80.04 (1C, C-5), 73.89 (1C, CH_2 benzyl), 73.39 (1C, C-3), 70.29 (1C, C-4), 69.32 (1C, C-6, 64.78 (1C, C-1), 55.68 (1C, OMe), 52.91 (1C, C-2), 20.91 (1C, $CO\underline{C}H_3$).

3-*O*-Acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl cyanide (31a)

To **29** (2.0 g ,4.18 mmol) in acetonitrile:water (9:1, 30 mL), was added DDQ (190 mg, 0.2 eq.) and the deep-red solution was stirred at 45 °C for 1 h. The reaction mixture was diluted with ethyl acetate (80 mL) and washed with a 1:1 mixture of sat. sodium hydrogen carbonate solution and water. The aqueous phase was extracted with ethyl acetate (2x) and the combined organic phases dried over sodium sulfate. The crude product was purified by column chromatography on silica eluting with PE:EA 1:3 (v/v) to give **31a** (1.29 g, 85 %) as a white foam.

(31a) $C_{17}H_{16}N_2O_7$
R_f (PE:EA 1:3): 0.33
Yield: 58 % over 4 steps (from **3**)
whitefoam/powder



^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 7.93-7.76 (2m, 4H, phthaloyl), 5.61 (dd, 1H, $^3J_{3,4}$ = 9.0 Hz, H-3), 5.39 (d, 1H, $^3J_{1,2}$ = 10.7 Hz, H-1), 4.57 (t, 1H, $^3J_{2,3}$ = 10.3 Hz, H-2), 4.01 (dd, 1H, $^2J_{6a,6b}$ = 12.3 Hz, $^3J_{6a,5}$ = 2.9 Hz, H-6a), 3.93-3.85 (m, 2H, H-6b, H-4), 3.69-3.63 (m, 1H, H-5), 1.95 (s, 3H, $COCH_3$).

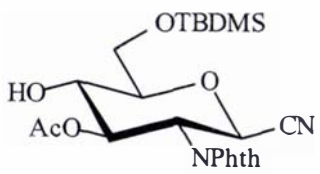
^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 171.58 (1C, $\underline{C}OCH_3$), 167.80 (2C, CO, phthaloyl), 135.23, 131.45, 124.41 (6C, phthaloyl), 155.31 (1C, CN), 81.19 (1C, C-5), 73.39 (1C, C-3), 69.26 (1C, C-4), 64.74 (1C, C-1), 62.05 (1C, C-6), 52.96 (1C, C-2), 20.92 (1C, $CO\underline{C}H_3$).

Microanalysis calculated for $C_{17}H_{16}N_2O_7$ (0.5 H_2O): C, 55.28; H, 4.64; N, 7.56. Found: C, 55.05; H, 4.59; N, 7.51.

3-*O*-Acetyl-6-*O*-*tert*-butyl-dimethylsilyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl cyanide (31b)

Imidazole (605 mg, 2.5 eq.) and TBDMSCl (589 mg, 1.1 eq.) were added to a solution of **31a** in DMF (10 mL). The mixture was stirred at r.t for 1 h. The solvent was evaporated and the residue taken up in dichloromethane (50 mL) and washed with brine and water. The aqueous phase was extracted with dichloromethane (2x) and the combined organic phases dried over sodium sulfate. The crude product was purified by column chromatography on silica using PE:EA 1:1 (v/v) as the eluent to give **31b** (1.52 g, 90 %) as a white foam.

(31b) $C_{23}O_7H_{30}N_2Si$
R_f (PE:EA 1:1): 0.6
Yield: 1.52 g, 90 %
white foam/colourless syrup



^1H NMR (300 MHz, $CDCl_3$): δ ppm: 7.77-7.62 (2m, 4H, phthaloyl), 5.49 (t, 1H, $^3J_{3,4}$ = 9.2 Hz, H-3), 5.23 (d, 1H, $^3J_{1,2}$ = 10.8 Hz, H-1), 4.39 (t, 1H, $^3J_{2,3}$ = 10.1 Hz, H-2), 3.88 (dd, 1H, $^2J_{6a,6b}$ = 10.9 Hz, $^3J_{6a,5}$ = 4.5 Hz, H-6a), 3.80-3.70 (m, 2H, H-6b, H-4), 3.52-3.46 (m, 1H, H-5), 1.80 (s, 3H, OAc), 0.79 (s, 9H, *tert*-butyl), 0.00 (s, 3H, Si-CH_3), -0.01 (s, 3H, Si-CH_3).

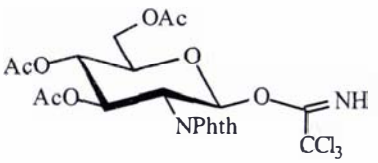
^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 171.19 (1C, $\underline{C}OCH_3$), 135.08, 131.55, 124.34 (6C, phthaloyl), 115.08 (1C, CN), 79.82 (1C, C-5), 73.27 (1C, C-3), 71.32 (1C, C-4), 64.68 (1C, C-1), 64.08 (1C, C-6), 52.86 (1C, C-2), 26.22 (3C, $C(\underline{C}H_3)_3$, *t*-butyl), 20.9 (1C, $CO\underline{C}H_3$), 18.67 (1C, $\underline{C}(CH_3)_3$, *t*-butyl), -5.05 (2 C, $Si(CH_3)_2$).

Microanalysis calculated for $C_{23}H_{30}N_2O_7Si$ (0.25 H_2O): C, 57.66; H, 6.42; N, 5.84. Found: C, 57.45; H, 6.32; N, 5.79.

_O_-(3,4,6-Tri-_O_-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl) trichloroacetimidate (32b)

The donor **32b** [109] was synthesised in a two step procedure from *1,3,4,6-tetra-O-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranoside* **2** [141]. Compound **2** (5.96 g, 12.5 mmol) in dimethylformamide (30 mL) was treated with hydrazinium acetate (1.27 g, 1.1 eq.) at r.t. The reaction was complete after dissolution of the hydrazinium acetate. The mixture was diluted with water (50 mL) and extracted with ethyl acetate (4x 50 mL), the combined organic phases dried over sodium sulfate and azeotroped with toluene to give the intermediate *3,4,6-tri-O-Acetyl-2-deoxy-2-phthalimido-β-D-glucopyranoside* **32a** . The intermediate (4.91 g) was then treated with trichloroacetonitrile (11.22 mL, 10 eq.) and DBU (0.16 mL, 0.1 eq.) in dichloromethane (40 mL) at 0 °C. The mixture was warmed to r.t. and stirring continued for 2 h. The solvent was evaporated and the donor purified by column chromatography using PE:EA 1:1 (v/v) as the eluent giving **32b** (4.35 g, 67 %) as a white foam which could then be crystallised from diethyl ether.

(32b) $C_{22}Cl_3H_{21}N_2O_{10}$
R_f (PE:EA 1:1): 0.47
Yield: 4.35g, 67 %
white foam, powder



^1H NMR (300 MHz, $CDCl_3$): δ ppm: 8.66 (s, 1H, NH), 7.92-7.81 (m, 2H, phthaloyl), 7.77-7.70 (m, 2H, phthaloyl), 6.63 (d, 1H, $^3J_{1,2}$ = 8.9 Hz, H-1), 5.93 (dd, 1H, $^3J_{3,2}$ =

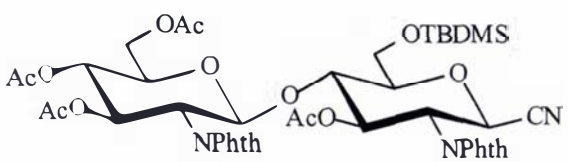
10.7 Hz, $^{3}J_{3,4}$ = 9.1 Hz, H-3), 5.28 (dd, 1H, $^{3}J_{4,3}$ = 9.1 Hz, $^{3}J_{4,5}$ = 10.1 Hz, H-4), 4.64 (dd, 1H, $^{3}J_{2,1}$ = 8.9 Hz, $^{3}J_{2,3}$ =10.7 Hz, H-2), 4.41 (dd, 1H, $^{2}J_{6a,6b}$ = 12.5 Hz, $^{3}J_{6a,5}$ = 4.3 Hz, H-6a), 4.21 (dd, 1H, $^{2}J_{6b,6a}$ = 12.5 Hz, $^{3}J_{6b,5}$ = 2.2 Hz, H-6b), 4.08 (ddd, 1H, $^{3}J_{5,4}$ = 10.1 Hz, H-5), 2.12, 2.05 and 1.89 (3s, 3H each, OAc).

^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 171.03, 170.43, 169.78 (3C, 3 x $\underline{C}OCH_3$), 160.96 (1C, C=NH, imine), 134.80, 131.64, 124.06 (6C, phthaloyl), 94.00 (1C, C-1), 73.22 (1C, C-5), 70.85 (1C, C-3), 68.89 (1C, C-4), 61.97 (1C, C-6), 53.99 (1C, C-2), 21.11, 20.97, 20.80 (3C, 3 x $\underline{C}OCH_3$).

4-*O*-(3,4,6-Tri-*O*-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-3-*O*-acetyl-6-*O*-*tert*-butyl-dimethylsilyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl cyanide (33)

Trichloroacetimidate **32b** (2.18 g, 3.78 mmol, 1.2 eq.) and the acceptor **31b** (1.49 g, 3.15 mmol, 1.0 eq.) were combined and azeotroped twice with toluene. The mixture was dissolved in dichloromethane (20 mL) at 0 °C and TMSOTf (~ 60 μL, 0.1 eq.) was added with vigorous stirring. The pink mixture was kept at 0 °C for 1 h (precipitate of trichloroacetamide was observed) and then neutralised with triethylamine (~ 40 μL, the colour changes from pink to light brown/green). The solvent was evaporated and the product purified by column chromatography on silica eluting with PE:EA 1:1 (v/v) to give **33** (2.4 g, 90 %) as a white foam.

(33) $C_{43}H_{49}N_3O_{16}Si$
R_f (PE:EA 1:1): 0.43
Yield: 2.4 g, 90 %
white foam/powder



^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 7.80-7.74 (m, 4H, phthaloyl), 7.69-7.64 (m, 4H, phthaloyl), 5.65 (dd, 1H, $^{3}J_{3',2'}$ = 10.5 Hz, $^{3}J_{3',4'}$ = 9.1 Hz, H-3'), 5.52 (dd, 1H, $^{3}J_{3,2}$ = 10.1 Hz, $^{3}J_{3,4}$ = 9.3 Hz, H-3), 5.41 (d, 1H, $^{3}J_{1',2'}$ = 8.4 Hz, H-1'), 5.12 (d, 1H, $^{3}J_{1,2}$ =

10.9 Hz, H-1), 5.01 (t, 1H, $^{3}J_{4',5'}$ = 10.1 Hz, H-4'), 4.39-4.30 (m, 2H, H-2, H-6'a), 4.14-3.96 (m, 3H, H-2', H-4, H-6'b), 3.71 (m, 1H, H-5'), 3.62-3.58 (m, 1H, H-6a), 3.37 (dd, 1H, $^{2}J_{6b,6a}$ = 12.3 Hz, $^{3}J_{6b,5}$ = 2.8 Hz, H-6b), 3.28-3.24 (m, 1H, H-5), 2.00, 1.91, 1.84 and 1.73 (4s, 3H each, $COCH_3$), 0.86 (s, 9H, tert-butyl), 0.02 (s, 3H, Si-CH_3), 0.00 (s, 3H, Si-CH_3).

^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 170.82, 170.40, 170.35, 169.81 (4C, 4 x $\underline{C}OCH_3$), 167.71 (4C, 4 x CO, phthaloyl), 135.07, 134.85, 131.65, 131.49, 124.30, 124.05 (12C, 2 x phthaloyl), 115.05 (1C, CN), 97.13 (1C, C-1'), 80.45 (1C, C-5'), 72.82 (1C, C-4), 72.21 (1C, C-5), 71.01 (1C, C-3'), 70.63 (1C, C-3), 69.02 (1C, C-4'), 64.41 (1C, C-1), 62.12 (1C, C-6'), 61.01 (1C, C-6), 55.26 (1C, C-2'), 53.30 (1C, C-2), 26.24 (3C, $C(\underline{C}H_3)_3$, t-butyl), 21.07, 20.98, 20.93, 20.72, (4C, 4 x $CO\underline{C}H_3$), 18.67 (1C, $\underline{C}(CH_3)_3$, t-butyl), -4.89 (2C, $Si(CH_3)_2$).

Microanalysis calculated for $C_{43}H_{49}N_3O_{16}Si$ (1 H_2O): C, 56.76; H, 5.65; N, 4.62. Found: C, 56.57; H, 5.49; N, 4.47.

4-*O*-(3,4,6-Tri-*O*-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-3,6-di-*O*-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl cyanide (34)

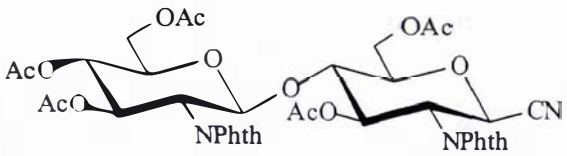
A solution of compound **33** (1.95 g, 2.19 mmol) in a mixture of glacial acetic acid (25 mL) and water (5 mL) was stirred at 80 °C for 6 h. After cleavage of the TBDMS-group, the solvent was evaporated and the product azeotroped with toluene. The residue was dissolved in a mixture of pyridine (15 mL) and acetic anhydride (15 mL) and stirred at r.t. for 12 h. After azeotroping with toluene the crude product was purified by column chromatography on silica eluting with PE:EA 1:1 (v/v) to give **34** (1.56 g, 81 %) as a white foam.

(34) $C_{39}H_{37}N_3O_{17}$

R_f (PE:EA 1:1): 0.25

Yield: 1.56 g, 81 %

white foam/powder



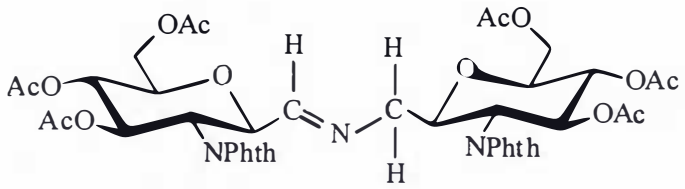
^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 7.89-7.83 (m, 4H, phthaloyl), 7.79-7.75 (m, 4H, phthaloyl), 5.75-5.68 (m, 2H, H-3, H-3'), 5.46 (d, 1H, $^3J_{1',2'}$ = 8.5 Hz, H-1'), 5.28 (d, 1H, $^3J_{1,2}$ = 10.9 Hz, H-1), 5.14 (t, 1H, $^3J_{4',5'}$ = 9.9 Hz, H-4'), 4.51 (t, 1H, $^3J_{2,3}$ = 10.5 Hz, H-2), 4.41 (dd, 1H, $^2J_{6'a,6'b}$ = 12.5 Hz, $^3J_{6'a,5'}$ = 4.2 Hz, H-6'a), 4.35-4.31 (m, 1H, H-6a), 4.25 (dd, 1H, $^3J_{2',3'}$ = 10.6 Hz, H-2'), 4.10 (dd, 1H, $^2J_{6'b,6'a}$ = 12.5 Hz, $^3J_{6'b,5'}$ = 2.3 Hz, H-6'b), 3.97 (t, 1H, $^3J_{4,5}$ = 9.5 Hz, H-4), 3.83 (m, 1H, H-5'), 3.74-3.67 (m, 2H, H-5, H-6b), 2.11, 2.00, 1.99, 1.92, 1.82 (5s, 3H each, $COCH_3$).

^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 170.87, 170.42, 170.20, 170.09, 169.70 (5C, 5 x $\underline{C}OCH_3$), 167.90 (4C, 4 x CO, phthaloyl), 135.15, 134.95, 131.53, 124.36, 124.26 (12C, phthaloyl), 114.70 (1C, CN), 97.67 (1C, C-1'), 75.09 (1C, C-4), 72.45 (1C, C-5'), 70.87 (1C, C-3), 70.56 (1C, C-3'), 68.73 (1C, C-4'), 64.62 (1C, C-1), 62.02 (1C, C-6), 61.92 (1C, C-6'), 55.13 (1C, C-2'), 53.21 (1C, C-2), 21.05, 20.92, 20.68 (5C, 5 x $CO\underline{C}H_3$).

Microanalysis calculated for $C_{39}H_{37}N_3O_{17}$ (1 H_2O): C, 55.92; H, 4.69; N, 5.02. Found: C, 55.69; H, 4.60; N, 4.92.

_C_-(3,4,6-Tri-_O_-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-_N_-(3,4,6-tri-_O_-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosylmethyl)formamidine (35)

The aldimine was unintentionally obtained during the hydrogenation of nitrile **34** as described in chapter 3. The compound was clearly identified by its NMR data.

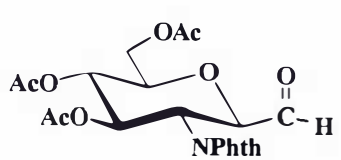


^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 7.91-7.73 (m, 8H, phthaloyl), 7.51 (d, 1H, $^3J_{1,CH=N}$ = 3.6 Hz, CH imine), 5.87-5.78 (m, 1H, H-3'), 5.70 (t, 1H, $^3J_{3,4}$ = 9.5 Hz, H-3), 5.15-5.09 (m, 1H, H-4'), 5.00 (t, 1H, $^3J_{4,5}$ = 9.8 Hz, H-4), 4.59 (dd, 1H, $^3J_{1,2}$ = 10.5 Hz), 4.40-4.26 (m, 2H, H-2, H-6a), 4.20-3.99 (m, 5H, H-6b, H-2', H-6'a, H-6'b, H-1'), 3.77-3.71 (m, 1H, H-5'), 3.63-3.59 (m, 1H, H-5), 3.37-3.34 (m, 1H, ½ CH_2), 3.17-3.10 (m, 1H, ½ CH_2), 2.12, 2.11, 2.03, 2.01,1.86, 1.81 (3H each, 6 x $COCH_3$).

^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 169.68, 169.20, 168.52, 166.27 (6C, 6x $\underline{C}OCH_3$), 162.75 (1C, CH=N), 133.83, 133.26, 130.54, 122.72 (12C, 2 x phthaloyl), 74.64 (1C, C-5'), 74.36 (1C, C-5), 73.41 (1C, C-1), 72.41 (1C, C-1'), 70.39 (1C, C-3), 70.23 (1C, C-3'), 68.02 (1C, C-4), 67.89 (1C, C-4'), 61.20 (2C, C-6, C-6'), 60.38 (1C, CH_2), 51.80 (1C, C-2'), 51.07 (1C, C-2), 19.77, 19.62, 19.45 (6C, 6 x $CO\underline{C}H_3$).

1-Formyl-3,4,6-tri-*O*-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranoside (36)

The aldehyde was unintentionally obtained from the imine **35** as described in chapter 3. The compound was clearly identified by its NMR data.



^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 9.54 (d, 1H, $^3J_{1,CHO}$ = 1.0 Hz, CHO), 7.80-7.77 (m, 2H, phthaloyl), 7.70-7.65 (m, 2H, phthaloyl), 5.79 (dd, 1H, $^3J_{3,4}$ = 9.3 Hz, H-3),

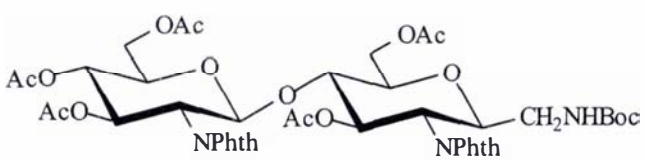
5.21 (t, 1H, $^{3}J_{4,5}$ = 9.8 Hz, H-4), 4.75 (dd, 1H, $^{3}J_{1,2}$ = 10.8 Hz, H-1), 4.38 (t, 1H, $^{3}J_{2,3}$ = 10.6 Hz, H-2), 4.28 (dd, 1H, $^{2}J_{6a,6b}$ = 12.5 Hz, $^{3}J_{6a,5}$ = 4.8 Hz, H-6a), 4.19 (dd, 1H, $^{2}J_{6b,6a}$ = 12.5 Hz, $^{3}J_{6b,5}$ = 2.1 Hz, H-6b), 3.93-3.87 (m, 1H, H-5), 2.06, 1.97, 1.79 (3s, 3H each, 3 x $COCH_3$).

^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 196.58 (1C, CHO), 169.62, 168.44, 166.53 (3C, 3 x $\underline{C}OCH_3$), 133.45, 130.32, 122.71 (6C, phthaloyl), 74.97 (1C, C-1), 74.23 (1C, C-5), 70.12 (1C, C-3), 67.56 (1C, C-4), 61.00 (1C, C-6), 49.52 (1C, C-2), 19.72, 19.57, 19.39 (3C, 3 x $CO\underline{C}H_3$).

N-tert-Butoxycarbonyl-4-_O_-(3,4,6-tri-_O_-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl)-(3,6-di-_O_-acetyl-2-deoxy-2-phthalimido-β-D-glucopyranosyl) methylamine (37)

Triethylamine (2.21 mL, 10 eq.), di-*tert*-butyl-dicarbonate (1.04 g, 3eq.) and 10 % Pd/C (600 mg) were added sequentially to a solution of glucopyranosylcyanide **34** (1.27 g, 1.54 mmol, 1 eq.) in a mixture of EtOH (25 mL) and THF (15 mL). The mixture was stirred under an atmosphere of hydrogen at 30 °C for 6 h. The progress of the reaction was monitored by TLC ($CHCl_3$:MeOH 35:1). After completion the mixture was filtered and and the solvent removed from the filtrate by rotary evaporation. The residue was taken up in dichloromethane (100 mL), washed successively with sat. sodium hydrogen carbonate solution, brine, water, dried over sodium sulfate and subjected to column chromatography on silica using PE:EA 1:1 (v/v) as the eluent to give title compound **37** as a white foam (1.04 g, 73 %).

(37) $C_{44}H_{49}N_3O_{19}$
R_f ($CHCl_3$:MeOH, 35:1): 0.64
R_f (PE:EA 1:1): 0.27
Yield: 1.04 g, 73 %
white foam/powder



^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 7.88-7.85 (m, 4H, phthaloyl), 7.77-7.74 (m, 4H, phthaloyl), 5.78-5.69 (m, 2H, H-3, H-3'), 5.46 (d, 1H, $^3J_{1',2'}$ = 8.4 Hz, H-1'), 5.14 (t, 1H, $^3J_{4',5'}$ = 9.8 Hz, H-4'), 4.69 (m, 1H, NHBoc), 4.44-4.40 (m, 2H, H-1, H-6'a), 4.32-4.23 (m, 2H, H-2', H-6a), 4.17 (t, 1H, $^3J_{2,3}$ = 10.4 Hz, H-2), 4.05 (m, 1H, H-6'b), 3.92 (t, 1H, $^3J_{4,5}$ = 9.3 Hz, H-4), 3.80 (m, 1H, H-5'), 3.70-3.63 (m, 2H, H-5, H-6b), 3.21 (m, 2 H, CH_2N), 2.13,1.99,1.96,1.87,1.82 (5s, 3H each, 5 x $COCH_3$), 1.54 (s, 9H, Boc).

