University of Oxford
St Edmund Hall

# THE MECHANISM OF ACTION OF IMINOSUGARS AS ANTIRETROVIRALS

by Simon Spiro

Submitted for the Examination of Doctor of Philosophy

OCTOBER 1, 2014

# i. Table of Contents

## ii. Acknowledgments

I have been very lucky to receive excellent support from a wide variety of people over the course of my DPhil. Particular mention must go to my supervisor, Prof Nicole Zitzmann, for accepting me into her lab and then providing the vital direction and guidance that have made this science possible. I am also indebted to Prof Raymond Dwek for his continuous encouragement and his ability to see my work in the context of the decades of glycobiology that he has experienced, as well as for his support for the Glycobiology space programme. I would also like to thank Profs Fiona Powrie and Keith Gull for accepting me onto the IITM programme and for their support and the many other opportunities they have made available to me over my time in Oxford. I particularly wish to thank Dr Joanna Miller for providing day-to-day support both practically and in discussing theory, as well as for reading this thesis and providing useful comments.

In Glycobiology I have benefitted from discussions with almost everyone, but I must single out Laura Pritchard and Drs Chris Scanlan, Max Crispin and Mark Wormald for their particular insight into the HIV envelope, and Michelle Hill and Snezana Vasiljevic for their practical assistance and trouble shooting of several experiments. I would also like to thank all of the students and post-docs of the Zitzmann group for making the second floor such an interesting, rewarding and fun place to work. In the rest of the university, Drs Quentin Sattentau, Becky Moore and Omer Dushek have provided invaluable assistance with specific problems. In the Netherlands, Dr Rogier Sanders of the University of Amsterdam and Prof Ineke Braakman of the University of Utrecht welcomed me into their labs and allowed me to perform experiments that would not have been possible in Oxford, and also provided me with some fascinating discussions. In their labs, I would particularly like to thank Ronald Derking and Nick McCaul for the experiments they performed on my behalf and for making me feel so welcome in Holland. I am also indebted to the staff of Public Health England's Porton Down facility, especially Dr Stuart Dowall and Irene Taylor, for making the Ebola experiments possible, and to my cousin, Dr Michael Spiro, for advice on critical care in haemorrhagic fevers. I would also like to thank the Wellcome Trust for funding my work and United Therapeutics for their general support and the provision of iminosugars.

Finally, I would like to thank Athena who has been so very supportive of me and all of my work over the last four years and has listened to me gripe endlessly; if anyone needs a barrister with unusually detailed knowledge of protein folding and ER quality control please contact the clerks of Crown Office Chambers.

## iii. Table of abbreviations

| | |
|---|---|
| 2-AA | 2-Aminobenzoic Acid |
| Ab | Antibody |
| AIDS | Acquired Immunodeficiency Syndrome |
| ALL | Acute Lymphoblastic Leukaemia |
| ANOVA | Analysis of Variance |
| ATP | Adenosine Triphosphate |
| AZT | Azidothymidine |
| BDBV | Bundibugyo Ebolavirus |
| BHK | Baby Hamster Kidney |
| bid | *Bia in* die(twice daily) |
| BiP | Immunoglobulin Heavy-Chain Binding Protein |
| bnAb | Broadly Neutralising Antibody |
| BSA | Bovine Serum Albumin |
| BSL | Biosafety Level |
| BVDV | Bovine Viral Diarrhoea Virus |
| C1-5 | Constant regions of HIV gp120 1-5 |
| CA | Capsid |
| Cas | Castanospermine |
| CCR5 | C-C Chemokine Receptor 5 |
| CD | Cluster of Differentiation |
| CD4i | CD4 induced |
| CDC | Centre for Disease Control |
| CDG-IIb | Congenital Disorder of Glycosylation type IIb |
| cDNA | Complementary Deoxyribonucleic Acid |
| CHO | Chinese Hamster Ovary |
| CI | Confidence Interval |
| CNX | Calnexin |
| ConA | Concavalin-A |
| COP | Cytosolic Coatamer Protein |
| CPY | Carboxypeptidase |
| CRT | Calreticulin |
| CXCR4 | C-X-C Chemokine Receptor 4 |
| Da | Dalton |
| DC-SIGN | Dendritic Cell Specific Intracellular adhesion molecule 3 Grabbing Non-Integrin |
| $ddH_2O$ | Double Distilled Water |
| DMEM | Dulbecco's Modified Eagle Medium |
| DMJ | 1-Deoxymannojirimycin |
| DMSO | Dimethyl Sulphoxide |
| DNA | Deoxyribonucleic Acid |
| DNJ | 1-Deoxynojirimycin |
| DNM | 1-Deoxynojirimycin |
| DTT | Dithiothreitol |
| EBOV | Zaire Ebolavirus |
| $EC_{50}$ | The concentration of a drug or antibody which causes 50% of the maximal effect or binding |
| EDEM | Endoplasmic Reticulum Degradation Enhancing α-Mannosidase-like Proteins |
| EeyI | Eeyarestatin I |
| ELISA | Enzyme-Linked Immunosorbent Assay |

| | |
|---|---|
| EM | Electron Microscopy |
| $E_{max}$ | The upper plateau of an ELISA curve |
| $E_{min}$ | The lower asymptote of an ELISA curve |
| Endo-H | Endoglycosidase H |
| ER | Endoplasmic Reticulum |
| ERAD | Endoplasmic Reticulum Associated Degradation |
| ERGIC | Endoplasmic Reticulum-Golgi Intermediate Compartment |
| ERQC | Endoplasmic Reticulum Quality Control |
| EU | Emission Units |
| EVD | Ebola Virus Disease |
| $F_{ab}$ | Fragment (Antigen Binding) |
| FCS | Foetal Calf Serum |
| FDA | Food and Drug Administration |
| FOS | Free Oligosaccharides |
| GAV | Guinea Pig Adenovirus |
| GCS | Glucosyl-Ceramide Synthetase |
| GH | Glycoside Hydrolase |
| Glc | Glucose |
| GlcI | Endoplasmic Reticulum α-glucosidase I |
| GlcII | Endoplasmic Reticulum α-glucosidase II |
| GlcNAc | *N*-acetyl Glucosamine |
| GnTI | *N*-acetyl glucosamine transferase I |
| GRP | Glucose Regulated Protein |
| GSL | Glucosphingolipid |
| H&E | Haematoxylin and Eosin |
| HA | Haemagglutinin |
| HBSS | Hank's Buffered Saline Solution |
| HEK | Human Embryonic Kidney |
| HEPES | (4-(2-Hydroxyethyl)-1-Piperazineethanesulfonic acid |
| HIV | Human Immunodeficiency Virus |
| HPLC | High Performance Liquid Chromatography |
| HRP | Horse Radish Peroxidase |
| HSP | Heat Shock Protein |
| $IC_{50}$ | The concentration of a drug that inhibits a specific process by 50% |
| ICTV | International Committee for the Taxonomy of Viruses |
| IgG | Immunoglobulin G |
| IN | Integrase |
| ip | Intraperitoneal |
| iv | Intravenous |
| Kd | Dissociation Constant |
| LB | Luria-Bertani |
| LTR | Long Terminal Repeat |
| MA | Matrix |
| mAb | Monoclonal Antibody |
| Man | Mannose |
| ManI | Mannosidase I |
| ManII | Mannosidase II |
| MDL | 2,6-dideoxy-2,6,-imino-7-*O*-(β-D-glucopyranosyl) -D-glycerol-L-gulohepitol |

| | |
|---|---|
| MEM | Minimal Essential Medium |
| MOI | Multiplicity of Infection |
| M*O*N-DNJ | *N*-methoxynonyl Deoxynojirimycin |
| MSX | Methionine Sulphoxime |
| MTS | 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl) -2-(4-sulfophenyl)-2H-tetrazolium |
| NACWO | Named Animal Care and Welfare Officer |
| *N*B-DNJ | *N*-butyl Deoxynojirimycin |
| NC | Nucleocapsid |
| NHP | Non-Human Primate |
| *N*M-DNJ | *N*-methyl Deoxynojirimycin |
| *N*N-DNJ | *N*-nonyl Deoxynojirimycin |
| NPC1 | Niemann-Pick C 1 |
| NP-HPLC | Normal Phase High Performance Liquid Chromatography |
| $OD_{450}$ | Optical Density at 450nm |
| ORF | Open Reading Frame |
| OST | Oligosaccharide Transferase |
| PBMC | Peripheral Blood Mononuclear Cell |
| PBS | Phosphate Buffered Saline |
| PCP | *Pneumocystis carinii* Pneumonia |
| PCR | Polymerase Chain Reaction |
| PDI | Protein Disulphide Isomerase |
| PEI | Polyethylenimine |
| PHE | Public Health England |
| PMSF | Phenylmethanesulfonylfluoride |
| PNGS | Potential *N*-linked Glycosylation Site |
| po | *Per os* (by mouth) |
| PR | Protease |
| qid | *Quarter in die* (four-times daily) |
| $R_C$ | Reduced gp120 following signal peptide cleavage |
| RESTV | Reston Ebolavirus |
| rgp120 | Recombinant gp120 |
| RLU | Relative Luminescence Units |
| RNA | Ribonucleic Acid |
| RT | Reverse Transcriptase |
| $R_U$ | Reduced gp120 including the signal peptide |
| SARS | Severe Acute Respiratory Syndrome |
| sCD4 | Soluble CD4 |
| SDS | Sodium Dodecyl Sulphate |
| SDS-PAGE | Sodium Dodecyl Sulphate Polyagarose Gel Electrophoresis |
| siRNA | Small Interfering Ribonucleic Acid |
| $SIV_{Mac}$ | Macaque Simian Immunodeficiency Virus |
| SRT | Substrate Reduction Therapy |
| SU | Surface Unit |
| SUDV | Sudan Ebolavirus |
| TAFV | Tai Forest Ebolavirus |
| TBS | Tris-Buffered Saline |
| $TCID_{50}$ | Tissue Culture Infectious Dose, 50% |
| $T_h$ cell | Helper T-cell |
| tid | *Ter in die* (thrice daily) |
| TM | Transmembrane |

| | |
|---|---|
| TMB | Tetramethylbenzidine |
| UGGT | UDP-Glucose:Glycoprotein Glucosyltransferase |
| V1-5 | Variable regions of HIV gp120 1-5 |
| vRNA | Viral Ribonucleic Acid |
| VSV | Vesicular Stomatitis Virus |
| WT | Wild Type |

# iv. Abstract

Iminosugars are a class of molecules that resemble sugars but for the substitution of a nitrogen for the oxygen in the hemiacetal ring. Several iminosugars are inhibitors of cellular glycosidase enzymes in the *N*-linked glycan processing pathway. Iminosugars inhibiting endoplasmic reticulum (ER) α-glucosidases are known to be antiviral against a broad range of enveloped viruses including human immunodeficiency virus (HIV), hepatitis C virus, dengue virus and influenza virus. This antiviral effect is believed to be due to the indispensability of the calnexin/calreticulin pathway, entry into which is dependent on the trimming of glucoses from *N*-linked glycans, for the correct folding of viral glycoproteins. Thus cells treated with these drugs are unable to correctly fold the envelope glycoproteins of viruses with the result that the secreted virions have diminished infectivity.

A long standing question arising from this hypothesis is how such a mechanism can prove to be damaging to such a wide range of viral glycoproteins yet not show any significant cytotoxicity. This thesis answers that question by demonstrating that, at antiviral concentrations, only a small proportion of viral glycoproteins are misfolded suggesting an amplification effect from protomer oligomerisation and clustering. This thesis also shows how α-glucosidase inhibition results in abnormal disulphide bond formation and isomerisation of the HIV envelope protein, gp120, during folding, as well as faster transit times through the secretory pathway, both potentially explaining the observed misfolding. A specific glycan, N241, is identified as the potential key for gp120/calnexin interaction.

α-glucosidase inhibitors have two effects on glycoproteins; inhibition of calnexin mediated folding but also on the structure of the resultant glycan. This thesis uses mannosidase inhibiting iminosugars to examine the comparative significance of each of these to the antiviral effects of iminosugars. This thesis confirms that failure to process glycans beyond glucose trimming can be substantially antiviral in at least one HIV isolate.

Finally this thesis looks at the potential for using α-glucosidase inhibiting iminosugars as antivirals against Ebola virus in a guinea pig model. These experiments show that iminosugars are safe, even at very high doses, when given intravenously in animals and that one of them, *N*B-DNJ, may offer partial protection against Ebola virus, allowing 1/4 guinea pigs to survive a lethal dose of the virus. Although preliminary, this is the first time that a licensed drug has been shown to be antiviral against Ebola virus and shows the potential utility of these drugs as protection against emerging viral infections for which specific therapies are not yet available.

# 1 Introduction

## 1.1 Iminosugars

In the 1930s oncologists noticed that patients with acute lymphoblastic leukaemia (ALL) were deficient in folic acid and hypothesised that this was the cause of the disease. They had it hopelessly backwards; the cancer was, in fact, depleting folic acid at such a rate that it had become a limiting factor in its growth, leading to some tragic clinical trials in which injection with folic acid led to accelerated progression of the disease [1]. This disaster led Sidney Farber to consider using recently synthesised folic acid antagonists to treat the disease- molecules that looked a lot like folic acid but that were different enough that they would inhibit folic acid metabolism in the tumour [2]. One such analogue, aminopterin, differed from folic acid by the simple substitution of a hydroxyl with a primary amine (Figure 1-1), and yet became the first drug to ever induce remission in children with ALL [3,4].

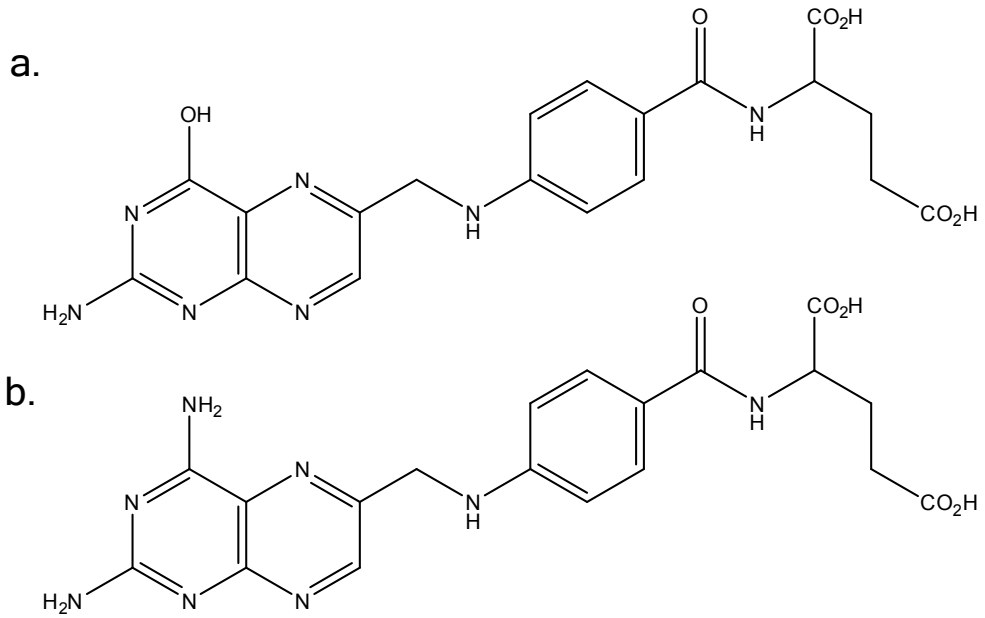



**Figure 1-1** Spot the difference: folic acid (a), a B-vitamin vital for nucleotide synthesis and tumour growth, and aminopterin (b), the first successful anti-cancer chemotherapeutic.

Aminopterin was a key success story in the history of chemotherapy, and the principle of inhibiting processes vital to diseases using substrate analogues was exploited further over the coming decades, including the development of iminosugars. In 1965 the replacement of the hemiacetal oxygen in sugars, initially xylose, with sulphur was found to result in compounds capable of inhibiting glycosidases [5]. This work was quickly expanded to show that similar results could be obtained in a range of sugars, also using nitrogen as the heteroatom [6]. Such sugars, with nitrogen substitutions, would later come to be known as iminosugars.

In the coming years, the naturally occurring iminosugars nojirimycin, a sugar indistinguishable from D-glucose but for its hemiacetal nitrogen, and the closely related 1-deoxynojirimycin (DNJ, also known as DNM) were isolated from a number of bacterial species and the white mulberry tree [7-12]. In 1982 DNJ was shown to inhibit the formation of complex glycans in *Saccharomyces cerevisiae* by inhibiting the action of the Endoplasmic Reticulum (ER) α-glucosidases I and II (GlcI and GlcII), the first enzymes in the N-glycan processing pathway (see section 1.3) [13]. Subsequently, *N*-alkyl derivatives of DNJ were shown to have even greater potency at inhibiting these enzymes *in cellula* [13,14]. However, the real interest began when α-glucosidase inhibitors were shown to inhibit a new virus that had recently begun a devastating pandemic- the Human Immunodeficiency Virus (HIV) [15-18].

## 1.2 HIV and AIDS

### 1.2.1 Acquired Immunodeficiency Syndrome

Human Immunodeficiency Virus (HIV) is the aetiological agent of Acquired Immunodeficiency Syndrome (AIDS) in humans [19-23]. Two HIV viruses have been discovered, HIV-1 and HIV-2, both of which are retroviruses of the family lentiviruses and both of which cause AIDS [19-25]. This work will focus primarily on HIV-1 and any reference to HIV should be assumed to refer to HIV-1 unless otherwise specified.

AIDS is characterised by the progressive destruction of the host's immune system, in particular the loss of $CD4^+$ helper T-cells ($T_h$ cells). $T_h$ cells play a central role in the adaptive immune response, detecting foreign and/or dangerous xenobiotics via peptide fragments displayed on professional antigen presenting cells such as dendritic cells. They then have numerous effector roles including secreting cytokines that will define the nature of the immune response as well as stimulating the effector functions of cytotoxic T-cells and B lymphocytes. Therefore, the immunosuppression induced by $T_h$ cell loss is particularly profound and leads to a high susceptibility to the opportunistic pathogens that typify AIDS.

Since its discovery in 1981 as the cause of a cluster of cases of rare, infectious diseases including Kaposi's Sarcoma, *Pneumocystis carinii* pneumonia (PCP), herpes simplex and mucosal candidiasis, in homosexual American males [26-29], AIDS has caused the deaths of over 30 million people worldwide [30] with approximately 35.3 million people currently living with HIV infection [31]. There is no licensed vaccine.

### 1.2.2 HIV structure

The HIV virion consists of a lipid envelope separated from an internal protein capsid by a matrix. The capsid contains two copies of the RNA genome (vRNA) and two enzymes; the viral reverse transcriptase and integrase (Figure 1-2).

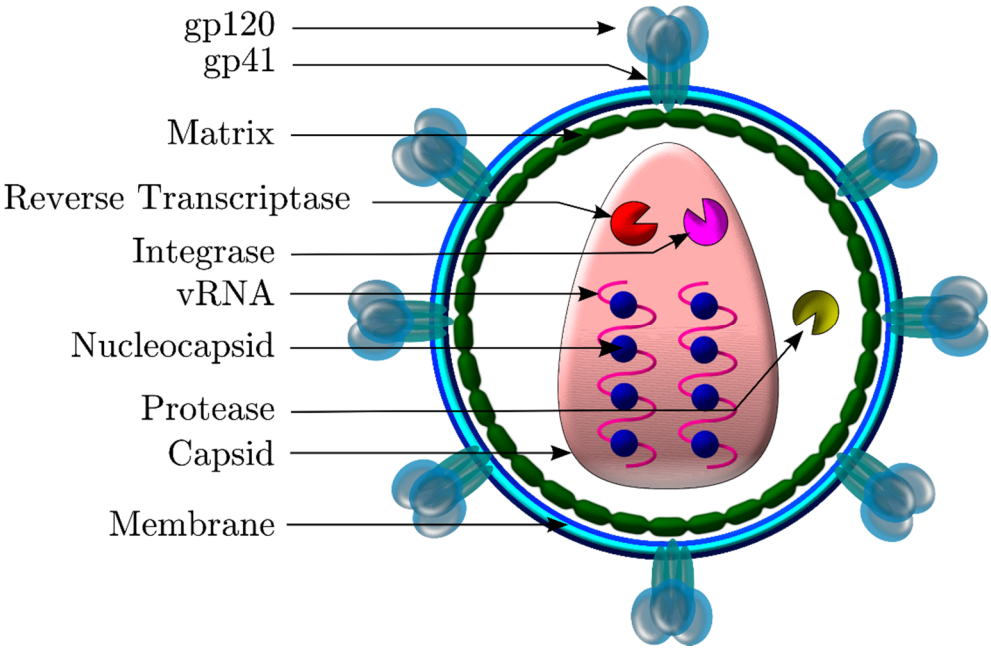



**Figure 1-2** Schematic diagram of the HIV virion.

All of the structural proteins of the virus are translated from the three major open reading frames (ORFs) which are produced as large polyproteins that are cleaved into the individual proteins by host and viral proteases, while a number of regulatory proteins are produced from their own ORFs. The three major ORFs are named Group-Specific Antigen (*gag)*, Polymerase *(pol)* and Envelope (*env)*.

The *gag* gene encodes the 55 kDa Gag polyprotein, which is post-translationally cleaved during virion maturation to form three major and three minor cleavage products. The three major proteins are the nucleocapsid p7 (NC) which binds and stabilises the RNA genomes, the capsid protein p24 (CA), roughly 2000 of which form the conical inner capsid and the matrix protein p17 (MA), roughly 2000 of which line the inner surface of the viral envelope [32].

The envelope itself is budded from the previous host cell's plasma membrane and hence is decorated with a variety of host proteins as well as the major viral envelope proteins derived from the 160 kDa Env polyprotein. This envelope spike consists of a homotrimer of heterodimers of the surface unit gp120 (SU) and the transmembrane protein gp41 (TM) [32]. None of these interactions

are covalent and the trimer is, at best, metastable, a property that is vital for driving cell/virion fusion, discussed in detail later (see Section 1.2.7).

The Pol polyprotein encodes the three essential enzymes that are carried within the virion; protease (PR), reverse transcriptase (RT) and integrase (IN). Three accessory proteins; Nef, Vif and Vpr are also packaged into the virion while another three; Rev, Tat and Vpu appear to be non-structural [32].

### 1.2.3 HIV replication cycle

HIV replicates in $CD4^+$ cells; primarily $T_h$ cells but also in macrophages. The cycle begins when the virion binds to CD4 via gp120. This induces a conformational change in gp120 exposing a second receptor site which allows binding of a co-receptor- either CCR5 or CXCR4. This binding leads to further conformational changes which induce shedding of gp120, insertion of gp41 into the host cell membrane and then further rearrangements and clustering leading to cell/virion fusion and capsid release into the cytoplasm. As the fusion process is central to this thesis it is discussed further in Sections 1.2.5-7.

Following fusion the capsid is unpacked to release the viral nucleoprotein complex. An RNA-dependent-DNA-polymerase (reverse transcriptase), transcribes the vRNA genome into linear, double stranded cDNA which is then trafficked to the nucleus and integrated into the host cell's genome by integrase, whereupon it becomes a provirus. Viral transcripts can then be produced by the cell's own RNA polymerase via a promoter region in the 5' long terminal repeat (LTR) of the provirus. The accessory protein Tat can significantly enhance the rate of provirus expression. Following splicing, the transcripts are exported to the cytoplasm and either translated (in the early stages of virion production) or packaged into nascent virions (later in the cycle when structural proteins are available), a process regulated by Rev [33].