^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 170.95, 170.48, 170.21, 169.71 (5C, 5 x $COCH_3$), 156.29 (1C, CO, Boc), 134.85, 131.61, 124.22, 123.72 (12C, 2 x phthaloyl), 97.81 (1C, C-1'), 76.20 (1C, C-4), 74.65 (1C, C-1), 72.31 (1C, C-5'), 72.08 (1C, C-3'), 70.95 (1C, C-3), 68.81 (1C, C-4'), 62.69 (1C, C-6), 61.95 (1C, C-6'), 55.24 (1C, C-2'), 52.21 (1C, C-2), 41.49 (1C, CH_2N), 28.65 (3C, $C(CH_3)_3$, Boc), 21.05, 20.93, 20.83, 20.71 (5C, 5 x $COCH_3$).

Microanalysis calculated for $C_{44}H_{49}N_3O_{19}$ (1.5 H_2O): C, 55.57; H, 5.51; N, 4.42. Found: C, 55.62; H, 5.35; N, 4.38.

N-tert-Butoxycarbonyl-4-_O_-(2-acetamido-3,4,6-tri-_O_-acetyl-2-deoxy-β-D-glucopyranosyl)-(2-acetamido-3,6-di-_O_-acetyl-2-deoxy-β-D-glucopyranosyl) methylamine (38)

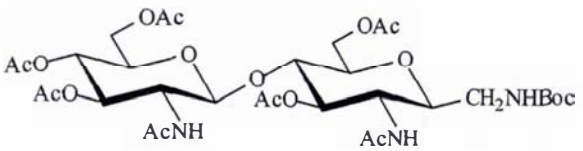
Hydrazine monohydrate (0.8 mL, 14 eq) was added to a solution of compound **37** (1.08 g, 1.17 mmol, 1eq.) in EtOH (30 mL) and the mixture was refluxed at 80 °C for 1 h (a white precipitate forms). The solvent was evaporated followed by azeotroping with toluene to remove trace hydrazine and water. In the next step the residue was dissolved in a mixture of pyridine (20 mL) and acetic anhydride (20 mL) and stirred at r.t for 12 h. The solvent was evaporated, the residue azeotroped with toluene and taken up in chloroform (50 mL). The organic phase was washed successively with 0.1 M hydrochloric acid, sat. sodium hydrogen carbonate solution and water, then dried over sodium sulfate. After column chromatography on silica, using $CHCl_3$:MeOH 20:1 (v/v) as the eluent, **38** (439 mg, 50 %) was obtained as a white powder.

(38) $C_{32}H_{49}N_3O_{17}$

R_f ($CHCl_3$:MeOH 20:1): 0.14

Yield: 439 mg, 50 %

white flakes/powder



^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 6.09 (d, 1H, $^3J_{NH',2'}$ = 8.4 Hz, NHAc'), 5.95 (d, 1H, $^3J_{NH,2}$ = 9.1 Hz, NHAc), 5.17 (t, 1 H, $^3J_{3',4'}$ = 9.7 Hz, H-3'), 5.05-4.93 (m, 3H, NHBoc, H-3, H-4'), 4.54 (d, 1H, $^3J_{1',2'}$ = 8.3 Hz, H-1'), 4.33-4.27 (m, 2H, H-6a, H-6'a), 4.18 (dd, 1H, $^2J_{6b,6a}$ = 12.0 Hz, $^3J_{6b,5}$ = 4.6 Hz H-6b), 3.96 (dd, 1H, $^2J_{6'b,6'a}$ = 12.4 Hz, $^2J_{6'b,5'}$ = 2.0 Hz, H-6'b), 3.86 (t, 1H, $^3J_{2,3}$ = 10.0 Hz, H-2), 3.74 (q, 1H, $^3J_{2',3'}$ = 10.1 Hz, H-2'), 3.66-3.53 (m, 2H, H-4, H-5'), 3.48-3.43 (m, 2H, ½ CH_2, H-5), 3.27 (m, 1H, H-1), 2.96-2.87 (m, 1H, ½ CH_2), 2.07, 2.01, 1.98, 1.94, 1.93, 1.88 and 1.86 (7s, 3H each, 7 x $COCH_3$), 1.36 (s, 9H, Boc).

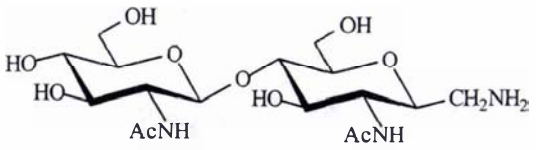
^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 170.33, 170.14, 169.73, 169.52, 168.41 (7C, 7 x $\underline{C}OCH_3$), 154.82 (1C, CO, Boc), 100.05 (1C, C-1'), 76.94 (1C, C-1), 75.23 (1C, C-4), 73.09 (1C, C-3), 71.41 (1C, C-3'), 70.85 (1C, C-5'), 67.29 (1C, C-4'), 61.44 (1C, C-6), 60.90 (1C, C-6'), 53.94 (1C, C-2'), 50.66 (1C, C-2), 40.84 (1C, CH_2), 27.38 (3C, $C(\underline{C}H_3)_3$, Boc), 22.13, 19.96, 19.67, 19.59 (7C, 7 x $CO\underline{C}H_3$).

Microanalysis calculated for $C_{32}H_{49}N_3O_{17}$: C, 51.40; H, 6.60; N, 5.62. Found: C, 51.53; H, 6.56; N, 5.74.

4-*O*-(2-Acetamido-2-deoxy-β-D-glucopyranosyl)-(2-acetamido-2-deoxy-β-D-glucopyranosyl)methylamine (39)

Hydrochloric acid in 1,4-dioxane (4 M, 5 mL) was added to a solution of compound **38** (218 mg, 0.29 mmol) in dichloromethane (20 mL) and the mixture stirred at r.t. for 1 h. After the Boc group cleavage was complete the solvent was evaporated and the residue azeotroped with toluene. The residue was taken up in methanol (20 mL) and the solution saturated with ammonia at 0 °C. The flask was stoppered and stirring continued for 24 h at r.t. After evaporation a colourless syrup was obtained.

(39) $C_{17}H_{31}N_3O_{10}$
positive ninhydrin stain
colourless syrup



^{1}H NMR (300 MHz, MeOH-d_4): δ ppm: 4.41 (d, 1H, $^3J_{1',2'}$ = 8.4 Hz, H-1'), 3.82 (dd, 1H, $^2J_{6'a,6'b}$ = 11.8 Hz, $^3J_{6'a,5'}$ = 2.1 Hz, H-6'a)*, 3.73 (dd, 1H, $^2J_{6a,6b}$ = 11.7 Hz, $^2J_{6a,5}$ = 2.0 Hz, H-6a)*, 3.63-3.17 (m, carbohydrate CH-signals), 2.94-2.91 (dd, 1H, $^2J_{CH,CH'}$ = 13.7 Hz, $^3J_{CH,1}$ = 2.6 Hz, CH), 2.76 (dd, 1H, $^2J_{CH',CH}$ = 13.7 Hz, $^3J_{CH',1}$ = 6.2 Hz, CH'), 1.92 and 1.90 (2s, 3H each, NHAc).
*assignments of 6'a and 6a may be reversed

^{13}C NMR (300 MHz, MeOH-d_4): δ ppm: 174.95 (2C, 2 x $\underline{C}OCH_3$), 103.64 (1C, C-1'), 82.09 (1C), 80.31 (1C), 78.64 (1C), 77.77 (1C), 76.05 (1C), 74.82 (1C), 72.45 (1C), 63.00 (1C), 62.39 (1C), 57.84 (1C, C-2'), 53.93(1C, C-2), 43.09 (1C, CH_2), 23.45, 22.99 (2C, 2 x $CO\underline{C}H_3$).

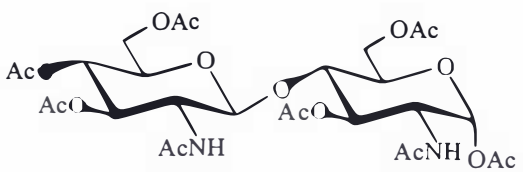
ES-MS : m/z for $C_{17}H_{31}N_3O_{10}$ $(M+H)^+$: 437.8; calculated: 438.5

4-*O*-(2-Acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranosyl)-2-acetamido-1,3,6-tri-*O*-acetyl-2-deoxy-α-D-glucopyranoside (chitobiose-octaacetate) (40a) and 4-*O*-(2-Acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranosyl)-(1-4)-(2-acetamido-3,6-di-*O*-acetyl-2-deoxy-β-D-glucopyranosyl)-2-acetamido-1,3,6-tri-*O*-acetyl-2-deoxy-α-D-glucopyranoside (chitotriose-undecaacetate) (41a)

The title compounds were prepared from chitin (crab shells) using variations of known literature procedures [142-143]. Concentrated sulphuric acid (160 mL) was added to cooled (ice bath) acetic anhydride (1.5 L). Chitin (150 g, previously processed to a fine powder using a knife mill) was added to the cooled liquid with vigorous stirring. The flask was transferred to a water bath and stirring continued at 55 °C. After 4 h the temperature was adjusted to 36 °C and stirring continued for 24 h. The reaction was then quenched by pouring it into an aqueous solution of sodium

acetate (750 g in 2.5 L). Once the mixture had returned to r.t. it was neutralised by the addition of solid ammonium hydrogen carbonate accompanied by vigorous foaming. The aqueous phase was extracted with chloroform (3 x 1 L) and the extracts dried over anhydrous sodium sulfate. After filtration and removal of the solvent, the dark brown syrup was applied in chloroform to a column of silica (60 x 18 cm) and eluted with $CHCl_3$:MeOH 20:1 (v/v). Fractions of chitobiose and chitotriose, which showed no UV activity in comparison to the large amounts of aldehydes which developed during the acetolysis, were collected separately. After removal of the solvent, both the chitobiose and the chitotriose fractions were each recrystallised from ethyl acetate (1.5 L and 400 mL respectively). Overlapping fractions of chitobiose and chitotriose were rechromatographed on silica to achieve separation. Chitobiose-octaacetate **40a** (12.5 g) and chitotriose-undecaacetate **41a** (2.5 g) were both obtained as white powders.

(40a) $C_{28}H_{40}N_2O_{17}$
R_f (EA:MeOH 9:1): 0.54
R_f ($CHCl_3$:MeOH 20:1): 0.2
Yield (weight %): 8.3 %
white powder



^{1}H NMR (300 MHz, $CDCl_3$) for **40a**: δ ppm: 6.17 (d, 1H, $^3J_{2',NH'}$ = 9.2 Hz, NHAc'), 6.11 (d, 1H, $^3J_{1,2}$ = 3.6 Hz, H-1), 5.82 (d, 1H, $^3J_{2,NH}$ = 9.0 Hz, NHAc), 5.23 (dd, 1H, $^3J_{3,2}$ = 11.0 Hz, $^3J_{3,4}$ = 9.0 Hz, H-3), 5.15 (t, 1H, $^3J_{3',2'}$ = 10.1 Hz, $^3J_{3',4'}$ = 9.4 Hz, H-3'), 5.05 (t, 1H, $^3J_{4',5'}$ = 9.6 Hz, H-4'), 4.52 (d, 1H, $^3J_{1',2'}$ = 8.4 Hz, H-1'), 4.44-4.33 (m, 3H, H-6a, H-6'a, H-2), 4.22 (dd, 1H, $^2J_{6b,6a}$ =12.2 Hz, $^3J_{6b,5}$ = 1.8 Hz, H-6b), 4.03 (dd, 1H, $^2J_{6'b,6'a}$ = 12.6 Hz, $^3J_{6'b,5'}$ = 2.3 Hz, H-6'b), 3.97-3.90 (m, 2H, H-2', H-5), 3.77 (t, 1H, $^3J_{4,5}$ = 9.8 Hz, H-4), 3.65 (m, 1H, H-5'), 2.17, 2.14, 2.09, 2.06, 2.01, 2.01, 1.95, 1.93 (8 s, 3H each, 8x $COCH_3$).

^{13}C NMR (300 MHz, $CDCl_3$) for **40a**: δ ppm: 171.80, 171.59, 171.14, 170.87, 170.79, 170.58, 169.65, 169.28 (8C, 8x $\underline{C}OCH_3$), 102.07 (1C, C-1'), 90.83 (1C, C-1), 76.29 (1C, C-4), 72.97 (1C, C-3'), 72.32 (1C, C-5'), 71.13 (1C, C-5), 71.07 (1C, C-3), 68.40

(1C, C-4'), 62.13 (1C, C-6''), 61.98 (1C, C-6), 54.85 (1C, C-2''), 51.55 (1C, C-2), 23.50-20.93 (8C, 8x $CO\underline{C}H_3$).

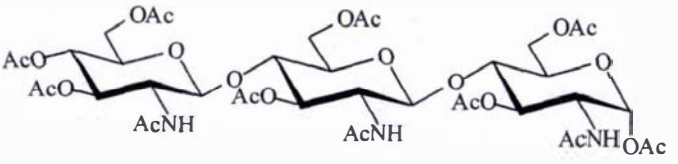
(41a) $C_{40}H_{57}N_3O_{24}$

R_f (EA:MeOH 9:1): 0.33

R_f ($CHCl_3$:MeOH 20:1): 0.12

Yield (weight %): 1.7 %

white powder



^{1}H NMR (300 MHz, DMSO-d_6) for **41a**: δ ppm: 7.97 (d, 1H, $^3J_{NH'',2''}$ = 9.1 Hz, N<u>H</u>Ac''), 7.89 (d, 2H, $^3J_{NH',2' \text{ and } NH,2}$ = 9.2 Hz, N<u>H</u>Ac', N<u>H</u>Ac), 5.82 (d, 1H, $^3J_{1,2}$ = 3.5 Hz, H-1), 5.13 (t, 1H, $^3J_{3'',2''}$ = 10.0 Hz, $^3J_{3'',4''}$ = 9.8 Hz, H-3''), 5.07-4.97 (m, 2H, H-3, H-3'), 4.81 (t, 1H, $^3J_{4'',5''}$ = 9.6 Hz, H-4''), 4.65 (d, 1H, $^3J_{1'',2''}$ = 8.4 Hz, H-1''), 4.56 (d, 1H, $^3J_{1',2'}$ = 8.4 Hz, H-1'), 4.38-4.34 (m, 1H, H-6a), 4.29-4.23 (m, 2H, H-6'a, H-6''a), 4.17-4.09 (m, 1H, H-2), 4.06-3.99 (m, 2H, H-6b, H-6'b), 3.92-3.88 (m, 1H, H-6''b), 3.84-3.79 (m, 3H, H-5, H-5'', H-4), 3.72 (t, 1H, $^3J_{4',3' \text{ and } 4',5'}$ = 9.1-9.4 Hz, H-4'), 3.60-3.47 (m, 3H, H-5, H-2'', H-2'), 2.14, 2.09, 2.05, 2.00 (4 s, 3 H each, 4x $COCH_3$), 1.95-1.94 (m, 9H, 3x $COCH_3$), 1.90,1.78, 1.75, 1.73 (4 s, 3 H each, 4x $COCH_3$).

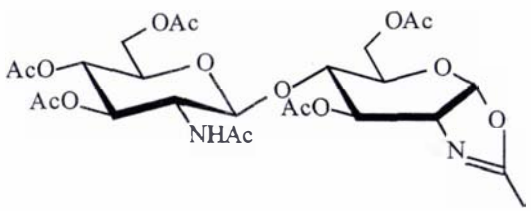
^{13}C NMR (300 MHz, DMSO-d_6) for **41a**: δ ppm: 170.50, 170.32, 169.97, 169.62, 169.52 (11C, 11 x $\underline{C}OCH_3$), 100.48 (1C, C-1''), 100.30 (1C, C-1'), 90.04 (1C, C-1), 76.01 (1C, C-4'), 75.14 (1C, C-4), 73.68 (1C, C-3'), 72.67 (1C, C-3''), 72.00 (1C, C-5'), 70.82 (1C, C-5''), 70.54 (2C, C-3, C-5), 68.58 (1C, C-4''), 62.96 (1C, C-6'), 61.97 (2C, C-6, C-6''), 54.23 and 54.10 (2C, C-2', C-2''), 50.39 (1C, C-2), 22.89, 22.55, 21.19, 21.00, 20.72 (11C, 11 x $CO\underline{C}H_3$).

4-*O*-(2-Acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranosyl)-(1-4)-2-methyl-(3,6-di-*O*-acetyl-1,2-dideoxy-α-D-glucopyrano)-[2,1-d]-2-oxazoline (42)

The title compound was synthesised from chitobiose-octaacetate using the method of S.-I. Nishimura and H. Kuzuhara [144].

TMSOTf (2.3 mL, 12.7 mmol) was added to a solution of chitobiose-octaacetate **40a** (7.7 g, 11.4 mmol) in 1,2-dichloroethane (65 mL). The mixture was heated to 50 °C with stirring for 12 h. Further TMSOTf (0.6 mL) was added to drive the reaction to completion. The mixture was then neutralised with triethylamine (5 mL) and the solvent removed by rotary evaporation. The residue was taken up in a minimal amount of chloroform and applied to a column of silica, then eluted with toluene:ethylacetate:triethylamine 100: 200:1 (v/v/v) followed by toluene:ethylacetate:triethylamine 100:400:1 (v/v/v) to give the title compound (4.38 g, 62 %) as a white powder. The relatively low yield and the prolonged reaction time, which required the addition of more TMSOTf in comparison with the literature procedure, was caused by the poor quality of the only batch of TMSOTf available at the time of the reaction.

(42) $C_{26}H_{36}N_2O_{15}$
R_f (Tol:EA 1:4): 0.18
Yield: 4.38 g, 62 %
white powder



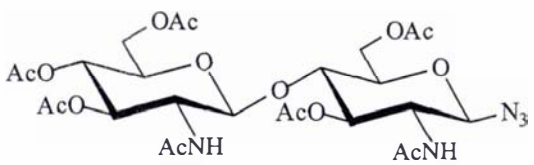
^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 5.91 (d, 1H, $^3J_{1,2}$ = 7.3 Hz, H-1), 5.88 (d, 1H, $^3J_{NH',2'}$ = 8.9 Hz, NHAc'), 5.66 (d, 1H, $^3J_{3,2}$ = 2.3 Hz, H-3), 5.22 (t, 1H, $^3J_{3',2'}$ = 10.3 Hz, $^3J_{3',4'}$ = 9.4 Hz, H-3'), 5.07 (t, 1H, $^3J_{4',3'}$ = 9.4 Hz, $^3J_{4',5'}$ = 9.8 Hz, H-4'), 4.76 (d, 1H, $^3J_{1',2'}$ = 8.5 Hz, H-1'), 4.32 (dd, 1H, $^2J_{6a,6b}$= 12.1 Hz, $^3J_{6a,5}$ = 4.6 Hz, H-6a), 4.27 (dd, 1H, $^2J_{6'a,6'b}$ = 12.3 Hz, $^3J_{6'a,5'}$ = 4.5 Hz, H-6'a), 4.15-4.10 (m, 3 H, $^2J_{6b,6a}$ =12.1 Hz, $^3J_{6b,5}$ = 2.3 Hz, $^2J_{6'b,6'a}$ = 12.3 Hz, $^3J_{6'b,5'}$ = 2.3 Hz, H-6b, H-6'b, H-2), 3.93 (m, 1H, H-2'), 3.78-3.72 (m, 1H, H-5'), 3.56 (d, 1H, $^3J_{4,5}$= 9.6 Hz, H-4), 3.48-3.43 (m, 1H, H-5), 2.13, 2.10, 2.09, 2.08, 2.02, 2.01, 1.94 (7 s, 3 H each, 7 x $COCH_3$).

^{13}C NMR (300 MHz, $CDCl_3$): δ ppm:171.56, 171.20, 170.79, 169.72 (6C, 6 x $\underline{C}OCH_3$), 167.34 (1C, N=C), 102.82 (1C, C-1'), 99.56 (1C, C-1), 78.06-76.96 (1C, C-4 under $CDCl_3$-triplett), 73.24 (1C, C-3'), 72.44 (1C, C-5'), 70.83 (1C, C-3), 68.78 (1C, C-4'), 68.15 (1C, C-5), 65.32 (1C, C-2), 63.47 (1C, C-6), 62.45 (1C, C-6'), 54.91 (1C, C-2'), 23.53, 21.36, 21.03 (6C, 6 x $COCH_3$), 14.36 (1C, NC$\underline{C}H_3$).

4-*O*-(2-Acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranosyl)-2-acetamido-3,6-tri-*O*-acetyl-2-deoxy-β-D-glucopyranosyl azide (43a)

Trimethylsilyl azide (30 μL, 0.2 eq.) was added to a solution of **42** (780 mg, 1.26 mmol) in dichloromethane (10 mL). Tin tetrachloride (0.34 mL, 2eq.) was added and the solution stirred for 12 h. The mixture was quenched with sat. sodium hydrogen carbonate solution and the compound extracted with dichloromethane. The organic phase was dried over sodium sulfate and the product purified on silica using $CHCl_3$:MeOH 20:1 (v/v) as the eluent to give title compound **43a** (710 mg, 85 %) as a white powder.

(43a) $C_{26}H_{37}N_5O_{15}$
R_f ($CHCl_3$:MeOH 20:1): 0.18
Yield: 710 mg, 85 %
white flakes/powder



^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 6.18 (d, 1H, $^3J_{NH',2'}$ = 8.9 Hz, NHAc'), 6.06 (d, 1H, $^3J_{NH,2}$ = 9.4 Hz, NHAc), 5.20 (t, 1H, $^3J_{3',4'}$ = 9.6 Hz, H-3'), 5.13-5.02 (m, 2H, H-3, H-4'), 4.58 (t, 2H, $^3J_{1',2'}$ = 7.5 Hz, $^3J_{1,2}$ = 8.6 Hz, H-1', H-1), 4.43-4.28 (m, 3H, H-6'a, H-6a, H-6b), 4.12-4.01 (m, 2H, H-2, H-6'b), 3.85 (q, $^3J_{2',3'}$ = 10.2 Hz, H-2'), 3.79-3.73 (m, 2H, H-4, H-5), 3.66 (m, 1H, H-5'), 2.16, 2.09, 2.07, 2.01, 2.00, 1.99, 1.95 (7s, 3H each, 7 x $COCH_3$).