The Gag polyproteins are post-translationally modified with myristic acid at their *N*-termini which allows them to associate with the inner leaflet of the plasma membrane [34]. The Pol polyprotein is found covalently associated with some copies of the Gag protein and so is similarly localized. The Env polyprotein trimerises in the ER before being exported to the Golgi body where cleavage of gp120 from gp41 is mediated by the host protease furin. It is then exported to the plasma membrane as a membrane protein. To prevent the gp120 immediately associating with CD4 on the host cell surface, the accessory proteins Vpu and Nef work to reduce surface expression the receptor [33].

vRNA associates with Gag which, in an unrelated function, induces curvature of the plasma membrane allowing virions to bud from the cell. Following budding the virion undergoes a process of maturation, whereupon the viral protease cleaves Gag and Pol into their constituent proteins, leading to the formation of a mature, infectious virion capable of now infecting a new $CD4^+$ cell.

This whole process, from infection of the first cell through to infection and production of new virions by a second cell, takes an average of 2.6 days, with the intracellular phase of the lifecycle (from virus in to virus out) lasting a minimum of 0.9 days [35]. Each productively infected cell (i.e. those actively secreting virus) has an average lifespan of roughly 1 day, in which time it produces on the order of $10^4$ new viruses [36].

### 1.2.4 Pathogenesis of HIV infection

The primary routes of HIV infection are by sexual transmission (vaginal or anal intercourse), blood to blood transmission (transfusions, intravenous drug use) and vertical transmission (breast milk and neonatal exposure to maternal blood). One to three weeks following the point of infection, individuals experience symptoms of acute HIV infection, including a combination of headache, retro-orbital pain, myalgia, fever, sore throat and lymphadenopathy. In addition to these relatively non-specific symptoms (easily misdiagnosed as influenza, mononucleosis or other acute viral infections) some patients experience a macular, erythematous rash which can be more pathognomonic. These symptoms subside quickly but the lymphadenopathy can persist for months. Following this acute phase there then follows a latency period of many years during which the patient as asymptomatic [37]. This long latency can explain the high prevalence of HIV as it provides a very large window for further transmission to occur.

Due to the presence of large numbers of effector and memory T-cells in the tissues exposed to HIV during infection (e.g. the lamina propria of the rectum), virions which are selective for these cells tend to form the founder population to a greater degree than those targeting naïve T-cells [38]. This selectivity is conferred by the co-receptor tropism of the virus. HIV can use either of the two cytokine receptors CCR5 and CXCR4 to mediate cell fusion; CCR5 is found only on effector T-cells and thus so-called "R5 tropic" strains tend to predominate early in HIV infection. "X4 tropic" strains evolve later in the infection and are associated with a precipitous drop in circulating $CD4^+$ cells. Destruction of T-cells is via a number of mechanisms, including direct viral cytopathy, destruction of infected $CD4^+$ cells by $CD8^+$ cells and the induction of apoptosis by viral interference with cell function [37].

The following latent period sees a slow decline in peripheral $CD4^+$ cell numbers which ultimately accelerates to a critical level leading to immunosuppression. However, although it appears slow this is a highly dynamic process with large $CD4^+$ losses being matched by the high regenerative capacity of the bone marrow in maintaining homeostasis. However, this capacity for replacement appears to be limited, hence the progression to AIDS [38]. This capacity for regeneration is important as it means that antiviral chemotherapy can quickly lead to a recovery in T-cell numbers [35].

However, not all infected cells die; due to the nature of the retrovirus lifecycle it is possible for a cell to harbour an integrated provirus that does not immediately express virus proteins but can instead remain quiescent for many years before recrudescing. These cells, notably memory $T_h$ cells but potentially also macrophages, brain cells and or haematopoietic stem cells, then act as a reservoir for resumption of the infection should the actively infected cells be destroyed [39].

### 1.2.5 HIV envelope

Attachment and the subsequent fusion of the HIV virion to the target cell are key stages in the viral lifecycle and thus a potential targets for therapy. This has traditionally taken the form of developing vaccines designed to elicit neutralising antibodies that could bind to viral ligands and block their association with cellular receptors, but they are also a target for antiviral chemotherapeutics. Such a therapy will be the subject of much of this thesis and so the fusion mechanism deserves further elaboration.

Fusion is mediated by the HIV envelope spike, a homotrimer of heterodimers of gp120 and gp41, the products of furin cleavage of the Env (gp160) polyprotein [40]. Despite the popular image of HIV as being coated in a dense thicket of spikes like some sort of pathogenic sea urchin, cryo-electron tomography studies have revealed that the average number of spikes is 14 per virion [41] and super-resolution fluorescence microscopy puts the average figure nearer 7-8 spikes [42].

gp120 is a globular protein that associates atop gp41 via non-covalent interactions and is responsible for binding to the CD4 and chemokine receptors. Despite its name implying a 120 kDa mass, roughly half of this mass comes from the *N*-linked glycans that populate its surface [43]. The exact number is strain dependent and can vary from 18 to as many as 33, with a median of 24 [44,45]. The protein is divided into 5 variable and 5 (relatively) constant regions (termed V1-5 and C1-5, respectively) [46,47]. Four of the variable regions form separate folds to the rest of the protein by means of 9 disulphide bonds that bunch them into loops, namely the V1, V2, V3 and V4 loops, with the first two further bunched to form the so-called V1/V2 loop [48]. The protein can be further subdivided into "inner" and "outer" domains. The inner domain is orientated towards the centre of

the trimer while the outer is exposed on the surface; as well as into distal and proximal regions, with the two termini being held proximally to the virion surface while V1/V2 and V3 are distally located at the tip of the spike [49]. These factors help explain HIV's extreme capacity for antigenic escape as the constant regions, which maintain much of the functionality of the protein, can remain masked in the core of the globule while the glycans and variable loops form the antigenic surface of the protein [49].

Beneath this cap of gp120s sits a trimer of gp41. Each gp41 possesses a single trans-membrane domain with a long (˜150 amino acid) *C*-terminal cytoplasmic tail, a single disulphide bond and 4 glycans. The extracellular domains of the three gp41 protomers form a 6-helical bundle formed of a central coiled-coil surrounded by three anti-parallel helices, held together by strong hydrophobic interactions [50]. The 28 *N*-terminal amino acids are all hydrophobic and thus form the fusion peptide [51].

### 1.2.6 SOSIP envelope trimers

HIV envelope is the only exposed antigen on the viral surface and so has been the major focus for the development of antibody-based vaccines. Inoculation with recombinant envelope proteins does result in the production of antibodies but they are often very narrow in their neutralisation profile, being only substantially effective against the inoculum strain [52,53], and even then the virus can quickly evolve resistance [54]. This has led to a lot of focus on engineering of envelope protein immunogens to best elicit broadly neutralising antibodies (bnAbs), as well as structural surveys of the protein to best understand these broadly neutralising epitopes. One feature of bnAbs is that they tend to associate better with trimeric envelope than with monomeric gp120 [55]. These effects are attributed to the immunodominance of non-neutralising epitopes which are normally buried in the trimer but are exposed when the trimers break down. As the gp120/gp41 interaction is so unstable the majority of envelope protein in a patient's serum is defective and exposes these epitopes, potentially distracting the immune system from more significant, broadly neutralising epitopes on the trimer [56].

These problems have led to the invention of recombinant, stabilised, soluble envelope trimers that closely mimic the structure of the native viral spike but do not dissociate as readily. Solubilisation is a simple matter of deleting the cytoplasmic and transmembrane domains of gp41, leaving only the external $gp41_{ECTO}$ domain. The easiest way to prevent $gp120/gp41_{ECTO}$ dissociation is to mutate out the cleavage site, but this results in trimers which are antigenically very different from the native virion. Instead, an extra disulphide bond can be introduced between $gp41_{ECTO}$ and

gp120 to prevent dissociation while still allowing cleavage [55]. The fact that these possessed an extra disulphide ("*S-S*") bond led to the nomenclature "SOS trimer"; "SS trimer" was rejected due to unfortunate fascist connotations (Rogier Sanders, personal communication). This interaction can be further optimised by introducing other mutations in gp41 to stabilise its pre-fusogenic form [57]. As the key mutation was I559P these further stabilised trimers became known as "SOSIP". Cryoelectron tomography revealed that SOSIP closely resembled the native virion associated trimer and underwent similar conformational changes on CD4 binding [58].

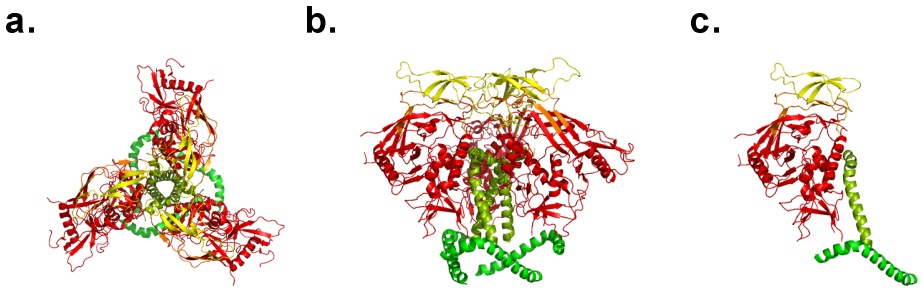



**Figure 1-3** Crystal structure of BG505 SOSIP.664, a soluble, cleaved HIV envelope trimer. The core gp120 is shown in red with the V1/V2 and V3 loops in yellow and orange, respectively. The membrane proximal helix of $gp41_{ECTO}$ is shown in grass-green while the distal helix is shown in yellow-green. The trimer is shown from above, looking down towards the virion surface along the axis of symmetry (a), and from the side (b). A single protomer is also shown, in the same orientation as in b (c). The majority of the V2 loop and the fusion peptide are unresolved. Glycans have been removed for clarity. This image is based on the crystal structure described in *Julien et al.* [59] *(PDB: 4NCO)*

SOSIP has been in development for over a decade and the current generation BG505 SOSIP.664 produces homogenous trimers with high efficiency [60]. This trimer was so homogenous that it was possible to perform both crystallographic and cryoelectron microscopy (cryo-EM) structural studies of the trimers (Figure 1-3) [59,61]. Unlike previous structures of deglycosylated, core gp120 [49] these structures included the V1/V2 loop and certain glycans as well as the $gp41_{ECTO}$ domain. Nonetheless, the core gp120 structure largely mapped well into the trimer structure, confirming the utility of monomeric gp120 as a model for envelope structure and function.

### 1.2.7 HIV fusion

As well as binding T-cell receptors, gp120 is also capable, via its glycans, of binding DC-SIGN, a dendritic cell lectin that allows transportation of the virus from the periphery to lymph nodes where it facilitates the infection in *trans* of $T_h$ cells [62]. The fusion process only begins, however, when gp120 binds CD4, thus determining HIV's tropism and, hence, its pathology. This occurs via a deep cavity at the junction of the inner and outer domains of gp120, which interacts with the

crucial Phe43 residue of CD4 [56,63]. This binding triggers a conformational rearrangement as the phenylalanine stabilises a hydrophobic pocket of gp120 around itself [64,65], as well as creating a number of new antigenic sites, termed CD4i (CD4 induced) antigens [66].

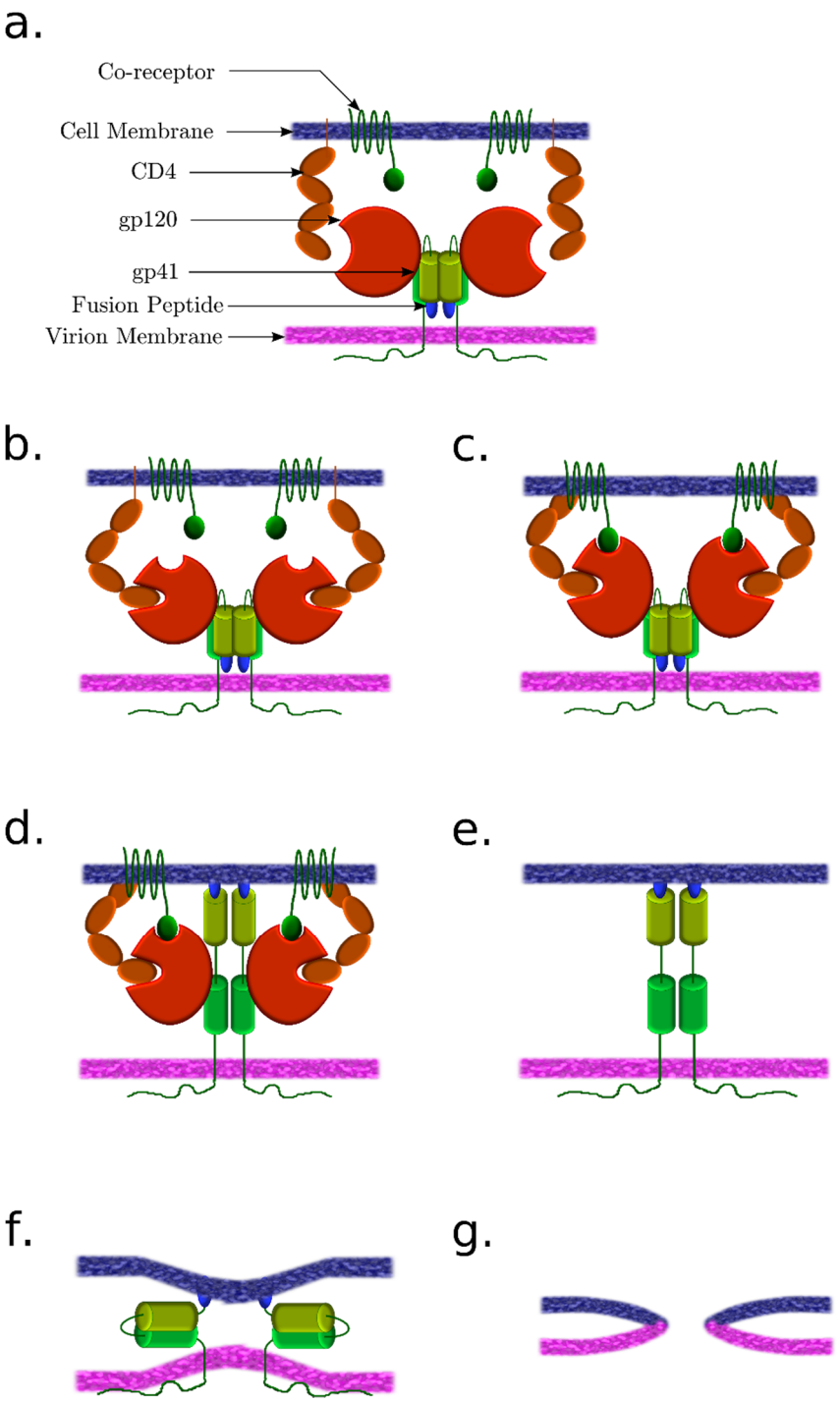



**Figure 1-4** The HIV fusion process. The HIV fusion machinery comprises a trimer of gp41/gp120 dimers (only two pictured for clarity) (a). gp120 binds CD4 on the target cell triggering exposure of the co-receptor binding site (b). The co-receptor, either CCR5 or CXCR4, binds the gp120 (c) triggering the extension of the gp41 *N*-terminal fusion peptide, which can then insert into the target cell membrane (d). gp120 may or may not be shed at this point; for clarity it and its receptors will be excluded from this diagram anyway (e). The helices of gp41 rearrange into six helical bundles (only two shown, for clarity), pulling the two membranes closer together (f) ultimately resulting in the formation of a fusion pore between virus and cell (g).

This rearrangement seems to involve movement of the V1/V2 and V3 loops leading to exposure and stabilisation of the bridging sheet, a collection of β-sheets that join V1/V2 to the core gp120, forming the co-receptor binding site [49,64]. This triggers further rearrangement resulting in the insertion of the gp41 fusion peptide into the target membrane [67-69]. It was initially proposed that prior to insertion gp120 had to dissociate from gp41, a process termed shedding [70,71], but the biological significance of this is questionable [72]. The six helices of the three gp41s then coil into a six-helical bundle, dragging the two membranes together and thus breaking the energy barrier for membrane fusion [63,67-69] (Figure 1-4). It is likely that several trimers need to work together to overcome this energy barrier [63,73] and trimers are observed to cluster together on the virion surface prior to fusion [42].

Thus the HIV fusion process is controlled solely, from the virus's perspective, by the envelope glycoprotein. This protein has to be able to perform a number of different functions, including maintaining stability, binding at least three different receptors, controlled insertion of the fusion peptide and as a motor to draw the membranes together. Not only that, but it must maintain all of these functions while under constant attack from the adaptive immune system, driving extremely high levels of sequence variability [44]. It is this complex structure and function relationship that is disrupted by iminosugar treatment, but to understand why we must first look at how folding and glycosylation pathways interact in the host endoplasmic reticulum (ER).

## 1.3 *N*-linked glycosylation and ER quality control

*N*-linked glycosylation is one of the most common post-translational (actually co-translational) modifications that a protein can receive and has been observed in every domain of life from the Archaea to the Eukarya [74,75]. Not only are *N*-linked glycans (hereafter, "glycans") common, they are also massive. To go back to gp120, two of its glycans sit at the base of the V3 loop, with a combined mass of approximately 4 kDa. The V3 loop itself has a mass of just 3 kDa [45]!

*N*-linked glycans are added to proteins *en bloc* at a potential *N*-linked glycosylation site (PNGS). A PNGS comprises of the motif Asn-X-Ser/Thr (where X is any amino acid except proline), termed a "sequon", with the subsequent amino acid [76] and local protein structure [77] both having an influence on whether the sequon is glycosylated or not. Initially, a "core" oligosaccharide consisting of $Glc_3Man_9GlcNAc_2$, where Glc is glucose, Man is mannose and GlcNAc is *N*-acetyl glucosamine (Figure 1-5), is added to the sequon from a dolichol carrier by the enzyme complex oligosaccharide transferase (OST) [78,79]. As the glycoprotein progresses through the ER and Golgi

compartments it has sugar residues removed ("trimmed") and potentially new residues added to form complex and hybrid type glycans.

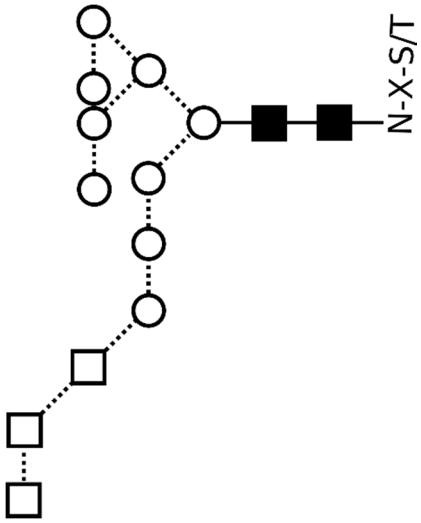



**Figure 1-5** The structure of the core *N*-linked glycan displayed using the standardised system described in *Harvey et al* [80]. Briefly, squares denote glucose residues and circles mannoses; filled shapes denote *N*-acetyl amino variants (therefore, filled squares are *N*-acetyl glucosamine). Connecting lines indicate the nature of the linkage- solid being β linkages and dashed α, and the angle determines the linkage position- 180° is 1,2; 225° 1,3; 270° 1,4; 315° 1,6 and, for completeness 360° is 1,8 (not shown). By convention, the three "arms" are termed D1, D2 and D3 (in order from the bottom of the figure). The sequon is shown using the single-letter amino acid code where X is any amino acid except proline and the third position can be serine or threonine.

In the ER, the glycans serve a key role in glycoprotein folding. This can be shown by treatment with tunicamycin, which blocks synthesis of the core glycan and thus prevents glycosylation, and quickly leads to the accumulation of unfolded proteins in the ER. Part of this can be ascribed to the physical effects of glycan removal- they may be a necessary folding scaffold or serve to bury hydrophobic residues. However, glycans also have a specific role in glycoprotein folding through the so-called calnexin/calreticulin cycle (Figure 1-6).

### 1.3.1 Glucosylation and the calnexin cycle

Calnexin (CNX) and calreticulin (CRT) are structurally related ER lectins that are preferentially associated with misfolded glycoproteins [81,82]. Calnexin is bound to the ER membrane while calreticulin is found in the lumen [82]. As calnexin is the better characterised of the two, and more relevant to the proteins in this study, only calnexin will be discussed in detail but the two can generally be considered analogous. Although calnexin was initially only recognised to bind misfolded glycoproteins, it became clear that it actually very specifically bound $Glc_1Man_9GlcNAc_2$ glycoproteins, regardless of the folding state [83].

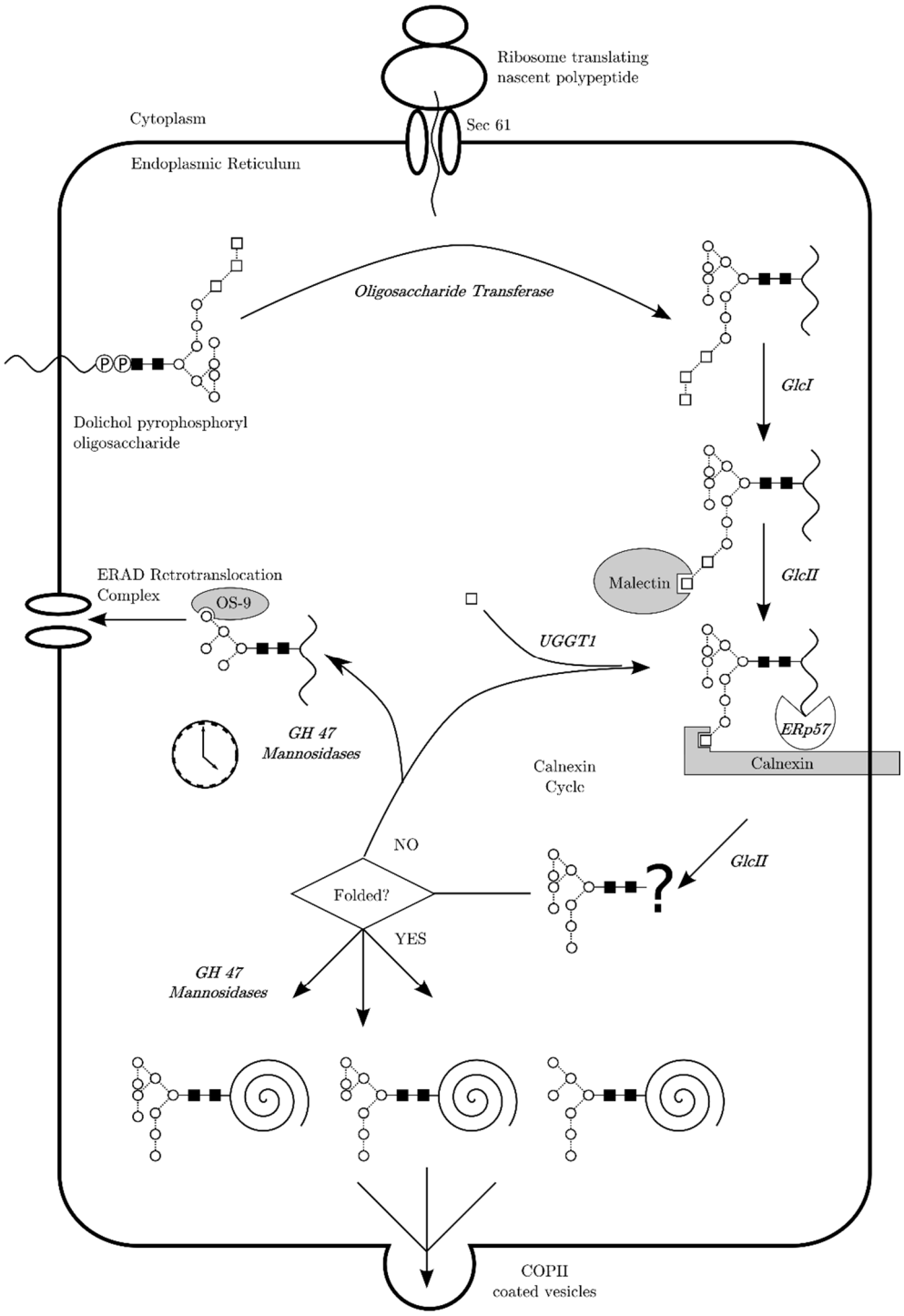



**Figure 1-6** Glycan trimming and the calnexin cycle in the ER. The core oligosaccharide is added to nascent polypeptides, co-translationally, by OST. GlcI and GlcII trim the two terminal glucoses. $Glc_1Man_9GlcNAc_2$ is the optimal substrate for calnexin (or its soluble analogue, calreticulin), which holds the protein while associated chaperones, such ERp57, fold it. GlcII cleaves the remaining glucose freeing the protein. If the protein is correctly folded it is sent to COPII coated vesicles for export to the Golgi body, with or without further trimming of mannose residues by GH 47 family α-1,2 mannosidases. If it is misfolded then this is sensed by UGGT1 and the protein is re-glucosylated, allowing it to re-associate with calnexin and have a further chance at folding. This cycle cannot continue forever and terminally misfolded proteins can escape the cycle by the time dependent removal of mannoses. Removal of mannoses from the D3 arm is recognised by OS-9 which then targets the protein for degradation through the ERAD pathway. For full description see text. *Nota bene*: Calnexin substrates are usually transmembrane proteins while calreticulin usually associates with soluble proteins; for clarity the transmembrane domains are excluded from the proteins shown above. Enzymes are shown in *italics* and lectins in grey.

The $Glc_1Man_9GlcNAc_2$ species is generated by the sequential removal of the two terminal glucoses by GlcI and GlcII [84]. The first glucose is removed by GlcI within seconds, leading to the suggestion that it exists only to bias oligosaccharide transferase (which is highly specific for $Glc_3Man_9GlcNAc_2$) towards addition, and not removal, of the glycan [85]. The resulting $Glc_2Man_9GlcNAc_2$ species is not of any known significance, though the discovery of malectin, a lectin specific for this intermediate, suggests that it may have undiscovered functions [86]. The removal of the next glucose by GlcII is dependent on the presence of a second glycan, either on the same polypeptide or on an associated protomer. The mannoses from this glycan bind to the β subunit of GlcII and activate its glucosidase activity allowing efficient removal of the glucose [87].

Calnexin does not seem to have any catalytic activity of its own but instead facilitates folding by retaining the protein in the ER. It also directly associates with the protein disulphide isomerase (PDI) family member ERp57 [88] which is then responsible for disulphide isomerisation of calnexin-bound glycoproteins [89]. Dissociation from calnexin is also controlled by GlcII, this time removing the last remaining glucose [83]. Even though, as previously, this is an α-1,3 linkage, it is in a different orientation thus preventing GlcII from removing it immediately after the first cleavage event and thus allowing time for calnexin association. In this orientation the glycan's own mannoses can bind the β subunit thus obviating the need for a second glycan [87]. Calnexin also has an independent role in $Ca^{2+}$ signalling (hence the name); this job appears to be limited to palmitoylated calnexin, while the unmodified form fulfils the lectin role [90]. Under ER stress calnexin can be depalmitoylated to increase that available for folding; this is worth being aware of as it means that the functional concentration of calnexin can vary in a way that would be invisible by mRNA or protein quantification.

As well as assisting in initial folding, the calnexin and GlcII also play a role in glycoprotein quality control through the action of a folding sensor termed UGGT1 (UDP-glucose:glycoprotein glucosyltransferase 1). The nature of UGGT1's ability to detect misfolding is uncertain, but the two leading hypotheses are the exposure of hydrophobic residues that would normally be buried [91] and the exposure of disulphide bonds that are either potentially incorrectly exposed or paired, as evidenced by its association with the redox-active protein Sep15 [92]. However it does it, once UGGT1 senses a misfolded glycoprotein it reglucosylates it, reforming the $Glc_1Man_9GlcNAc_2$ species and allowing reassociation with calnexin and another chance at correctly folding [83,93].

### 1.3.2 Mannosylation and degradation

There is no theoretical limit to how many times a protein can go through the calnexin cycle, but in reality some proteins will be irrevocably misfolded and no amount of cycling will be able to recover them. Thus the cell needs an alternative exit from the cycle for terminally misfolded proteins. While UGGT1 is a folding sensor it cannot distinguish transiently from terminally misfolded proteins, so it is unable to control this alternative exit. Instead, degradation of proteins appears to be time dependent; proteins that linger in the ER, failing to fold, become increasingly at risk of being targeted for destruction by being retrotranslocated to the cytoplasm for proteasomal degradation- ER Associated Degradation (ERAD).

As in the calnexin cycle, this process is dependent on protein glycosylation (although glycan independent degradation exists as well), with the glycan acting to target the protein to specific pathways. In this case, the trimming of the mannose residues is key [94]. Work in yeast with carboxypeptidase Y* (CPY*), a commonly used misfolded protein model, identified that $Man_8GlcNAc_2$ and, to a substantially lesser extent, $Man_5GlcNAc_2$ species were more readily degraded than either glucosylated or alternatively mannosylated glycoproteins [95]. This, combined with the knowledge that ER α-1,2 mannosidase I (ManI) trimming was slower than that of the glucosidases [94,95], led to the "mannose timer" hypothesis. This hypothesised a specific lectin that could recognise $Man_8GlcNAc_2$ glycoproteins and target them for degradation. As ER ManI had a relatively long half-life (10 minutes for CPY* cf. < 2 minutes for the glucosidases) only proteins that persisted in the ER would be trimmed, and thus targeted, in this way [85,95]. A family of mammalian genes that resembled mannosidases but seemed not to have catalytic activity, ER degradation enhancing α-mannosidase-like proteins (EDEMs), were soon proposed as these lectins [96].

However, later work in mammalian cells showed that ERAD substrates were more extensively trimmed, down to $Man_6GlcNAc_2$ and $Man_5GlcNAc_2$ [97]. This was significant because, unlike $Man_8GlcNAc_2$, these species lack the UGGT1 reglucosylation site and thus are permanently lost to the calnexin cycle with no hope of refolding. Although ER ManI could only efficiently generate $Man_8GlcNAc_2$ *in vitro,* it was shown to be highly concentrated in a specific region of the ER, the ER quality control compartment (ERQC), permitting it to trim down to these species [97,98]. OS-9 was identified as the lectin responsible for binding these glycoproteins and sending them to ERAD; it specifically recognises the lack of a terminal mannose on the D3 arm [99].

Although great progress has been made in the last fifteen years understanding ERAD targeting, which mannosidases are responsible for the trimming have been highly debated. As well as having lectin abilities, there is also evidence that EDEMs 1 and 3 do, in fact, participate in mannose trimming, either directly or indirectly [100,101], while a recent paper was even bolder in claiming that all 3 EDEMs have mannosidase activity and that it is EDEM 2 that is responsible for the first trimming step, the one step that most authors agreed was performed by ER ManI [102]! Even Golgi mannosidases have been implicated with the discovery that they are capable of cycling back to the ER [103]. All of these mannosidases belong to the Glycoside Hydrolase 47 (GH 47) family [85] and this will be used as a general term to describe these mannosidases, such as in Figure 1-6.

### 1.3.3 From the ER to the Golgi

Correctly folded glycoproteins are able to leave the ER via cytosolic coatamer protein (COP) II coated vesicles. These vesicles transport the proteins to the ER-Golgi intermediate compartment (ERGIC) from which they leave for the Golgi in COPI coated vesicles [85]. Entry into COPII vesicles is facilitated by L-type lectins termed VIPL and ERGIC53 and entry into COPI by another termed VIP36. These bind the D1 branch of the glycan [104], further supporting the notion that it is mannosidase trimming of this branch that targets glycoproteins for ERAD. $Man_{8/9}GlcNAc_2$ and, to a lesser extent, $Man_7GlcNAc_2$ are the highest affinity ligands for VIPL and VIP36 while, importantly for the study of glucosidase inhibition, ERGIC53 is less discerning and binds a range of glycans, though with low affinity, including monoglucosylated species [104]. It is reasonable to assume that, under normal circumstances, the majority of glycoproteins will arrive at the Golgi bearing $Man_{8/9}GlcNAc_2$ glycans.

As the glycoproteins transit the three Golgi compartments they can potentially undergo further mannose trimming followed by the addition of new sugars such as *N*-acetyl glucosamine, fucose, galactose and *N*-acetylneuraminic acid; these are known as complex glycans. In other cases steric hindrances prevent these enzymes from accessing the glycan, causing it to be unmodified from the form in which it left the ER- these are known as high mannose glycans. When only part of the glycan is trimmed (for example, only one branch) a glycan can be formed with a mixture of mannosylated and modified branches on the same oligosaccharide; a hybrid glycan [105]. Unlike ER processing, this process is highly variable. As a result, a single glycoprotein can find itself decorated with a range of glycans of varying structures; a pattern that is both protein and cell specific, probably with a random element as well [105].

### 1.3.4 Alternative ER folding pathways

Protein folding is obviously not exclusively glycan dependent, as there are a large number of proteins that are not glycosylated at all. These pathways involve the activity of oxidoreductases, peptidyl-prolyl *cis-trans* isomerases and other miscellaneous chaperones [106], the two most important of which will be considered briefly below.

#### *Oxidoreductases*

Oxidoreductases are chaperones that assist in the correct formation of disulphide bonds. In the ER, these are generally members of the protein disulphide isomerase (PDI) family. These proteins contain thioredoxin domains, characterised by the C-X-X-C motif (where X is any amino acid), with which they form transient mixed disulphides with the protein they are folding [107-109]. In this way they can retard the disulphide based aggregation of proteins [110] and break and reform disulphides that have mispaired [111,112]. There are a number of PDIs, including tissue specific PDIs and ERp57, the calnexin associated oxidoreductase discussed previously, but the most common of them is simply referred to as "PDI" and is one of the most common chaperones in the ER, second only to the immunoglobulin heavy-chain binding protein, BiP [106].

#### *Heat Shock Protein (HSP) analogues*

BiP (also known as Glucose Regulated Protein 78, GRP78) is the mammalian ER luminal analogue of the bacteria heat shock protein HSP70 [113] and to many (non-glycobiologists) it is synonymous with ER protein folding. It possesses an *N*-terminal ATPase domain and a *C*-terminal domain which is able to bind hydrophobic domains, such as those that might be exposed in an unfolded protein. Unfolded proteins go through repeated cycles of binding and release from BiP in an ATP dependent manner [114]. BiP does not function on its own but instead has many interaction partners that enhance its activity by helping bind substrates or enhance ATP hydrolysis [106]. It does not appear to have any "foldase" activity of its own but instead serves to prevent nascent peptides from aggregating and maintain an unfolded state long enough for PDI and peptidyl-prolyl *cis-trans* isomerases, with which it is often found in complex, to function [115,116].

Another abundant HSP analogue is GRP94, which shares substantial similarity with bacterial HSP90. It is also responsible for chaperoning unfolded proteins but it differs from BiP in two key ways. Firstly, although it has ATP hydrolysing activity, structural studies suggest that this activity alone may not be sufficient to drive the conformational changes that regulate peptide binding and release, inferring the existence of unidentified cofactors [106,117]. Secondly, GRP94 binds a much

smaller subset of unfolded proteins than BiP, although the nature of this selectivity is somewhat mysterious [118].

## 1.4 The glycosylation and folding of HIV envelope protein

Having looked at glycoprotein folding in general, it is now time to apply these principles to our molecule of interest, HIV envelope gp160. While gp160 mostly exploits the systems described above similarly to any other glycoprotein, it does have several unique features worthy of note.

### 1.4.1 Signal peptide cleavage

Signal peptides are short *N*-terminal leader sequences that target proteins to the secretory pathway during translation [119]. Signal peptides are usually cleaved co-translationally [120] but that of gp160 is retained for at least 15 minutes, and potentially many hours, after translocation is completed [121]. Calnexin association occurs while the peptide is still attached and the removal of the peptide correlates with the completion of protein folding, as measured by a gain of CD4 binding ability [121]. There is some argument as to the teleology of this, with Li *et al.* suggesting that the signal peptide purposely retards folding while Braakman *et al.* favour the hypothesis that lack of proper folding retards peptide cleavage [121,122]

### 1.4.2 Glycosylation

gp160 is heavily glycosylated; with a median of 28 sequons (4 on gp41 and 24 on gp120) glycans account for roughly half of the protein's mass [43,45]. Studies of recombinant gp120 show these glycans to be a mix of ˜75% complex and ˜25% high-mannose/hybrid forms, with the variable loops generally bearing the complex species [48]. The constant regions are buried within the core of the protein to prevent exposure to the immune system; it is likely that this location also shields their glycans from processing in the Golgi. However, studies on virus derived envelope show almost no ($< 3\%$) complex glycosylation with $Man_{5\text{-}9}GlcNAc_2$ dominating the glycan profile regardless of viral clade or cell type [123].

Despite some early suggestions that *O*-glycosidases could reduce the apparent molecular weight of gp120 [124], evidence for *O*-linked glycosylation of the envelope is weak. Despite the discovery of carbohydrate binding broadly-neutralising antibodies revitalising the study of HIV glycosylation, no further evidence has been found. Also, the recent electron micrograph and crystallographic structures of HIV envelope proteins provide extensive evidence of *N*-linked glycosylation but none of *O*-linked [49,59,61].

*N*-linked glycosylation is vital for correct gp160 folding; if glycosylation is prevented by, for example, tunicamycin then the resultant protein is unable to bind CD4 [125]. However, gp120 that has had its glycans removed enzymatically post-secretion is still able to bind CD4 [125]. This difference is almost certainly due to the role of glycans as calnexin/calreticulin ligands in the ER. Indeed, gp160 binds both of these lectins soon after translation begins, with calnexin being the major lectin though calreticulin is also found bound to calnexin-linked gp160 [126].

### 1.4.3 Oxidation and folding

gp160 contains ten disulphide bonds, nine in gp120 and one in gp41 [48]. Many of these are non-sequential (i.e. each cysteine does not necessarily pair with the next in sequence) [48,127] making it likely that non-native disulphides will form during translocation into the oxidising environment of the ER. As a result, disulphide intermediates are formed before the final conformation is reached up to four hours after translation begins. These intermediates are evidence of extensive disulphide isomerisation by PDI family members in the ER [128]. Only half of these disulphide bonds are vital for gp160 folding, as defined by signal peptide cleavage and secretion from the ER, although deletion of all but two of these abrogates infectivity in virions [129]. Interestingly, the V1/V2 loop appears to fold independently of the rest of the protein [129].

The envelope protein also interacts with BiP immediately upon translocation, meaning that this interaction covers the same period as disulphide bond formation [130], an unsurprising finding considering the cooperativity of the chaperones. Computer based predictions show that most BiP binding sites in gp160 are located in the conserved regions, perhaps indicating that BiP binding is a vital, and hence conserved, step in gp160 folding [131]. There are no published data on GRP94 having any role to play in HIV envelope folding*.

Interestingly, whilst slow, the folding of gp160 does appear to be highly efficient, with the level of protein sent for proteasomal degradation being undetectable [128]. However, combined with the evidence that non-functional envelope can still be secreted [129] this may reflect poor quality control rather than high efficiency of folding.

### 1.4.4 Oligomerisation, cleavage and export

Non-covalent oligomerisation of gp160 into trimers occurs in the ER while cleavage into gp120/gp41 heterodimers does not occur until later [130]. Monomeric gp160 has a half-life of around 30 minutes [130] and trimerisation only seems to occur after the protein has reached its native

* Based on Scopus and PubMed searches using the keywords “HIV” and “GRP94”.

structure as oligomeric gp160 is not found in association with BiP or calnexin/calreticulin [126,132]. Oligomerisation appears to be dependent on the transmembrane domain of gp41, as recombinant gp140 lacking this region is monomeric, but the introduction of a single point mutation in the gp41 ecto-domain is enough to restore trimerisation in SOSIP gp140 [57].

The exact location of the compartment in which cleavage occurs is controversial with conflicting evidence existing as to whether it occurs in the Golgi itself or the trans-Golgi network [40,133]. Cleavage is vital for HIV infectivity and occurs at a conserved REKR sequence at the *N*-terminal of C5 [134,135]. A number of convertases of the kexin/subtilisin family are able to catalyse cleavage but furin is considered the most likely candidate *in vivo* [136,137]. After cleavage, gp120/gp41 trimers are exported to the plasma membrane from which they are incorporated into nascent virions.

## 1.5 Iminosugars as glycosylation inhibitors

Having seen how glycans play such a large role in glycoprotein folding, compartmentalisation and function we can start to see how glycosidase inhibitors like iminosugars could have significant effects on viral glycoproteins both in the ER and beyond. We will look in detail at the effects of two classes of iminosugar, those inhibiting glucosidases and those inhibiting mannosidases.

### 1.5.1 Mannosidase inhibitors

As we have seen, high mannose glycans are the "prototype" glycans originally produced in the ER, but they can undergo trimming in both the ER and Golgi to remove mannose residues and subsequently replace them with alternative sugars to create complex glycans. Thus mannosidase inhibitors would bias cells away from producing complex glycans towards producing various isomers of high mannose and hybrid glycans depending on which specific mannosidases were affected.

The first mannosidase inhibitor to be characterised was swainsonine. This is an alkaloid produced by a variety of plant and fungal genera around the world and is recognised as the cause of the wonderfully evocatively named neurological disorders locoism and slobbers [138]. While these diseases are caused by the inhibition of a lysosomal α-mannosidase [139], swainsonine also has activity against mannosidases in the secretory pathway, preventing the formation of complex glycans [140,141]. While initially thought to inhibit early mannosidase processing and thus causing a high mannose phenotype, swainsonine was soon identified as a Golgi mannosidase II (ManII) inhibitor, thus allowing the formation of hybrid, but not complex, glycans [142,143].

To produce a pure high mannose phenotype one would have to inhibit either ER or Golgi mannosidases I (ManI), all of which are responsible for the initial trimming of the $Man_9GlcNAc_2$

precursor. Such an inhibitor was synthesised by creating an enantiomer of DNJ with the D-mannose stereochemistry- deoxymannonojirimycin (DMJ) [144]. However, DMJ is specific for the Golgi ManI and has no significant activity against ER ManI, meaning that it induces the production of a range of high mannose isomers from $Man_5GlcNA_2$ up to $Man_9GlcNAc_2$ [145,146].

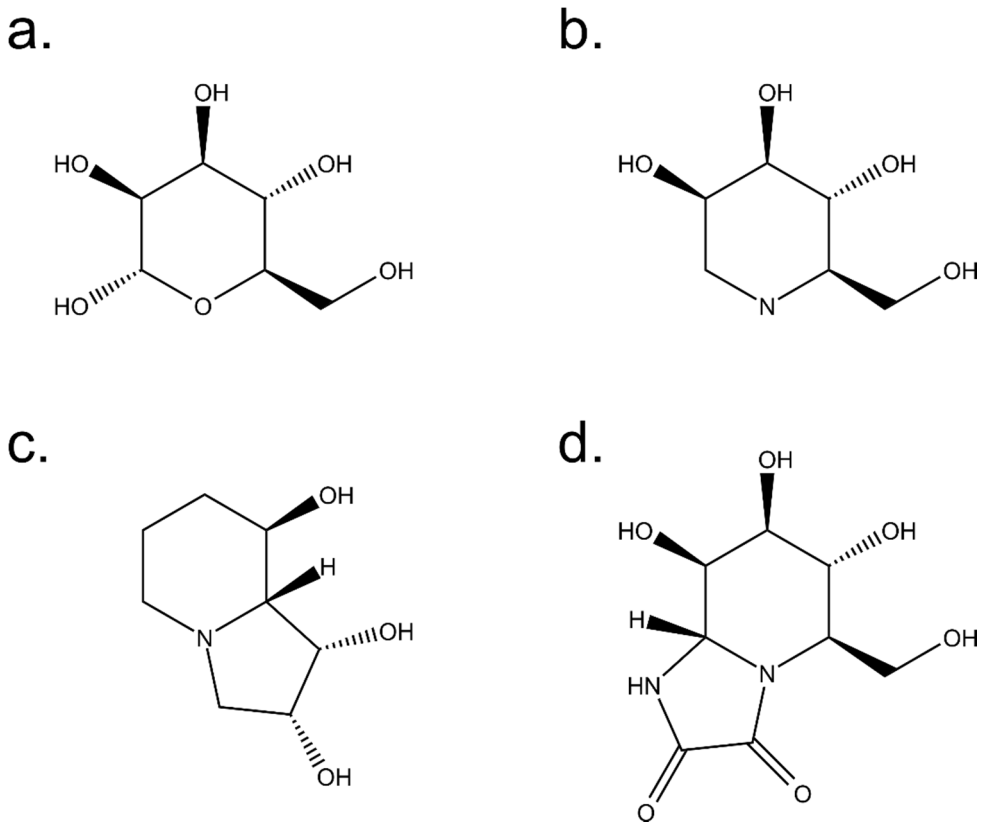



**Figure 1-7** D-mannose (a) and the mannosidase inhibitors DMJ (b), swainsonine (c) and kifunensine (d).

Kifunensine is a bicyclic iminosugar, isolated from bacteria, that has similar activity to DMJ but is between 50-100 times more potent [147]. Kifunensine was initially reported to only inhibit Golgi ManI enzymes and the fact that only $Man_{8-9}GlcNAc_2$ was produced was attributed to a lack of ER ManI activity in the cell line used [147]. However, there are a number of ER mannosidases, not all of which have been characterised, and it appears that there are both kifunensine resistant and susceptible ER mannosidases [96,148]. Functionally, however, kifunensine is the strongest mannosidase inhibitor in general use and appears to restrict glycosylation to an almost complete $Man_9GlcNAc_2$ phenotype, with a small component of $Man_8GlcNAc_2$ isomers due to the activity of resistant ER ManII and endomannosidase [148-150]. Kifunensine has now become a popular reagent for structural biologists as its potent ability to prevent complex glycosylation can be used to make proteins 100% susceptible to endoglycosidase-H (EndoH, an enzyme specific for high-mannose glycans) thus allowing for easy removal of glycans prior to crystallisation [151]. Kifunensine is also able to inhibit

ERAD [96] by inhibiting the various mannosidases involved in targeting proteins to that pathway (see Section 1.3.2). The structures of all of these inhibitors are shown in Figure 1-7, while Figure 1-8 shows the enzymes they inhibit and, therefore, the glycans expected to be formed by their inhibition.