^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 171.52, 171.39, 171.21, 170.83, 169.73 (7C, 7 x $\underline{C}OCH_3$), 101.67 (1C, C-1'), 89.03 (1C, C-1), 76.11 (1C, C-5), 75.40 (1C, C-4), 72.81 (2C, C-3, C-3'), 72.35 (1C, C-5'), 68.53 (1C, C-4'), 62.50 (1C, C-6), 62.16 (1C, C-6'), 55.16 (1C, C-2'), 53.60 (1C, C-2), 23.52, 21.29, 21.03, 20.95 (7C, 7 x $CO\underline{C}H_3$).

Microanalysis calculated for $C_{26}H_{37}N_5O_{15}$: C, 47.34; H, 5.65; N, 10.62. Found: C,47.26; H, 5.68; N, 10.33.

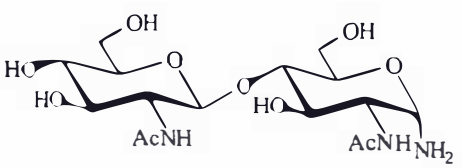
4-*O*-(2-Acetamido-2-deoxy-β-D-glucopyranosyl)-2-acetamido-2-deoxy-α-D-glucopyranosylamine (44)

A suspension of glycopyranosyl azide **43a** (710 mg, 1.08 mmol) in methanol (50 mL) was saturated with ammonia at 0 °C. The suspension gave way to a homogeneous solution. The flask was stoppered and stirring continued at r.t. for 48 h. TLC (EA:MeOH, 9:1) was used to show the disappearance of the starting material and when this was complete the solvent was removed by rotary evaporation to give the intermediate 4-*O*-(2-acetamido-2-deoxy-β-D-glucopyranosyl)-2-acetamido-2-deoxy-β-D-glucopyranosyl azide **43b** as a white powder.

The intermediate was dissolved in methanol (20 mL), 10 % palladium on charcoal (100 mg) added and the mixture stirred under an atmosphere of hydrogen for 12 h. After filtration and evaporation of the filtrate white needle shaped crystals appeared upon drying under vacuum.

Comparative NMR studies showed that the β-azide **43b** had anomerised during the hydrogenation procedure to yield solely the α-configured amine **44** (no yield was calculated for the crude material).

(44) $C_{16}H_{29}N_3O_{10}$
R_f ($CHCl_3$:MeOH, 1:1): 0.18
positive ninhydrin stain
syrup/white needle shaped crystals



^{1}H NMR (300 MHz, D_2O): δ ppm: 4.60 (1H, d, $^3J_{1',2'}$ = 8.3 Hz, H-1'), 3.94-3.32 (m, carbohydrate CH), 2.06, 2.04 (2s, 3H each, 2 x $COCH_3$). The spectrum also shows minor signals of unidentified side products.

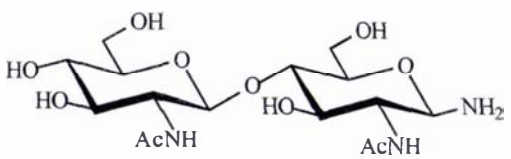
^{13}C NMR (300 MHz, D_2O): δ ppm: 175.53, 175.25 (2C, 2 x $\underline{C}OCH_3$), 101.44 (1C, C-1'), 78.91 (1C), 76.15 (1C), 73.93 (2C), 71.51 (1C), 70.32 (1C), 69.89 (1C), 62.20, 61.07 (2C, C-6, C-6'), 56.09 (2C, C-2, C-2'), 22.57, 21.70 (2C, 2 x $CO\underline{C}H_3$).

ES-MS: m/z for $C_{16}H_{29}N_3O_{10}$: $(M+H)^+$: 424.2; calculated: 424.4

4-*O*-(2-Acetamido-2-deoxy-β-D-glucopyranosyl)-2-acetamido-2-deoxy-β-D-glucopyranosylamine (45)

A suspension of chitobiose-octaacetate **40a** (2.0 g, 2.96 mmol) in methanol (150 mL) was saturated with ammonia at 0 °C. Stirring was continued at r.t. for 48 h. Once TLC analysis (EA:MeOH 9:1) showed that the starting material had all been consumed the solvent was evaporated by rotary evaporation to give the deacetylated intermediate 4-*O*-(2-acetamido-2-deoxy-β-D-glucopyranosyl)-2-acetamido-2-deoxy-α-D-glucopyranoside **40b** (1.23 g, 98 %). A portion of the triol (1.0 g, 2.36 mmol) was dissolved in a saturated aqueous ammonium hydrogencarbonate solution (100 mL) and processed following the standard procedure (general procedures) to give title compound **45** in an almost quantitative yield (957 mg, 96 %).

(45) $C_{16}H_{29}N_3O_{10}$
R_f ($CHCl_3$:MeOH 1:1): 0.18
positive ninhydrin stain
white powder



^{1}H NMR (300 MHz, D_2O): δ ppm: 4.55 (d, 1H, $^3J_{1',2'}$ = 8.4 Hz, H-1'), 4.12 (d, 1H, $^3J_{1,2}$ = 8.5 Hz, H-1), 3.91-3.44 (m, carbohydrate CH), 2.04, 2.01 (2s, 3H each, 2 x $COCH_3$).

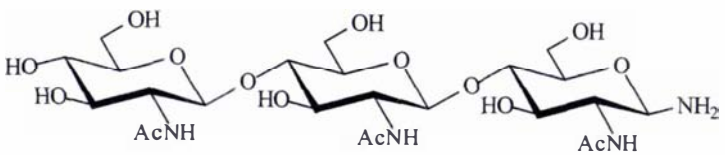
^{13}C NMR (300 MHz, D_2O): δ ppm: 175.00 (2C, 2 x $\underline{C}OCH_3$), 101.90 (1C, C-1'), 84.53 (1C, C-1), 80.22 (1C, C-4), 76.32 (1C, C-5'), 75.82 (1C, C-5), 73.89 (1C, C-3'), 73.57 (1C, C-3), 70.13 (1C, C-4'), 60.96 (1C, C-6'), 60.73 (1C, C-6), 56.17 (1C, C-2), 56.01 (1C, C-2'), 22.68, 22.54 (2C, 2 x $COCH_3$).
These ^{13}C NMR assignments are in agreement with Ref. [119]

ES-MS: m/z for $C_{16}H_{29}N_3O_{10}$: $(M+H)^+$: 424.2; calculated: 424.4

4-*O*-(2-Acetamido-2-deoxy-β-D-glucopyranosyl)-(1-4)-(2-acetamido-2-deoxy-β-D-glucopyranosyl)-2-acetamido-2-deoxy-β-D-glucopyranosylamine (46)

A suspension of chitotriose-undecaacetate **41a** (1g, 1.04 mmol) in methanol (65 mL) was saturated with ammonia at 0 °C. Stirring was continued at r.t. for 48 h. The progress of the reaction was monitored by TLC (EA:MeOH 9:1). When there was no longer evidence of starting material, the solvent was evaporated to give the deacetylated intermediate 4-*O*-(2-acetamido-2-deoxy-β-D-glucopyranosyl)-(1-4)-(2-acetamido-2-deoxy-β-D-glucopyranosyl)-2-acetamido-2-deoxy-α-D-glucopyranoside **41b** (638 mg, 98 %). This residue was dissolved in a saturated aqueous ammonium hydrogen carbonate solution (100 mL) and subjected to the the standard procedure (general procedures) to give title compound **46** in an almost quantitative yield (611 mg, 96 %).

(46) $C_{24}H_{42}N_4O_{15}$
R_f ($CHCl_3$:MeOH 1:1): 0.13
positive ninhydrin stain
white powder



^{1}H NMR (300 MHz, D_2O): δ ppm: 4.56 (d, 2H, $^3J_{1',2' \text{ and } 1'',2''}$ = 8.5 Hz, H-1', H-1''), 4.11 (d, 1H, $^3J_{1,2}$ = 8.6 Hz, H-1), 3.91-3.40 (m, carbohydrate CH), 2.04, 2.01 (m, 9H, 3 x $COCH_3$).

^{13}C NMR (300 MHz, D_2O): δ ppm: 175.02 (3C, 3 x $\underline{C}OCH_3$), 101.88, 101.69 (2C, C-1', C-1''), 79.60 (2C, C-4, C-4'), 76.32, 75.85, 74.91 (3C, C-5, C-5', C-5''), 73.86, 73.52, 72.57 (3C, C-3, C-3', C-3''), 70.13 (1C, C-4''), 60.96, 60.40 (3C, C-6, C-6', C-6''), 56.00, 55.44 (3C, C-2, C-2', C-2''), 22.68, 22.53 (3C, 3 x $COCH_3$).

ES-MS: m/z for $C_{24}H_{42}N_4O_{15}$: $(M+H)^+$: 628.2; calculated: 627.6

8.4 Peptide Synthesis

N-tert-Butoxycarbonyl-β-benzyl-L-aspartyl-L-alanyl-L-serine methyl ester (Boc-Asp(OBn)Ala-Ser-OMe) (47a)

Boc-L-alanine (1.89 g ,10 mmol) was coupled to L-Ser-OMe (1.56 g, 10 mmol) using the standard procedure (general procedures). The dipeptide was isolated by column chromatography on silica using $CHCl_3$:EtOH 25:1 (v/v) as the mobile phase to give Boc-L-Ala-L-Ser-OMe (2.68 g, 92 %) as a colourless syrup. After cleavage of the Boc-group (general procedures), the hydrochloride salt (663 mg, 2.93 mmol) was coupled to Boc-L-Asp(OBn)-OH (1.42 g, 4.39 mmol, 1.5 eq.) using the standard procedure and subjected to column chromatography on silica using $CHCl_3$:EtOH 30:1 (v/v) as the mobile phase to give Boc-L-Asp(OBn)-L-Ala-L-Ser-OMe **47a** (780 mg, 54 %) as a colourless syrup.

(47a) $C_{23}H_{33}N_3O_9$
R_f ($CHCl_3$:EtOH 30:1): 0.13
Yield: 780 mg, 50 % (two steps)
colourless syrup

Boc-Asp(OBn)-Ala-Ser-OMe

^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 7.34 (s, 5H, benzyl), 7.17 (d, 1H, $^3J_{NH,\alpha CH}$ = 7.8 Hz, NH Ser), 7.03 (d, 1H, $^3J_{NH,\alpha CH}$ = 7.4 Hz, NH Ala), 5.61 (d, 1H, $^3J_{NH,\alpha CH}$ = 8.0 Hz, NH Asp), 5.12 (m, 2H, $PhCH_2$), 4.64-4.59 (m, 1H, α CH Ser), 4.53-4.44 (m, 2H, α CH Ala, Asp), 3.98-3.88 (m, 2H, β CH_2 Ser), 3.76 (s, 3H, OMe), 2.97 (dd, 1H, $^2J_{\beta CH',\beta CH}$ = 17.2 Hz, $^3J_{\beta CH',\alpha CH}$ = 4.6, ½ β CH_2 Asp), 2.82 (dd, 1H, $^2J_{\beta CH,\beta CH'}$ = 17.2 Hz, $^3J_{\beta CH,\alpha CH}$ = 6.6 Hz, ½ β CH_2 Asp), 1.44 (s, 9H, Boc), 1.39 (d, 3H, $^3J_{CH_3,\alpha CH}$ = 7.0 Hz, CH_3 Ala).

^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 172.23, 172.11, 171.42, 171.05 (4C, 4 x CO), 129.00, 128.83, 128.66 (5C, benzyl), 67.44 (1C, CH_2 benzyl), 63.07 (1C, CH_2OH Ser), 55.27 (1C, α CH Ser), 53.00 (1C, OMe), 51.25 (1C, α CH Ala), 49.65 (1C, α CH Asp), 36.40 (1C, β CH_2 Asp), 28.63 (3 C, $C(\underline{C}H_3)_3$ Boc), 17.77 (1C, CH_3 Ala).

N-tert-Butoxycarbonyl-L-aspartyl-L-alanyl-L-serine methyl ester (Boc-Asp(COOH)-Ala-Ser-OMe) (47b)

10 % Palladium on charcoal (200 mg) was added to a solution of compound **47a** (620 mg, 1.25 mmol) in EtOH (15 mL) and THF (3 mL). The mixture was stirred under an atmosphere of hydrogen for 12 h. The mixture was then filtered through glass fibre filter paper and the filtrate evaporated to give a colourless syrup which was triturated with diethyl ether to give **47b** as a white powder (490 mg, 97 %).

(47b) $C_{16}H_{27}N_3O_9$
R_f ($CHCl_3$:MeOH 5:1): 0.36 **Boc-Asp(COOH)-Ala-Ser-OMe**
Yield: 490 mg, 97 %
white powder

^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 7.56 (d, 2H, $^3J_{NH,\alpha CH}$ = 7.8 Hz, NH Ser, Asp), 5.90 (d, 1H, $^3J_{NH,\alpha CH}$ = 7.2 Hz, NH Ala), 4.67-4.53 (m, 3H, α CH Ser, Asp, Ala), 4.00-3.80 (m, 2H, β CH_2 Ser), 3.76 (s, 3H, OMe), 2.84-2.76 (m, 2H, β CH_2 Asp), 1.44 (s, 9H, Boc), 1.39 (d, $^3J_{CH3,\alpha CH}$ = 7.0 Hz, CH_3 Ala).

^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 174.43, 172.99, 171.85, 171.13 (4C, 4 x CO), 156.04 (1C, CO Boc), 62.71 (1C, CH_2OH Ser), 55.16 (1C, α CH Ser), 53.05 (1C, OMe), 51.21 (1C, α CH Ala), 49.60 (1C, α CH Asp), 36.81 (1C, β CH_2 Asp), 28.66 (3 C, C($\underline{C}H_3$)$_3$ Boc), 17.88 (1C, CH_3 Ala).

ES-MS: *m/z* for $C_{16}H_{27}N_3O_9$: $(M+H)^+$: 406.5; calculated: 406.4

N-tert-Butoxycarbonyl-γ-benzyl-L-glutaminyl-L-alanyl-L-serine methyl ester (Boc-Glu(OBn)Ala-Ser-OMe) (48a)

Boc-L-Glu(OBn)-OH (2.0 g, 5.93 mmol, 1.5 eq) was coupled to HCl·L-Ala-L-Ser-OMe (895 mg, 3.95 mmol) using the standard procedure (general procedures) to give

title compound **48a** (973 mg, 48 %) as a white foam after column chromatography on silica using $CHCl_3$:EtOH 30:1 (v/v) as the eluent.

(48a) $C_{24}H_{35}N_3O_9$

Boc-Glu(OBn)-Ala-Ser-OMe

R_f ($CHCl_3$:EtOH 30:1): 0.11

Yield: 973 mg, 48 %

white foam

^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 7.39-7.30 (m, 6H, benzyl, NH Ser), 7.02 (d, 1H, $^3J_{NH,\alpha CH}$ = 7.4 Hz, NH Ala), 5.57 (d, 1H, $^3J_{NH,\alpha CH}$ = 7.5 Hz, NH Glu), 5.12 (m, 2H, $PhCH_2$), 4.67-4.52 (m, 2H, α CH Ser, Ala), 4.17 (m, 1H, α CH Glu), 3.94 (dd, 1H, $^2J_{\beta CH',\beta CH}$ = 11.5 Hz, $^3J_{\beta CH,\alpha CH}$ = 3.3 Hz , 1/2 β CH_2 Ser), 3.87 (dd, 1H, $^2J_{\beta CH,\beta CH'}$ = 11.5 Hz, $^3J_{\beta CH',\alpha CH}$ = 3.6 Hz , 1/2 β CH_2 Ser) 3.75 (s, 3H, OMe), 2.51 (t, 2H, $^3J_{\gamma CH_2,\beta CH_2}$ = 7.1 Hz, γ CH_2 Glu), 2.18-2.10 (m,1H, ½ β CH_2 Glu), 2.01-1.90 (m, 1H, ½ β CH_2 Glu), 1.42 (s, 9H, Boc), 1.39 (d, 3H, $^3J_{CH_3,\alpha CH}$ = 7.0 Hz, CH_3 Ala).

^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 173.75, 172.52, 172.27, 171.07 (4C, 4 x CO), 156.21 (1C, CO Boc), 128.97, 128.71, 128.64 (5C, benzyl), 67.08 (1C, CH_2 benzyl), 62.98 (1C, CH_2OH Ser), 55.24 (1C, α CH Ser), 54.46 (1C, α CH Glu), 52.97 (1C, OMe), 49.41 (1C, α CH Ala), 34.24 (1C, γ CH_2 Glu), 28.66 (3C, $C(\underline{C}H_3)_3$ Boc), 27.95 (1C, β CH_2 Glu), 18.21 (1C, CH_3 Ala).

N-tert-Butoxycarbonyl-L-glutaminyl-L-alanyl-L-serine methyl ester (Boc-Glu-Ala-Ser-OMe) (48b)

10 % Palladium on charcoal (300 mg) was added to a solution of compound **48a** (880 mg, 1.73 mmol) in EtOH (15 mL) and THF (3 mL). The mixture was stirred under an atmosphere of hydrogen for 12 h. The mixture was then filtered through glass fibre filter paper and the filtrate evaporated to give title compound **48b** (703 mg, 97 %) as a white foam/colourless syrup.

(48b) $C_{17}H_{29}N_3O_9$

R_f ($CHCl_3$:MeOH 5:1): 0.35 **Boc-Glu(COOH)-Ala-Ser-OMe**

Yield: 97 %

white foam/colourless syrup

^{1}H NMR (300 MHz, MeOH-d_4): δ ppm: 4.52 (t, 1H, $^3J_{\alpha CH, \beta CH_2}$ = 4.3 Hz, α CH Ser), 4.45 (q, 1H, $^3J_{\alpha CH, CH_3}$ = 7.1 Hz, α CH Ala), 4.13-4.08 (m, 1H, α CH Glu), 3.91 (dd, 1H, $^2J_{\beta CH', \beta CH}$ = 11.3 Hz, ½ β CH_2 Ser), 3.81 (dd, 1H, $^2J_{\beta CH, \beta CH'}$ = 11.3 Hz, ½ β CH_2 Ser), 3.76 (s, 3H, OMe), 2.42 (t, 2H, $^3J_{\gamma CH_2, \beta CH_2}$ = 7.5 Hz, γ CH_2 Glu), 2.14-2.03 (m, 1H, ½ β CH_2 Glu), 1.94-1.82 (m, 1H, ½ β CH_2 Glu), 1.46 (s, 9H, Boc), 1.41 (d, 1H, $^3J_{CH_3, \alpha CH}$ = 7.1 Hz, CH_3, Ala).

^{13}C NMR (300 MHz, MeOH-d_4): δ ppm: 63.18 (1C, CH_2OH Ser), 56.60 (1C, α CH Ser), 55.67 (1C, α CH Glu), 53.19 (1C, OMe), 50.65 (1C, α CH, Ala), 31.69 (1C, γ CH_2 Glu), 29.07 (4C, $C(CH_3)_3$ Boc, β CH_2 Glu), 18.49 (1C, CH_3 Ala).

ES-MS: *m/z* for $C_{17}H_{29}N_3O_9$: $(M+H)^+$: 420.1; calculated: 420.4

N-tert-Butoxycarbonyl-β-benzyl-L-aspartyl-L-alanyl-L-threonine methyl ester (Boc-Asp(OBn)Ala-Thr-OMe) (49a)

p-Tolouenesulfonic acid (7.6 g, 40 mmol) was added to a suspension of L-threonine (2.38 g, 20 mmol) in methanol (100 mL). The resulting solution was heated at reflux for 24 h. After evaporation of the solvent, the oily residue (consisting of more than 95 % L-threonine methyl ester *p*-toluenesulfonate by NMR) was coupled to Boc-Ala-COOH following the standard procedure (general procedures). After column chromatography on silica using $CHCl_3$:EtOH 25:1 (v/v) as the eluent, the product fractions were pooled, the solvent evaporated and the residue treated with diethyl ether and filtered to remove any insoluble material. The filtrate was rechromatographed using the same conditions as above to give the dipeptide Boc-L-Ala-L-Thr-OMe (5.65 g, 93 %) as a white solid. After cleavage of the Boc-group (general procedures) the resulting HCl·L-Ala-L-Thr-OMe (1.0 g, 4.16 mmol) was

coupled to Boc-L-Asp(OBn) (2.02 g, 6.25 mmol, 1.5 eq) using the standard procedure (general procedures). The product was isolated by column chromatography using $CHCl_3$:EtOH 30:1 (v/v) as the eluent to yield the title compound **49a** (1.44 g, 68 %) as a white foam.

(49a) $C_{24}H_{35}N_3O_9$
R_f ($CHCl_3$:EtOH 30:1): 0.14
Yield: 1.44 g, 60 % (4 steps)
white foam/powder

Boc-Asp(OBn)-Ala-Thr-OMe

^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 7.34 (s, 5H, benzyl), 7.11 (m, 2H, NH Thr, Ala), 5.63 (d, 1H, $^3J_{NH, \alpha CH}$ = 8.5 Hz, NH Asp), 5.12 (m, 2H, $PhCH_2$), 4.59-4.50 (m, 3H, α CH Ala, Thr, Asp), 4.31 (dq, 1H, $^3J_{\beta CH, CH_3}$ = 6.4 Hz, $^3J_{\beta CH, \alpha CH}$ = 2.8 Hz, β CH Thr), 3.74 (s, 3H, OMe), 2.98 (dd, 1H, $^2J_{\beta CH', \beta CH}$ = 17.0 Hz, $^3J_{\beta CH', \alpha CH}$ = 4.9 Hz, ½ β CH_2 Asp), 2.82 (dd, 1H, $^2J_{\beta CH, \beta CH'}$ = 17.0 Hz, $^3J_{\beta CH, \alpha CH}$ = 6.6 Hz ½ β CH_2 Asp), 1.44 (s, 9H, Boc), 1.41 (d, 3H, $^3J_{CH_3, \alpha CH}$ = 7.0 Hz, CH_3 Ala), 1.18 (d, 3H, $^3J_{CH_3, \alpha CH}$ = 6.4 Hz, CH_3 Thr).

^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 172.68, 172.09, 171.57, 171.22 (4C, 4 x CO), 155.88 (1C, CO Boc), 128.98, 128.80, 128.65 (5C, benzyl), 68.52 (1C, β CH Thr), 67.40 (1C, CH_2 benzyl), 57.98 (1C, α CH Thr), 52.87 (1C, OMe), 51.17 (1C, α CH Ala), 49.64 (1C, α CH Asp), 36.48 (1C, β CH_2 Asp), 28.64 (9C, $C(CH_3)_3$ Boc), 20.35 (1C, CH_3 Thr), 18.12 (1C, CH_3 Ala).

N-tert-Butoxycarbonyl-L-aspartyl-L-alanyl-L-threonine methyl ester (Boc-Asp-Ala-Thr-OMe) (49b)

10 % Palladium on charcoal (200 mg) was added to a solution of compound **49a** (600 mg, 1.18 mmol) in EtOH (15 mL) and THF (6 mL). The mixture was stirred under an atmosphere of hydrogen for 12 h. The mixture was then filtered through glass fibre filter paper and the filtrate was evaporated to give title compound **49b** (484 mg, 98 %) as a white foam/colourless syrup.