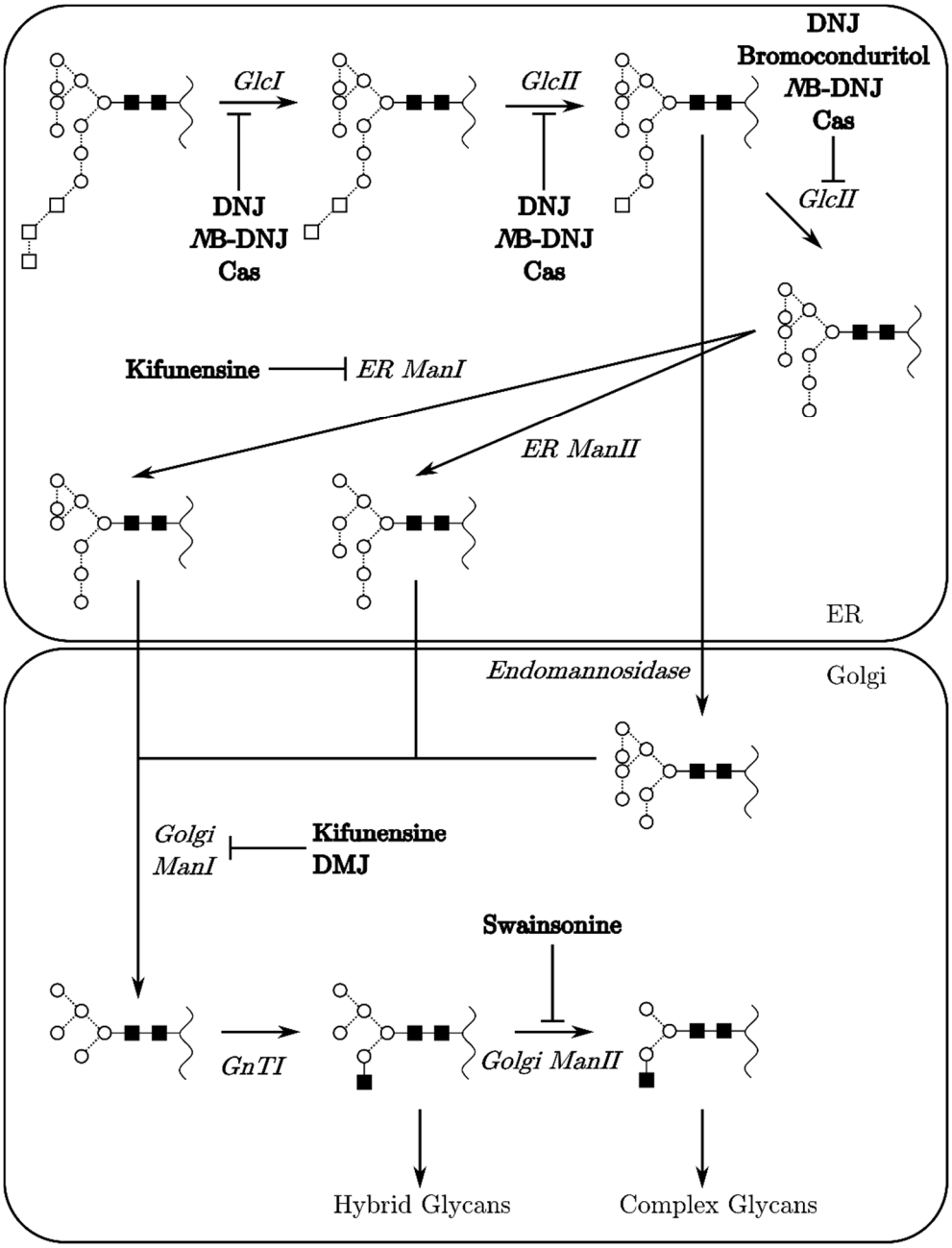



**Figure 1-8** Glycan trimming enzymes in the ER and Golgi and their inhibitors. All glycans are shown using notation described in *Harvey et al.* [80]. Enzymes are shown in *italics* and inhibitors in **bold**, with blunt ending arrows indicating the step inhibited. For the sake of clarity the calnexin cycle has been omitted, "Golgi ManI" is used to describe three enzymes, Golgi ManI A-C, which are shown only in the Golgi despite also being known to cycle to and from the ER. For full descriptions see text. Abbreviations not described previously: GnTI, *N*-acetyl glucosamine transferase I; *N*B-DNJ, *N*-butyl deoxynojirimycin; cas, castanospermine.

### 1.5.2 Glucosidase inhibitors

The most studied ER glucosidase inhibitors are castanospermine (Cas), DNJ and its more potent, alkylated derivative *N*-butyl DNJ (*N*B-DNJ) (Figure 1-9). These are all non-selective, inhibiting both ER α-glucosidase I and II, though $Glc_{1\text{-}2}Man_9GlcNAc_2$ species can be seen at lower doses suggesting a slight preference for GlcII over GlcI [13] (Figure 1-8). Specific inhibitors of each glucosidase have been developed- australine for GlcI and bromoconduritol and the snappily named 2,6-dideoxy-2,6,-imino-7-*O*-(β-D-glucopyranosyl)-D-glycerol-L-gulohepitol (MDL) for GlcII [152,153], however these compounds have not been as intimately studied as those listed above. As well as their activity against ER α-glucosidases, DNJ, *N*B-DNJ and castanospermine also have effects on a variety of intestinal digestive enzymes including sucrases and isomaltases [154,155]. Inhibition of intestinal enzymes is not seen with mannosidase inhibitors [155] presumably due to the relative unimportance of mannose as a dietary carbohydrate. *N*B-DNJ is also known to be an inhibitor of glucosyl-ceramide synthase (GCS), a glucosyltransferase, an effect that is not dependent on mimicking glucose stereochemistry [156,157].

The effects of α-glucosidase inhibitors are two-fold. Firstly, by preventing the production of the $Glc_1Man_9GlcNAc_2$ glycan they rob glycoproteins of their ligand for calnexin/calreticulin binding [83]. Secondly, the retained glucose residues prevent access of ManI enzymes to the D1 arm of the glycan, thus making them indirect mannosidase inhibitors as well, theoretically preventing the formation of complex glycans [13].

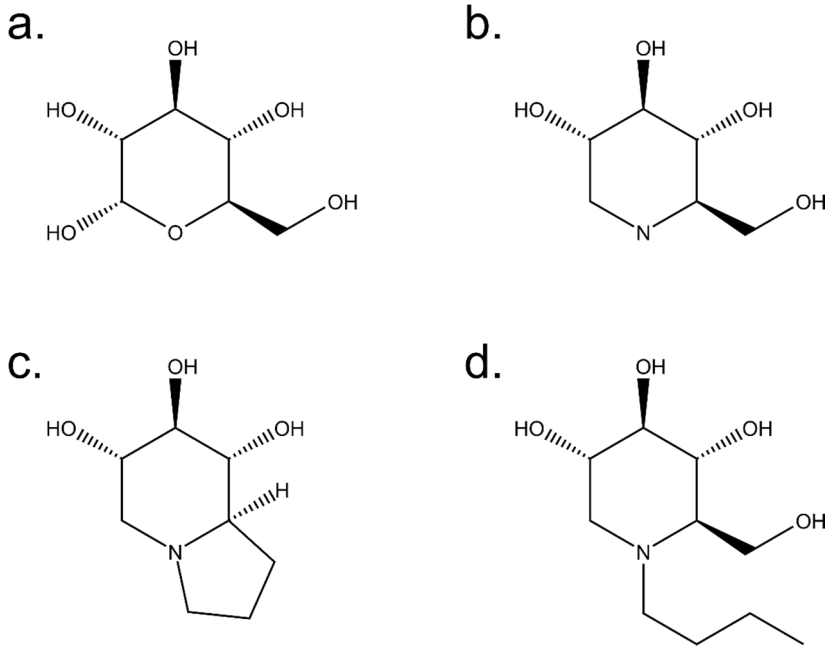



**Figure 1-9** D-glucose (a) and the glucosidase inhibitors DNJ (b), castanospermine (c) and *N*B-DNJ (d).

### 1.5.3 The Golgi endomannosidase

As α-glucosidases are the first enzymes in the glycan processing pathway it was thought for a long time that their inhibition should lead to total suppression of complex glycan synthesis. However, small proportions of complex and de-glucosylated high mannose glycans continued to be detected even in the presence of high concentrations of inhibitors [13,14]. This anomaly was eventually explained by the laboratory of Robert Spiro (no relation), who discovered an endomannosidase in the Golgi that was able to cleave the α1,2 glycosidic bond linking the two most distal mannose residues of the D1 arm [158]. This enzyme was able to cleave $Glc_1Man_9GlcNAc_2$ and, to a lesser extent, $Glc_{2\text{-}3}Man_9GlcNAc_2$ to create $Man_8GlcNAc_2$, thus allowing for continued processing even in the presence of glucosidase inhibitors [159,160]. However, as this salvage pathway occurs only in the Golgi, and can only generate $Man_8GlcNAc_2$, it cannot create the $Glc_1Man_9GlcNAc_2$ species needed for calnexin interaction. Therefore, glucosidase inhibition can theoretically affect a protein's folding without affecting its glycosylation. However, effects on glycosylation do still occur, as the endomannosidase is far from 100% efficient, especially with a $Glc_3Man_9GlcNAc_2$ substrate [159]. Also, the activity of endomannosidase appears to be highly cell line dependent with some lines, such as Chinese hamster ovary (CHO) cells, having no detectable activity at all [161]. Endomannosidase is unique in the secretory pathway as the only endoglycosidase and mediates the only step in which more than one sugar is removed at a time; this uniqueness means that none of the inhibitors previously described can affect its activity [162].

## 1.6 Iminosugars as antivirals

Thus we have seen how iminosugars can have a variety of effects on both glycoprotein glycosylation and quality control (Table 1-1). As mentioned previously, research into the antiviral capabilities of iminosugars began in earnest in the 1980s, a decade before the discovery of the calnexin cycle [83] and two decades before the discovery of mannosidases' role in ERAD [96]. Therefore, early hypotheses were that complex glycans were potentially vital for the functioning of viral envelope glycoproteins and their inhibition by iminosugars may, therefore, be antiviral.

| Inhibitor | Target(s) | Effects on Quality Control | Predicted Glycosylation |
|---|---|---|---|
| DNJ<br>*N*B-DNJ<br>Cas | GlcI & GlcII | Calnexin/Calreticulin cycle bypass | Triglucosylated high mannose oligomers, potentially some hybrid and complex |
| Bromo-conduritol | GlcII (second reaction) | Calnexin/Calreticulin cycle arrest in endomannosidase$^{-/-}$ cells | Monoglucosylated high mannose oligomers, potentially some hybrid and complex |
| DMJ | Golgi ManI | Weak ERAD bypass | High mannose oligomers |
| Kifunensine | EDEM, ER & Golgi ManI | ERAD bypass | Majority $Man_9GlcNAc_2$, some $Man_8GlcNAc_2$ |
| Swainsonine | Golgi ManII | None | Hybrid and high mannose oligomers |

**Table 1-1** Summary of the effects of various iminosugars on the quality control and glycosylation pathways of treated cells.

### 1.6.1 Early experiments with iminosugars and viruses

Some of the earliest work on the antiviral effects of glycosidase inhibitors was done on influenza virus. Initially, bromoconduritol was shown to inhibit the production of complex glycans on the haemagglutinin (HA) of Fowl Plague Virus (High Pathogenicity Avian Influenza) with a 90-98% reduction plaque formation [153]. This work was followed up by investigations of other glycosylation inhibitors, including the GlcI and GlcII inhibitors castanospermine and *N*-methyl DNJ (*N*M-DNJ) and the complete glycosylation inhibitor tunicamycin. Tunicamycin was antiviral, appearing to inhibit complete HA cleavage, a vital step for viral infectivity, though all this proves is that glycosylation is important for folding and secretion of certain glycoproteins [163]. *N*M-DNJ and castanospermine were both able to suppress complex glycosylation but neither were antiviral-indeed, they seemed to increase plaque size and number suggesting that there was an increase in infectious particle production [164,165]. However, this positive effect was small and insufficient replicates were performed (or, at least, published) to demonstrate statistical significance.

This apparent paradox may be because bromoconduritol leads to a preponderance of $Glc_1Man_{7\text{-}9}GlcNAc_2$ glycosylation, implying that it is a specific inhibitor of the second trimming action of GlcII (Figure 1-8) [153]. As the calnexin cycle had not been elucidated at this point the authors could not have known that this activity was required to prevent glycoproteins becoming

trapped in association with calnexin. Thus, in retrospect, it seems possible that the decrease in infectious particle production was due to decreased secretion of HA. *N*M-DNJ and castanospermine inhibit GlcI and GlcII (although GlcII inhibition is not apparent in the face of GlcI inhibition) and thus bypass the calnexin cycle altogether, avoiding this trap. Indeed, DNJ was later shown to have no effect on HA secretion, incorporation into viral particles or on the infectivity of those particles [166]. Such a trapping effect would probably also make bromoconduritol toxic as the retention of proteins in the calnexin cycle could lead to ER stress; indeed the authors raise the possibility of toxicity in a later paper, though never say whether they actually tested for it.

These findings, that (non-specific) α-glucosidase inhibitors are not antiviral against influenza, are of some historical note as a decade later it was influenza HA that was used as a model by Hammond, Braakman and Helenius to elucidate the calnexin cycle [83]. Although these experiments showed that HA did bind calnexin they did not show that this was required for HA folding. Hence the conclusion appears to be that α-glucosidase inhibiting iminosugars are not antiviral against influenza viruses as HA neither has a requirement for complex glycosylation nor for calnexin association. However, recent, unpublished work by the biotech company United Therapeutics has shown antiviral effects of a novel iminosugar, *N*-methoxynonyl DNJ (M*O*N-DNJ) in a variety of human-pathogenic influenza strains (personal communication, United Therapeutics). Their work suggests that the effect may be strain dependent, potentially explaining why these earlier experiments did not detect any activity.

Following the influenza work, other groups began experimenting with glycosidase inhibitors in a range of viruses. DNJ and castanospermine were able to decrease the infectivity of the rhabdovirus, vesicular stomatitis virus (VSV), but only in one specific strain and this effect was abrogated if the virus was grown at 30°C. Production of the envelope glycoprotein, G, was inhibited by DNJ but a similar amount was still present on the cell surface. However, virion production and infectivity were still decreased, suggesting that the protein was no longer functional [167]. The same group followed up this study by looking at the effects of DNJ and castanospermine on the togavirus, Sindbis virus. Both compounds inhibited the growth of this virus, apparently by preventing the proteolytic cleavage of its envelope glycoprotein E2 from the precursor PE2. Crucially, the authors used *N*-acetyl glucosamine transferase I (GnTI) negative cells to express the virus, thus preventing complex glycan formation in both the treated and untreated groups, thus controlling for the "indirect mannosidase inhibitor" effects of the drugs.

The authors hypothesise that the antiviral effects are due to a conformational change in the protein preventing proteolytic cleavage [168].

Meanwhile, another group were looking at the effects of *N*M-DNJ and castanospermine on the coronavirus, mouse hepatitis virus A59, and showed that these inhibitors decreased the secretion of viral particles. They also seemed to affect the fusion ability of the envelope glycoprotein E2, as cells expressing it no longer formed syncytia. However, further investigation showed that this was not due to a decrease in the functionality of E2 but rather a long delay in the time it took for the glycoprotein to reach the cell surface, with large quantities ending up sequestered in unidentified intracellular compartments instead. Immunofluorescence suggested that these treated E2 proteins were also in an aberrant conformation, as they were no longer recognised by certain antibodies. They too controlled for the effect of glycosylation changes, this time using swainsonine and DMJ to show that blocking complex glycosylation alone was not antiviral [169].

These three studies established many of the roads of enquiry that were to become standard for iminosugar research over the coming decade and yet received almost no citations over that period. The reasons for this oversight are unclear, but it may be that they were overlooked as they were all based around relatively clinically insignificant viruses. Nonetheless, these were the first papers to suggest that glucosidase inhibitors may be antiviral by inducing a conformational change in envelope glycoproteins and that such a change was not simply due to the absence of complex glycans, and much of their work would be re-invented when iminosugar research moved to looking at a very clinically significant pathogen, HIV.

### 1.6.2 Glucosidase inhibitors and HIV

In contrast to these elegant studies on Sindbis and mouse hepatitis virus, the earliest experiments on HIV revert to the hypothesis that it is the sugars themselves which are key ligands in HIV envelope mediated fusion and make little attempt to divine a mechanism. Both *Gruters et al.* and *Tyms et al.* examined the effects of DNJ and castanospermine on syncytium formation in HIV infected cells and the production of virions *in vitro* [15,16]. *Gruters* found that neither compound had any effect on the production of HIV virions or on envelope cleavage, but did find decreased infectivity of said virions, as determined by decreased tissue culture infectious dose and syncytium formation [15]. *Tyms* showed similar results but also claimed that the inhibitors caused a decrease in virion secretion, something *Gruters* had denied [16]. However, as *Tyms* ran their experiment for longer (up to 11 days) it is likely that the decreased infectivity of the secreted virus led to a slower spread of the infection through the cells, thus resulting in fewer infected cells to produce virus, hence the

discrepancy. Indeed, the difference in viral secretion between treated and untreated cells is far greater at day 11 than at day 3.

*Walker et al.*, working solely with castanospermine and independently of the above, showed similar effects but made more effort to determine the mechanism. They showed that although castanospermine reduced the secretion of gp120/gp41 to the cell surface it did not alter the concentration of envelope in the actual virion, thus showing that this could not be the mechanism of action [18]. Although not cited, this is significant as the only previous mechanism of action study of glucosidase inhibitors in a retrovirus had shown the opposite: *Pinter et al.* had shown that DNJ inhibited the proteolytic cleavage and virion incorporation of the envelope glycoprotein of the Friend Murine Leukaemia Virus [170]. *Walker* was also the first to actually show that glucosidase inhibitors did, in fact, affect gp120 and gp41 glycosylation rather than just assuming so. It was also significant in that it showed that treated virions were still able to bind to $CD4^+$ cells via specific, gp120 mediated interactions, thus proving that neither retained glucose residues nor the absence of complex glycans interfered with gp120/CD4 interactions, and that it was a later stage of the fusion process that was being inhibited [18]. *Montefiori et al.* also tested bromoconduritol, the GlcII second-step specific inhibitor, and found that it was not antiviral [171]. This may be because this inhibitor still allows production of the calnexin ligand $Glc_1Man_9GlcNAc_2$ species and endomannosidase could compensate to remove the last glucose. This hypothesis is supported by the fact that gp120 produced in the presence of bromoconduritol migrated at the same rate as untreated gp120, suggesting normal glycosylation.

*Fleet et al.* took a structural chemistry approach, testing the chemical space around DNJ and castanospermine to see what features could improve efficacy and decrease toxicity. *Fleet* found that an ability to inhibit α-glucosidase *in vitro* was a key predictor of efficacy against HIV, and also that *N*-alkylation of DNJ could increase potency, especially *N*B-DNJ which was also less toxic than both DNJ and castanospermine [17]. They then went on to show that long term treatment of infected cell cultures with *N*B-DNJ could eventually lead to clearance of the virus due to cell turnover, thus raising the possibility that *N*B-DNJ could become a viable treatment for HIV [172].

### 1.6.3 Towards a mechanism of action

By the start of the 1990s it had become uncontroversial that α-glucosidase inhibitors were antiretroviral, with *N*B-DNJ becoming the lead candidate for clinical trials. As these trials approached research shifted to focus on the mechanism of action. In 1991 two papers were published that claimed that α-glucosidase inhibitors could change the conformation of the gp120 V3 loop.

*Fenouillet and Gluckman* showed that the anti-V3 monoclonal antibody (mAb) 110-4 bound to DNJ treated rgp120 with lower affinity than to untreated gp120 [173]. Importantly, they confirmed that binding was equally diminished even if both proteins were enzymatically deglycosylated and that binding was identical if the proteins were first denatured by binding to a nitrocellulose membrane. These two controls confirm that the differential binding is due to a change in the conformation of the protein and not merely due to a change in glycosylation. A similar finding, but with none of the necessary controls, was claimed by *Jones and Jacob* using the anti-V3 mAb ADP358, although their drop in affinity was a more modest 50% compared to *Fenouillet's* 90% [174].

*Ratner et al.* claimed that recombinant gp120 (rgp120) from *N*B-DNJ treated cells was less susceptible to thrombin cleavage at a site in the V3 loop and used this to claim that that region may have adopted an alternate conformation as a result of the treatment [175]. However, their data absolutely does not support this conclusion, as they clearly state in their own results section. They use radiolabelled rgp120 incubated with varying concentrations of thrombin then use SDS-PAGE (sodium dodecyl sulphate polyagarose gel electrophoresis) to look for the presence of both the native rgp120 and a 70 kDa cleavage fragment. It is true that the presence of the fragment is diminished on *N*B-DNJ treatment, but its density on the gel is somewhat variable on the untreated bands anyway, making it hard to say whether the changes are significant. More importantly, although the fragment is not visible the gp120 does disappear to the same extent whether treated or untreated, so either it is being cleaved and the fragment is invisible or there was simply less gp120 to be cleaved in the first place. The authors seem to agree with these conclusions as in their results section they say of this gel:

> *"No significant differences in the susceptibility to cleavage* in vitro *with thrombin in the presence or absence of [soluble] CD4 were noted in comparisons of envelope proteins from untreated and [NB-DNJ] treated cells"* [175]

This conclusion is based on figure 3b, the only figure which claims to be looking at cleavage. However, in the discussion section the authors claim that their data support the hypothesis of reduced V3 cleavage and cite figure 3a. Figure 3a looks at CD4 binding of rgp120 and does not claim to look at cleavage at all. Thus, while this paper claims to have elucidated the mechanism of action of *N*B-DNJ as a modifier of the conformation of the V3 loop, I place no value on this finding.

### 1.6.4 Current understanding of the mechanism of action

Our current understanding of the mechanism of action of *N*B-DNJ (and, by extension, all α-glucosidase inhibitors) as antiretroviral drugs was established in the mid-90s by Per Fischer and colleagues at the Oxford Glycobiology Institute in a set of three papers published in the *Journal of Virology*. The first of these confirmed previous results that *N*B-DNJ did not interfere with gp120/CD4 interactions but that it did decrease infectivity of virions at later point in the entry process [176].

By this point Hammond, Braakman and Helenius had published their seminal paper divining the calnexin cycle and the role that monoglucosylated glycans played in correct glycoprotein folding [83]. Now there was a clear hypothesis for how *N*B-DNJ might work- by inhibiting the formation of this monoglucosylated species it could prevent gp160 from associating with calnexin and its associated chaperones and foldases. Without these it was possible that the protein failed to reach its native conformation, instead reaching a conformation that, while perfectly able to bind CD4, was less able to initiate the complex fusion process. The V3 loop had already been shown to have different mAb reactivity- could it be that this was the site of misfolding?

To investigate this hypothesis Fischer carried out further antibody binding experiments using a panel of 40 mAbs against a variety of gp120 domains [177]. These experiments showed that mAbs directed against the C1 region of the protein showed increased affinity, while those against the V1/V2 loop and discontinuous epitopes had decreased affinity. The gp120 used was a recombinant form produced in CHO cells. As previously mentioned, CHO cells lack endomannosidase activity so the treated rgp120 had no salvage pathway by which to form complex glycans. Fischer, therefore, showed that sialidase treatment abrogated the differences seen in the C1 region, suggesting that this was purely a glycan mediated effect, probably due to the negative charge of sialic acid residues. Neither sialidase nor glucosidases could counter the effect on the V1/V2 loop, however, suggesting that this was a real effect on the folding of the protein.

Interestingly, no changes were observed in the binding of antibodies to the V3 loop, in contention with previous studies. Fischer addresses this discrepancy by finding that there are two preparations of the ADP358 antibody and that only the low affinity preparation shows any change. The reasons for this are unclear but as Fischer tested eight different V3 mAbs, compared to the one each of *Fenouillet* and *Jones*, and found no effect of *N*B-DNJ on their binding, I am inclined to favour his conclusions as the more rigorous. Another interesting finding of this paper is that treated rgp120 was *more* susceptible to cleavage by thrombin than the untreated, but instead of forming the two 70 and 50 kDa products (indicating one cleavage site) many small bands were

formed on SDS-PAGE. This actually tallies with the data (but not the conclusions) of *Ratner et al.* and supports the notion of an aberrant conformation for the protein, just not one that is localised to the V3 loop specifically [175,177].

The final paper of this trilogy goes on to ask what effect this misfolding might have on the function of the envelope protein [178]. Fischer shows that *N*B-DNJ is able to diminish the cleavage of gp160 into gp120/gp41 in the infected cells, but that cleaved envelope is selectively enriched in the produced virions. Nonetheless, modest decreases in total and cleaved envelope were observed in treated virions, although only at 2 mM *N*B-DNJ and not at the lower (but still antiviral) concentration of 0.5 mM, suggesting that this alone could not be the cause of the antiviral effect.

As described in Section 1.2.5, following CD4 binding the gp120/gp41 trimers undergo substantial conformational changes to first bind the co-receptor and then to insert the gp41 fusion peptide. Fischer measured the ability of treated gp120/gp41 to undergo these changes by using antibodies that bound either the induced co-receptor binding site or gp41. Following addition of soluble CD4 (sCD4), those binding the co-receptor site bound equally well whether the envelope was treated or not. However, exposure of gp41 epitopes was substantially reduced, as was exposure of V3 loop epitopes. Also reduced was shedding of gp120 from gp41 following sCD4 binding, though it is unlikely that this, of itself, is significant as shedding does not correlate with fusion [72]. However, it may be a reflection of underlying changes in gp120/gp41 that prevent both fusion and shedding.

### 1.6.5 Mannosidase inhibitors as antivirals

In many of the aforementioned studies, mannosidase inhibitors such as DMJ and swainsonine were also tested for their antiviral effects. Unlike the α-glucosidase inhibitors, their effects were much more equivocal. DMJ was reported as having no effect on HIV mediated syncytia formation [15,17,179] but two studies also reported that DMJ was considerably more potent than DNJ at decreasing virion infectivity [171,179]. Swainsonine was not found to be antiviral in any studies, suggesting that mannosidase I inhibition was key [171,179]. The reasons for this inhibition were never divined, although it was interesting to note that pre-treating MT-2 lymphoblastoid cells with DMJ for two days prior to infection rendered the virus resistant to this effect, suggesting that the cells could compensate for mannosidase inhibition [171]. These contrasting reports are hard to resolve, and later papers seem to ignore them, simply describing DMJ as not having antiviral properties [176,177].

As kifunensine and GnTI$^{-/-}$ cells became popular for the production of recombinant SOSIP trimers for structural biology, *Eggink et al.* tested whether HIV molecular clones produced in the GnTI$^{-/-}$ HEK (human embryonic kidney) 293S cells had altered infectivity, and found that they did

not [180]. This is not directly comparable as GnTI inhibition arrests glycans at $Man_5GlcNAc_2$ and higher, while ManI inhibition tends to favour higher levels of mannosylation as it is earlier in the demannosylation pathway. However, these results confirm that if DMJ does have anti-HIV activity it is not by inhibiting complex or hybrid glycan formation and must be due to ManI inhibition specifically.

## 1.7 Iminosugars as pharmaceuticals

Having established that *N*B-DNJ had both efficacy against HIV and low cytotoxicity pharmaceutical companies began to take an interest in developing it for clinical use. Iminosugars had first been given to people in the early 1980s, when volunteers had been given 50 mg *N*-hydroxy-ethyl DNJ (miglitol) orally to assess the role of glucosidase inhibition on post-prandial glucose and insulin levels [181]. In 1996, miglitol became the first iminosugar to be licensed for clinical use in humans, as an oral anti-diabetic therapy [182]. Given at doses of up to 300 mg per day, the main side effect of miglitol use is diarrhoea and flatulence associated with retained intestinal disaccharides [183]. Simultaneously to the development of miglitol, the pharmaceutical company Searle acquired the rights to *N*B-DNJ and began phase I and II clinical trials for the treatment of HIV infection.

### 1.7.1 Clinical trials of *N*B-DNJ

Searle's phase I and II clinical trials were carried out with both *N*B-DNJ and a per-butyrylated pro-drug version in both humans and rhesus macaques in the first half of the 1990s. These trials were not all published and many were published "out of order" (for example, the results of a phase I trial were not published until a year after the results of the phase II) so I will attempt to describe them in a logical, rather than chronological, order.

For the phase I trial, HIV positive patients in the advanced stages of AIDS were recruited into a dose escalation study to test the safety of oral *N*B-DNJ [184]. Anticipating that inhibition of intestinal glucosidases would cause an accumulation of complex sugars in the gut leading to flatulence and diarrhoea (as was reported with miglitol) patients were fed a special diet high in monosaccharides. Patients were given doses ranging from 8 to 64 mg/kg/day and observed for adverse effects as well as giving blood samples to assess the drug's pharmacodynamics. Doses up to 165 mg/kg were planned but these were cancelled due to poor enrolment. In fact, the trial was stopped prematurely as parallel (unpublished) trials in rats suggested that the drug may cause cataracts, though this was reportedly only a problem in one specific strain of rats.

23 out of 27 of the patients (not all data were reported as there were patients lost to follow up) suffered from diarrhoea and 19/27 from flatulence. These effects were reported even on the lowest dose of 8 mg/kg/day, where 6/7 and 5/7 of the patients had diarrhoea and flatulence, respectively. Other adverse effects included neutropaenia (7/27) and leukopaenia (4/27). The significance of these findings is hard to estimate as this was not a randomised, double-blind, placebo-controlled trial and it is hard to say whether these would have occurred anyway seeing as the patients were in the advanced stages of an immunodeficiency disease.

The elimination half-life of *N*B-DNJ was approximately seven hours and steady state levels of the drug were observed after 30 days of four time daily (qid) dosing. However, the steady state concentrations were not impressive, with 48 mg/kg/day (a total daily dose of 3.36 grams in a 70 kg person) achieving only an 8-9 μg/ml plasma concentration.

A phase II efficacy trial was also carried out, this time of *N*B-DNJ in combination with azidothymidine (AZT), a combination which had already been shown to be synergistic *in vitro* [175,185]. 118 adults with symptomatic AIDS were randomly assigned to receive either AZT or AZT plus *N*B-DNJ for up to 22 weeks. *N*B-DNJ was given as a 1000 mg dose every 8 hours (tid), equivalent to 43 mg/kg/day in a 70 kg person. The safety and pharmacokinetic data agreed well with the results of the phase I study, with mild-to-moderate diarrhoea being the major side-effect and the maximal steady state plasma concentration being 9.38 ± 1.35 μg/ml.

The study did not find any statistically significant differences between the two treatment groups initially, with CD4 counts and viral titres being similar between the two groups. If one solely looks at patients who had detectable serum p24 at the start of the study and observed the proportion of those patients who subsequently became p24 negative then there does appear to be a greater effect of *N*B-DNJ and AZT over AZT alone. However, I am reluctant to read too much into this as it is clearly a sub-group analysis and no reduction in viral load was seen when it was analysed by reverse transcriptase activity or quantitative culture. It seems likely that the main reason for this failure was the low serum concentrations of *N*B-DNJ that were achieved. 9 μg/ml corresponds to just 50 μM, a dose below that shown to be antiviral *in vitro*.

Realising that intestinal side-effects were limiting the dose that could be safely given to humans, Searle tried to produce a pro-drug form of *N*B-DNJ, termed SC-49483, in which the hydroxyl groups were modified with butyrate esters, which would inhibit the drug's effects in the gut but be cleaved off releasing the active metabolite following absorption. This drug was tested in macaques infected with the HIV model macaque Simian Immunodeficiency Virus 251 ($SIV_{mac}251$). These results were

never published but a conference abstract claims that antiviral serum concentrations were achieved without diarrhoea (dosing was up to 750 mg/kg/day via gastric catheter) and that the drug was at least as good, if not better, as AZT at reducing viral load and increasing the CD4 cell count [186]. It goes without saying that the failure of this study to be published, or for the drug to go much further in clinical trials, means that these grand claims should be treated with considerable scepticism.

A clinical trial of SC-49483 in humans was carried out, and apparently completed, according to the United States register of clinical trials (identifier code NCT00000791) and the title of another conference abstract [187], but no information is available about the results. According to people who were working on *N*B-DNJ in collaboration with Searle at the time the trial was discontinued as the butyrylation made the drug exceptionally unpalatable and failed to prevent diarrhoea (Raymond Dwek, Terry Butters, personal communication).

No further clinical trials were carried out by Searle or anyone else into the use of *N*B-DNJ (or any other α-glucosidase inhibitor) as an antiretroviral, but the drug was not dead. *Platt et al.* described how, as well as being an inhibitor of α-glucosidases, *N*B-DNJ was also an inhibitor of ceramide glucosyltransferase, a Golgi enzyme that was vital for the production of glucosphingolipids (GSLs) [157]. GSL accumulation is the pathogenic feature of a number of lysosomal storage diseases. In type 1 Gaucher Disease, an inherited condition, the enzyme glucosylceramidase, which is responsible for the catabolism of GSLs, is deficient, leading to the accumulation of GSLs in lysosomes. By inhibiting ceramide glucosyltransferase *N*B-DNJ could reduce the amount of GSL being produced and thus reduce the strain on the mutant glucosylceramidase, limiting lysosomal storage; so called substrate reduction therapy (SRT) [188,189].

*N*B-DNJ entered clinical trials for the treatment of type I Gaucher Disease. 29 patients were given doses of 100 mg tid (equivalent to just 4.3 mg/kg/day in a 70 kg adult), a dose far lower than even the lowest dose used in the HIV trial (8 mg/kg/day). After 4 weeks, steady state plasma concentrations were reached, peaking at 1.5 μg/ml (6.9 μM), with troughs at around 0.8 μg/ml (3.7 μM) [190]. Theses concentrations were adequate to inhibit the ceramide glucosyltransferase *in vitro*, and *N*B-DNJ went on to be licensed for the treatment of type I Gaucher in the UK, USA and Israel. The major side effect, even at this low dose, is still diarrhoea but this appears to resolve with time; other than this, the drug appears to be safe, having now been used in patients for over a decade [188].

### 1.7.2 Other iminosugars, other viruses

The clinical trials of *N*B-DNJ seemed to converge on two main conclusions. Firstly, that the drug was relatively safe, although the side effects were unpleasant. Secondly, the molecule was not adequately potent at obtainable plasma concentrations. *N*B-DNJ had been derived from DNJ on the basis that increased alkyl chain length improved efficacy and decreased cytotoxicity, so further alkyl derivatives of DNJ were investigated to see if more a more active compound could be derived. By this point, a large number of anti-HIV drugs had been licensed and so attention switched to the infectious hepatitises B and C, for which there were no specific antivirals at the time.

The nonyl derivative *N*-nonyl DNJ (*N*N-DNJ) showed 46x greater efficacy than *N*B-DNJ at reducing the cytopathic effect of bovine viral diarrhoea virus (BVDV), a model for hepatitis C virus (HCV) which was, at the time, unable to be cultured [191]. Although the antiviral mechanism of *N*B-DNJ in BVDV was subsequently found to be similar to that in HIV (α-glucosidase inhibition, calnexin cycle bypass, misfolded envelope proteins, decreased infectivity) [192] further experiments revealed that long alkyl chain iminosugars such as *N*N-DNJ (but not *N*B-DNJ) were also potential inhibitors of an ion channel present in BVDV, potentially explaining the enhanced efficacy [193]. This result potentially explained why *N*N-DNJ was more efficacious than *N*B-DNJ *in cellula* as its ability to inhibit α-glucosidase *in vitro* is similar [194].

Meanwhile, tests on castanospermine against another flavivirus, dengue virus, showed that it was efficacious (through the classic calnexin based mechanism) both *in vitro* [195] and in an *in vivo* mouse model [196]. A butylated pro-drug of castanospermine, termed Celgosivir, entered phase Ib clinical trials for treatment of dengue fever but failed to show a decrease in fever duration or viral load [197]. Unlike the failure of *N*B-DNJ, the poor results in this trial can mostly be attributed to design, as most of the patients only received the drug late in the course of the (highly acute) disease when fever and viral load were expected to soon fall anyway. The antiviral effects *in vivo* were also hard to assess as the team only quantified viral concentration, not the infectivity of the virions, so it is possible that although viral load did not decrease the more significant infectious viral load may have.

*N*N-DNJ was also developed for the treatment of dengue, a virus that does not have any ion channels, but was superseded by its less toxic derivative, *N*-methoxynonyl DNJ (M*O*N-DNJ). M*O*N-DNJ is comparable to *N*B-DNJ in its anti-dengue effects when tested in human primary monocyte derived macrophages [198] but is disproportionately effective *in vivo*, rescuing mice from a lethal dengue challenge when *N*B-DNJ could not [199]. Pharmacokinetics for M*O*N-DNJ is only published in mice but suggests that a 200 mg/kg oral dose results in a peak plasma concentration

of 100 μg/ml which is rapidly eliminated to just 2 μg/ml within 4 hours; steady state data are not available. Pharmacokinetics of *N*B-DNJ were not performed in this study but previous studies of *N*B-DNJ in mice had shown that doses of 600 mg/kg/day (equivalent to the dose of M*O*N-DNJ taken by mice in this study) resulted in steady state plasma concentrations of 4 μg/ml [200].

These studies are not perfectly comparable, as one looks at steady state concentrations while another looks at only a single dose, *N*B-DNJ was given *ad lib* in food while M*O*N-DNJ was dosed by oral gavage, and they use different strains of mouse. However, they do seem to support the hypothesis that there is not a substantial difference in the pharmacokinetics of the two compounds. Therefore, pharmacokinetics cannot explain the enhanced availability or the enhanced effects of M*O*N-DNJ *in vivo*, raising the possibility that M*O*N-DNJ has as yet unelucidated additional mechanisms of action *in vivo*. M*O*N-DNJ is currently in phase I clinical trials in humans (identifier code NCT02061358).

### 1.7.3 Congenital disorder of glycosylation type IIb: a human model for α-glucosidase I inhibition

There have been just three reported cases of humans with congenital α-glucosidase I deficiency in the history of medicine, termed congenital disorder of glycosylation type IIb (CDG-IIb). The first of these died aged 74 days [201] but the other two are still alive, though severely mentally and physically disabled, and are now aged 11 and 6 years [202]. These cases, while tragic for the families affected, provide a rare opportunity to model what would happen if glucosidase inhibitors could reach maximal inhibitory concentrations in humans for long periods of time.

*Sadat et al.* investigated claims that the two living patients, despite having hypogammaglobulinaemia, had suffered from an unusually low incidence of viral infections over the course of their lives, and examined whether the GlcI deficiency was providing them with protection [202]. First they confirmed by western blot of cultured, Epstein-Barr Virus transformed B-cell lines that the children do indeed have no GlcI present, though they do have GlcII. The blot they publish appears to show increased GlcII over the control but without loading controls it is hard to say whether this is significant- if it is, it could suggest that GlcII is overexpressed to compensate for GlcI deficiency. They also perform mass-spectrographic analysis of the glycans from immunoglobulin G (IgG) purified from one of the patient's serum. This profile only has two labelled peaks but clearly shows an (unquantified) decrease in complex glycans in favour of triglucosylated high mannose oligomers. This suggests that endomannosidase activity, while present, is not able to

completely bypass GlcI blockade, an unsurprising result as it is considerably less able to cleave triglucosylated than monoglucosylated glycans.

The children had received the normal range of childhood vaccines prescribed in the United States, a regime that includes both recombinant antigens and live, attenuated viruses. The children showed normal responses to the recombinant antigens but developed no antibodies to the live viruses, including measles, mumps, rubella and varicella-zoster. As these inoculations require a productive infection in order to work, this further contributed to the hypothesis that the children had an innate resistance to viral infections.

When cells cultured from these children were challenged with HIV and influenza virus they were equally susceptible as control cells. However, virions produced in those cells were less infectious than those produced when the GlcI gene was complemented back in. No effect was seen on the unenveloped viruses adenovirus type 5 and poliovirus 1, nor was any effect seen in vaccinia virus, an enveloped virus that does not use glycoproteins for cell fusion. These findings led the authors to conclude that α-glucosidase inhibitors may have a role in the treatment of viral infections.

Sadly, the authors appear to be ignorant of the significance of α-glucosidase in protein folding, with no mention of the word calnexin in the entire paper, resulting in a paper that is largely descriptive without using this opportunity to discuss the implications to the wider field. Nonetheless, armed with this knowledge this paper does address some unanswered questions. It has always seemed suspicious to me that ER α-glucosidase activity should be so vital for the proper functioning of enveloped viruses and yet so disposable for the correct functioning of the cell. If this were the case one would expect glucosidase deficient mutants to have a massive competitive advantage over their kin causing the gene to quickly be eliminated from human populations by natural selection.

The severely detrimental phenotype of CDG-IIb explains why this might be the case- all three patients were born with severe abnormalities suggesting that while this enzyme may not be so important post-natally it is vital for the correct development of the foetus. It is likely that many more cases of this disorder have occurred in history but have either aborted or quickly died as infants before diagnosis could be established. Interestingly, the parents of the first child both are heterozygous for the mutant allele and hence have only 50% of the normal GlcI activity [201]. It would be fascinating to investigate the distribution of this mutation in the general population and correlate this to resistance to viral infections, though this is beyond the scope of this thesis.

Another interesting observation is that the HIV produced in the GlcI deficient cells is still infectious and, at most, complementation of the enzyme only doubles this infectivity. Thus it seems

likely that even if iminosugars could be given at doses so massive that α-glucosidase inhibition was 100% infectious virus would still be produced. This suggests that the calnexin cycle is helpful, but not vital, for correct glycoprotein folding. Finally, the existence of these children does provide evidence that α-glucosidase inhibiting drugs are likely to be safe in non-gravid adults though extra care must be taken if they are ever licensed for females planning to become pregnant.

## 1.8 Scope of this thesis

Since the failure of *N*B-DNJ as an anti-HIV drug iminosugar research has focussed mainly on making new, more active derivatives (such as the long alkyl chain iminosugars) and finding new applications (HCV, dengue virus and others). However, despite the continued use of these molecules in substrate reduction therapy and as experimental antivirals, no effort has been made to answer one of the key questions that has hung over iminosugar research ever since the calnexin/misfolding theory was formulated. That question asks why α-glucosidase inhibition is so damaging to the function of such a wide variety of viral envelope glycoproteins while having no toxicity against cellular proteins. Chapter 3 and 4 of this thesis will seek to answer this question and will provide evidence that the oligomeric nature of viral glycoproteins serves to amplify a very small effect on protein conformation. These chapters will also serve to further clarify the nature of *N*B-DNJ induced misfolding on gp120 and will try to explain why calnexin is required for correct envelope function.

Another unresolved question is the antiviral activity of mannosidase inhibitors, which will be examined in Chapter 5 using kifunensine, a far stronger inhibitor than DMJ and one known to inhibit ERAD as well. This study will investigate whether kifunensine is indeed antiviral and whether it is through either the envelope glycan or ERAD mechanism. Finally, Chapter 6 will examine whether the enhanced *in vivo* activity of M*O*N-DNJ against dengue virus also applies to another acute febrile disease, Ebola haemorrhagic fever, using a guinea pig model.

# 2 Materials and Methods

## 2.1 Iminosugars and chemicals

*N*B-DNJ was a gift from Oxford Glycosciences (Oxford, UK) and M*O*N-DNJ (UV-4) and its HCl salt (UV-4B) were gifts from United Therapeutics (Silver Springs, MD). Kifunensine, swainsonine, 1-deoxymannojirimycin, methionine sulfoxime (MSX), indinavir, DEAE-dextran and eeyarestatin I (EeyI) were from Sigma-Aldrich (St Louis, MO). Polyethyenimine MAX (PEI-MAX) with a molecular weight of 40,000 kDa was from Polysciences Inc. (Warrington, PA). All chemicals were dissolved in phosphate buffered saline (PBS), except EeyI which was solubilised in dimethyl sulfoxide (DMSO), and sterilised by passing through a 0.2 μm filter (Millipore, Cork, Ireland). When concentrations of UV-4 greater than 25 mM were required hydrochloric acid was used to lower the pH of the solution, improving solubility; UV-4B does not have this requirement. Where cytotoxicity was unknown chemicals were tested for their ability to inhibit cell growth using the CellTiter 96 Aqueous One Assay (Promega, Madison, WI) according to the manufacturer's protocol and by visual inspection of treated cells.

## 2.2 Expression of recombinant gp120

Recombinant gp120 (rgp120) from the HIV-1 strains IIIB, HXB2, LAI and BaL were expressed in mammalian cell lines. Cells were maintained at 37 ºC in 5% $CO_2$ unless otherwise stated.

### 2.2.1 Expression of rgp120$_{IIIB}$

CHO (Chinese hamster ovary) cells stably transfected with pEE6HCMVgp120GS, thus constitutively expressing gp120 from HIV$_{LAI}$ IIIB (rgp120$_{IIIB}$), were obtained from P Stevens (MRC AIDS Directed Programme Reagent Project, UK). These cells were maintained in Minimal Essential Medium (MEM, General Electric, Fairfield, CT) supplemented with 10% foetal calf serum (FCS), 100 U/ml penicillin, 100 μg/ml streptomycin and 2mM L-glutamine ("MEM-10"). When cells were at roughly 30% confluency rgp120$_{IIIB}$ production was induced by removing L-glutamine supplementation and adding 200 μM MSX, as described in Davis *et al.* [203]. Cells were incubated for 2 days after which the supernatant was removed, clarified by 5 min centrifugation at 4000 rpm and either used immediately or stored at -20 ºC for short periods.

### 2.2.2 Expression of rgp120$_{BaL}$ and rgp120$_{HXB2}$

rgp120$_{BaL}$ and rgp120$_{HXB2}$ were produced by transient transfection of HEK 293T cells. Both proteins were expressed from the vector pHLsec [204], gifts from Chris Scanlan. These plasmids were transformed into Zymo-5α cells according to the manufacturer's instructions (Zymo Research Corporation, Irvine, CA). Cultures were grown overnight in Luria-Bertani medium (LB,) supplemented with 0.1 mg/ml carbenicillin (both from Melford, Ipswich, UK) after which plasmid DNA was purified using the Qiagen Maxi-Prep kit according to the manufacturer's protocol (Qiagen, Venlo, The Netherlands) and quantified using a Nanodrop spectrophotometer (Thermo Scientific, Walford, MA).

To produce rgp120$_{BaL}$ for glycan analysis, cells were grown in 225 cm$^2$ cell culture flasks until approximately 80% confluent. Immediately prior to transfection 78 μg of plasmid was incubated with 273 μg PEI-MAX, molecular weight 40,000 kDa, (Polysciences Incorporated, Warrington, PA) in 15ml of serum-free DMEM for 15 minutes on a roller. The medium was removed from the HEK 293T cultures and replaced with the DNA/PEI-MAX solution plus a further 45 ml serum-free DMEM. Expression was allowed to continue for 5-7 days whereupon the supernatant was removed, centrifuged at 4000 rpm for 5 minutes to remove particulates and filtered through a 2 μm filter and degassed under vacuum.

rgp120$_{BaL}$ was then purified from the supernatant by Ni-NTA affinity. Briefly, HisPur Ni-NTA resin (Thermo Scientific, Walford, MA) was equilibrated in equilibration buffer; 20mM imidazole in phosphate buffered saline (PBS) at a pH of 7.4. Supernatant was diluted 1 in 2 in equilibration buffer and passed over the resin in a gravity column at roughly 1 ml/min until the whole volume had passed twice. The resin was washed in equilibration buffer and the protein then eluted with an elution buffer comprising 500 mM imidazole/PBS at pH 7.4, both times ceasing wash or elution phases when no further protein could be detected in the flowthrough by Nanodrop.

The protein was concentrated using an Amicon UIltra 30 kDa centrifugal filter (Millipore, Cork, Ireland) and monomeric rgp120 was separated from oligomeric forms using an Akta size exclusion chromatography machine (General Electric, Fairfield, CT). A Superdex 200 10/300 GL column (General Electric, Fairfield, CT) was equilibrated in PBS at pH 7.4. Subsequently the rgp120 solution was applied to the column then eluted in PBS at 0.5 ml/min. The monomer could then be identified by its elution profile, its concentration normalised by Nanodrop, aliquoted and stored at -20 ºC.

rgp120$_{HXB2}$ was expressed similarly but in 6-well plates (Corning, Corning[†] , NY). Plates were seeded with 2 x $10^5$ HEK 293T cells/well in DMEM-10 and incubated overnight. The following day the medium was removed and replaced with serum free medium containing 3.33 µg of the pHLsec$_{HXB2}$ plasmid preincubated for 15 minutes with 11.53 µg PEI-MAX, all per well. This is the same ratio (1 µg DNA to 3.5 µg PEI-MAX) as used above and as recommended in *Kirschner et al.* [205] just scaled for the growing area. The cells were incubated for two days in a total volume of 2 ml/well after which the medium was removed, clarified by 5 min centrifugation at 4000 rpm and either used immediately or stored at -20 ºC for short periods.

### 2.2.3 Expression of BG505 SOSIP.664

BG505 SOSIP.664 gp140 trimers were expressed in HEK 293T cells as described in *Sanders et al.* [60]. Briefly, the BG505 SOSIP.664 and furin expression plasmids were co-transfected into 293T cells in a 3:1 ratio using PEI-MAX as described above. The co-expression of furin ensures efficient gp140 cleavage. After 72 hours the supernatant was harvested and the relative concentrations of BG505 SOSIP.664 were determined by SDS-PAGE (sodium dodecyl sulphate polyagarose gel electrophoresis) with Coomassie blue staining and be blue-native PAGE [60].