(49b) $C_{17}H_{29}N_3O_9$

R_f ($CHCl_3$:MeOH, 5:1): 0.39 **Boc-Asp(COOH)-Ala-Thr-OMe**

Yield: 484 mg, 98 %

white foam

^{1}H NMR (300 MHz, $CDCl_3$): δ ppm: 7.53 (d, 1H, $^3J_{NH, \alpha CH}$ = 7.0 Hz, NH Ala), 7.40 (d, 1H, $^3J_{NH, \alpha CH}$ = 8.6 Hz, NH Thr), 5.87 (m, 1H, NH Asp), 4.67 (m, 3H, α CH Ala, Thr, Asp), 4.33 (m, 1H, β CH Thr), 3.75 (s, 3H, OMe), 1.44 (9H, Boc), 1.40 (d, 1H, $^3J_{CH3, \alpha CH}$ = 6.9 Hz, CH_3 Ala), 1.19 (d, 1H, $^3J_{CH3, \alpha CH}$ = 6.1 Hz, CH_3 Thr).

^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 174.45, 173.32, 171.62 (4C, 4 x CO), 156.01 (1C, CO Boc), 68.59 (1C, β CH Thr), 58.08 (1C, α CH Thr), 53.00 (1C, OMe), 51.18 (1C, α CH Ala), 49.80 (1C, α CH, Asp), 36.70 (1C, β CH_2 Asp), 28.66 (3C, $C(CH_3)_3$ Boc), 20.12 (1C, CH_3 Thr), 18.04 (1C, CH_3, Ala).

ES-MS: *m/z* for $C_{17}H_{29}N_3O_9$: $(M+H)^+$: 420.1; calculated: 420.4

_N_-Acetyl-L-tyrosyl-L-isoleucyl-L-aspartyl-L-alanyl-L-serine amide (Ac-Tyr-Ile-Asp-Ala-Ser-NH_2) (50)

The title compound was synthesised following the protocol for solid phase peptide synthesis outlined in the general procedures. After freeze drying, pentapeptide **50** was obtained as a white powder (119 mg, 20 %) .

(50) $C_{27}H_{40}N_6O_{10}$ **Ac-Tyr-Ile-Asp(COOH)-Ala-Ser-NH_2**

Yield: 119 mg, 20 %

white powder

^{1}H NMR (300 MHz, DMSO-d_6): δ ppm: 8.26 (d, 1H, $^3J_{NH, \alpha CH}$ = 7.5 Hz, NH Asp), 7.98 (d, 1H, $^3J_{NH, \alpha CH}$ = 8.1 Hz, NH Tyr), 7.88 (d, 1H, $^3J_{NH, \alpha CH}$ = 7.1 Hz, NH Ala), 7.82-7.76 (m, 2H, NH Ser, Ile), 7.02 (d, 2H, $^3J_{CH\ meta, CH\ ortho}$ = 8.2 Hz, 2 x CH Tyr),

6.63 (d, 2H, $^{3}J_{CH\ ortho,CH\ meta}$ = 8.2 Hz, 2 x CH Tyr), 4.57-4.43 (m, 3H, α CH Asn, Tyr, OH), 4.26-4.15 (m, 4H, α CH, Ala, Ser, Ile, OH), 3.58 (m, 2H, β CH_2 Ser), 2.89-2.55 (m, 4 H β CH_2 Tyr, Asp), 1.75 (m, 4H, $COCH_3$, β CH Ile), 1.42 (m, 1H, ½ γ CH_2 Ile), 1.20 (d, 3H, $^{3}J_{CH3,\ \alpha CH}$ = 7.0 Hz, CH_3 Ala), 1.06 (m, 1H, ½ γ CH_2 Ile), 0.82 (m, 6H, β CH_3, γ CH_3 Ile).

^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 172.17, 171.85, 171.27, 170.59, 169.59 (6C, 6 x CO), 156.02 (1C, COH Tyr), 130.36, 128.44, 115.16 (5C, Tyr), 61.91 (1C, β CH_2 Ser), 57.10 (1C, α CH Ile), 55.47 (1C, α CH Ser), 54.51 (1C, α CH Tyr), 49.76 (1C, α CH Asp), 48.98 (1C, α CH Ala), 37.28 (1C, β CH lle), 36.69 (1C, β CH_2 Tyr), 36.06 (1C, β CH_2 Asp), 24.55 (1C, γ CH_2 Ile), 22.77 (1C, $CO\underline{C}H_3$), 18.30 (1C, CH_3 Ala), 15.54 (1C, CH_3 Ile), 11.43 (1C, CH_3 Ile).

ES-MS: *m/z* for $C_{27}H_{40}N_6O_{10}$: $(M+H)^+$: 609.6; calculated: 609.7

_N_-Acetyl-L-tyrosyl-L-isoleucyl-L-glutaminyl-L-alanyl-L-serine amide (Ac-Tyr-Ile-Glu-Ala-Ser-NH_2) (51)

The title compound was synthesised following the protocol for solid phase peptide synthesis as outlined in the general procedures. After freeze drying pentapeptide **51** (71 mg, 11 %) was obtained as a white powder.

(51) $C_{28}H_{42}N_6O_{10}$
Yield: 71 mg, 11%
white powder

Ac-Tyr-Ile-Glu(COOH)-Ala-Ser-NH_2

^{1}H NMR (300 MHz, DMSO-d_6): δ ppm: 8.04-7.93 (m, 3H, NH Tyr, Ala, Glu), 7.87 (d, 1H, $^{3}J_{NH,\ \alpha CH}$ = 8.5 Hz, NH Ser or Ile), 7.77 (d, 1H, $^{3}J_{NH,\ \alpha CH}$ = 7.9 Hz, NH Ser or Ile), 7.07 (d, 2H, $^{3}J_{CH\ meta,CH\ ortho}$ = 8.4 Hz, 2 x CH Tyr), 6.63 (d, 2H, $^{3}J_{CH\ ortho,CH\ meta}$ = 8.4 Hz, 2 x CH Tyr), 4.46 (m, 1H, α CH Tyr), 4.30-4.26 (m, 2H, α CH Glu, Ala), 4.20-4.15 (m, 2H, α CH Ser, Ile), 3.56 (m, 2H, β CH_2 Ser), 2.86 (dd, 1H, $^{2}J_{\beta CH,\ \beta CH'}$ =

14.0 Hz, $^{3}J_{\beta CH, \alpha CH}$ = 3.7 Hz, ½ β CH_2 Tyr), 2.64-2.56 (m, 1H, ½ β CH_2 Tyr), 2.26 (m, 2H, γ CH_2 Glu), 1.92-1.86 (m,1H, ½ β CH_2 Glu), 1.74 (m, 5H, ½ β CH_2 Glu, $COCH_3$, β CH Ile), 1.43 (m, 1H, ½ γ CH_2 Ile), 1.21 (d, 3H, $^{3}J_{CH_3, \alpha CH}$ = 7.0 Hz, CH_3 Ala), 1.08 (m, 1H, ½ γ CH_2 Ile), 0.83 (m, 6H, β CH_3, γ CH_3 Ile).

^{13}C NMR (300 MHz, $CDCl_3$): δ ppm: 174.32, 172.20, 171.87, 171.32, 171.03, 169.49 (6C, 6 x CO), 156.02 (1C, COH Tyr), 130.36, 128.46, 115.14 (5C, Tyr), 61.97 (1C, β CH_2 Ser), 57.19 (1C, α CH Ile), 55.31 (1C, α CH Ser), 54.47 (1C, α CH Tyr), 52.15 (1C, α CH Glu), 48.69 (1C, α CH Ala), 36.91 (2C, β CH_2 Tyr, β CH Ile), 30.41 (1C, γ CH_2 Glu), 24.67 (1C, γ CH_2 Ile), 22.77 (1C, $CO\underline{C}H_3$), 18.34 (1C, CH_3 Ala), 15.65 (1C, CH_3 Ile), 11.37 (1C, CH_3 Ile).

ES-MS: *m/z* for $C_{28}H_{42}N_6O_{10}$: $(M+H)^+$: 623.3; calculated: 623.7

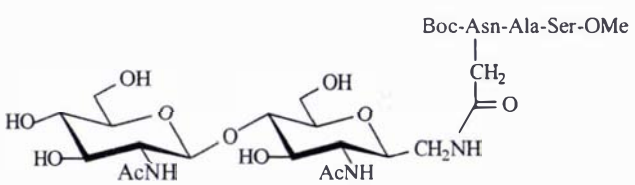
8.5 Synthesis of *N*-Linked Glycopeptides and Glycopeptide Mimetics

8.5.1 *C*-Glycopeptides

N^4-[4-*O*-(2-Acetamido-2-deoxy-β-D-glucopyranosyl)-(2-acetamido-2-deoxy-β-D-glucopyranosyl)methyl]-N^2-(*tert*-butoxycarbonyl)-L-asparaginyl-L-alanyl-L-serine methyl ester (Boc-Asn-(GlcNAc-β-1-4-GlcNAc-β-CH_2)-Ala-Ser-OMe) (52)

To glycopyranosyl methylamine **39** (20.5 mg, 47 μM, 1.5 eq) in DMSO (400 μL) and DMF (250 μL) was added sequentially *N*,*N*-diisopropylethylamine (8 μL, 1eq.), tripeptide **47b** (12.7 mg, 1eq.) in DMF (250 μL), HBTU (35.6 mg, 3eq.) in DMF (500 μL) and HOBt (4.2 mg, 1 eq.) in DMF (50 μL). The mixture was processed using the standard procedure (general procedures). Compound **52** was isolated as a white solid (12.5 mg, 48 %).

(52) $C_{33}H_{56}N_6O_{18}$
Yield: 12.5 mg, 48 %
white powder



^{1}H NMR (300 MHz, D_2O): δ ppm: 4.56-4.52 (m, 2H, H-1', α CH Ser), 4.42-4.31 (m, 2H, α CH Ala, Asn), 3.98-3.83 (m, 4H, β CH_2 Ser, H-6'a, H-6a), 3.75 (s, 3H, OMe), 3.71-3.39 (m, 12H, H-6'b, H-6b, 9 x carbohydrate CH, ½ CH_2N), 3.25-3.23 (m, 1H, ½ CH_2N), 2.77-2.54 (m, 2H, β CH_2 Asn), 2.04, 2.01 (s, 3H each, $COCH_3$), 1.40 (s, 12H, CH_3 Ala, Boc).

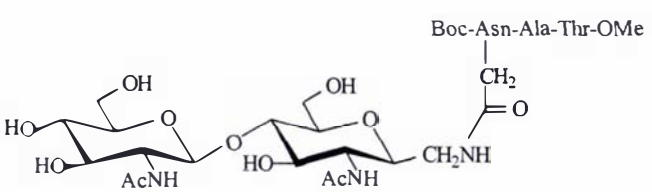
^{13}C NMR (300 MHz, D_2O): δ ppm: 174.99-172.39 (6C,CO), 101.89 (1C, C-1'), 80.15, 78.40, 76.34, 74.04, 73.88, 70.13 (6C, carbohydrate), 61.36 (1C, β CH_2 Ser), 60.96, 60.77 (2C, C-6, C-6'), 56.02 (1C, carbohydrate), 55.32 (1C, α CH Ser), 53.45 (1C, OMe), 52.64 (1C, C-2'), 52.15 (1C, carbohydrate), 50.15 (2C, α CH Ala, Asn), 40.66 (1C, CH_2N), 37.84 (1C, β CH_2 Asn), 28.01 (3C, $C(\underline{C}H_3)_3$ Boc), 22.55 (2C, 2 x $CO\underline{C}H_3$), 17.00 (1C, CH_3 Ala).

ES-MS: *m/z* for $C_{33}H_{56}N_6O_{18}$: $(M+H)^+$: 825.4; calculated: 825.8

N^4-[4-*O*-(2-Acetamido-2-deoxy-β-D-glucopyranosyl)-(2-acetamido-2-deoxy-β-D-glucopyranosyl)methyl]-N^2-(*tert*-butoxycarbonyl)-L-asparaginyl-L-alanyl-L-threonine methyl ester (Boc-Asn-(GlcNAc-β-1-4-GlcNAc-β-CH_2)-Ala-Thr-OMe) (53)

To glycopyranosyl methylamine **39** (27.5 mg, 63 μM, 1.5 eq) in DMSO (800 μL) and DMF (500 μL) was added sequentially *N*,*N*-diisopropylethylamine (9 μL, 1.2 eq), tripeptide **49b** (17.6 mg, 1eq) in DMF (1 mL), HBTU (47.7 mg, 3eq) in DMF (500 μL) and HOBt (5.7 mg, 1 eq) in DMF (50 μL). The mixture was processed using the standard procedure (general procedures). Compound **53** was isolated as a white solid (3.3 mg, 8 %) .

(53) $C_{34}H_{58}N_6O_{18}$
Yield: 18.5 mg, 53 %
white powder



^{1}H NMR (300 MHz, D_2O): δ ppm: 4.55 (1H, $^3J_{1',2'}$ = 8.3 Hz, H-1'), 4.96-4.27 (m, 3H, β CH Thr, α CH Ala, Asn), 3.91-3.78 (m, 2H, H-6'a, H-6a), 3.75 (s, 3H, OMe), 3.71-3.32 (m, 12H, 11x carbohydrate CH, ½ CH_2N), 3.25 (m, 1H, ½ CH_2N), 2.77-2.53 (m,

2H, β CH_2 Asn), 2.03, 2.01 (s, 3H each, $COCH_3$), 1.43-1.40 (m, 12H, CH_3 Ala, Boc), 1.17 (d, 3H, $^3J_{CH3, \beta CH}$ = 4.3 Hz, CH_3 Thr).

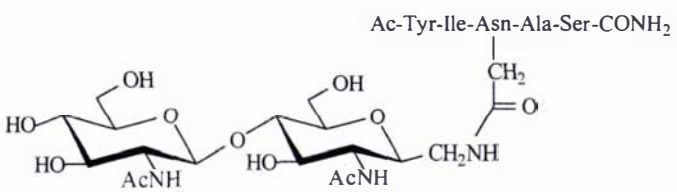
^{13}C NMR (300 MHz, D_2O): δ ppm: 180.37-173.64 (6C, CO), 106.59 (1C, C-1'), 84.85, 83.10, 81.73, 81.04, 78.74, 78.59, 74.83 (7C, carbohydrate), 72.25 (1C, β CH Thr), 65.65, 65.48 (2C, C-6, C-6'), 63.19 (1C, α CH Thr), 60.73 (1C, C-2'), 58.13 (1C, OMe), 57.36 (1C, C-2), 56.86 (1C, α CH Asn), 54.91 (1C, α CH Ala), 45.38 (1C, CH_2N), 42.53 (1C, β CH_2 Asn), 32.72 (3C, $C(\underline{C}H_3)_3$ Boc), 27.26 (2C, 2 x $CO\underline{C}H_3$), 23.83 (1C, CH_3 Thr), 21.71 (1C, CH_3 Ala).

ES-MS: *m/z* for $C_{34}H_{58}N_6O_{18}$: $(M+H)^+$: 839.7; calculated: 839.9

N^4-\|4-*O*-(2-Acetamido-2-deoxy-β-D-glucopyranosyl)-(2-acetamido-2-deoxy-β-D-glucopyranosyl)methyl\|-N^2-(*N*-acetyl-L-tyrosyl-L-isoleucyl)-L-asparaginyl-L-alanyl-L-serine amide (Ac-Ile-Tyr-Asn-(GlcNAc-β-1-4-GlcNAc-β-CH_2)-Ala-Ser-NH_2) (54)

To glycopyranosyl methylamine **39** (25 mg, 57 μM, 1.5 eq) in DMSO (800 μL) and DMF (500 μL) was added sequentially *N*,*N*-diisopropylethylamine (8 μL, 1.2 eq.), pentapeptide **50** (23.2 mg, 1eq.) in DMF (1 mL), HBTU (43.4 mg, 3eq.) in DMF (500 μL) and HOBt (5.2 mg, 1 eq.) in DMF (50 μL). The mixture was processed using the standard procedure (general procedures). Compound **54** was isolated as a white solid (18.2 mg, 46 %).

(54) $C_{44}H_{69}N_9O_{19}$
Yield: 18.2 mg, 46 %
white powder



^{1}H NMR (300 MHz, DMSO-d_6): δ ppm: 9.13 (s, 1H, OH Tyr), 8.19-8.17 (m, 2H, 2 x NH), 8.99-7.89 (m, 3H, N^{4}H Asn, 2 x NH), 7.77-7.72 (m, 3H, NH, NHAc, NHAc'), 7.06-6.91 (m, 4H, 2 x CH Tyr, unknown signals), 6.63 (d, 2H, $^{3}J_{CHortho,CHmeta}$ = 8.3 Hz, 2 x CH Tyr), 5.04-4.96 (m, 2H, 2 x OH), 4.76-4.69 (m, 3H, 3 x OH), 4.53-4.41 (m, 3H, α CH Asn, Tyr, unknown signal), 4.35 (d, 1H, $^{3}J_{1',2'}$ = 8.3 Hz, H-1'), 4.22-4.16 (m, 3H, α CH Ile, Ala, Ser), 3.75-3.63 (m, 4H, β CH_2 Ser, H-6a, H-6'a), 3.51-3.17 (m, signals buried under or distorted by water signal, H-6b, H-6'b, H-2, H-2', ½ CH_2N, 6 x carbohydrate CH), 3.07-3.04 (m, 1H, carbohydrate CH), 2.88-2.85 (m, 2H, ½ CH_2N, ½ β CH_2 Tyr), 2.63-2.49 (m, signals distorted by or buried under DMSO signal, ½ β CH_2 Tyr, β CH_2 Asn), 1.82 (s, 6H, 2 x $COCH_3$), 1.75 (s, 3H, $COCH_3$), 1.69 (m, 1H, β CH Ile), 1.40 (m, 1H, ½ γ CH_2 Ile), 1.22 (d, 3H, $^{3}J_{CH3,\alpha CH}$ = 7.1 Hz, CH_3 Ala), 1.06 (m, 1 H, ½ γ CH_2 Ile), 0.82-0.80 (m, 6H, 2 x CH_3 Ile).

^{13}C NMR (300 MHz, DMSO-d_6): δ ppm: 172.32, 171.78, 171.09, 169.85, 169.46 (9C, 9 x CO), 156.02 (1C, COH Tyr), 130.34, 128.46, 115.17 (5C, Tyr), 102.39 (1C, C-1'), 82.14, 78.79, 77.30, 74.31, 73.51, 71.01 (7C, carbohydrate), 61.86 (1C, β CH_2 Ser), 61.36 (2C, C-6, C-6'), 56.98 (1C, α CH), 55.95 (2C, C-2', α CH), 54.54 (1C, α CH Tyr), 52.61 (1C, C-2), 50.08 (1C, α CH Asn), 49.34 (1C, α CH), 40.73-39.06 (2C, CH_2N, β CH Ile under DMSO septett), 37.39 (1C, β CH_2 Asn), 36.71 (1C, γ CH_2 Ile), 23.37, 23.15, 22.79 (3C, 3 x $COCH_3$), 17.87 (1C, CH_3 Ala), 15.54, 11.47 (2C, 2 x CH_3 Ile).

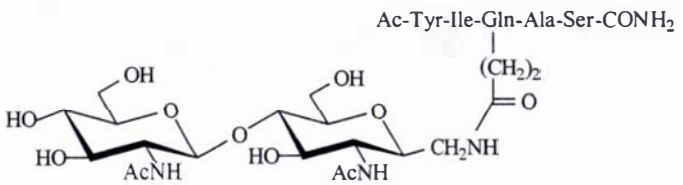
ES-MS: *m/z* for $C_{44}H_{69}N_9O_{19}$: $(M+H)^+$: 1028.5; calculated: 1029.1

N^5-[4-*O*-(2-Acetamido-2-deoxy-β-D-glucopyranosyl)-(2-acetamido-2-deoxy-β-D-glucopyranosyl)methyl]-N^2-(*N*-acetyl-L-tyrosyl-L-isoleucyl)-L-glutaminyl-L-alanyl-L-serine amide (Ac-Ile-Tyr-Gln-(GlcNAc-β-1-4-GlcNAc-β-CH_2)-Ala-Ser-NH_2) (55)

To glycopyranosyl methylamine **39** (24 mg, 55 μM, 1.5 eq) in DMSO (800 μL) and DMF (500 μL) was added sequentially *N*,*N*-diisopropylethylamine (6 μL, 1 eq.),

pentapeptide **51** (22.8 mg, 1 eq.) in DMF (1 mL), HBTU (41.6 mg, 3 eq.) in DMF (500 μL) and HOBt (4.9 mg, 1 eq.) in DMF (50 μL) and the mixture was processed using the standard procedure (general procedures). Compound **55** was isolated as a white solid (5.9 mg, 15 %).

(55) $C_{45}H_{71}N_9O_{19}$
Yield: 5.9 mg, 15 %
white grains



^{1}H NMR (300 MHz, DMSO-d_6): δ ppm: 9.06 (s, 1H, OH Tyr), 8.03 (s, 1H, ^3J = 7.3 Hz, NH), 7.95-7.90 (m, 2H, 2 x NH), 7.72-7.61 (m, 4H, NHAc, NHAc', 2 x NH), 7.12 (s, 1H, unknown signal), 6.96 (m, 3H, 2 x CH Tyr, unknown signal), 6.56 (d, 2H, $^3J_{CHmeta,CHortho}$ = 8.4 Hz, 2 x CH Tyr), 6.44 (s, 1H, unknown signal), 4.97 (d, 1H, ^3J = 5.3 Hz, OH), 4.88 (d, 1H, ^3J = 5.2 Hz, OH), 4.78 (t, 1H, ^3J = 5.3 and 5.5 Hz, OH), 4.61-4.59 (m, 2H, OH), 4.38 (m, 1H, α CH Tyr), 4.28 (d, 1H, $^3J_{1',2'}$ = 8.4 Hz, H-1'), 4.23-4.12 (m, 4H, 4 x α CH), 3.67-3.64 (m, 1H, H-6), 3.55-3.47 (m, 3H, β CH_2 Ser, H-6), 3.45-3.32 (m, 6H, ½ CH_2N, 2 x H-6, H-2, H-2', carbohydrate CH), 3.24 (m, 1H, carbohydrate CH under water signal), 3.14-2.93 (m, 5 x carbohydrate CH), 2.82-2.78 (m, 1H, ½ β CH_2 Tyr), 2.66 (m, 1H, ½ CH_2N), 2.58-2.54 (m, 1H, ½ β CH_2 Tyr), 2.07 (m, 2H, γ CH_2 Gln), 1.82-1.68 (m, 12H, 3 x $COCH_3$, β CH_2 Gln, β CH Ile), 1.35 (m, 1H, ½ γ CH_2 Ile), 1.14 (d, 3H, $^3J_{CH_3,\alpha CH}$ = 6.9 Hz, CH_3 Ala), 1.01 (m, 1H, ½ γ CH_2 Ile), 0.77-0.72 (m, 6H 2 x CH_3 Ile).