### 2.2.4 Expression of rgp120$_{LAI}$

rgp120$_{LAI}$ was expressed in HeLa cells for pulse-chase experiments. HeLa cells were maintained in MEM-10 until needed whereupon they were split 1/6 into 3.5 cm$^2$ tissue culture plates and incubated overnight. rgp120$_{LAI}$ was expressed from the pMQ plasmid: 2 µg of plasmid was incubated with 250 µl serum-free MEM at room temperature for 10 minutes after which 5 µg PEI-MAX was added. The tube was then vortexed and incubated at room temperature for a further 15 minutes, after which another 1 ml serum-free MEM was added. The cells were washed with Hank's Buffered Saline Solution (HBSS) whereupon 1.25 ml of the transfection mixture was added. The cells were incubated with this mixture for 4 hours after which the medium was aspirated, the cells washed with HBSS and 1 ml MEM-10 added instead.

[†] Sic- Corning Inc. is named for city in which it was founded.

## 2.3 Glycan analyses

### 2.3.1 Analysis of gp120 glycans

$gp120_{BaL}$ was purified as described above and 10 μg run into a NuPAGE 4-12% Bis-Tris protein gel (Life Technologies, Carlsbad, CA). The gp120 band was visualised with Coomassie stain and then destained with a buffer of 63% water, 30% methanol and 7% acetic acid. The band was then cut out of the gel and washed three times with 200 mM ammonium bicarbonate with 40% acetonitrile. The gel piece was then washed with 100% acetonitrile and dried using a vacuum centrifuge. Glycans were then removed from the protein by overnight PNGase F (New England BioLabs, Ipswich, MA) digestion, as per manufacturer's protocol, followed by extraction from the gel in double-distilled water ($ddH_2O$). The glycans were labelled with the fluorescent marker 30 mg/ml 2-aminobenzoic acid (2-AA) in a buffer of 4% w/v sodium acetate trihydrate and 2% w/v boric acid in methanol, along with 45 mg/ml sodium cyanoborohydride, at 80 °C for 1 hour.

Glycans were then purified using solid phase extraction Spe-ed 2 cartridges (Applied Separations, Allentown, PA) in 97% acetonitrile. Cartridges were washed with 95% acetonitrile and glycans extracted with $ddH_2O$. Samples were then diluted in acetonitrile to get a final acetonitrile concentration of 65% and then analysed normal-phase high performance liquid chromatography (NP-HPLC). When required, endomannosidase digests were performed as described in *Kukushkin et al.* [206].

### 2.3.2 Quantification of Free Oligosaccharides (FOS)

FOS from cultured cells and guinea pig liver tissue were quantified using the methods described in *Alonzi et al. 2008* [207]. Briefly, cells/tissue were lysed using cycles of freeze-thawing in double distilled water and concentrations of protein were normalised by Bradford assay. The samples were then de-salted and de-proteinated using ion-exchange chromatography and FOS were extracted using porous graphitised carbon chromatography. The FOS were labelled with 2-AA and further purified using concavalin-A (ConA) sepharose chromatography. The pure, labelled FOS were then analysed using NP-HPLC.

### 2.3.3 Normal-Phase High Performance Liquid Chromatography (NP-HPLC)

NP-HPLC was performed using an Alliance 2695 separation module with 2475 detector and an X-bridge Amide ethylene bridged hybrid column (130 Å, 3.5 μm, 250 x 4.6 mm) (all Waters, Milford,

MA). A linear gradient between 35% 50 mM ammonium formate (pH 4.4) and 65% acetonitrile transitioning between the following conditions:

| Time (minutes) | Ammonium formate (%) | Acetonitrile (%) |
|---|---|---|
| 0 | 35 | 65 |
| 6 | 35 | 65 |
| 46 | 46 | 54 |
| 48 | 80 | 20 |
| 50 | 80 | 20 |
| 52 | 35 | 65 |
| 75 | 35 | 65 |

## 2.4 Enzyme Linked Immunosorbent Assays (ELISAs)

### 2.4.1 rgp120 and BG505 SOSIP.664 ELISAs

ELISAs were used to assess the affinity of various monoclonal antibodies (detailed in Table 2-1) to the rgp120 proteins. Half area 96-well plates were coated overnight with D7324 (Aalto Bioreagents, Dublin, Ireland) at a concentration of 5 μg/ml in coating buffer (15mM $Na_2CO_3$, 35mM $NaHCO_3$, pH 9.2) at room temperature. The plates were then blocked with 2% w/v bovine serum albumin (BSA, Sigma Aldrich, St Louis, MO) in Tris-buffered saline (TBS) for 30 min at room temperature, then washed in TBS. Supernatants were diluted 1/3 in TBS/10% FCS and added to the plate at 50 μl /well before being incubated overnight with gentle shaking. Dilutions of the primary antibodies were made in 2% BSA and incubated in the washed plates for 2 hrs at 37 ºC. After washing, the plates were then incubated for 1 hour at room temperature with one of the following horse radish peroxidase (HRP) linked secondary antibodies, all diluted in 2% BSA: 1/1000 ECL sheep anti-human IgG, 1/5000 ECL sheep anti-mouse IgG (both Amersham biosciences, Amersham, UK) or 1/1000 goat anti-rat IgG (Millipore, Cork, Ireland). After this incubation the plates were washed thoroughly in TBS/0.05% Tween 20 and then incubated with tetramethylbenzidine (TMB) ELISA substrate until a clear response could be seen with the naked eye, at which point the reaction was stopped with an equal volume of 0.5 M sulphuric acid. Paired readings were always matched such that incubation times were the same. After quenching, the absorbance of each well was immediately assessed at 450 nm and 570 nm on a SpectraMax M5 plate reader (Molecular Devices, Sunnyvale, CA).

The readings at 570nm were subtracted from those at 450nm to adjust for slight differences in volume between the wells and the results were plotted using GraphPad Prism v6.0 (GraphPad Software, San Diego, CA). Results were plotted as the $\log_{10}$ of the dilution factor (exact

concentrations were not known for most antibodies as they are supplied as supernatant fluid) versus the absorbance at 450-570nm, termed $OD_{450}$ (optical density at 450 nm).

ELISAs of BG505 SOSIP.664 material were carried out as described in *Sanders et al.* [60]. Primary antibodies were kindly provided by Rogier Sanders but were originally sourced from various laboratories as detailed in the reference.

### 2.4.2 p24 ELISA

p24 ELISA was used to quantify HIV capsid protein p24 in infected cell supernatants. Half-area 96-well ELISA plates were coated overnight at room temperature with 10 μg/ml sheep polyclonal anti-p24 (D7320, Aalto Bioreagents, Dublin, Ireland) in 100 mM $NaHCO_3$ coating buffer (pH 8.5). After this they were washed in TBS and 50 μl/well of Empigen treated supernatant (see Section 2.6.2), diluted appropriately in TBS/10% FCS was applied to each well. A standard curve was produced using recombinant bacterial p24 (AG6054, Aalto Bioreagents, Dublin, Ireland) diluted similarly.

After overnight incubation at room temperature the plates were washed in TBS and the primary antibody, biotinylated mouse monoclonal anti-p24 (BC 1071-BIOT, Aalto Bioreagents, Dublin, Ireland) was added at 1/1000 dilution in TBS/2% BSA/20% FCS and incubated for 2 hours at 37 ºC. The plates were washed then streptavidin-conjugated HRP was added at 1/10,000 dilution for 1 hour at room temperature. Finally, the plates were washed extensively in TBS/0.05% Tween 20 and then incubated with TMB ELISA substrate until a response could be seen in the highest dilution of the standard, at which point the reaction was stopped with an equal volume of 0.5 M sulphuric acid. Paired readings were always matched such that incubation times were the same. After quenching, the absorbance of each well was immediately assessed at 450 nm and 570 nm on a SpectraMax M5 plate reader (Molecular Devices, Sunnyvale, CA).

The 570 nm readings were subtracted from the 450 nm readings to adjust for small differences in volume between the wells. The linear range of the standard curve was identified and a curve fitted using linear regression. The equation of this curve was then used to interpolate the concentration of p24 in the samples from the absorbance readings. Multiple dilutions of the supernatant were used (typically 1/5, 1/10, 1/20 and 1/40) and the dilutions that gave absorbances outside of the linear range of the standards were extirpated from further analyses.

| Designation | Host | Isotype | Original designation | Target (where known) | Conf? | Form | Source |
|---|---|---|---|---|---|---|---|
| **ARP 3016** | human | IgG1κ | 1.5E | Discontinuous | Y | P | JR |
| **ARP 388** | rat | IgG2b | ICR38.1a | C3 | N | AF | JC & CD |
| **ARP3035** | rat | IgG2a | 11/4C | V1/V2 | N | SNF | JC & CD |
| **ARP3036** | rat | IgM | 8/64b | V3, CD4i | N | SNF | CS & CD |
| **ARP3037** | rat | IgG2a | 11/75a | V3 | Y | SNF | CS & CD |
| **ARP3038** | rat | IgG1 | 10/540.w | V3 | N | SNF | CS & CD |
| **ARP3039** | rat | IgG2a | 8/38 | V3 | N | SNF | CS & CD |
| **ARP3074** | rat | IgG1 | 66a | V2 | Y | SNF | JM & CS |
| **ARP3075** | rat | IgG1 | 62c | V2 | Y | SNF | JM & CS |
| **ARP3076** | rat | IgG2b | 11/65a | C1 | N | SNF | JM & CS |
| **ARP3077** | rat | IgG2a | 10/76b | V2 | N | SNF | JM & CS |
| **ARP3078** | human | IgG1 | 1.7B (PG20) | Discontinuous | Y | P | JR |
| **ARP3203** | rat | | 68/77a | V1/V2 & V3 | Y | SNF | JM & CS |
| **ARP3204** | rat | | 68/91 | Discontinuous, V1/V2/V3 independent | Y | SNF | JM & CS |
| **ARP3218** | human | IgG1λ | 697-D | V2 | Y | SNF | SZP |
| **ARP3220** | human | IgG1λ | 654-D | CD4bs, discontinuous, glycan dependent | Y | SNF | SZP |
| **ARP390** | rat | IgG2b | ICR39.13g | C2 & C3 | Y | AF | JC & CD |
| **ARP391** | rat | IgG2b | IRC39.3b | C2 & C3 | Y | AF | JC & CD |
| **EVA3206** | rat | | 63/10 | C1 | N | SNF | JM & CS |
| **EVA3207** | rat | | 63/69a,45a,84,91a | C1 | N | SNF | JM & CS |
| **EVA3208** | rat | | 63/49 | V1 | Y | SNF | JM & CS |
| **EVA3209** | rat | | 63/11b | V1/V2 | Y | SNF | JM & CS |
| **EVA3210** | rat | | 63/42 | C1 | Y | SNF | JM & CS |
| **EVA3211** | rat | | 63/45b | V2 | Y | SNF | JM & CS |

**Table 2-1** Anti gp120 monoclonal antibodies used in this study. All antibodies were obtained from the Centre for AIDS Reagents. For brevity, original references for each antibody have been listed only when specifically requested by the depositor. Blank cells indicate unknown or unavailable data. Abbreviations: Conf, conformation sensitive; Y, yes; N, no; SNF, supernatant fluid; AF, ascites fluid; P, purified, JR, J Robinson; JC, J Cordell; CD, C Dean; JM, J McKeating; SZP, S Zolla Pazner

## 2.5 Thrombin sensitivity assay

rgp120$_{BaL}$, produced in the presence of 20 μM kifunensine plus or minus 2 mM *N*B-DNJ, was dialysed into a citrate buffer (50 mM sodium citrate, 150 mM sodium chloride, pH 6.5) and then divided into 10 μg aliquots. Each aliquot was incubated overnight at 37 ºC in the presence of either 5, 2.5, 0.5 or 0 units of thrombin from human plasma (Sigma Aldrich, St Louis, MO). NuPAGE LDS sample buffer (Life Technologies, Carlsbad, CA) was added to the samples which were then divided into two aliquots, one of which had dithiothreitol (DTT) added to a final concentration of 50 mM. Both sets of samples were boiled at 95 ºC for 5 minutes then analysed by SDS-PAGE using a NuPAGE 4-12% Bis-Tris protein gel (Life Technologies, Carlsbad, CA) and Coomassie staining.

## 2.6 HIV molecular clones

The HIV molecular clone pLAI.2 (formerly pBRU2), which expresses the isolate HIV LAI [208], was obtained from the Centre for AIDS Reagents (CFAR, Potters Bar, UK). Mutant molecular clones of HIV LAI with altered V1/V2 loop structure were a kind gift from Ilja Bontjer and Rogier Sanders of the Academic Medical Centre, Amsterdam. The tested mutants were C119A/C205A, ΔV1/V2.2, ΔV1/V2.4.DNGSEK, ΔV1/V2.9.VE and ΔV1/V2.9.VK, the production and sequences of which can be found in the two studies of *Bontjer et al.* [209,210]. Glycan deletion mutants were derived from pLAI.2 as described below.

Unless stated otherwise, all plasmids were transformed into Zymo 5α cells and grown in LB as described above for pHLsec. Plasmids were purified from bacterial cultures using the GenElute HP MaxiPrep kit (Sigma Aldrich, St Louis, MO) according to the manufacturer's instructions.

### 2.6.1 Generation of glycan mutants

To investigate the role of glycans in HIV infectivity and *N*B-DNJ susceptibility a number of glycan sequons were deleted from the HIV LAI clone. Mutants were generated of each glycan site in the C1, V1/V2, C2 and V3 regions of gp120. Mutations were introduced to the pLAI.2 plasmid

using the QuikChange II XL site-directed mutagenesis kit (Agilent, Santa Clara, CA) using the following primers:

| Mutation | Primers | |
|---|---|---|
| N88D | Fwd | aaaattttctgtcacatctaccaatactacttcttgtgggttggg |
| | Rev | cccaacccacaagaagtagtattggtagatgtgacagaaaatttt |
| N136Q | Fwd | attactactattggtattagtagcctgccccaaatcagtgcactttaaact |
| | Rev | agtttaaagtgcactgatttggggcaggctactaataccaatagtagtaat |
| N141Q | Fwd | gctactactattggtattactactctgggtattagtagcattccccaaatc |
| | Rev | gatttggggaatgctactaatacccagagtagtaataccaatagtagtagc |
| N146Q | Fwd | catcatcatttccccgctactactctgggtattactactattggtattagta |
| | Rev | tactaataccaatagtagtaatacccagagtagtagcggggaaatgatgatg |
| N161D | Fwd | atattgaaagagcagtcttttatctctcctttctccatcatcatttcc |
| | Rev | ggaaatgatgatggagaaaggagagataaaagactgctctttcaatat |
| N165D | Fwd | cctcttatgcttgtgctgatatcgaaagagcagtttttatctct |
| | Rev | agagataaaaaactgctctttcgatatcagcacaagcataagagg |
| N191D | Fwd | cgtatagctggtagtatcatcatctattggtattatatcaagtttataaaaaaatgcat |
| | Rev | atgcatttttttataaacttgatataataccaatagatgatgatactaccagctatacg |
| N202D | Fwd | tgactgaggtgtcacaacttgtcaacgtatagctgg |
| | Rev | ccagctatacgttgacaagttgtgacacctcagtca |
| N234D | Fwd | ccattgaacgtcttatcattacattttagaatcgcaaaaccagcc |
| | Rev | ggctggttttgcgattctaaaatgtaatgataagacgttcaatgg |
| N239S | Fwd | acatggtcctgttccactgaacgtcttattattacattttagaatcgc |
| | Rev | gcgattctaaaatgtaataataagacgttcagtggaacaggaccatgt |
| N246S | Fwd | attgtactgtgctgacacttgtacatggtcctgttcc |
| | Rev | ggaacaggaccatgtacaagtgtcagcacagtacaat |
| N267D | Fwd | ttctgctagactgccatccaacagcagttgagttg |
| | Rev | caactcaactgctgttggatggcagtctagcagaa |
| N281D | Fwd | gttttagcattgtctgtgaaatcggcagatctaattactacctct |
| | Rev | agaggtagtaattagatctgccgatttcacagacaatgctaaaac |
| N294D | Fwd | tacaattaatttctacagattggtccagctgtactattatggttttagc |
| | Rev | gctaaaaccataatagtacagctggaccaatctgtagaaattaattgta |
| N300Q | Fwd | ttgttgggtcttgtacactgaatttctacagattggttcagctgtactattat |
| | Rev | ataatagtacagctgaaccaatctgtagaaattcagtgtacaagacccaacaa |
| N306Q | Fwd | ggatacggatacttttttcttgtattctggttgggtcttgtacaattaatttct |
| | Rev | agaaattaattgtacaagacccaaccagaatacaagaaaaagtatccgtatcc |

The polymerase chain reaction (PCR) was performed using the following mixture and conditions:

| Reagent | Quantity |
|---|---|
| 10x Reaction Buffer | 5 µl |
| pLAI.2 | 10 ng |
| Forward Primer | 125 ng |
| Reverse Primer | 125 ng |
| Deoxynucleotide triphosphate mix | 1 µl |
| QuikSolution | 3 µl |
| *PfuUltra* HF DNA Polymerase | 2.5 U |
| Double distilled water | To 50 µl total volume |

| Segment | Cycles | Temperature | Time |
|---|---|---|---|
| 1 | 1 | 95 ºC | 1 minute |
| 2 | 18 | 95 ºC | 50 seconds |
| | | 60 ºC | 50 seconds |
| | | 68 ºC | 14 minutes |
| 3 | 1 | 68 ºC | 7 minutes |

Following PCR the mixture was cooled to room temperature and 10 U of *DpnI* was added to each reaction, mixed and incubated for 1 hour at 37 ºC; this enzyme digests the methylated (template) DNA to prevent transformation of the unmutated plasmid. Meanwhile, XL10-Gold ultracompetent cells (Agilent, Santa Clara, CA) were thawed on ice and incubated with 2 µl β-mercaptoethanol per 45 µl cells for 10 minutes before 2 µl of the post-digestion PCR products were added to each 45 µl cell aliquot and incubated on ice for 30 minutes. The aliquots were then heat-shocked at 42 ºC for exactly 30 seconds before being returned to the ice bath. Cells were allowed to recover in 0.5 ml NZY+ broth for 1 hour at 37 ºC, with shaking, and then were plated on LB agar plates containing 100 µg/ml carbenicillin.

After 24 hours incubation at 37 ºC colonies were picked from each plate and grown in 3 ml cultures of LB overnight. 0.5 ml of each culture was used to prepare glycerol stocks (final volume 40% glycerol) and stored at -80 ºC. The plasmids were purified from the remaining volume using by the QiaPrep Spin Miniprep kit (Qiagen, Venlo, The Netherlands). These plasmids were then sequenced using the sequencing primer V1-3seq (aacaaatgctctccctggtcc) by Source BioScience (Nottingham, UK). If the sequencing showed that mutagenesis had failed alternative colonies were

picked and the plasmid purification/sequencing process repeated until a successful clone was identified. Successful mutants were grown from their glycerol stocks in large cultures (150 ml) and the plasmids purified by GenElute HP Maxiprep kit (Sigma Aldrich, St Louis, MO) according to manufacturer's instructions.

### 2.6.2 Expression of HIV molecular clones

2 x $10^5$ HEK 293T cells were seeded into each well of a 6-well tissue culture plate in 2 ml DMEM-10 and incubated overnight. The following day the transfection mixtures were made up. 5 μg plasmid DNA and 12.5 μl lipofectamine 2000 were pre-incubated with 0.25 ml reduced serum medium (both Thermo Scientific, Walford, MA) for 5 minutes at room temperature, per well. These were then mixed at a 1:1 ratio and incubated at room temperature with occasional shaking for 20 minutes.

The medium was removed from the cells, which were washed once in PBS, and replaced with 2 ml DMEM-10 and 0.5 ml of the transfection mix. Cells were then allowed to express the virus for 2 days, after which the supernatant was removed and centrifuged at 4000 rpm to remove cellular debris. The clarified supernatants were aliquoted with some aliquots being inactivated using a final volume of 0.1% Empigen (Millipore, Darmstadt, Germany) and incubation at 56 ºC for 30 minutes, before being frozen at -80 ºC or used immediately. This inactivation rendered the virus non-infectious allowing these aliquots to be removed from the biosafety level (BSL) 3 laboratory for ELISA. For infectivity analysis aliquots of supernatant were left "as is" and frozen at -80 ºC.

### 2.6.3 Quantification of virion infectivity

Infectivity of viral preparations was determined using the TZM-bl reporter cell line obtained from the Centre for AIDS Reagents courtesy of John Kappes and Tranzyme Incorporated. These cells were maintained in DMEM-10. For infectivity assays these cells were plated into white, opaque, 96-well tissue culture plates and incubated overnight in 200 μl DMEM-10/well. The following day 100 μl medium was removed from each well and a volume of supernatant equivalent to 500 pg p24 (as determined by p24 ELISA of the paired, inactivated aliquot) was added to the well and the volume brought back up to 200 μl. Also in the medium was 1 μg/ml indinavir sulphate and 20 μg/ml DEAE-dextran (both Sigma Aldrich, St Louis, MO). DEAE-dextran serves to enhance the sensitivity of the assay while indinavir is an HIV protease inhibitor. By preventing the

maturation of HIV particles produced in the infected TZM-bl cells indinavir inhibits multiple rounds of infection, thus ensuring that only the infectivity of the initial inoculum is being measured.

The cells were incubated with the virus for 2 days after which 100 µl of supernatant was removed and the cells allowed to equilibrate to room temperature. 100 µl of room-temperature Bright-Glo Luciferase Assay reagent (Promega, Madison, WI) was added to each well. Following a 5 minute incubation the luminescence of each well was recorded using a Novostar plate reader (BMG Labtech, Ortenburg, Germany). One set of wells per plate were left uninfected; these values are averaged and subtracted to control for background luminescence.

## 2.7 Analysis of oxidative folding

### 2.7.1 Radioactive pulse chase

To investigate the effects of *N*B-DNJ and kifunensine of the oxidative folding of rgp120 radioactive pulse-chase labelling was employed. HeLa cells were transfected with $rgp120_{LAI}$ as described above and incubated at 37 ºC overnight. Pulse-chase labelling includes a number of different stages, each of which requires the cells to be incubated in a different medium (abbreviations explained in text, below):

| Name | Volume per dish | Components | Temperature |
|---|---|---|---|
| Pre-treatment | 1 ml | MEM-10 | 37 ºC |
| Depletion | 1 ml | Cysteine and methionine deficient MEM with 1/100 HEPES | 37 ºC |
| Pulse | 0.3 ml | Depletion medium + 55 µCi radiolabelled cysteine and methionine | 37 ºC |
| Chase | 1 ml | MEM-10 with 5mM cysteine and 5mM methionine, 1/100 HEPES | 37 ºC |
| Stop | 1 ml | HBSS with 20 mM iodoacetamide | Ice cold |
| Lyse | 0.3 ml | MNT buffer with 0.5% Triton, 20 mM iodoacetamide and protease inhibitors (CLAP and PMSF) | Ice cold |

The pre-treatment phase is the period of 30 minutes in which the cells are incubated with the drugs under test (e.g. 2 mM *N*B-DNJ) in MEM-10. This period is designed to give the drugs adequate time to reach and inhibit their target enzymes before the experiment begins. The drug is then also included, at the same concentration, in all subsequent steps up until the "stop" phase.

Following the pre-treatment the cells are washed once in HBSS and the depletion medium is added. Depletion is a stage in which the cells are starved of methionine and cysteine. This causes them to use up their supplies so when small quantities are added in the pulse stage they will be readily incorporated into the proteome. From this stage onwards the cells are being handled regularly and so it becomes impractical to keep them in an incubator. Instead they are kept atop a 37 °C water bath such that the base of the plate is in contact with the water. HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) is included in the media to buffer against pH changes caused by the lack of $CO_2$ supplementation in these phases.