^{13}C NMR (300 MHz, DMSO-d_6): δ ppm: 172.26, 171.76, 171.21, 169.68 (9C, 9 x CO), 156.03 (1C, COH Tyr), 130.34, 128.46, 115.18 (5C, Tyr), 102.41 (1C, C-1'), 82.25, 78.77, 77.74, 77.31, 74.34, 73.56, 71.00 (7C, carbohydrate), 61.99 (1C, CH_2 Ser), 61.36 (2C, C-6, C-6'), 57.13 (1C, α CH), 55.89 (1C, C-2), 55.28 (1C, α CH), 54.57 (1C, α CH Tyr), 52.99 (1C, C-2), 52.59 (1C, α CH), 48.73 (1C, α CH), 40.73-39.07 (1C, CH_2N under DMSO septett), 39.07 (1C, β CH Ile), 36.71 (1C, β CH_2 Tyr), 32.03 (1C, γ CH_2 Gln), 28.14 (1C, β CH_2 Gln), 24.61 (1C, γ CH_2 Ile), 23.36, 23.19, 22.78 (3C, 3 x $COCH_3$), 18.30 (1C, CH_3 Ala), 15.64, 11.42 (2C, 2 x CH_3 Ile).

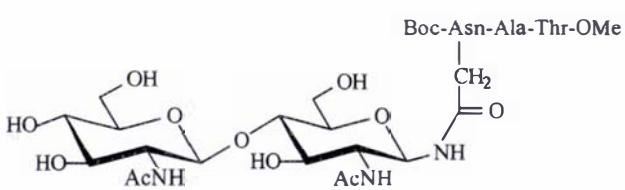
ES-MS: m/z for $C_{45}H_{71}N_9O_{19}$: $(M+H)^+$: 1042.5; calculated: 1043.1

8.5.2 *N*-Glycopeptides

N^4-[4-*O*-(2-Acetamido-2-deoxy-β-D-glucopyranosyl)-2-acetamido-2-deoxy-β-D-glucopyranosyl]-N^2-(tert-butoxycarbonyl)-L-asparaginyl-L-alanyl-L-threonine methyl ester (Boc-Asn-(GlcNAc-β-1-4-GlcNAc-β)-Ala-Thr-NH_2) (56)

To the β-glycopyranosylamine **45** (30 mg, 70.85 μM, 1.5 eq.) in DMSO (450 μL) and DMF (250 μL) was added sequentially *N*,*N*-diisopropylethylamine (10 μL, 1.2 eq.), tripeptide **49b** (19.8 mg, 1 eq.) in DMF (500 μL), HBTU (53.7 mg, 3 eq.) in DMF (500 μL) and HOBt (6.4 mg, 1 eq.) in DMF (50 μL) and the mixture was processed using the standard procedure (general procedures). Compound **56** was isolated as a white solid (7.4 mg, 19 %).

(56) $C_{33}H_{56}N_6O_{18}$
Yield: 7.4 mg, 19 %
white powder



^{1}H NMR (300 MHz, D_2O): δ ppm: 5.04 (d, 1H, $^3J_{1,2}$ = 9.5 Hz, H-1), 4.57 (d, 1H, $^3J_{1',2'}$ = 8.4 Hz, H-1'), 4.50 (d, 1H, $^3J_{\alpha CH, \beta CH}$ = 3.0 Hz, α CH Thr), 4.43-4.35 (m, 3H, α CH Asn, Ala, β CH Thr), 3.92-3.78 (m, 3H, H-6a, H-6'a, H-2), 3.75 (s, 3H, OMe), 3.73-3.41 (m, 9H, H-6b, H-6'b, H-5, H-5', H-4, H-4', H-3, H-3', H-2'), 2.78 (dd, 1H, $^2J_{CH, CH'}$ = 15.9 Hz, $^3J_{CH, \alpha CH}$ = 4.6 Hz, ½ β CH_2 Asn), 2.62 (dd, 1H, $^2J_{CH, CH'}$ = 15.9 Hz, $^3J_{CH', \alpha CH}$ = 8.9 Hz, ½ β CH_2 Asn), 2.04, 2.00 (2s, 3H each, 2 x $COCH_3$), 1.41 (m, 12H, Ala CH_3, Boc), 1.17 (d, 1H, $^3J_{CH3, \beta CH}$ = 6.44 Hz, CH_3 Thr).

^{13}C NMR (300 MHz, D_2O): δ ppm: 175.66, 175.19, 175.02, 173.61, 172.90, 172.68 (6C, 6 x CO), 101.81 (1C, C-1'), 79.24 (1C, C-4'), 78.62 (1C, C-1), 76.62, 76.34 (2C, C-5, C-5'), 73.86, 73.22 (2C, C-3, C-3'), 70.14 (1C, C-4), 67.53 (1C, β CH Thr),

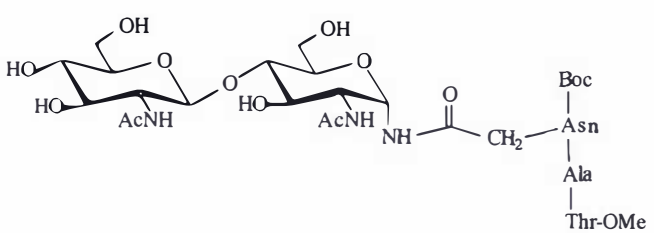
60.98, 60.33 (2C, C-6, C-6'), 58.45 (1C, α CH Thr), 56.01 (1C, C-2'), 54.18 (1C, C-2), 53.42 (1C, OMe), 51.50 (1C, α CH Asn), 50.21 (1C, α CH Ala), 37.65 (1C, β CH_2 Asn), 27.98 (3C, C($\underline{C}H_3$)$_3$ Boc), 22.54 (2C, 2 x CO$\underline{C}H_3$), 19.10 (1C, CH_3 Thr), 17.03 (1C, CH_3 Ala).

ES-MS: *m/z* for $C_{33}H_{56}N_6O_{18}$ $(M+H)^+$: 825.5; calculated: 825.8

N^4-[4-*O*-(2-Acetamido-2-deoxy-β-D-glucopyranosyl)-2-acetamido-2-deoxy-α-D-glucopyranosyl]-N^2-(*tert*-butoxycarbonyl)-L-asparaginyl-L-alanyl-L-threonine methyl ester (Boc-Asn-(GlcNAc-β-1-4-GlcNAc-α)-Ala-Thr-NH_2) (57)

To the α-glycopyranosylamine **44** (40 mg, 94 μM, 1.5 eq.) in DMSO (450 μL) and DMF (250 μL) was added sequentially *N*,*N*-diisopropylethylamine (13 μL, 1.2 eq.), tripeptide **49b** (26.38 mg, 1 eq.) in DMF (500 μL), HBTU (71.56 mg, 3 eq.) in DMF (500 μL) and HOBt (8.50 mg, 1 eq.) in DMF (50 μL) and the mixture was processed using the standard procedure (general procedures). Compound **57** was isolated as a white solid (13 mg, 25 %).

(57) $C_{33}H_{56}N_6O_{18}$
Yield: 13 mg, 25 %
white powder



^{1}H NMR (300 MHz, D_2O): δ ppm: 4.60 (d, 1H, $^3J_{1',2'}$ = 8.4 Hz, H-1'), 4.52-4.48 (m, 1H, α CH Thr), 4.41-4.33 (m, 3H, α CH Asn, Ala, β CH Thr), 3.91-3.78 (m, 5H, H-6a, H-6b, H-1, 2 x carbohydrate CH), 3.75 (m, 3H, OMe), 3.74-3.68 (m, 2H, H-6'a, H-2'), 3.56-3.38 (m, 5H, carbohydrate CH), 2.71 (dd, 1H, $^2J_{CH, CH'}$ = 15.4 Hz, $^3J_{CH, \alpha CH}$ = 4.5 Hz, ½ β CH_2 Asn), 2.04, 2.00 (s, 3H each, $COCH_3$), 1.41 (m, 12H, CH_3 Ala, Boc), 1.17 (m, 3H, CH_3 Thr).

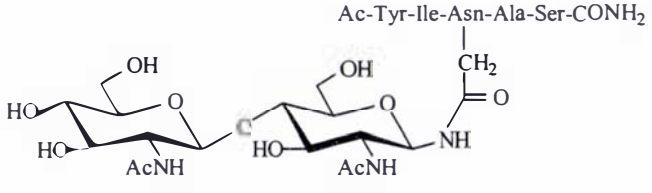
^{13}C NMR (300 MHz, D_2O): δ ppm: 175.67, 175.18, 174.67, 173.67, 172.69, 172.38 (6C, 6 x CO), 101.78 (1C, C-1'), 79.97 (1C, C-1), 76.10, 74.08, 71.86, 70.07, 69.45 (5C, carbohydrate), 67.53 (1C, β CH Thr), 62.19, 60.93 (2C, C-6, C-6'), 58.46 (1C, α CH Thr), 56.14 (1C, C-2'), 53.41 (1C, OMe), 51.90 (1C, α CH Asn), 51.12 (1C, C-2), 50.19 (1C, α CH Ala), 40.94 (1C, unknown), 37.72 (1C, β CH_2 Asn), 27.98 (3C, C($\underline{C}H_3)_3$ Boc), 22.60 (2C, 2 x CO$\underline{C}H_3$), 19.10 (1C, CH_3 Thr), 17.03 (1C, CH_3 Ala).

ES-MS: *m/z* for $C_{33}H_{56}N_6O_{18}$ $(M+H)^+$: 825.3; calculated: 825.8

N^4-[4-*O*-(2-Acetamido-2-deoxy-β-D-glucopyranosyl)-2-acetamido-2-deoxy-β-D-glucopyranosyl]-N^2-(*N*-acetyl-L-tyrosyl-L-isoleucyl)-L-asparaginyl-L-alanyl-L-serine amide (Ac-Tyr-Ile-Asn-(GlcNAc-β-1-4-GlcNAc-β)-Ala-Ser-NH_2) (58)

To the β-glycopyranosylamine **45** (30 mg, 71 μM, 1.5 eq.) in DMSO (450 μL) and DMF (250 μL) was added sequentially *N*,*N*-diisopropylethylamine (10 μL, 1.2 eq.), pentapeptide **50** (28.7 mg, 1 eq.) in DMF (500 μL), HBTU (53.7 mg, 3 eq.) in DMF (500 μL) and HOBt (6.4 mg, 1 eq.) in DMF (50 μL) and the mixture was processed using the standard procedure (general procedures). Compound **58** was isolated as a white solid (19 mg, 40 %).

(58) $C_{43}H_{67}N_9O_{19}$
Yield: 19 mg, 40 %
white powder



^{1}H NMR (300 MHz, DMSO-d_6): δ ppm: 9.11 (s, 1H, OH Tyr), 8.31 (d, 1H, $^3J_{NH,1}$ = 8.9 Hz, N^4H Asn), 8.18 (d, 1H, $^3J_{NH,\alpha CH}$ = 7.5 Hz, N^2H Asn), 7.91 (m, 2H, NH Tyr, NH), 7.82-7.76 (m, 4H, NH Ala, Ser, N$\underline{H}$Ac, N$\underline{H}$Ac'), 6.95 (d, 3H, $^3J_{CHortho,\ CHmeta}$ = 8.4 Hz, 2 x CH Tyr, unknown signal), 6.56 (d, 2H, $^3J_{CHmeta,\ CHortho}$ = 8.4 Hz, 2 x CH Tyr), 4.98 (d, 1H, ^3J = 5.4 Hz, OH), 4.88 (d, 1H, ^3J = 5.3 Hz, OH), 4.77-4.72 (m, $^3J_{1,2}$ = ~8.5 Hz, H-1, unknown signal), 4.62 (m, 1H, OH), 4.55 (m, 1H, OH), 4.49-4.47 (m,

1H, α CH Asn), 4.39-4.35 (m, 1H, α CH Tyr), 4.30 (d, 1H, $^{3}J_{1',2'}$ = 8.3 Hz, H-1'), 4.14-4.08 (m, 3H, α CH Ile, Ala, Ser), 3.67 (m, 1H , H-6), 3.56-3.52 (m, 2H, β CH_2 Ser, H-2, H-6), 3.49-3.36 (m, 3H, 2 x H-6, H-2', carbohydrate CH), 3.12 (m, 2H, 2 x carbohydrate CH), 2.97 (m, 1H, carbohydrate CH), 2.83-2.79 (m, 1H, ½ β CH_2 Tyr), 2.67-2.50 (m, 2H, ½ β CH_2 Tyr, ½ β CH_2 Asn), 2.38-2.33 (m, 1H, ½ β CH_2 Asn), 1.76, 1.73, 1.69 (3s, 3H each, 3 x $COCH_3$), 1.64 (m, 1H, β CH Ile), 1.34 (m, 1H, ½ γ CH_2 Ile), 1.13 (d, 3H, $^{3}J_{CH_3, \alpha CH}$ = 7.1 Hz, CH_3 Ala), 0.98 (m, 1H, ½ γ CH_2 Ile), 0.73 (m, 6H, 2 x CH_3 Ile).

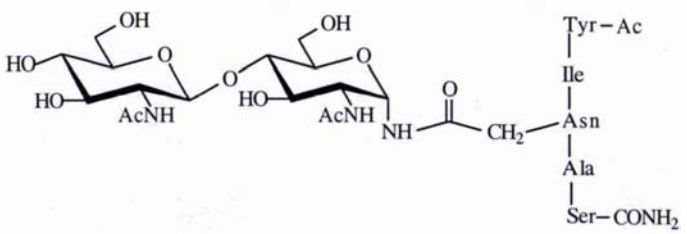
^{13}C NMR (300 MHz, DMSO-d_6): δ ppm: 172.24, 171.75, 171.19, 170.71, 170.27, 169.60 (9C, 9 x CO), 156.03 (1C, COH Tyr), 130.35, 128.43, 115.16 (5C, Tyr), 102.44 (1C, C-1'), 81.67 (1C, carbohydrate), 79.00 (1C, C-1), 77.29, 74.34, 73.16, 71.04 (5C, carbohydrate), 61.85 (1C, β CH_2 Ser), 61.40, 60.09 (2C, C-6, C-6'), 56.98 (1C, α CH Ala), 55.89 (2C, α CH Ser, C-2'), 54.48 (2C, α CH Tyr, C-2), 49.21 (1C, α CH Ile, Asn), 37.33 (1C, α CH, Ile), 36.73 (1C, β CH_2 Asn), 24.57 (1C, γ CH_2 Ile), 23.35, 23.13, 22.78 (3C, 3 x $COCH_3$), 18.12 (1C, CH_3 Ala), 15.53, 11.43 (2C, 2 x CH_3 Ile).

ES-MS: m/z for $C_{43}H_{67}N_9O_{19}$: $(M+H+Na)^{2+}$: 1037.5; calculated: 1038.0

N^4-[4-O-(2-Acetamido-2-deoxy-β-D-glucopyranosyl)-2-acetamido-2-deoxy-α-D-glucopyranosyl]-N^2-(N-acetyl-L-tyrosyl-L-isoleucyl)-L-asparaginyl-L-alanyl-L-serine amide (Ac-Tyr-Ile-Asn-(GlcNAc-β-1-4-GlcNAc-α)-Ala-Ser-NH_2) (59)

To the α-glycopyranosylamine **44** (19.1 mg, 45 μM, 1.5 eq.) in DMSO (450 μL) and DMF (250 μL) was added sequentially *N*,*N*-diisopropylethylamine (8 μL, 1 eq.), pentapeptide **50** (18.4 mg, 1 eq.) in DMF (250 μL), HBTU (34.5 mg, 3 eq.) in DMF (500 μL) and HOBt (4.1 mg, 1 eq.) in DMF (50 μL) and the mixture was processed using the standard procedure (general procedures). Compound **59** was isolated as a white solid (12.5 mg, 41 %).

(59) $C_{43}H_{67}N_9O_{19}$
Yield: 12.5 mg, 41 %
white powder



^{1}H NMR (300 MHz, DMSO-d_6): δ ppm: 9.15 (s, 1H, OH Tyr), 8.17 (d, 1H, $^3J_{NH, \alpha CH}$ = 7.5 Hz, N^{2}H Asn), 8.06 (d, 1H, ^{3}J = 6.6Hz, NH), 7.99 (d, 2H, $^3J_{NH, \alpha CH}$ = 8.4 Hz, NH Ser, Tyr), 7.94-7.91 (m, 2H, NH, unknown signal), 7.79 (d, 1H, $^3J_{NH, \alpha CH}$ = 8.6 Hz, NH Ile), 7.71 (d, 1H, $^3J_{NH', 2'}$ = 8.3 Hz, N<u>H</u>Ac'), 7.42 (d, 1H, ^{3}J = 8.8 Hz, NH), 7.03 (m, 4H, 2 x CH Tyr, unknown signals), 6.64 (d, 2H, $^3J_{CHortho, CHmeta}$ = 8.4 Hz, 2 x CH Tyr), 5.00-4.95 (m, 2H, 2 x OH), 4.83 (t, 1H, ^{3}J = 5.8 Hz, OH), 4.63 (t, 1H, ^{3}J = 5.6 Hz, OH), 4.52 (m, 2H, α CH Asn, unknown signal), 4.44 (m, 3H, $^3J_{1', 2'}$ = 8.4 Hz, H-1', α CH Tyr, Ser), 4.24-4.13 (m, 3H, α CH Ile, Ala, unknown signal), 4.06 (m, 1H, unknown signal), 3.73-3.48 (m, 8H, 3 x H-6, β CH_2 Ser, 3 x carbohydrate CH), 3.20 (m, 2H, unknown signal), 3.11 (m, 2H, 2 x carbohydrate CH), 2.87 (dd, 1H, $^2J_{\beta CH', \beta CH}$ = 14.1 Hz, $^3J_{\beta CH', \alpha CH}$ = 3.5 Hz, ½ β CH_2 Tyr), 2.64-2.54 (m, 2H, ½ β CH_2 Tyr, ½ β CH_2 Asn), 2.52-2.44 (m, ½ β CH_2 Asn under DMSO septett), 1.86, 1.82 (2s, 6H, 2 x $COCH_3$), 1.75 (m, 4H, $COCH_3$, β CH Ile), 1.41 (m, 1H, ½ γ CH_2 Ile), 1.22 (d, 3H, $^3J_{CH3, \alpha CH}$ = 7.2 Hz, CH_3 Ala), 1.07 (m, 1H, ½ γ CH_2 Ile), 0.82 (m, 6H, 2x CH_3 Ile).

^{13}C NMR (300 MHz, DMSO-d_6): δ ppm: 172.36, 171.87, 171.18, 170.37, 170.22, 170.09, 169.60 (9C, 9 x CO), 156.03 (1C, COH Tyr), 130.35, 128.47, 115.18 (5C, Tyr), 102.37 (1C, C-1'), 82.64, 77.14, 74.47, 71.72, 70.81, 69.23 (6C, carbohydrate), 62.08 (1C, β CH_2, Ser), 61.83, 61.31 (2C, C-6, C-6'), 57.06 (1C, α CH Ile), 56.38 (1C, C-2'), 55.97 (1C, carbohydrate), 54.53 (2C, α CH Tyr, Ser), 50.40 (1C, unknown signal), 49.93 (1C, α CH Asn), 49.36 (1C, α CH Ala), 37.31 (2C, β CH_2 Asn, β CH Ile), 24.56 (1C, γ CH_2 Ile), 23.43, 23.19, 22.78 (3C, 3 x CO<u>C</u>H_3), 17.97 (1C, CH_3 Ala), 15.54, 11.44 (2C, 2 x CH_3 Ile).

^{1}H NMR (300 MHz, MeOH-d_4): δ ppm: 7.08 (d, 2H, $^3J_{CHortho, CHmeta}$ = 8.5 Hz, 2 x CH Tyr), 6.72 (d, 2H, $^3J_{CHmeta, CHortho}$ = 8.5 Hz, 2 x CH Tyr), 4.71-4.67 (m, 1H, α CH Asn), 4.62-4.67 (m, 2H, $^3J_{1',2'}$ = 8.4 Hz, H-1', α CH Tyr), 4.43-4.40 (m, 1H, α CH Ser), 4.29-4.21 (m, 2H, α CH Ala, Ile), 4.01-3.95 (m, 1H, β CH_2 Ser), 3.92-3.37 (m, 11H, carbohydrate), 3.09-3.02 (m, 1H, ½ β CH_2 Tyr), 2.84-2.74 (m, 3H, ½ β CH_2 Tyr, β CH_2 Asn), 2.03 (s, 6H, 2 x $COCH_3$), 1.93 (s, 3H, $COCH_3$), 1.83 (m, 1H, β CH Ile), 1.55 (m, 1H, ½ γ CH_2 Ile), 1.45 (d, 3H, $^3J_{CH_3, \alpha CH}$ = 7.3 Hz, CH_3 Ala), 1.21 (m, 1H, ½ γ CH_2 Ile), 0.94-.89 (m, 6H, 2 x CH_3 Ile).

^{13}C NMR (300 MHz, MeOH-d_4): δ ppm: 175.72-172.99 (9C, 9 x CO), 156.83 (1C, COH Tyr), 131.67, 129.53, 116.67 (5C, Tyr), 103.51 (1C, C-1'), 82.39, 78.45, 76.32, 73.78, 72.49, 70.94 (6C, carbohydrate), 63.66, 63.31 (2C, C-6, C-6'), 63.09 (1C, β CH_2 Ser), 59.72 (1C, α CH, Ile), 58.06 (2C, α CH Ser, C-2'), 56.75 (1C, α CH Tyr), 52.93 (1C, α CH Ala), 51.83 (1C, α CH Asn), 52.18 (1C, unknown signal), 39.02 (1C, β CH_2 Asn), 38.58 (1C, β CH Ile), 38.23 (1C, β CH_2 Tyr), 26.44 (1C, γ CH_2 Ile), 23.48, 22.79 (3C, 3 x $COCH_3$), 17.69 (1C, CH_3 Ala), 16.22, 11.80 (2C, 2 x CH_3 Ile).