Labelling begins as soon as the pulse medium replaces the depletion medium (no wash in between); this phase continues for 10 minutes after which the pulse is aspirated and 1 ml chase solution added. This solution is immediately removed and replaced with another 1 ml chase. This serves to wash off the radioactive methionine and cysteine in the pulse medium but also, more importantly, the high concentration of unlabelled methionine and cysteine massively outcompetes the remaining labelled amino acids, preventing further incorporation into nascent proteins.

The cells sit in their chase medium for either 15, 30, 60, 120 or 240 minutes at which point oxidative folding is stopped by aspirating the chase medium and placing the cells on a sheet of ice cold aluminium on top of an ice bath (to ensure even cooling), whereupon 1 ml of stop medium is added. The stop medium contains iodoacetamide, an agent which alkylates any sulphydryl groups preventing the creation or isomerisation of disulphide bonds thus "freezing" the proteins in their current conformation. Cells are then lysed in the lysis buffer and the rgp120 immunoprecipitated for analysis.

### 2.7.2 Immunoprecipitation and PAGE

To create an anti-rgp120 resin 8 µl of rabbit polyclonal anti-rgp120$_{LAI}$ is incubated with 50 µl of protein A sepharose beads at 4 °C with gentle end-over-end agitation for 15 minutes. 200 µl of lysate is then added to the beads and incubated at 4 °C overnight. The following day the tubes are spun at 14000 rpm in a bench-top centrifuge for 1 minute to pellet out the beads. The supernatant is removed and replaced with 15 µl of 100 mM sodium acetate buffer (pH 5.4) with 0.2 % sodium dodecyl sulphate (SDS), whereupon the tubes are incubated at 95 °C for 5 minutes to dissociate the protein from the beads.

The tubes are then cooled to room temperature and 10 µl of 100 mM sodium acetate buffer (pH 5.4) containing 2% Triton and CLAP and PMSF protease inhibitors was added. High mannose

glycans are removed from the protein by adding 250 U EndoH and incubating at 37 ºC for 2 hours. The samples are heated to 95 ºC with 7.5 µl 5x loading buffer for 5 minutes; following this, a 19 µl aliquot is removed to another tube and incubated with DTT to a final concentration of 25 mM to create paired reduced and non-reduced samples. Both samples are cooled to room temperature and have iodoacetamide added to a final concentration of 50 mM. This last step is necessary because it is possible that iodoacetamide could have dissociated from thiols after boiling.

The samples were then separated by electrophoresis on a 7.5% polyacrylamide gel with a superior 5% stacking gel. This gel was stained with Coomassie Brilliant Blue to confirm separation then destained with an aqueous solution of 20% methanol and 20% acetic acid. The gels were subsequently dried and used to expose phosphorescent screens.

## 2.8 *In vivo* tests of iminosugars

Outbred Hartley guinea pigs were supplied by Charles River Laboratories (Wilmington, MA). These animals were supplied with cannulated right-jugular veins to allow for dosing directly into central veins. Guinea pig adapted Zaire ebolavirus (EBOV) was previously produced by passaging EBOV ME718 through guinea pigs for five generations and titred on Vero E6 cells, as described previously [211]. $10^3$ $TCID_{50}$ (tissue culture infectious dose, 50%) of this strain was determined to be 100% lethal to Hartley guinea pigs within 14 days of infection [212].

Infection was performed by subcutaneous injection while dosing was performed via the jugular cannula in a volume of exactly 0.5 ml. Euthanasia at the termination of the experiment was performed either by asphyxiation with a rising concentration of carbon dioxide or via intravenous or intraperitoneal injection of pentobarbital sodium, in accordance with Schedule 1 of the Welfare of Animals (Scientific Procedures) Act 1986. Animals were weighed daily and assessed for clinical signs and temperature twice daily. Temperature readings were performed using idENTICHIP with Bio-Thermo (AnimalCare, York, UK) subcutaneous implants.

# 3 The effects of alpha-glucosidase inhibition on the structure of HIV gp120

## 3.1 Contributions

The glycan profiling presented in this chapter was performed by Laura Pritchard using material prepared by the author. The SOSIP trimers were expressed by Ronald Derking.

## 3.2 Introduction

As previously described, there is some disagreement in the literature as to whether glucosidase inhibitors induce misfolding in the V3 or the V1/V2 regions of the recombinant gp120 (rgp120) protein [173,174,177]. *Papandreou et al.* tried to reconcile these discrepancies and showed that this may be due to differences in the cell line used to express the rgp120 [213]. Specifically, they showed that altered V3 immunoreactivity was detectable in rgp120 from baby hamster kidney (BHK) but not from Chinese hamster ovary (CHO) cells, though they did not look at V1/V2 reactivity. This difference is significant because BHK cells have nearly full endomannosidase activity while CHO cells have none. This suggests that it may be glycosylation differences that affect the immunoreactivity rather than calnexin dependent changes in the tertiary structure of the underlying protein.

This chapter will attempt to clarify the nature of the misfolding induced by glucosidase inhibitors. Kifunensine will be used to block ManI processing in the rgp120 producing cells thus ensuring that all glycans are of a high-mannose phenotype. This control will allow us to separate the effects of complex versus high-mannose glycans on the immunoreactivity of the various domains. Immunoreactivity will then be assessed in two separate cell lines to look for consistency.

All previous work was performed in monomeric rgp120 but in the virion these proteins are trimeric. This could allow for inter-protomer interactions that could amplify the regional misfolding seen in the monomers into something much more significant. For example, small misfolds at the trimer interface could dramatically change the trimer conformation such that it can no longer undergo the necessary conformational changes for fusion. To investigate this, this chapter will look at the effects of *N*B-DNJ on the immunoreactivity of BG505.664 SOSIP trimers, as well as its capability to undergo CD4 induced conformational changes. Finally, this chapter will use deletion derivatives of gp120 in live virions to try and further localise the misfolded region(s).

## 3.3 Results

### 3.3.1 Effects of *N*B-DNJ on the glycan profile of recombinant gp120

The effects of glycosidase inhibitors on a protein's glycans can be observed using glycan profiling, in which the glycans are enzymatically detached from the purified protein using PNGase F and then labelled with the fluorescent marker 2-AA. The labelled glycans can then be separated by mass using HPLC. The resultant peaks can be assigned to the glycans by using similarly labelled dextrose to create a standard curve of equivalent masses of glucose ("glucose units"). However, as sugars such as mannose are just epimers of glucose, and hence have equal mass, it can be difficult to distinguish, for example, $Glc_3Man_8GlcNAc_2$ from $Glc_2Man_9GlcNAc_2$. In these cases, enzymatic digestion of specific linkages and observation of the resulting changes in the glycan profile can help identify these peaks.

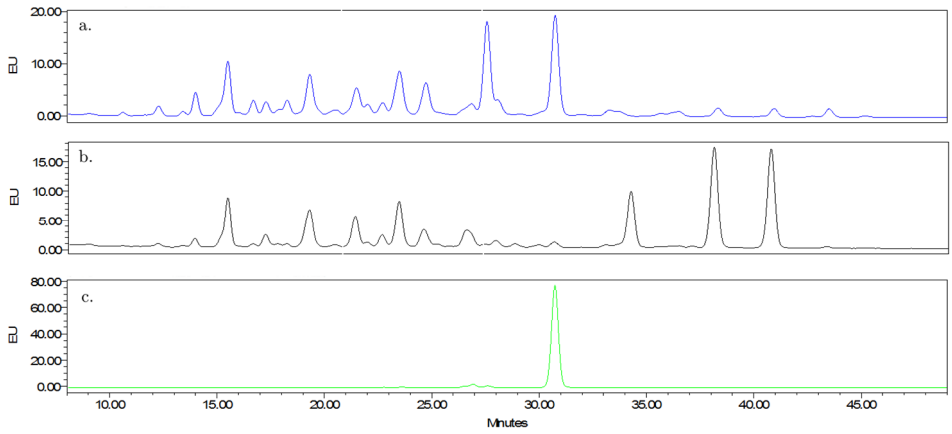



**Figure 3-1** HPLC glycan profiles of $gp120_{BaL}$ expressed in HEK293T cells (a), as well as in the presence of 2 mM NB-DNJ (b) or 20 μM kifunensine (c). Abbreviation: EU, emission units.

Figure 3-1 shows the glycan profiles of gp120 for the BaL strain of HIV produced in the presence and absence of 2 mM *N*B-DNJ or 20 μM kifunensine. Looking first at the untreated profile (a), we can assign the two tall peaks eluted at 26.5 and 30.5 minutes as being $Man_8GlcNAc_2$ ("$Man_8$") and $Man_9GlcNAc_2$ ("$Man_9$") respectively. The peaks eluting sooner than this are likely to be smaller mannosylated glycans, while those eluting later, such as the three small peaks at 38, 40.5 and 43.5 minutes, are most likely sialylated complex glycans. These assignments are based on comparison to the dextrose ladder and previous experience. Thus we can conclude that gp120 BaL is decorated with a mixture of complex and high-mannose glycans.

On treatment with *N*B-DNJ (b) three new peaks appear from 33 minutes onwards. Although they overlap with the complex glycans seen previously, they are most likely to be glucosylated high-

mannose glycans arising from the inhibition of the α-glucosidases. Although two of the complex peaks are obscured, the peak at 43.5 minutes is noticeably absent. This is consistent with the existing theory that retention of terminal glucose will prevent the access of exomannosidases to the α1,3 linked arm, thus preventing processing to complex glycans. Endomannosidases can bypass this block by cleaving the α1,2 linkage between the mannoses proximal to the glucose cap, thus removing the "terminal" mannose and any remaining glucoses. HEK 293T cells are not thought to have abnormal endomannosidase activity (unlike CHO cells) yet this enzyme does not appear to have had a significant effect, as the dominant phenotype is that of a mixture of glucosylated high-mannose glycans only. This effect of *N*B-DNJ could be a confounding factor in a study of the effects of the drug on protein structure, as the sialylated complex glycans have a significantly different structure and carry different charges to high mannose glycans. This substitution could, therefore, affect the binding of monoclonal antibodies and make it appear as though the underlying protein has folded differently when only the glycans have changed.

A possible solution to this is to use an exomannosidase inhibitor to control for the differences in glycosylation. Inhibiting exomannosidases has no effect on the removal of glucoses, and hence does not affect the calnexin cycle, but will prevent any further processing of glycans beyond $Man_9$ (with the possibility of generating $Man_8$ via endomannosidase). This effect can be seen in Figure 3-1 (c) where treatment with 20 μM kifunensine has resulted in a very large $Man_9$ peak at 30.5 minutes with only the merest hint of a $Man_8$ peak at 27.5 minutes, confirming the previous conclusion that the effect of endomannosidase is minimal. This may be because the dense packing of glycans on the gp120 surface make it difficult for endomannosidase to gain access.

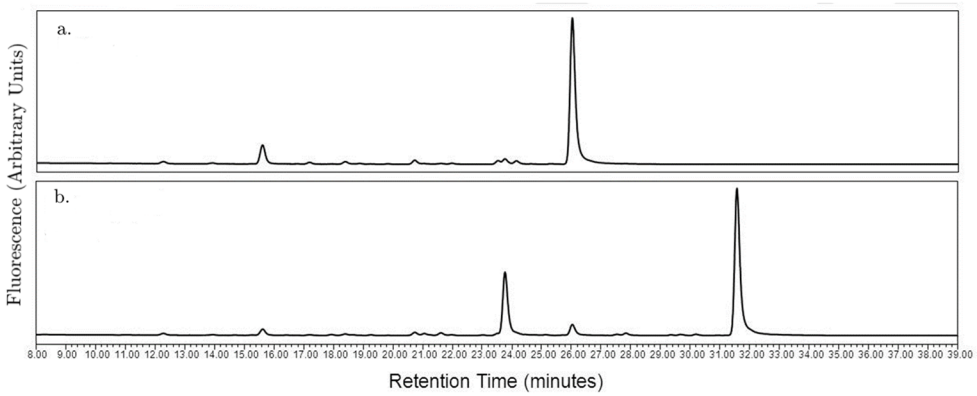



**Figure 3-2** HPLC glycan profiles of $gp120_{BaL}$ expressed in HEK 293T cells in the presence of 20 μM kifunensine (a) or 20 μM kifunensine and 2 mM *N*B-DNJ (b).

Thus one can predict co-treatment of *N*B-DNJ and kifunensine would result in almost exclusively $Glc_3Man_9GlcNAc_2$ glycans. Therefore one could compare kifunensine treated with

kifunensine and *N*B-DNJ treated gp120, knowing that while the glycans would be similar on both only the former will have gone through the calnexin cycle. If one tests this hypothesis (Figure 3-2 b) it can be seen that while the dominant peak in this instance is the $Glc_3Man_9GlcNAc_2$ at 31.5 minutes there is also a $Man_8$ at 23.5 minutes, which is absent in the kifunensine only treatment (a). This could suggest that retention of the glucose cap increases the activity of the endomannosidase. In an endomannosidase negative cell line, such as CHO cells, it is likely that only the $Man_9$ species (glucosylated or not) would be evolved. These observations are summarised in Table 3-1.

| Treatment | Calnexin Interaction? | Glycan Profile |
|---|---|---|
| Untreated | Yes | Mixture of high mannose and complex |
| *N*B-DNJ | No | Mixture of glucosylated and unglucosylated high mannose, potentially some complex |
| Kifunensine | Yes | $Man_9$ |
| *N*B-DNJ + Kifunensine | No | $Glc_3Man_9$ and $Man_8$ |

**Table 3-1** Summary of glycan profiles and likelihood of interacting with calnexin of $gp120_{BaL}$ expressed in HEK 293T cells and treated with glycosidase inhibitors.

### 3.3.2 *N*B-DNJ treatment induces altered conformation of the V2 loop of recombinant HIV $gp120_{IIIB}$

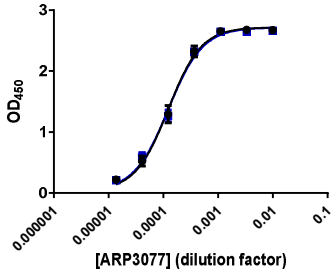
To investigate the effects of *N*B-DNJ treatment on the conformation of recombinant gp120 the binding of a panel of monoclonal antibodies (mAbs) was used to identify any regions whose affinity to the mAbs was altered by *N*B-DNJ treatment. Such an altered affinity could indicate that the epitope has changed, either conformationally or by becoming less accessible to the mAb. These two possibilities can be distinguished by using mAbs that target conformational and linear epitopes; only the former will be affected by a change in the conformation of the region while both will be affected by a change in its accessibility.

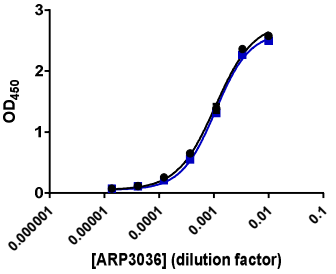
There exists a CHO cell line which constitutively expresses the recombinant gp120 (rgp120) from the IIIB strain of HIV-1 ($rgp120_{IIIB}$), a molecular clone of LAI, the original strain identified by Robert Gallo and a common laboratory strain, as a result of transfection with the pEE6HCMVgp120GS plasmid. Although CHO cells are known to be deficient in endomannosidase activity the convenience and reliability of constitutive expression caused me to select this system

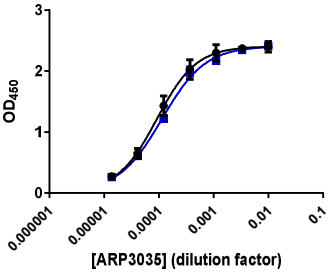
for the screening of mAbs. As the development of complex glycans is predicted to be impossible in an endomannosidase deficient cell line treated with glucosidase inhibitors, kifunensine will be used to control for *N*B-DNJ's effects, as set out above.

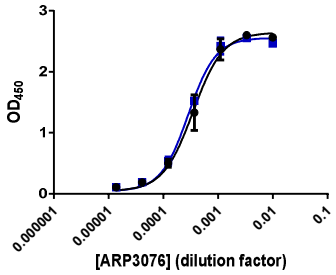
The transfected CHO cells were grown for two days in 20 μM kifunensine ± 2 mM *N*B-DNJ, after which the supernatant was collected and applied to Enzyme Linked Immunosorbent Assay (ELISA) plates coated with the antibody D7324, a sheep monoclonal which binds a linear epitope in the C5 region of $rgp120_{IIIB}$ and related strains. This antibody was selected as its epitope is at the *C*-terminus of the protein, near the gp41 cleavage site. This means that the proteins would be adsorbed to the plate in a similar orientation to that they possess when presented on a virion membrane, thus giving the best access to the mAbs, and because such a distal linear epitope would hopefully be unlikely to be affected by iminosugar induced misfolding. The mAbs could then be applied as the primary antibody and their binding detected with a suitable enzyme-linked secondary.

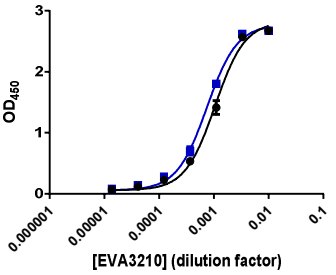
A potential pitfall of this method is that, by using unpurified supernatant, there could be differences in concentration of rgp120 between the two samples, which could confound ELISA readings of affinity. This problem can be overcome by using a range of dilutions of the primary mAb, thus generating a sigmoidal binding curve. The upper plateau of this curve ($E_{max}$) is proportional to the concentration of antigen in the ELISA well but the point half way between the $E_{max}$ and the background ($E_{min}$), known as the $EC_{50}$ (effective concentration, 50%; the concentration of antibody at which the binding is at 50% of its maximal) will be proportional to the affinity of the antibody. Thus a difference in concentration will be represented by a change in the $E_{max}$ while a change in affinity will affect the $EC_{50}$. A selection of antibody binding curves are presented in Figure 3-3. As can be seen from these curves, the majority converge on the same $E_{max.}$, suggesting that *N*B-DNJ does not affect the concentration of secreted rgp120. Some of the curves do not reach their $E_{max}$ due to the low affinity of the mAb in question; in these cases the curves have been interpolated with the constraint that they should arrive at the same $E_{max}$, since the supernatants under test are identical to those which do converge when tested with stronger mAbs. While one can immediately see that some curves overlap perfectly, e.g. those of ARP3077 or ARP3038, and some show a clear difference between the two curves, such as ARP3220 or ARP3074, it is important to assess whether these differences are statistically significant. This can be done using the $EC_{50}$.

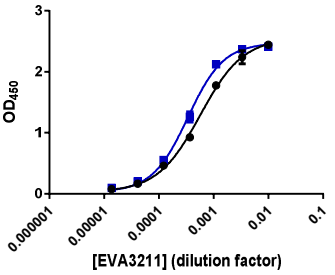

OD450
[ARP3077] (dilution factor)

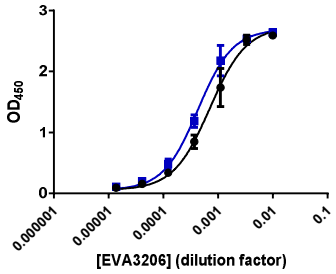

OD450
[ARP3036] (dilution factor)

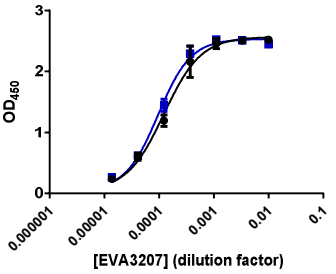

OD450
[ARP3035] (dilution factor)

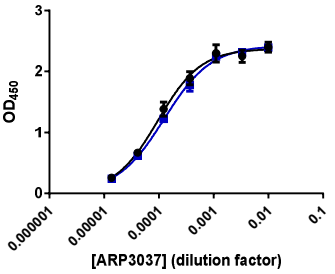

OD450
[ARP3076] (dilution factor)

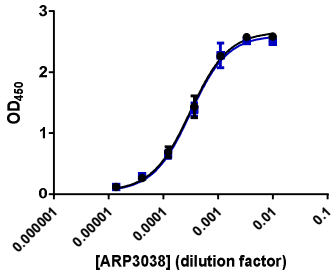

OD450
[EVA3210] (dilution factor)

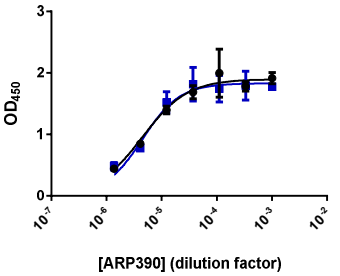

OD450
[EVA3211] (dilution factor)

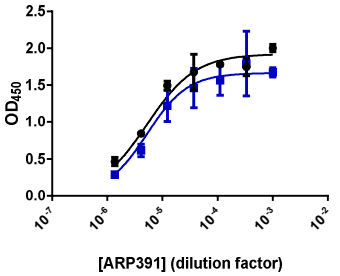

OD450
[EVA3206] (dilution factor)


OD450
[EVA3207] (dilution factor)


OD450
[ARP3037] (dilution factor)


OD450
[ARP3038] (dilution factor)


OD450
[ARP390] (dilution factor)


OD450
[ARP391] (dilution factor)

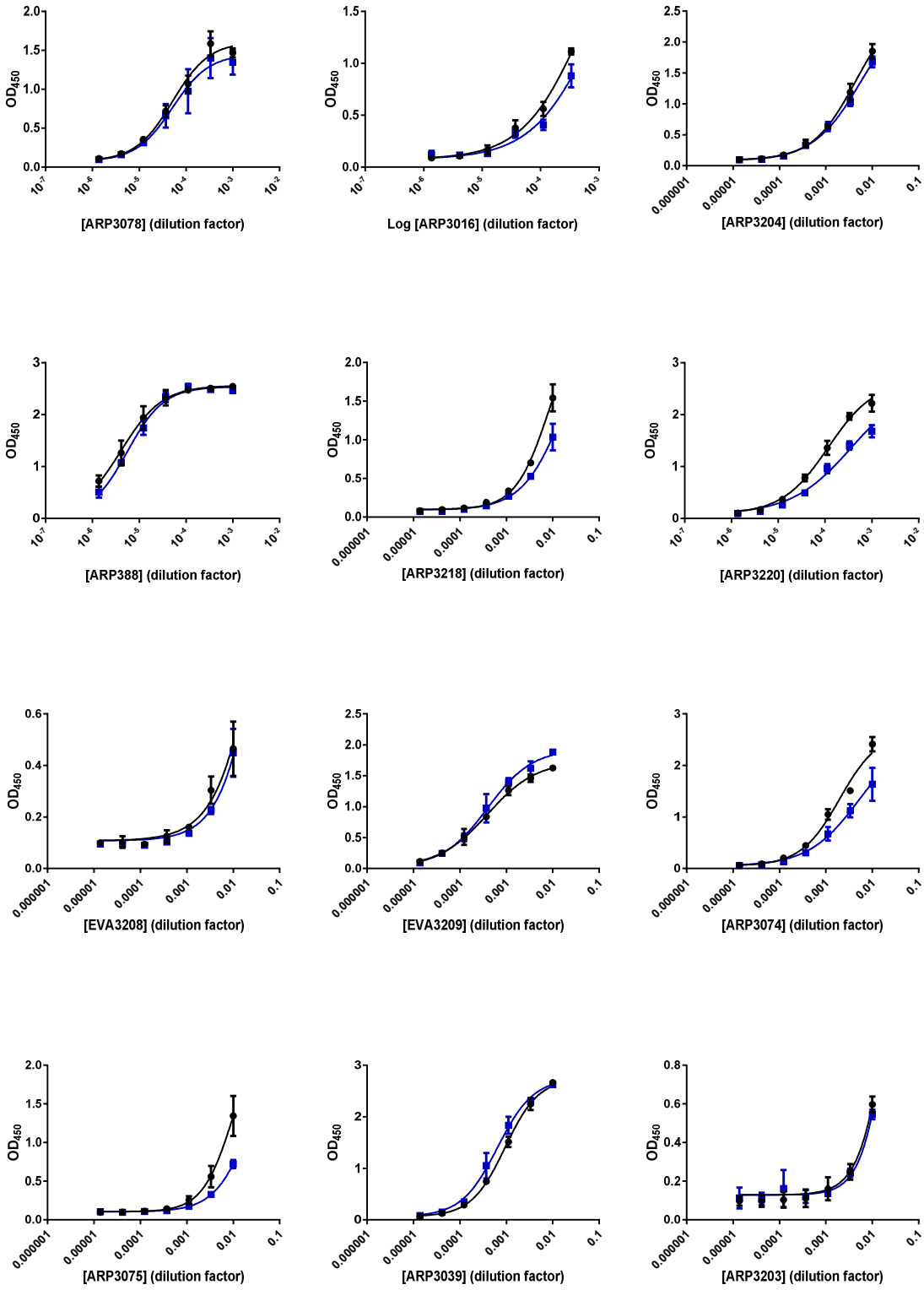



**Figure 3-3** The effects of *N*B-DNJ treatment on the binding of mAbs to rgp120$_{IIIB}$. The graphs show the binding of 24 mAbs (named in legends) to rgp120$_{IIIB}$ expressed in the presence of either 20 μM kifunensine (black squares) or 20 μM kifunensine and 2 mM *N*B-DNJ (blue circles). Bars show standard deviation, n = 3.