ES-MS: *m/z* for $C_{43}H_{67}N_9O_{19}$: $(M+H)^{3+}$: 1016.4; calculated: 1017.1

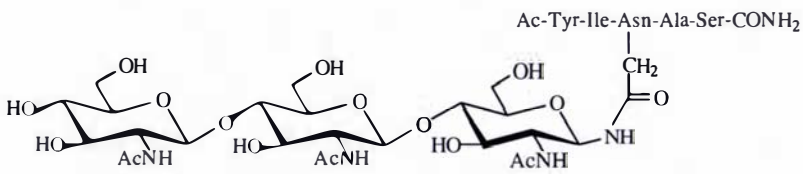
N^4-[4-*O*-(2-acetamido-2-deoxy-β-D-glucopyranosyl)-(1-4)-(2-acetamido-2-deoxy-β-D-glucopyranosyl)-2-acetamido-2-deoxy-β-D-glucopyranosyl]-N^2-(*N*-acetyl-L-tyrosyl-L-isoleucyl)-L-asparaginyl-L-alanyl-L-serine amide (Ac-Tyr-Ile-Asn-(GlcNAc-β-1-4- GlcNAc-β-1-4-GlcNAc-β)-Ala-Ser-NH_2) (60)

To the β-glycopyranosylamine **46** (40 mg, 64 μM, 1.5 eq.) in DMSO (2 mL) was added sequentially *N*,*N*-diisopropylethylamine (9 μL, 1 eq.), pentapeptide **50** (25.9 mg, 1 eq.) in DMF (500 μL), HBTU (48.4 mg, 3 eq.) in DMF (500 μL) and HOBt (5.8 mg, 1 eq.) in DMF (50 μL) and the mixture was processed using the standard procedure (general procedures). Compound **60** was isolated as a white solid (21.4 mg, 41 %).

(60) $C_{51}H_{80}N_{10}O_{24}$

Yield: 21.4 mg, 41 %

white powder



^{1}H NMR (300 MHz, DMSO-d_6): δ ppm: 9.11 (m, 1H, OH Tyr), 8.31 (d, 1H, $^3J_{NH,1}$ = 8.9 Hz, N^4H Asn), 8.18 (d, 1H, $^3J_{NH,\alpha CH}$ = 7.7 Hz, N^2H Asn), 7.90 (d, 2H, ^3J = 8.2 Hz, NH Tyr, NH), 7.80-7.70 (m, 4H, 2 x NH, 2 x NHAc), 7.65 (d, 1H, $^3J_{NH,2}$ = 9.0 Hz, NHAc), 6.95 (m, 4H, 2 x CH Tyr, 2 x unknown), 6.56 (d, 2H, $^3J_{CHmeta,\ CHortho}$ = 8.4 Hz, 2 x CH Tyr), 4.76 (t, 1H, $^3J_{1,2}$ = 9.3 Hz, H-1), 4.63 (m, 2H, 2 x OH), 4.48 (m, 1H, α CH Asn), 4.38 (m, 1H, α CH Tyr), 4.33 (d, 1H, $^3J_{1',2'}$ = 7.4 Hz, H-1' *), 4.27 (d, 1H, $^3J_{1'',2''}$ = 8.4 Hz, H-1'' *), 4.14-4.05 (m, 3H, α CH Ile, Ala, Ser), 3.68-3.41 (m, 8H, 4 x H-6, 2 x H-2, β CH_2 Ser), 2.99-2.93 (m, 2H, 2x carbohydrate CH), 2.80 (m, 1H, ½ β CH_2 Tyr), 2.67-2.50 (m, 2H, ½ β CH_2 Asn, ½ β CH_2 Tyr), 2.34 (m, 1H, ½ β CH_2 Asn), 1.75-1.73 (m, 9H, 3 x $COCH_3$), 1.69 (s, 3H, $COCH_3$), 1.64 (m, 1H, β CH Ile), 1.34 (m, 1H, ½ γ CH_2 Ile), 1.13 (d, 3H, $^3J_{CH3,\alpha CH}$ = 7.1 Hz, CH_3 Ala), 0.98 (m, 1H, ½ γ CH_2 Ile), 0.76-0.71 (m, 6H, 2 x CH_3 Ile).

* assignments of H-1' and H-1'' might be reversed

^{13}C NMR (300 MHz, DMSO-d_6): δ ppm: 172.23, 171.75, 171.18, 170.86, 170.70, 170.20, 169.48 (10C, 10 x CO), 156.03 (1C, COH Tyr), 130.35, 128.43, 115.16 (5C, Tyr), 102.44, 102.19 (2C, C-1', C-1''), 82.10, 81.47 (2C, carbohydrate), 78.94 (1C, C-1), 77.29, 76.95, 75.09, 74.36, 73.19, 72.82, 71.02 (8C, carbohydrate), 61.85 (1C, β CH_2 Ser), 61.37, 60.52, 60.09 (3C, C-6, C-6', C-6''), 56.97 (1C, α CH), 55.97 (2C, α CH, C-2), 55.06 (1C, C-2), 54.50 (1C, α CH Tyr), 54.29 (1C, C-2), 49.20 (2C, α CH Asn, α CH), 37.34 (1C, β CH Ile), 36.74 (2C, β CH_2 Asn, Tyr), 24.57 (1C, γ CH_2 Ile), 23.37, 23.15, 22.79 (4C, 4 x $COCH_3$), 18.14 (1C, CH_3 Ala), 15.54, 11.43 (2C, 2 x CH_3 Ile).

ES-MS: *m/z* for $C_{51}H_{80}N_{10}O_{24}$ $(M+H)^+$: 1217.6; calculated: 1218.2

8.6 Synthesis of FITC-Labelled Ovalbumin Glycopeptide Substrate

The protocol for FITC-labelling of proteins and peptides was adapted from Hentz *et al.* [104] with some modifications as noted below.

Material:

Ovalbumin glycopeptide was isolated from grade II chicken egg albumin (Sigma) after a procedure described by Norris *et al.* [63]. Fluorescein isothiocyanate isomer I (Aldrich), acetone reagent grade (Aldrich).

Chromatographic media used includes: Jupiter C_{18} column (5μ, 250 x 10 mm) with Jupiter Security Guard column (50 x 10 mm) both from Phenomenex.

Method:

Purified ovalbumin glycopeptide (18 mg) was dissolved in Na_2HPO_4 buffer (100 mM, pH 7.0, 5mL).

FITC isomer I (4 mg) was dissolved in reagent grade acetone (500 μL).

The FITC-solution in acetone was added dropwise to the ovalbumin glycopeptide in the darkness under stirring.

After reverse phase HPLC showed completion of the labelling reaction the product was purified using preparative reverse phase HPLC (Solvent A: water 0.1 % TFA, Solvent B: MeCN 0.08 % TFA) with the following gradient: 1) isocratic flow at 80 % A, 20 % B for 2 min.; 2) gradient to 60 % A, 40 % B from starting conditions over 15 min.; 3) gradient to 30 % A, 70 % B over 10 min.; 4) gradient to 100 % B over 5 min.; 5) isocratic flow at 100 % B for 5 min.; 6) gradient to 80 % A, 20 % B over 10 min. Eluents were monitored at λ 214 nm. For the use of a fluorescence detector the exitation wavelength was 495 nm and the emission wavelength 520 nm.

The reaction should have resulted in a single labelled product. However, it was shown that two labelling products occured in different ratios varying between reactions.

A possible explanation for this observation was that the homoserine-lactone on the *C*-terminal side of the peptide could exist in both its lactone or the "open" form.

The labelling product was therefore totally converted to the "open" form by boiling in NH_4HCO_3-buffer (100 mM, pH 8.5) followed by lyophilisation and a final purification step using reverse phase HPLC.

Sometimes, only one peak developed during the labelling reaction and in these cases the following test had to be carried out to confirm the presence of the stable form: A small amount of the labelling product was collected from the HPLC, freeze-dried, taken up in a small volume of NH_4HCO_3-buffer (100 mM, pH 8.5) and the yellow solution boiled for approximately 30 minutes on a water bath. After cooling to r.t. the sample was subjected to reverse phase HPLC. If there was a decrease of the retention time (approx. 0.7 min), the labelled product was present in the lactone form. If the retention time did not change the 'open' form had been obtained.

After HPLC purification (figure 42) the product was lyophilised to give a yellow powder that could be stored at – 20 °C. Using the above procedure yielded typically 6 mg of material.

The FITC-labelled ovalbumin glycopeptide was unstable in biological buffers at fridge temperature over a longer period of time but perfectly stable at -20 °C.

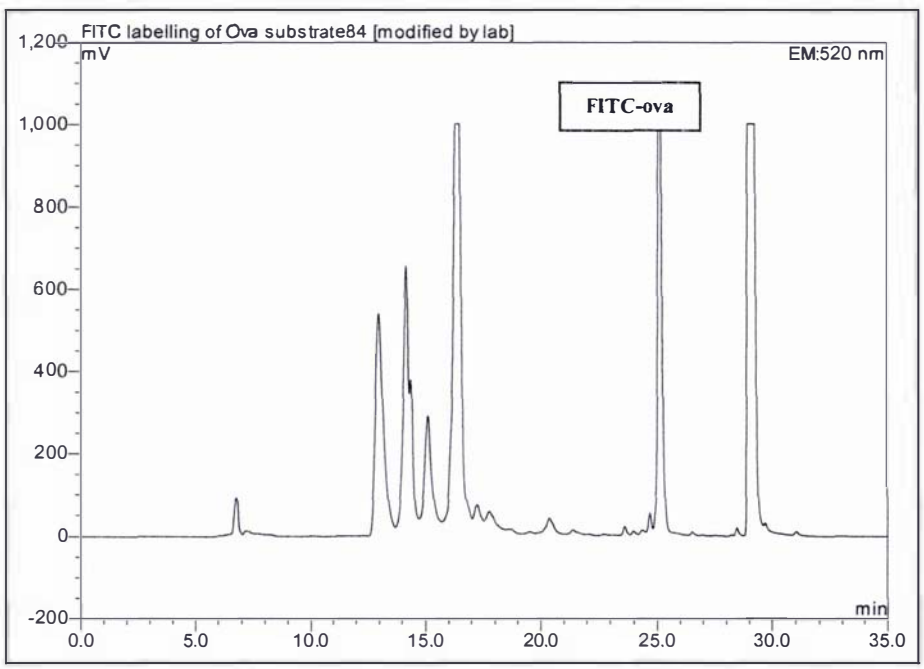


Figure 42: HPLC-chromatogram for the purification of the labelling product (FITC-ova) using preparative HPLC.

CHAPTER 9

Literature

[1] Clarke,R.M. (1998) *Citroen DS & ID 1955-1975*. Brooklands Books.

[2] Varki,A. (1993) Biological Roles of Oligosaccharides. *Glycobiology*, **3**, 97-130.

[3] Suzuki,T., Kitajima,K., Inoue,S., & Inoue,Y. (1994) Occurence and biological roles of 'proximal glycanases' in animal cells. *Glycobiology*, **4**, 777-789.

[4] Tarentino,A.L., Quinones,G., Trumble,A., Changchien,L.M., Duceman,B., Maley,F., & Plummer,T.H., Jr. (1990) Molecular cloning and amino acid sequence of peptide-N4-(N-acetyl-beta- D-glucosaminyl)asparagine amidase from *Flavobacterium meningosepticum*. *J. Biol. Chem.*, **265**, 6961-6966.

[5] Wang,L.-X., Tang,M., Suzuki,T., Kitajima,K., Inoue,Y., Inoue,S., Fan,J.-Q., & Lee,Y.C. (1997) Combined Chemical and Enzymatic Synthesis of a C-Glycopeptide and Its Inhibitory Activity toward Glycoamidases. *J. Am. Chem. Soc.*, **119**, 11137-11146.

[6] Deras,I.L., Takegawa,K., Kondo,A., Kato,I., & Lee,Y.C. (1998) Synthesis of a high-mannose-type glycopeptide analogue containing a glucose-asparagine linkage. *Bioorg. Med. Chem. Lett.*, **8**, 1763-1766.

[7] Haneda,K. & Tanabe,H. Complex glycopeptide derivatives stable to N-glycanases and their manufacture. JP 1999-88030[JP 2000319297], 1-11. 2001. Japan.
Ref Type: Patent

[8] Ferro,V., Weiler,L., Withers,S.G., & Ziltener,H. (1998) *N*-Glycosyl phosphonamidates: potential transition-state analogue inhibitors of glycopeptidases. *Can. J. Chem.*, **76**, 313-318.

[9] Shibata,S. & Nakanishi,H. (1980) The alpha-D configuration of the glycosyl-N linkage in the nephritogenic glycopeptide isolated from rat glomerular, basement membrane. *Carbohydr. Res.*, **86**, 316-320.

[10] Shibata,S., Takeda,T., & Natori,Y. (1988) The structure of nephritogenoside. A nephritogenic glycopeptide with alpha-N-glycosidic linkage. *J. Biol. Chem.*, **263**, 12483-12485.

[11] Imperiali,B. & O'Connor,S.E. (1999) Effect of N-linked glycosylation on glycopeptide and glycoprotein structure. *Curr. Opin. Chem. Biol.*, **3**, 643-649.

[12] Tai,V.W. & Imperiali,B. (2001) Substrate specificity of the glycosyl donor for oligosaccharyl transferase. *J. Org. Chem.*, **66**, 6217-6228.

[13] Kornfeld,R. & Komfeld,S. (1985) Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.*, **54**, 631-664.

[14] Cumming,D.A. (1991) Glycosylation of recombinant protein therapeutics: control and functional implications. *Glycobiology*, **1**, 115-130.

[15] Dwek,R.A. (1996) Glycobiology: Toward Understanding the Function of Sugars. *Chem. Rev.*, **96**, 683-720.

[16] Lis,H. & Sharon,N. (1993) Protein glycosylation. Structural and functional aspects. *Eur. J. Biochem.*, **218**, 1-27.

[17] O'Conner,S.E. & Imperiali,B. (1998) A molecular basis for glycosylation-induced conformational switching. *Chem. Biol.*, **5**, 427-437.

[18] Berger,S., Menudier,A., Julien,R., & Karamanos,Y. (1995) Do de-N-glycosylation enzymes have an important role in plant cells? *Biochimie*, **77**, 751-760.

[19] Risley,J.M. & Van Etten,R.L. (1985) 1H NMR evidence that almond "peptide: N-glycosidase" is an amidase. Kinetic data and trapping of the intermediate. *J. Biol. Chem.*, **260**, 15488-15494.

[20] Suzuki,T., Kitajima,K., Inoue,S., & Inoue,Y. (1994) Does an animal peptide: N-glycanase have the dual role as an enzyme and a carbohydrate-binding protein? *Glycoconj. J.*, **11**, 469-476.

[21] Suzuki,T., Kitajima,K., Emori,Y., Inoue,Y., & Inoue,S. (1997) Site-specific de-N-glycosylation of diglycosylated ovalbumin in hen oviduct by endogenous peptide: N-glycanase as a quality control system for newly synthesized proteins. *Proc. Natl. Acad. Sci. U. S. A*, **94**, 6244-6249.

[22] Suzuki,T., Park,H., Kitajima,K., & Lennarz,W.J. (1998) Peptides glycosylated in the endoplasmic reticulum of yeast are subsequently deglycosylated by a soluble peptide: N-glycanase activity. *J. Biol. Chem.*, **273**, 21526-21530.

[23] Suzuki,T., Park,H., Hollingsworth,N.M., Sternglanz,R., & Lennarz,W.J. (2000) PNG1, a yeast gene encoding a highly conserved peptide:N-glycanase. *J. Cell Biol.*, **149**, 1039-1052.

[24] Suzuki,T., Park,H., Till,E.A., & Lennarz,W.J. (2001) The PUB domain: a putative protein-protein interaction domain implicated in the ubiquitin-proteasome pathway. *Biochem. Biophys. Res. Commun.*, **287**, 1083-1087.

[25] Suzuki,T., Park,H., & Lennarz,W.J. (2002) Cytoplasmic peptide:N-glycanase (PNGase) in eukaryotic cells: occurrence, primary structure, and potential functions. *The FASEB Journal*, **16**, 635-641.

[26] Weng,S. & Spiro,R.G. (1997) Demonstration of a peptide:N-glycosidase in the endoplasmic reticulum of rat liver. *Biochem. J.*, **322 (Pt 2)**, 655-661.

[27] Selby,M., Erickson,A., Dong,C., Cooper,S., Parham,P., Houghton,M., & Walker,C.M. (1999) Hepatitis C virus envelope glycoprotein E1 originates in the endoplasmic reticulum and requires cytoplasmic processing for presentation by class I MHC molecules. *J. Immunol.*, **162**, 669-676.

[28] Skipper,J.C., Hendrickson,R.C., Gulden,P.H., Brichard,V., Van Pel,A., Chen,Y., Shabanowitz,J., Wolfel,T., Slingluff,C.L., Jr., Boon,T., Hunt,D.F., & Engelhard,V.H. (1996) An HLA-A2-restricted tyrosinase antigen on melanoma cells results from posttranslational modification and suggests a novel pathway for processing of membrane proteins. *J. Exp. Med.*, **183**, 527-534.

[29] Berger,S., Menudier,A., Julien,R., & Karamanos,Y. (1995) Endo-N-acetyl-beta-D-glucosaminidase and peptide-N4-(N-acetyl- glucosaminyl) asparagine amidase activities during germination of *Raphanus sativus*. *Phytochemistry*, **39**, 481-487.

[30] Lhernould,S., Karamanos,Y., Priem,B., & Morvan,H. (1994) Carbon starvation increases endoglycosidase activities and production of "unconjugated N-glycans" in *Silene alba* cell-suspension cultures. *Plant Physiol.*, **106**, 779-784.

[31] Priem,B., Gitti,R., Bush,C.A., & Gross,K.C. (1993) Structure of ten free N-glycans in ripening tomato fruit. Arabinose is a constituent of a plant N-glycan. *Plant Physiol.*, **102**, 445-458.

[32] Priem,B. & Gross,K.C. (1992) Mannosyl- and Xylosyl-Containing Glycans Promote Tomato (*Lycopersicon esculentum* Mill.) Fruit Ripening. *Plant Physiol.*, **98**, 399-401.

[33] Chang,T., Kuo,M.C., Khoo,K.H., Inoue,S., & Inoue,Y. (2000) Developmentally regulated expression of a peptide:N-glycanase during germination of rice seeds (*Oryza sativa*) and its purification and characterization. *J. Biol. Chem.*, **275**, 129-134.

[34] Wu,H.M., Wang,H., & Cheung,A.Y. (1995) A pollen tube growth stimulatory glycoprotein is deglycosylated by pollen tubes and displays a glycosylation gradient in the flower. *Cell*, **82**, 395-403.

[35] Sheldon,P.S., Keen,J.N., & Bowles,D.J. (1998) Purification and characterization of N-glycanase, a concanavalin A binding protein from jackbean (*Canavalia ensiformis*). *Biochem. J.*, **330 (Pt 1)**, 13-20.

[36] Vuylsteker,C., Cuvellier,G., Berger,S., Faugeron,C., & Karamanos,Y. (2000) Evidence of two enzymes performing the de-N-glycosylation of proteins in barley: expression during germination, localization within the grain and set-up during grain formation. *J. Exp. Bot.*, **51**, 839-845.

[37] Ftouhi-Paquin,N., Tarentino,A.L., & Plummer,T.H., Jr. (1998) Overexpression of PNGase At from *baculovirus*-infected insect cells. *Protein Expr. Purif.*, **14**, 302-308.

[38] Suzuki,T., Seko,A., Kitajima,K., Inoue,Y., & Inoue,S. (1993) Identification of peptide:N-glycanase activity in mammalian-derived cultured cells. *Biochem. Biophys. Res. Commun.*, **194**, 1124-1130.

[39] Suzuki,T., Seko,A., Kitajima,K., Inoue,Y., & Inoue,S. (1994) Purification and enzymatic properties of peptide:N-glycanase from C3H mouse-derived L-929 fibroblast cells. Possible widespread occurrence of post-translational remodification of proteins by N-deglycosylation. *J. Biol. Chem.*, **269**, 17611-17618.

[40] Suzuki,T., Kitajima,K., Inoue,Y., & Inoue,S. (1995) Carbohydrate-binding property of peptide: N-glycanase from mouse fibroblast L-929 cells as evaluated by inhibition and binding experiments using various oligosaccharides. *J. Biol. Chem.*, **270**, 15181-15186.

[41] Kitajima,K., Suzuki,T., Kouchi,Z., Inoue,S., & Inoue,Y. (1995) Identification and distribution of peptide:N-glycanase (PNGase) in mouse organs. *Arch. Biochem. Biophys.*, **319**, 393-401.

[42] Suzuki,T., Kitajima,K., Inoue,S., & Inoue,Y. (1995) N-Glycosylation/deglycosylation as a mechanism for the post-translational modification/remodification of proteins. *Glycoconj. J.*, **12**, 183-193.

[43] Seko,A., Kitajima,K., Inoue,Y., & Inoue,S. (1991) Peptide:N-glycosidase activity found in the early embryos of Oryzias latipes (Medaka fish). The first demonstration of the occurrence of peptide:N-glycosidase in animal cells and its implication for the presence of a de-N-glycosylation system in living organisms. *J. Biol. Chem.*, **266**, 22110-22114.

[44] Seko,A., Kitajima,K., Inoue,S., & Inoue,Y. (1991) Identification of free glycan chain liberated by de-N-glycosylation of the cortical alveolar

glycopolyprotein (hyosophorin) during early embryogenesis of the Medaka fish, *Oryzias latipes*. *Biochem. Biophys. Res. Commun.*, **180**, 1165-1171.

[45] Seko,A., Kitajima,K., Iwamatsu,T., Inoue,Y., & Inoue,S. (1999) Identification of two discrete peptide: N-glycanases in *Oryzias latipes* during embryogenesis. *Glycobiology*, **9**, 887-895.

[46] Taga,E.M., Waheed,A., & Van Etten,R.L. (1984) Structural and chemical characterization of a homogeneous peptide N- glycosidase from almond. *Biochemistry*, **23**, 815-822.

[47] Altmann,F., Paschinger,K., Dalik,T., & Vorauer,K. (1997) Characterisation of of peptide-N^4-(N-acetyl-β-glucosaminyl)asparagine amidase A and its N-glycans. *Eur. J. Biochem.*, **252**, 118-123.

[48] Sugiyama,K., Ishihara,H., Tejima,S., & Takahashi,N. (1983) Demonstration of a new glycopeptidase, from jack-bean meal, acting on aspartylglucosylamine linkages. *Biochem. Biophys. Res. Commun.*, **112**, 155-160.

[49] Kimura,Y. & Ohno,A. (1998) A new peptide-N4-(acetyl-beta-glucosaminyl)asparagine amidase from soybean (*Glycine max*) seeds: purification and substrate specificity. *Biosci. Biotechnol. Biochem.*, **62**, 412-418.

[50] Plummer,T.H., Jr., Phelan,A.W., & Tarentino,A.L. (1987) Detection and quantification of peptide-N4-(N-acetyl-beta- glucosaminyl)asparagine amidases. *Eur. J. Biochem.*, **163**, 167-173.

[51] Berger,S., Menudier,A., Julien,R., & Karamanos,Y. (1996) Regulation of de-N-Glycosylation Enzymes in Germinating Radish Seeds. *Plant Physiol.*, **112**, 259-264.

[52] Lhernould,S., Karamanos,Y., Lerouge,P., & Morvan,H. (1995) Characterization of the peptide-N^4-(N-acetylglucosaminyl)asparagine amidase (PNGase Se) from *Silene alba* cells. *Glycoconj. J.*, **12**, 94-98.

[53] Ftouhi-Paquin,N., Hauer,C.R., Stack,R.F., Tarentino,A.L., & Plummer,T.H., Jr. (1997) Molecular cloning, primary structure, and properties of a new glycoamidase from the fungus *Aspergillus tubigensis*. *J. Biol. Chem.*, **272**, 22960-22965.

[54] Plummer,T.H., Jr., Elder,J.H., Alexander,S., Phelan,A.W., & Tarentino,A.L. (1984) Demonstration of peptide:N-glycosidase F activity in endo-beta-N-acetylglucosaminidase F preparations. *J. Biol. Chem.*, **259**, 10700-10704.

[55] Kuhn,P., Guan,C., Cui,T., Tarentino,A.L., Plummer,T.H., Jr., & Van Roey,P. (1995) Active Site and Oligosaccharide Recognition Residues of Peptide-N^4-(N-acetyl-β-D-glucosaminyl)asparagine Amidase F. *J. Biol. Chem.*, **49**, 29493-29497.

[56] Altmann,F., Schweiszer,S., & Weber,C. (1995) Kinetic comparison of peptide: N-glycosidase F and A reveals several differences in substrate specificity. *Glycoconj. J.*, **12**, 84-93.

[57] Tarentino,A.L. & Plummer,T.H., Jr. (1982) Oligosaccharide accessibility to peptide:N-glycosidase as promoted by protein-unfolding reagents. *J. Biol. Chem.*, **257**, 10776-10780.

[58] Tretter,V., Altmann,F., & Marz,L. (1991) Peptide-N4-(N-acetyl-beta-glucosaminyl) asparagine amidase F cannot release glycans with fucose attached alpha 1----3 to the asparagine- linked N-acetylglucosamine residue. *Eur. J. Biochem.*, **199**, 647-652.

[59] Fan,J.-Q. & Lee,Y.C. (1997) Detailed Studies on Substrate Structure Requirements of Glycoamidases A and F. *J. Biol. Chem.*, **43**, 27058-27064.

[60] Katiyar,S., Suzuki,T., Balgobin,B.J., & Lennarz,W.J. (2002) Site-directed mutagenesis study of yeast peptide:N-glycanase. Insight into the reaction mechanism of deglycosylation. *J. Biol. Chem.*, **277**, 12953-12959.

[61] Takahashi,N. (1977) Demonstration of a new amidase acting on glycopeptides. *Biochem. Biophys. Res. Commun.*, **76**, 1194-1201.

[62] Bourgerie,S., Karamanos,Y., Berger,S., & Julien,R. (1992) Use of resorufin-labelled N-glycopeptide in a high-performance liquid chromatography assay to monitor endoglycosidase activities during cultivation of Flavobacterium meningosepticum. *Glycoconj. J.*, **9**, 162-167.

[63] Norris,G.E., Flaus,A.J., Moore,C.H., & Baker,E.N. (1994) Purification and crystallization of the endoglycosidase PNGase F, a peptide:N-glycosidase from *Flavobacterium meningosepticum. J. Mol. Biol.*, **241**, 624-626.

[64] Mussar,K.J., Murray,G.J., Martin,B.M., & Viswanatha,T. (1989) Peptide: N-glycosidase F: studies on the glycoprotein aminoglycan amidase from *Flavobacterium meningosepticum. J. Biochem. Biophys. Methods*, **20**, 53-68.

[65] Taylor,S., Ninjoor,V., Dowd,D.M., & Tappel,A.L. (1974) Cathepsin B2 measurement by sensitive fluorometric ammonia analysis. *Anal. Biochem.*, **60**, 153-162.

[66] Takahashi,N. & Nishibe,H. (1981) Almond glycopeptidase acting on aspartylglycosyl amine linkages. Multiplicity and substrate specificity. *Biochim. Biophys. Acta*, **657**, 457-467.

[67] Deras,I.L., Sano,M., Kato,I., & Lee,Y.C. (2000) Assay of glycoamidases and endo-beta-N-acetylglucosaminidases by lectin capture and dissociation-enhanced lanthanide fluorescence immunoassay. *Anal. Biochem.*, **278**, 213-220.

[68] Dubois,M., Gilles,K.A., Hamilton,J.K., Rebers,P.A., & Smith,F. (1956) Colorimetric Method for Determination of Sugars and Related Substances. *Anal. Chem.*, **28**, 350-356.

[69] Tarentino,A.L. & Plummer,T.H., Jr. (1987) Peptide-N4-(N-acetyl-beta-glucosaminyl) asparagine amidase and endo- beta-N-acetylglucosaminidase from *Flavobacterium meningosepticum. Methods Enzymol.*, **138**, 770-778.

[70] Lee,K.B., Matsuoka,K., Nishimura,S., & Lee,Y.C. (1995) A new approach to assay endo-type carbohydrases: bifluorescent-labeled substrates for glycoamidases and ceramide glycanases. *Anal. Biochem.*, **230**, 31-36.

[71] Lee,K.B. & Lee,Y.C. (1997) Preparation of bifluorescent-labeled glycopeptides for glycoamidase assay. *Methods Enzymol.*, **278**, 512-519.

[72] Ferro,V., Weiler,L., & Withers,S.G. (1998) Convergent synthesis of a fluorescence-quenched glycopeptide as a potential substrate for peptide:N-glycosidases. *Carbohydr. Res.*, **306**, 531-538.

[73] Meldal,M. & Breddam,K. (1991) Anthranilamide and nitrotyrosine as a donor-acceptor pair in internally quenched fluorescent substrates for endopeptidases: multicolumn peptide synthesis of enzyme substrates for subtilisin Carlsberg and pepsin. *Anal. Biochem.*, **195**, 141-147.

[74] Plummer,T.H., Jr. & Tarentino,A.L. (1981) Facile cleavage of complex oligosaccharides from glycopeptides by almond emulsin peptide: N-glycosidase. *J. Biol. Chem.*, **256**, 10243-10246.

[75] Plummer,T.H., Jr. & Tarentino,A.L. (1991) Purification of the oligosaccharide-cleaving enzymes of *Flavobacterium meningosepticum*. *Glycobiology*, **1**, 257-263.

[76] Barsomian,G.D., Johnson,T.L., Borowski,M., Denman,J., Ollington,J.F., Hirani,S., McNeilly,D.S., & Rasmussen,J.R. (1990) Cloning and expression of peptide-N4-(N-acetyl-beta-D- glucosaminyl)asparagine amidase F in *Escherichia coli*. *J. Biol. Chem.*, **265**, 6967-6972.

[77] Loo,T., Patchett,M.L., Norris,G.E., & Lott,J.S. (2002) Using Secretion to Solve a Solubility Problem: High-Yield Expression in *Escherichia coli* and Purification of the Bacterial Glycoamidase PNGase F. *Protein Expr. Purif.*, **24**, 90-98.

[78] Kuhn,P., Tarentino,A.L., Plummer,T.H., Jr., & Van Roey,P. (1994) Crystallization and preliminary crystallographic analysis of peptide-N4- (N-acetyl-beta-D-glucosaminyl)asparagine amidase PNGase F. *J. Mol. Biol.*, **241**, 622-623.

[79] Norris,G.E., Stillman,T.J., Anderson,B.F., & Baker,E.N. (1994) The three-dimensional structure of PNGase F, a glycosylasparaginase from *Flavobacterium meningosepticum*. *Structure*, **2**, 1049-1059.

[80] Kuhn,P., Tarentino,A.L., Plummer,T.H., Jr., & Van Roey,P. (1994) Crystal Structure of Peptide-N^4-(N-acetyl-β-D-glucosaminyl)asparagine Amidase F at 2.2-Å Resolution. *Biochemistry*, **33**, 11699-11706.

[81] Gosselin,S., Payie,K., & Viswanatha,T. (1993) Substrate specificity of peptide:N-glycosidase from *Flavobacterium meningosepticum*. *Glycobiology*, **3**, 419-421.

[82] Chu,F.K. (1986) Requirements of cleavage of high mannose oligosaccharides in glycoproteins by peptide N-glycosidase F. *J. Biol. Chem.*, **261**, 172-177.

[83] Nuck,R., Zimmermann,M., Sauvageot,D., Josi,D., & Reutter,W. (1990) Optimized deglycosylation of glycoproteins by peptide-N4-(N-acetyl-beta-glucosaminyl)-asparagine amidase from *Flavobacterium meningosepticum*. *Glycoconj. J.*, **7**, 279-286.

[84] Tarentino,A.L., Gomez,C.M., & Plummer,T.H., Jr. (1985) Deglycosylation of asparagine-linked glycans by peptide:N-glycosidase F. *Biochemistry*, **24**, 4665-4671.

[85] Mann,A.C., Self,C.H., & Turner,G.A. (1994) A general method for the complete deglycosylation of a wide variety of serum glycoproteins using peptide-*N*-glycosidase-F. *Glycosylation & Disease*, **1**, 253-261.

[86] Tarentino,A.L. & Plummer,T.H., Jr. (1993) Deglycosylation of Asparagine-linked Glycans by PNGase F. *Trends Glycosci. Glycotechnol.*, **5**, 163-170.

[87] Maley,F., Trimble,R.B., Tarentino,A.L., & Plummer,T.H., Jr. (1989) Characterization of glycoproteins and their associated oligosaccharides through the use of endoglycosidases. *Anal. Biochem.*, **180**, 195-204.

[88] Steube,K., Gross,V., Hösel,W., Tran-Thi,T.A., Decker,K., & Heinrich,P. (1986) Different Susceptibilities of Complex-, Hybrid- and High-mannose-

Type α_1-Proteinase Inhibitor and α_1-Acid Glycoprotein to Endo-β-*N*-acetylglucosaminidase F and Peptide:N-Glycosidase F. *Glycoconj. J.*, **3**, 247-254.

[89] Takegawa,K., Tabuchi,M., Yamaguchi,S., Kondo,A., Kato,I., & Iwahara,S. (1995) Synthesis of neoglycoproteins using oligosaccharide-transfer activity with endo-beta-N-acetylglucosaminidase. *J. Biol. Chem.*, **270**, 3094-3099.

[90] Fan,J.Q., Quesenberry,M.S., Takegawa,K., Iwahara,S., Kondo,A., Kato,I., & Lee,Y.C. (1995) Synthesis of neoglycoconjugates by transglycosylation with *Arthrobacter protophormiae* endo-beta-N-acetylglucosaminidase. Demonstration of a macro-cluster effect for mannose-binding proteins. *J. Biol. Chem.*, **270**, 17730-17735.

[91] Fan,J.Q., Takegawa,K., Iwahara,S., Kondo,A., Kato,I., Abeygunawardana,C., & Lee,Y.C. (1995) Enhanced transglycosylation activity of *Arthrobacter protophormiae* endo- beta-N-acetylglucosaminidase in media containing organic solvents. *J. Biol. Chem.*, **270**, 17723-17729.

[92] Lis,H. & Sharon,N. (1978) Soybean agglutinin-a plant glycoprotein. Structure of the carbohydrate unit. *J. Biol. Chem.*, **253**, 3468-3476.

[93] Lis,H., Sharon,N., & Katchalski,E. (1966) Soybean hemagglutinin, a plant glycoprotein. I. Isolation of a glycopeptide. *J. Biol. Chem.*, **241**, 684-689.

[94] Thiem,J. (1995) Applications of enzymes in synthetic carbohydrate chemistry. *FEMS Microbiol. Rev.*, **16**, 193-211.

[95] Yamamoto,K., Fujimori,K., Haneda,K., Mizuno,M., Inazu,T., & Kumagai,H. (1998) Chemoenzymatic synthesis of a novel glycopeptide using a microbial endoglycosidase. *Carbohydr. Res.*, **305**, 415-422.

[96] Haneda,K., Inazu,T., Yamamoto,K., Kumagai,H., Nakahara,Y., & Kobata,A. (1996) Transglycosylation of intact sialo complex-type oligosaccharides to the *N*-acetylglucosamine moieties of glycopeptides by *Mucor hiemalis* endo-β -*N*-acetylglucosaminidase. *Carbohydr. Res.*, **292**, 61-70.

[97] Copeland,R.A. (1996) Enzymes. A Practical Introduction to Structure, Mechanism and Data Analysis. (Copeland,R.A., ed), pp. 67-92. Wiley-VCH, New York.

[98] Morise,X., Savignac,P., Guillemin,J.C., & Denis,J.M. (1991) A Convenient Method for the Synthesis of α-Functionalized Chlorophosphonic Esters. *Synth. Commun.*, **21**, 793-798.

[99] Yamauchi,K., Kinoshita,M., & Imoto,M. (1972) Peptides Containing Aminophosphonic Acids. II. The Synthesis of Tripeptide Analogs. *Bull. Chem. Soc. Jpn.*, **45**, 2531-2534.

[100] McKenna,C.E., Higa,M.T., Cheung,N.H., & McKenna,M.-C. (1977) The Facile Dealkylation of Phosphonic Acid Dialkyl Esters by Bromotrimethylsilane. *Tetrahedron Lett.*, **2**, 155-158.

[101] Hirschmann,R., Yager,K.M., Taylor,C.M., Moore,W., Sprengeler,P.A., Witherington,J., Phillips,B.W., & Smith,A.B.I. (1995) Phosphonyltriethylammonium Salts: Novel Reactive Species for the Synthesis of Phosphonate Esters and Phosphonamides. *J. Am. Chem. Soc.*, **117**, 6370-6371.

[102] Edman,P. & Begg,G. (1967) A protein sequenator. *Eur. J. Biochem.*, **1**, 80-91.

[103] Udenfriend,S., Stein,S., Bohlen,P., Dairman,W., Leimgruber,W., & Weigele,M. (1972) Fluorescamine: a reagent for assay of amino acids, peptides, proteins, and primary amines in the picomole range. *Science*, **178**, 871-872.

[104] Hentz,N.G., Richardson,J.M., Sportsman,J.R., Daijo,J., & Sittampalam,G.S. (1997) Synthesis and Characterization of Insulin-Fluorescein Derivatives for Bioanalytical Applications. *Anal. Chem.*, **69**, 4994-5000.

[105] Taylor,C.M. (1998) Glycopeptides and Glycoproteins: Focus on the Glycosidic Linkage. *Tetrahedron*, **54**, 11317-11362.

[106] Shaban,M. & Jeanloz,R.W. (1971) The synthesis of O-α-D-mannopyranosyl-(1-6) -O-(2-acetamido-2-deoxy- β-D-glucopyranosyl)-(1-4)-2-acetamido-2-deoxy-D-glucose. *Carbohydr. Res.*, **19**, 311-318.

[107] Myers,R.W. & Lee,Y.C. (1986) Improved Preparations of some Per-O-Acetylated Aldohexopyranosylcyanides. *Carbohydr. Res.*, **154**, 145-163.

[108] Johansson,R. & Samuelsson,B. (1984) Regioselective Reductive Ring-opening of 4-Methoxybenzylidene Acetals of Hexopyranosides. Access to a Novel Protecting-group Strategy.Partl. *J. Chem. Soc. Perkin Trans. I*, 2371-2374.

[109] Grundler,G. & Schmidt,R.R. (1985) Anwendung des Trichloracetimidatverfahrens auf 2-desoxy-2-phthalimido-D-Glucose-Derivate. Synthese von Oligosacchariden der "Core-Region" von O-Glycoproteinen des Mucin-Typs. *Carbohydr. Res.*, **135**, 203-218.

[110] Tanemura,K., Suzuki,T., & Horaguchi,T. (1992) 2,3-Dichloro-5,6-dicyano-*p*-benzoquinone as a Mild and Efficient Catalyst for the Deprotection of Acetals. *J. Chem. Soc. Perkin Trans. I*, 979-980.

[111] Tanemura,K., Suzuki,T., & Horaguchi,T. (1994) Deprotection of Acetals and Silyl Ethers Using Some π-Acceptors. *Bull. Chem. Soc. Jpn.*, **67**, 290-292.

[112] Rylander,P.N. (1985) Hydrogenation of Nitriles and Oximes. In *Hydrogenation Methods* (Rylander,P.N., ed), pp. 94-103. London.

[113] Lenz,D.H., Norris,G.E., Taylor,C.M., & Slim,G.C. (2001) One pot transformation of glycopyranosylcyanides to N-(*t*-butoxycarbonyl)methylamines. *Tetrahedron Letters* , **42**, 4589-4591.

[114] Saito,S., Nakajima,H., Inaba,M., & Moriwake,T. (1989) One-Pot Transformation of Azido-Group to N-(t-Butoxycarbonyl)amino group. *Tetrahedron Letters*, **30**, 837-838.

[115] Barany,G. & Merrifield,R.B. (1980) *The Peptides. Analysis, Synthesis, Biology.* Academic Press, New York.

[116] Rink,H. (1987) Solid-Phase Synthesis of protected Peptide Fragments using a Trialkoxy-Diphenyl-Methylester Resin. *Tetrahedron Lett.*, **28**, 3787-3790.

[117] Anisfeld,S.T. & Lansbury,P.T., Jr. (1990) A Convergent approach to the Chemical Synthesis of Asparagine-Linked Glycopeptides. *J. Org. Chem.*, **55**, 5560-5562.

[118] Wong,S.Y.C., Guile,G.R., Rademacher,T.W., & Dwek,R.A. (1993) Synthetic glycosylation of peptides using unprotected saccharide β-glycosyl amines. *Glycoconj. J.*, **10**, 227-234.

[119] Likhosherstov,L.M., Novikova,O.S., Derevitskaja,V.A., & Kochetkov,N.K. (1986) A new simple synthesis of amino sugar β-D-glycosyl amines. *Carbohydr. Res.*, **146**, C1-C5.

[120] Manger,I.D., Rademacher,T.W., & Dwek,R.A. (1992) 1-N-glycyl beta-oligosaccharide derivatives as stable intermediates for the formation of glycoconjugate probes. *Biochemistry*, **31**, 10724-10732.

[121] Teshima,T., Nakajima,K., Takahashi,M., & Shiba,T. (1992) Total synthesis of nephritogenic glycopeptide, nephritogenoside. *Tetrahedron Lett.*, **33**, 363-366.

[122] Nakabayashi,S., Warren,C.D., & Jeanloz,R.W. (1988) The preparation of a partially protected heptasaccharide-asparagine intermediate for glycopeptide synthesis. *Carbohydr. Res.*, **174**, 279-289.

[123] Ogawa,T., Nakabayashi,S., & Shibata,S. (1983) Synthetic Studies on Nephritogenic Glycosides. Synthesis of *N*-(β-L-Aspartyl)-α-D-glucopyranosylamine. *Agric. Biol. Chem.*, **47**, 281-285.

[124] McDonald,F.E. & Danishefsky,S.J. (1992) A stereoselective route from glycals to asparagine-linked N-protected glycopeptides. *J. Org. Chem.*, **57**, 7001-7002.

[125] Kunz,H. & Unverzagt,C. (1988) Protecting-Group-Dependent Stability of Intersaccharide Bonds-Synthesis of a Fucosyl-Chitobiose Glycopeptide. *Angew. Chem. Int. Ed. Engl.*, **27**, 1697-1699.

[126] Thiem,J. & Wiemann,T. (1990) Combined Chemoenzymatic Synthesis of *N*-Glycoprotein Building Blocks. *Angew. Chem. Int. Ed. Engl.*, **29**, 80-81.

[127] Bayley,H., Standring,D.N., & Knowles,J.R. (1978) Propane-1,3-Dithiol: A selective reagent for the efficient reduction of alkyl and aryl azides to amines. *Tetrahedron Lett.*, **39**, 3633-3634.

[128] Unverzagt,C. (1996) Chemoenzymatic Synthesis of a Sialylated Undecasaccharide-Asparagine Conjugate. *Angew. Chem. Int. Ed. Engl.*, **35**, 2350-2353.

[129] Paulsen,H. & Lockhoff,O. (1981) Neue effektive β-Glycosidsynthese für Mannose-Glycoside. Synthesen von Mannose-haltigen Oligosacchariden. *Chem. Ber.*, **114**, 3102-3114.

[130] Chiesa,M.V. & Schmidt,R.R. (2000) Synthesis of an Asparagine-Linked Heptasaccharide - Basic Structure of *N*-Glycans. *Eur. J. Org. Chem.*, 3541-3554.

[131] Palmer,T. (1991) *Understanding Enzymes*, 4th edn. Ellis Horwood Ltd., London, New York, Toronto, Sydney, Tokyo, Singapore, Madrid, Mexiko City, Munich.

[132] Ishihara,H., Takahashi,N., Ito,J., Takeuchi,E., & Tejima,S. (1981) Either high-mannose-type or hybrid-type oligosaccharide is linked to the same asparagine residue in ovalbumin. *Biochim. Biophys. Acta*, **669**, 216-221.

[133] Juy,M., Amit,A.G., Alzari,P.M., Poljak,R.J., Claeyssens,M., Béguin,P., & Aubert,J.-P. (1992) Three-dimensional structure of a thermostable bacterial cellulase. *Nature*, **357**, 89-91.

[134] Rouvinen,J., Bergfors,T., Teeri,T., Knowles,J.K., & Jones,T.A. (1990) Three-dimensional structure of cellobiohydrolase II from Trichoderma reesei. *Science*, **249**, 380-386.

[135] Liu,Y., Guan,C., & Aronson,N.N., Jr. (1998) Site-directed mutagenesis of essential residues involved in the mechanism of bacterial glycosylasparaginase. *J. Biol. Chem.*, **273**, 9688-9694.

[136] Anderson,B.F., Baker,H.M., Norris,G.E., Rice,D.W., & Baker,E.N. (1989) Structure of human lactoferrin: crystallographic structure analysis and refinement at 2.8 A resolution. *J. Mol. Biol.* , **209**, 711-734.

[137] McPherson,A. (1990) Current approaches to macromolecular crystallization. *Eur. J. Biochem.*, **189**, 1-23.

[138] Arsenieva,D., Hardre,R., Salmon,L., & Jeffery,C.J. (2002) The crystal structure of rabbit phosphoglucose isomerase complexed with 5-phospho-D-arabinonohydroxamic acid. *Proc. Natl. Acad. Sci. U. S. A*, **99**, 5872-5877.

[139] Silverman,R.B. (2000) Enzymes as Catalysts. In *The Organic Chemistry of Enzyme-Catalyzed Reactions* (Silverman,R.B., ed), pp. 1-36. Academic Press, New York.

[140] Blow,D.M., Birktoft,J.J., & Hartley,B.S. (1969) Role of a buried acid group in the mechanism of action of chymotrypsin. *Nature*, **221**, 337-340.

[141] Dasgupta,F. & Garegg,P.J. (1988) Reductive dephthalimidation: A mild and efficient method for the N-phthalimido deprotection during oligosaccharide synthesis. *Carbohydr. Res.*, **7**, 701-707.

[142] Kurita,K., Tomita,K., Ishi,S., Nishimura,S.-I., & Shimoda,K. (1993) β-Chitin as a Convenient Starting Material for Acetolysis for Efficient Preparation of N-Acetylchitoologigosaccharides. *Journal of Polymer Science: Part A: Polymer Chemistry*, **31**, 2392-2395.

[143] Barker,S.A., Foster,A.B., Stacey,M., & Webber,J.M. (1958) Amino-sugars and Related Compounds.Part IV. Isolation and Properties obtained by Controlled Fragmentation of Chitin. *J. Chem. Soc.* , 2218-2227.

[144] Nishimura,S.-I., Kuzuhara,H., Takiguchi,Y., & Shimahara,K. (1989) Peracetylated Chitobiose: Preparation by Specific Degradations of Chitin, and Chemical Manipulations. *Carbohydr. Res.*, **194**, 223-231.

10. APPENDICES

Appendix I

Experimental Data Sheets

Experimental Data Sheet #1

[Substrate] ($mg*mL^{-1}$)	Rate ($\mu g*min^{-1}$)	Rate	Rate	Mean	Std. Deviation	S.E. Mean
2	4.96	6.07	-	5.515	0.784889	0.555
1.8	6.55	6.22	-	6.385	0.233345	0.165
1.6	5.51	6.66	6.62	6.26333	0.652712	0.376844
1.4	7.2	6.04	-	6.62	0.820244	0.58
1.2	6.34	7.6	-	6.97	0.890955	0.63
1	6.26	5.59	5.92	5.92333	0.335012	0.19342
0.5	5.28	5.58	6.38	5.74667	0.568624	0.328295
0.25	3.46	4.08	4.46	4	0.504777	0.291433
0.125	4.03	3.99	-	4.01	0.028284	0.02
0.0625	2.01	-	-	2.01	-	-
0.0313	1.15	1.198	-	1.174	0.033941	0.024

Data used for the determination of K_M and V_{max}. The table lists the initial rates for each substrate concentration followed by the mean value, the standard deviation, and the standard error of the mean. The software used for the calculations was GraFit Version 4.0 (Erithacus Software).

Experimental setup: 45 µl of the substrate (FITC-ova) in 10 mM BTP-buffer pH 8.5 was incubated with 5 µl of a recombinant PNGase F preparation in BTP-buffer pH 8.5 at 37 °C. The concentration of PNGase F was $9.09 * 10^{-6}$ mg/mL for all experiments. Reaction times were chosen which gave less than 10 % digestion of the substrate. The digest was stopped by boiling in a water bath for 3 min. The catalytic rates were determined using the discontinuous assay described in the general procedures (section 8.2).

Experimental Data Sheet #2

[Substrate] ($mg*mL^{-1}$)	Rate ($pmol*min^{-1}$)	Rate	Rate	Rate	Rate	Mean	Std. Deviation	S.E. Mean
1	8.88	6.9	8.99	8.57	-	8.335	0.973054	0.486527
0.9	8.88	8.26	8.03	-	-	8.39	0.439659	0.253837
0.8	7.83	9.45	6.12	-	-	7.8	1.6652	0.961405
0.7	8.53	6.93	9.23	7.47	6.49	7.73	1.13305	0.506715
0.5	6.16	5.79	5.33	6.41	-	5.9225	0.469991	0.234996
0.4	3.4	4.66	4.73	4.68	-	4.3675	0.645671	0.322836
0.3	3.12	3.25	-	-	-	3.185	0.091924	0.065
0.2	2.44	2.86	-	-	-	2.65	0.296985	0.21
0.1	1.21	1.24	-	-	-	1.225	0.021213	0.015

Data used for the determination of K_M and V_{max}. The table lists the initial rates for each substrate concentration followed by the mean value, the standard deviation, and the standard error of the mean. The software used for the calculations was GraFit Version 4.0 (Erithacus Software).

Experimental setup: 45 µl of the substrate (glycopeptide **58**) in 10 mM ammonium acetate buffer pH 8.5 was incubated with 5 µl of a recombinant PNGase F preparation in BTP-buffer pH 8.5 at 37 °C. The concentration of PNGase F was $2 * 10^{-6}$ mg/mL for all experiments. Reaction times were chosen which gave less than 10 % digestion of the substrate.The digest was stopped by boiling in a water bath for 3 min. The catalytic rates were determined using the discontinuous assay described in the general procedures (section 8.2).

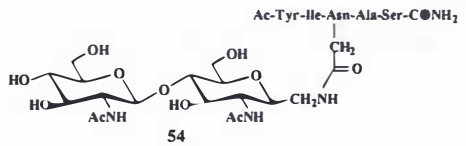
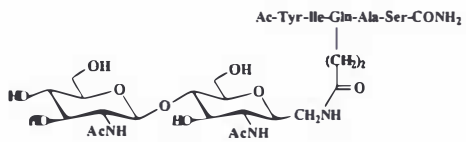
Experimental Data Sheet #3

[Substrate] ($mg*mL^{-1}$)	Rate ($pmol*min^{-1}$)	Rate	Mean	Std. Deviation	S.E.Mean
1	490	501	495.5	7.77817	5.5
0.9	494	501	497.5	4.94975	3.5
0.8	471	468	469.5	2.12132	1.5
0.7	449	437	443	8.48528	6
0.5	353	377	365	16.9706	12
0.4	362	370	366	5.65685	4
0.3	294	309	301.5	10.6066	7.5
0.2	224	207	215.5	12.0208	8.5
0.1	112	123	117.5	7.77817	5.5

Data used for the determination of K_M and V_{max}. The table lists the initial rates for each substrate concentration followed by the mean value, the standard deviation, and the standard error of the mean. The software used for the calculations was GraFit Version 4.0 (Erithacus Software).

Experimental setup: 45 µl of the substrate (glycopeptide **60**) in 10 mM ammonium acetate buffer pH 8.5 was incubated with 5 µl of a recombinant PNGase F preparation in BTP-buffer pH 8.5 at 37 °C. The concentration of PNGase F was $1 * 10^{-5}$ mg/mL for all experiments. Reaction times were chosen which gave less than 10 % digestion of the substrate. The digest was stopped by boiling in a water bath for 3 min. The catalytic rates were determined using the discontinuous assay described in the general procedures (section 8.2).

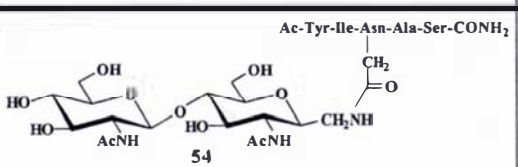
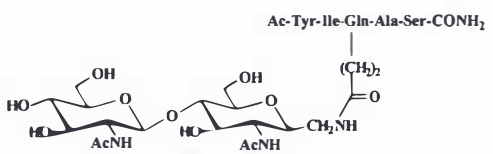
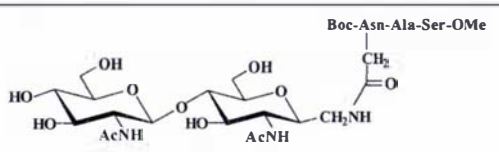
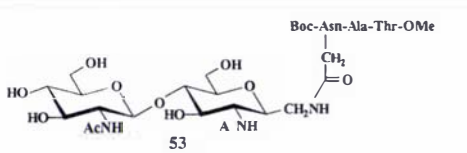
Experimental Data Sheet #4

C-glycopeptide present in the digest	Total percentage of hydrolysis of the FITC-labelled ovalbumin glycopeptide
 54	**9.79 %**
 55	**10.12 %**
nil (control)	**10.03 %**

Experimental setup: 90 µL of a mixture containing FITC-ova in 10 mM BTP- buffer pH 8.5 and an excess (~1.6 fold for **55** and ~2.5 fold for **54**) of one of the *C*-glycopeptides over [S] in the same buffer was incubated with 10 µL of a recombinant PNGase F preparation in BTP-buffer pH 8.5 at 37 °C (concentration of FITC-ova: $2.63 * 10^{-4}$ mmol/mL). The concentration of PNGase F was $2 * 10^{-6}$ mg/mL for all experiments. Reaction times were chosen so as to give approx. 10 % digest or less. The digest was stopped by boiling in a water bath for 3 min. The absolute percentage of hydrolysis was determined using the discontinuous assay described in the general procedures.

Only one set of data was collected as limited PNGase F was available and the intent of these experiments was only to identify compounds that needed more thorough testing.

Experimental Data Sheet #5

C-glycopeptide present in the digest	Total percentage of hydrolysis of the FITC-labelled ovalbumin glycopeptide
 54	**12.04 %**
 55	**17.30 %**
 52	**10.98 %**
 53	**12.61 %**
nil (control)	**14.51 %**

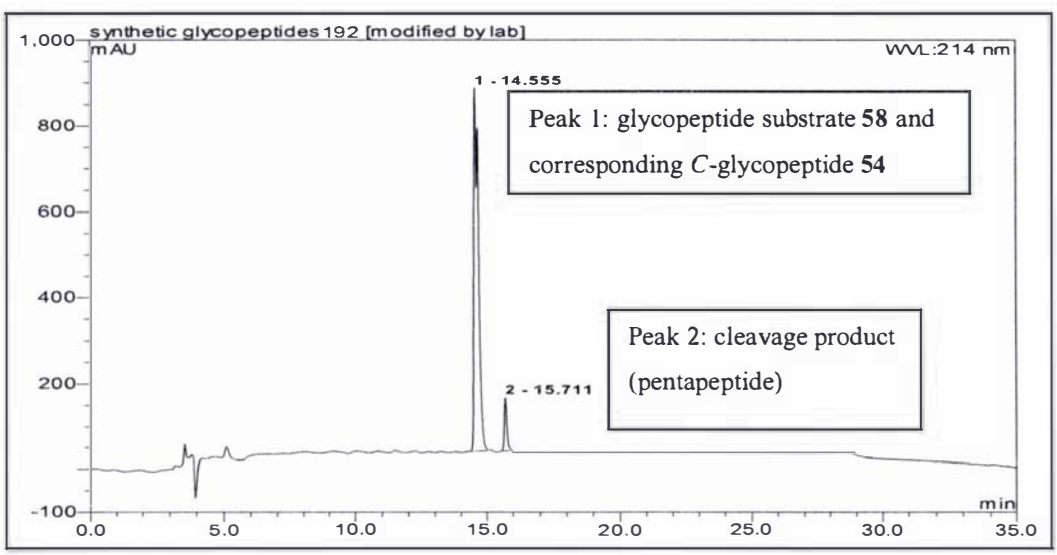
Experimental setup: In each experiment 90 µL of a saturated solution of one of the *C*-glycopeptides in BTP-buffer pH 8.5 and 10 µL of a PNGase F preparation in the same buffer were incubated for 24 h at r.t. The control contained BTP-buffer pH 8.5 instead of a saturated solution of a *C*-glycopeptide. Subsequently a solution of 50 µL of the FITC-labelled ovalbumin glycopeptide was digested at 37 °C with 5 µL of each batch of PNGase F preincubated with one of the *C*-glycopeptides (concentration of FITC-ova:

1.01 * 10^{-4} mmol/mL). The absolute percentage of hydrolysis was determined using the discontinuous assay described in the general procedures. The concentration of PNGase F was 9.09 * 10^{-6} mg/mL for all experiments.

Only one set of data was collected as limited PNGase F was available and the intent of these experiments was only to identify compounds that needed more thorough testing.

Experimental Data Sheet #6

For this experiment it was not possible to compare the absolute percentage of hydrolysis for the control runs and the inhibition trials as the glycopeptide substrate and the corresponding *C*-glycopeptide showed identical retention times in the HPLC-based assay rendering integration impossible (the HPLC-chromatogram below gives an example). For this reason the peak height and the peak area of the product peaks had to be compared.

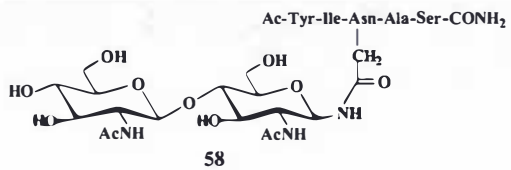
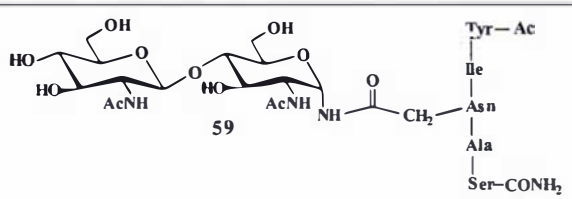
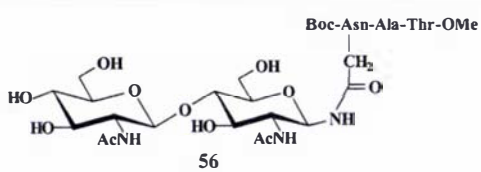
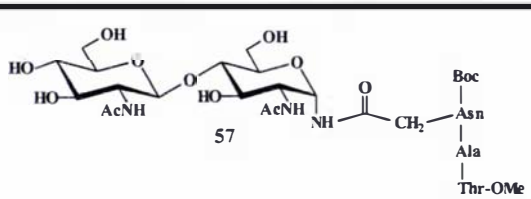


Experiment	[Substrate] ($mg*mL^{-1}$)	[*C*-glycopeptide] ($mg*mL^{-1}$)	Peak area of product	Peak height of product
1	0.5	nil	14.43	127.69
2	0.5	nil	14.44	125.85
3	0.5	0.5	14.22	125.64
4	0.5	0.5	15.41	131.23

The table lists the concentrations of substrate (glycopeptide 58) and *C*-glycopeptide 54 present and the dimensionless absolute values for the peak area and the peak height of the cleavage product (pentapeptide) for each experiment.

Experimental setup: In each experiment 45 µL of either a mixture of the substrate and *C*-glycopeptide or the substrate only (buffer: ammonium acetate pH 8.5) was incubated with a constant amount of PNGase F (5 µL) at 37 °C for a fixed time. The concentration of PNGase F was $1 * 10^{-4}$ mg/mL for all experiments. Analysis was performed as described above. Retention times are shown in minutes.

Experimental Data Sheet #7

Glycopeptide	Hydrolysis	
	PNGase F	PNGase A
 58	√	√
 59	nil	nil
 56	√	√
 57	nil	nil

Experimental setup: In each experiment 45 μL of the glycopeptide with the concentration of 1 mg/mL in 10 mM ammonium acetate buffer (pH 8.5 for PNGase F and pH 5.0 for PNGase A) was incubated with a constant amount of PNGase A or F for 14 h at 37 °C . The concentration of PNGase A was $7.8 * 10^{-4}$ mg/mL and that of PNGase F was $1 * 10^{-4}$ mg/mL. Analysis was performed using the discontinuous assay described in the general procedures. Hydrolysis was confirmed by collecting the fractions corresponding to the signals and determining their mass using ES-MS.

Experimental Data Sheet #8

Experiment	[Substrate] ($mg*mL^{-1}$)	[α-glycopeptide] ($mg*mL^{-1}$)	Rate ($pmol*min^{-1}$)	Rate ($pmol*min^{-1}$)	Mean ($pmol*min^{-1}$)
1	0.5	nil	340	341	340.5
2	0.5	0.5	330	324	327

The table lists the initial catalytic rates and the mean values for the inhibition trial (experiment 2) and the control (experiment 1) which had no α-linked glycopeptide present.

Experimental setup: In each experiment 45 μL of either a mixture of the substrate (glycopeptide **58**) and the α-linked glycopeptide **59** or the substrate only (buffer: ammonium acetate pH 8.5) was incubated with a constant amount of PNGase F (5 μL) at 37 °C for a fixed time. The concentration of PNGase F was $1 * 10^{-6}$ mg/mL for all experiments. Analysis was performed using the discontinuous assay described in the general procedures.

Experimental Data Sheet #9

Experiment	Substrate	Excess of di-chitobiose over [Substrate]	Rate ($mg*min^{-1}$)	Rate ($mg*min^{-1}$)	Mean ($mg*min^{-1}$)
1	FITC-ova	nil	$4.39 * 10^{-4}$	$4.12 * 10^{-4}$	$4.26 * 10^{-4}$
2	FITC-ova	~ 140 fold	$2.67 * 10^{-4}$	$2.61 * 10^{-4}$	$2.64 * 10^{-4}$

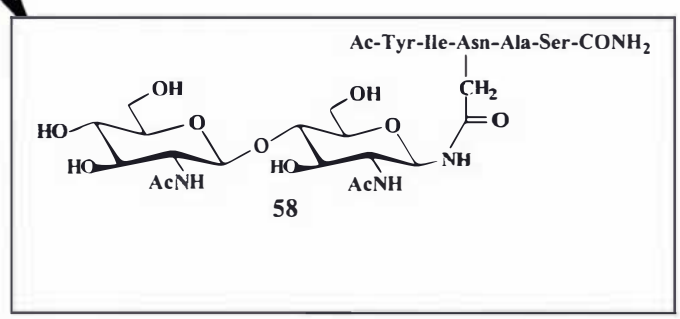
The table lists the initial catalytic rates and the mean values for the inhibition trial (experiment 2) and the control (experiment 1) which had no di-chitobiose present.

Experimental setup: In each experiment 45 µL of either a mixture of the substrate with an excess of di-chitobiose or the substrate only (buffer: ammonium acetate pH 8.5) was incubated with a constant amount of PNGase F (5 µL) at 37 °C for a fixed time. The concentration of FITC-ova in the experiments was $1.5 * 10^{-5}$ mmol/mL and the concentration of PNGase F was $5 * 10^{-6}$ mg/mL. Analysis was performed using the discontinuous assay described in the general procedures.

Experimental Data Sheet #10

Experiment	Substrate	Excess of di-chitobiose over [Substrate]	Rate (pmol*min^{-1})	Rate (pmol*min^{-1})	Mean (pmol*min^{-1})
1		nil	340	341	340.5
2		~ 2.4 fold	222	244	233
3		~ 4.8 fold	193	202	197.5
4		~ 7.2 fold	165	169	167

The table lists the initial catalytic rates and the mean values for the inhibition trial (experiments 2-4) and the control (experiment 1) which had no di-chitobiose present.



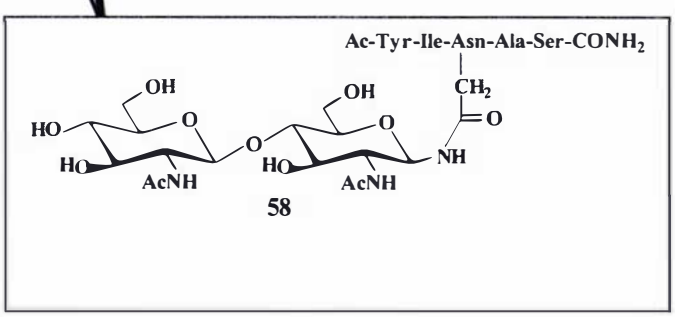
58

Experimental setup: In each experiment 45 µL of either a mixture of the substrate with an excess of di-chitobiose or the substrate only (buffer: ammonium acetate pH 8.5) was incubated with a constant amount of PNGase F (5 µL) at 37 °C for a fixed time. The substrate concentration in all experiments was 4.44 * 10^{-4} mmol/mL and the concentration of PNGase F was 1 * 10^{-4} mg/mL. Analysis was performed using the discontinuous assay described in the general procedures.

Experimental Data Sheet #11

Experiment	Substrate	% (w/v) PEG present in digest	Rate (pmol*min^{-1})	Rate (pmol*min^{-1})	Mean (pmol*min^{-1})
1		nil	357	365	361
2		12.5	130	128	129
3		25	54.6	47.5	51.1

The table lists the initial catalytic rates and the mean values for the inhibition trial (experiments 2 and 3) and the control (experiment 1) which had no PEG present.



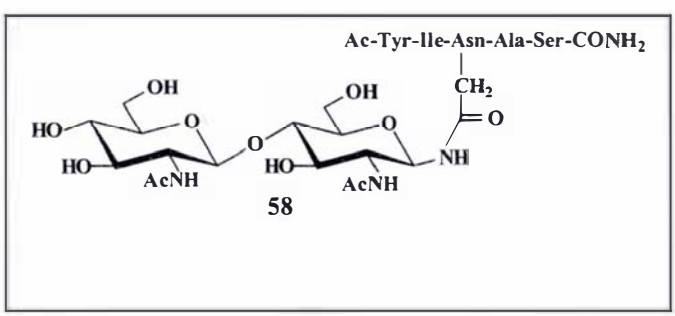
58

Experimental setup: In each experiment 45 µL of either a mixture of the substrate and PEG or the substrate only (buffer: ammonium acetate pH 8.5) was incubated with 5 µL of PNGase F at 37 °C for a fixed time. The substrate concentration for all experiments was 4.44 * 10^{-4} mmol/mL and the concentration of PNGase F was 1 * 10^{-4} mg/mL. Analysis was performed using the discontinuous assay described in the general procedures.

Experimental Data Sheet #12

Buffer	Substrate	Rate	Rate	Mean
NH_4OAc	FITC-ova	$4.39*10^{-4}$ mg*min^{-1}	$4.12*10^{-4}$ mg*min^{-1}	$4.26*10^{-4}$ mg*min^{-1}
NH_4OAc and BTP	FITC-ova	$3.29*10^{-4}$ mg*min^{-1}	$3.56*10^{-4}$ mg*min^{-1}	$3.43*10^{-4}$ mg*min^{-1}
NH_4OAc		340 pmol*min^{-1}	341 pmol*min^{-1}	340.5 pmol*min^{-1}
NH_4OAc and BTP		217 pmol*min^{-1}	219 pmol*min^{-1}	218 pmol*min^{-1}

The table lists the initial catalytic rates and the mean values for the digest of the two substrates in ammonium acetate buffer only and in the presence of BTP-buffer (see experimental setup).



58

Experimental setup: In a first set of experiments the initial catalytic rates of PNGase F with one of the two substrates in 10 mM ammonium acetate buffer (pH 8.5) only was assessed by incubating 45 µL of the substrate solution with a constant amount of PNGase F (5 µL) at 37 °C for a fixed time.

In a second set of experiments the substrate concentration was maintained but the buffer was changed to 5 mM BTP in 5 mM ammonium acetate (pH 8.5).

Using FITC-ova as the substrate the concentration of PNGase F was $5 * 10^{-6}$ mg/mL. In the experiments using **58** as the substrate the concentration of PNGase F was $1 * 10^{-4}$ mg/mL.

The substrate concentrations were $1.5 * 10^{-5}$ mmol/mL for FITC-ova and $4.44 * 10^{-4}$ mmol/mL for **58** in both experiments respectively.

Appendix II

Reproduction of a paper published in *Tetrahedron Letters* Journal