The $EC_{50}$s from these curves are interpolated and the ratio of the treated and untreated $EC_{50}$s calculated and plotted (Figure 3-4). An $EC_{50}$ ratio of 1 would indicate no change in binding while

values greater than or lower than 1 respectively indicate increased or decreased binding. The statistical significance of these ratios is calculated using a one-sample Student's *t*-test.

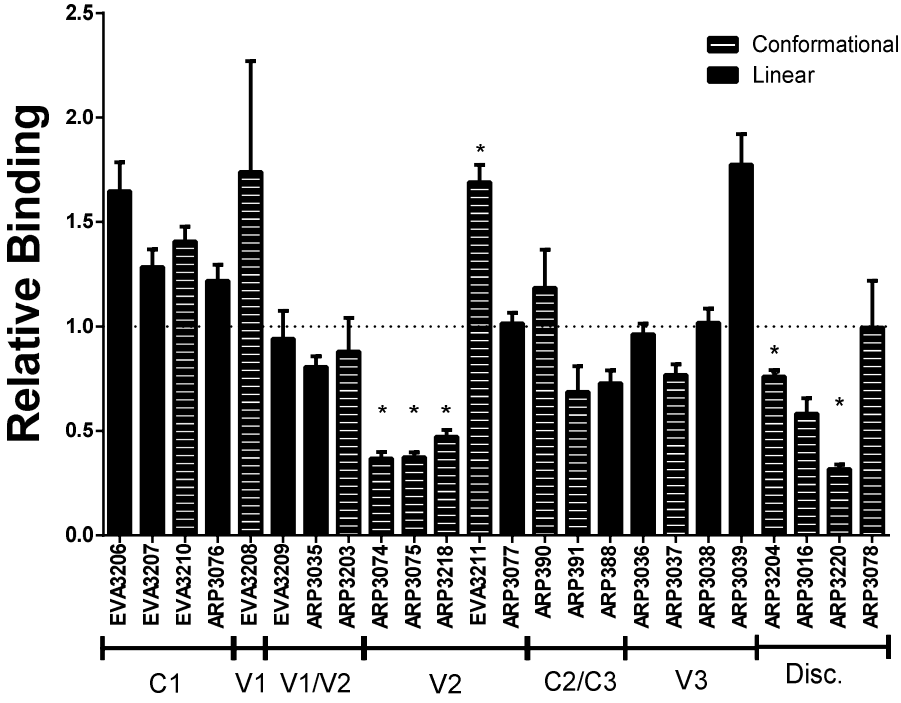



**Figure 3-4** Relative binding of mAbs to *N*B-DNJ and kifunensine treated rgp120$_{IIIB}$ compared to that treated with kifunensine only. Black bars represent mAbs targeting linear epitopes while hatched bars represent those targeting conformational epitopes. Brackets show the region of the gp120 protein in which the epitope is located. Bars show standard error, * indicates $p < 0.001$ (one sample Student's *t*-test against hypothetical value of 1, n = 6). Disc; discontinuous.

The significant differences are all located either in the V2 region of rgp120$_{IIIB}$ or in discontinuous epitopes, i.e. those spanning more than one region of the protein. Notably, only mAbs targeting conformational epitopes are affected; ARP3077 targeting a linear epitope in V2 is completely unaffected by the drug. This suggests that the V2 loop is in an aberrant conformation but is still as accessible to antibodies as normal. Interestingly, while binding of ARP3074, ARP3075 and ARP3218 to V2 is reduced, binding of EVA3211 is increased by *N*B-DNJ treatment. This is surprising, as one would expect a change in folding to decrease antibody affinity, but not impossible. It could be that this mAb targets a different region of the V2 loop to the other three and that the changes better expose that epitope to antibodies. A more interesting possibility is that this antibody was actually generated against misfolded gp120; it is possible that a small proportion of gp120 fails to bind calnexin naturally, especially when the protein is being overexpressed. On immunisation this population could induce antibodies which, while able to bind the regular protein, actually have

a higher affinity to the misfolded species. There would be advantages to the virus in evolving to exploit this in order to cause the host to produce less effective, non-neutralising antibodies, as HIV is already known to do.

ARP3204 targets a discontinuous epitope known to be independent of the V1/V2 and V3 loops, possibly indicating an additional site of misfolding. However, it is also possible that even small changes in the V2 loop could cause knock-on effects on the whole structure; having its epitope spread over many loops ARP3204 could be particularly sensitive to such a perturbation. Finally ARP3220 shows a significantly decreased binding to the treated protein. This mAb is known to be glycan dependent so it is possible that this difference is not a *bona fide* change in the protein structure but is instead a result of the remaining glucose residues interfering with this antibody's binding.

Although the changes in the V2 loop are clearly distinguishable in Figure 3-4, the non-significant changes in the other regions, especially C1, also stand out, and one has to remember that not being able to show a statistically significantly difference is not the same as being able to say there is a statistically significant identity. To address this question it is worth performing a second statistical test on the same curves, the F-test. The F-test asks whether it is possible to describe both curves (treated and untreated) with a single curve with a common $EC_{50}$: a low F statistic indicates that the two curves are similar and can be adequately described by one single curve while a high value of F describes two curves that differ substantially in $EC_{50}$ and thus are hard to describe with a single curve. These statistics are presented in Figure 3-5. When expressed as F statistics it becomes possible to quickly grasp that the statistically significant changes are occurring in the V2 and discontinuous epitopes only.

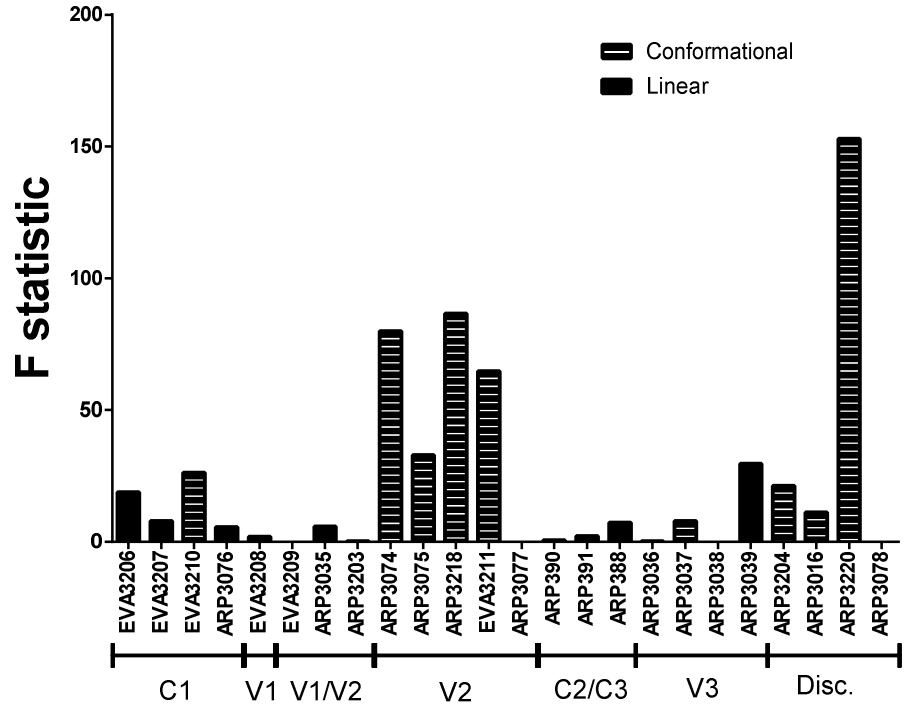



**Figure 3-5** F statistics describing the difference in binding of linear (black) and conformational (hatched) epitope targeting mAbs to *N*B-DNJ treated and untreated rgp120$_{IIIB}$ in the presence of kifunensine. Brackets show the region of the gp120 protein in which the epitope is located. Disc; discontinuous.

### 3.3.3 The misfolding of the V2 loop is not cell line dependent

Although the CHO cell line was useful for screening large numbers of antibodies, it is not necessarily a good model of gp120 folding. It is already known that the cell line lacks the endomannosidase, a deficiency we controlled for with kifunensine, but one cannot control for the fact that it is a line derived from the cells of hamsters, a decidedly non-native host for HIV. Therefore, antibodies against the V2 loop were re-tested against gp120 produced in *N*B-DNJ treated and untreated HEK 293T cells. The 293T cell line is derived from human cells and is not reported to have any glycosylation deficits. While a T-cell line might be a better model for the folding of HIV proteins 293T cells are renowned for their ease of transfection and high protein yields, making them more suitable for this sort of work. While they do have endomannosidase activity, it is not complete (as can be seen in Figure 3-1), so kifunensine was also used once again. A plasmid expressing rgp120$_{IIIB}$ was not available so rgp120$_{HXB2}$, a closely related strain differing by just a 6 amino acid duplication in V1, was used.

ELISA curves for antibodies ARP3074 and ARP3075, targeting two conformational epitopes in V2, and ARP3077, targeting a linear epitope in V2, were obtained using the same methodology as previously (Figure 3-6).

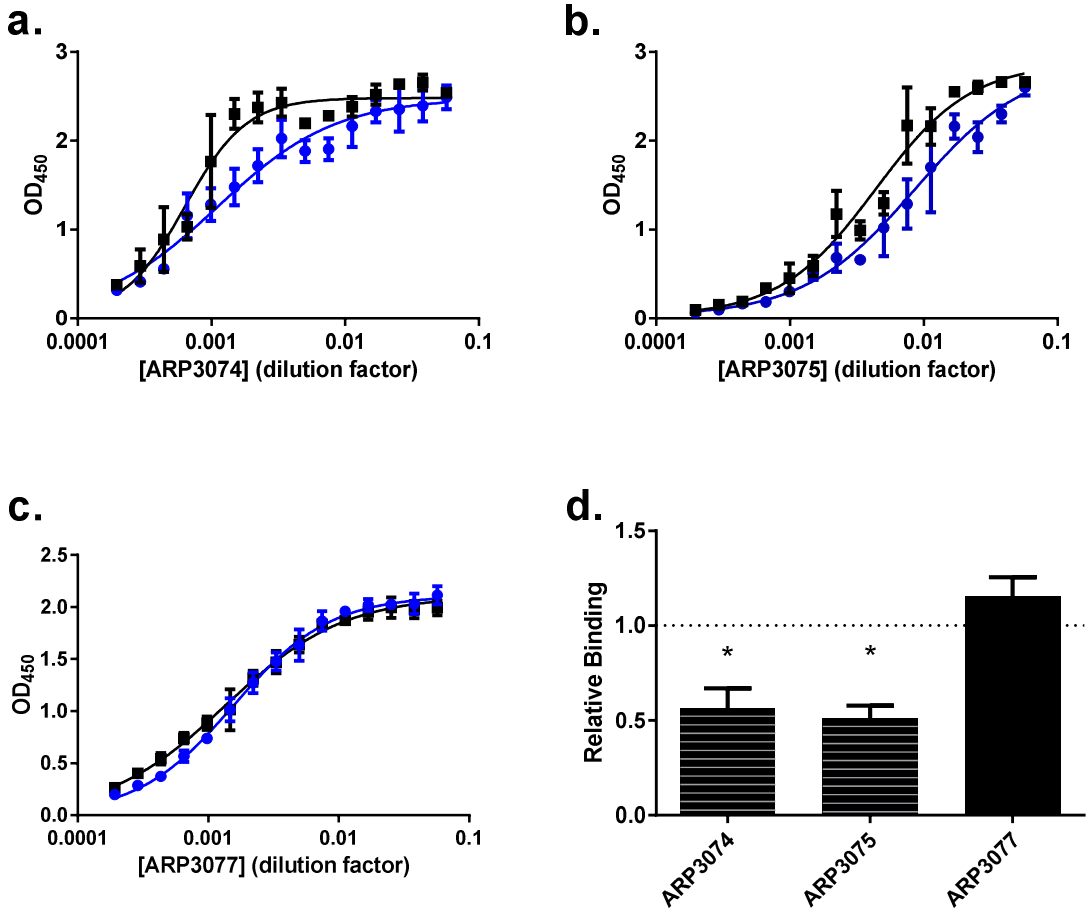



**Figure 3-6** The binding of mAbs to rgp120$_{HXB2}$ produced in HEK 293T cells treated with 20 μM kifunensine (black squares) or 20 μM kifunensine and 2 mM *N*B-DNJ (blue circles) for antibodies ARP3074 and ARP3075 (a and b) which target conformational epitopes in V2 and antibody ARP3077 (c) which targets a linear epitope in V2. The ratio of the $EC_{50}$s of each pair of curves is used to calculate the relative binding between treated and untreated rgp120 (d). In d hatched bars indicate conformational epitopes while solid bars indicate linear epitopes. Bars show standard deviation (n = 6), * indicates significance in a single sample *t* test against a theoretical value of 1, $p < 0.01$.

Figure 3-6 a and b show a clear right shift, indicating less binding, while the two curves in c overlap, all while retaining the same $E_{max}$ that is our control for equal antigen concentrations. However, when one calculates the $EC_{50}$ ratios (d) it appears to suggest that ARP3077, targeting a linear epitope and previously unaffected by the drug, binds slightly more strongly to the treated than untreated protein (single sample *t*-test against 1, $p = 0.0138$). However, the $p$ values for ARP3074 and ARP3075 are considerably lower, both being less than 0.0005, the F statistics considerably higher (31.16 and 66.84 versus 9.86) and a left shift is not at all apparent when looking at the original curves.

### 3.3.4 *N*B-DNJ treatment does not affect gp120's sensitivity to cleavage by thrombin

It has previously been reported by *Fischer et al.* that *N*B-DNJ treated rgp120$_{IIIB}$ is more sensitive to digestion by thrombin than untreated rgp120 [177]. *Ratner et al.* have claimed that the drug decreases thrombin sensitivity, although their data in no way support this claim (see Section 1.6.3 for full discussion) [175]. To settle this debate monomeric rgp120$_{BaL}$ was purified from the supernatant of transfected HEK 293T cells treated with either 20 μM kifunensine plus 2 mM *N*B-DNJ or kifunensine alone. These proteins were normalised to identical concentrations and incubated overnight with varying concentrations of human thrombin at 37 °C. SDS-PAGE and Coomassie blue stain was then used to analyse each protein's susceptibility to thrombin cleavage (Figure 3-7).

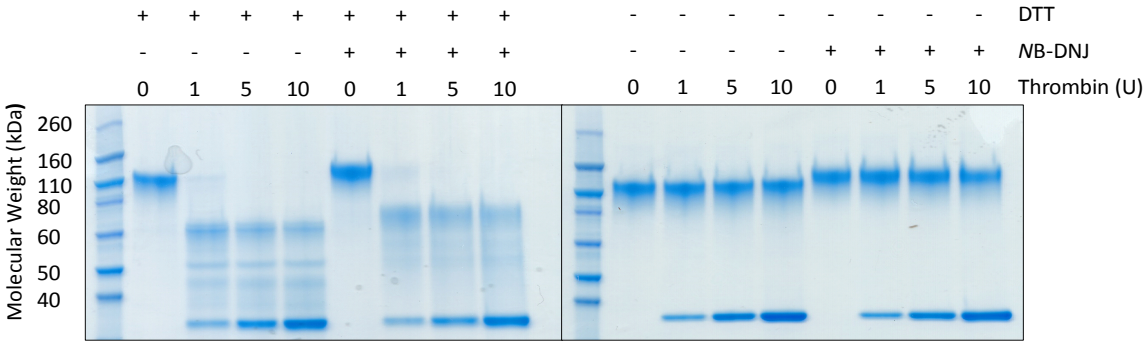



**Figure 3-7** The effects of thrombin on the electrophoretic mobility of rgp120BaL produced in either the presence or absence of 2 mM NB-DNJ. Both proteins were produced in the presence of 20 μM kifunensine to control for glycosylation differences. Samples were reduced (DTT +) by incubation with 50 mM dithiothreitol (DTT) for 5 minutes at 95 °C.

Looking first at the unreduced, uncleaved proteins it can be seen that *N*B-DNJ causes the rgp120 to run at a slightly higher weight; this is due to retained glucose residues. It has previously been reported that thrombin cleaves at a site in the V3 loop producing cleavage products of 50 kDa and 70 kDa. These bands are clearly visible in the reduced samples (but not in the unreduced as the disulphide bonds retain the two fragments together), as well as another band at 45 kDa. If one looks closely one can also see that the 70 kDa band is actually two close but distinct bands. It is possible that these distinct bands are due to differences in glycosylation. The thrombin itself is also visible on the gels at the predicted weight of 37 kDa.

Although the cleavage bands for the *N*B-DNJ treated rgp120 are fainter than those of the untreated, they are recognisable and there is no sign of the "ladder" of cleavage products observed by *Fischer et al.* Therefore, there is no evidence that rgp120 is any more or less susceptible to cleavage by thrombin. This is consistent with evidence from the previous section that the conformation of the V3 loop is unaffected by the drug.

### 3.3.5 Antibody binding differences cannot be detected in a stabilised gp140 trimer

HIV envelope is initially translated as a 160 kDa protein (gp160) and is not cleaved into gp120 and gp41 until it reaches the Golgi, by which stage most of the folding and oligomerisation has already occurred. Thus one must question whether monomeric recombinant gp120 is an adequate model for envelope folding. Luckily, the long running quest for an HIV vaccine has led to many attempts to create a recombinant HIV envelope protein that closely mimics the gp41/gp120 trimer as found on the virion surface. One such mimic is the SOSIP family of antigens; these are gp140 proteins (meaning that they lack the transmembrane and cytoplasmic domains, thus making them soluble) that possess an additional disulphide bond linking the (truncated) gp41 and gp120 subunits after cleavage. This bond stabilises the trimer, creating an antigen that closely resembles native HIV envelope by cryo-EM and has very similar antigenicity.

The current lead SOSIP in development is that based on the envelope of the BG505 strain, a clade A virus (cf. LAI/IIIB, which are clade B). The significant sequence disparity between IIIB and BG505 means that the antibodies used in the previous experiments were unusable, so the ELISAs were performed using bnAbs instead. Unlike in the previous experiments, kifunensine treatment was applied separately to the control and *N*B-DNJ treatments. This is because some bnAbs, including PGT145 and 2G12, are known to be dependent on the glycans on the antigen surface, specifically mannose residues at the tips of high-mannose glycans. Thus kifunensine is known to affect the binding of these antibodies already, so it was desirable to observe the kifunensine effect separately from any *N*B-DNJ effect.

The binding curves for these antibodies, as well as a CD4/antibody fusion protein, are presented in Figure 3-8. Binding of CD4 to gp120 induces conformational changes that expose the co-receptor binding site; to test whether *N*B-DNJ interferes with this rearrangement, antibodies that specifically target these newly exposed regions (known as CD4 induced, or CD4i) were tested in the presence and absence of soluble CD4 (Figure 3-9). The $EC_{50}$ ratios of the kifunensine and *N*B-DNJ treated SOSIP against the PBS treated controls were elucidated and subjected to single sample *t*-tests against 1 (Figure 3-10 a and b) and the data were then collated onto a single pair of axes and a two-way ANOVA was used to compare the $EC_{50}$ ratios between kifunensine/PBS and *N*B-DNJ/PBS (Sidak correction for multiple comparisons).

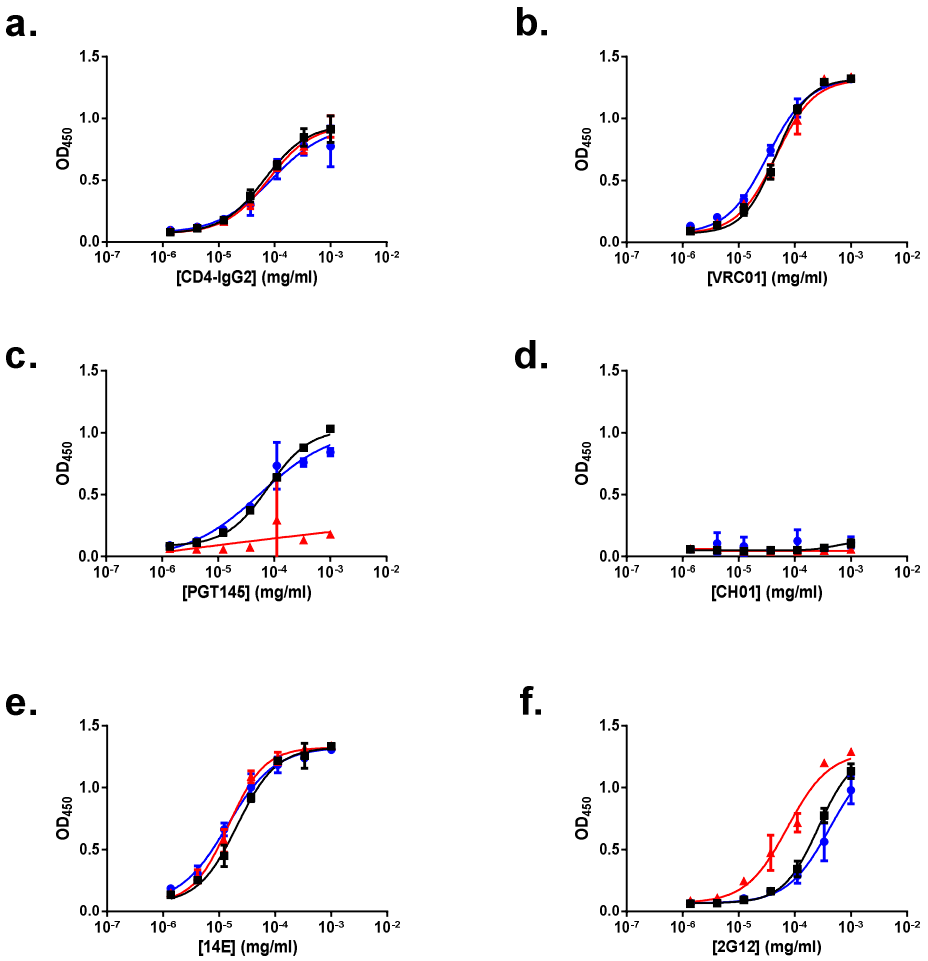



**Figure 3-8** The binding of CD4 (a) and 5 mAbs (b-f) to BG505 SOSIP produced in HEK 293T cells treated with PBS (black squares), 2 mM *N*B-DNJ (blue circles) or 20 µM kifunensine (red triangles). Bars show standard deviation (n=3).

Looking first at Figure 3-8 a we can see that *N*B-DNJ does not affect the binding of CD4 (in this case a CD4-IgG2 fusion protein for easy ELISA detection) to BG505 SOSIP. This is consistent with previous studies that show that NB-DNJ treated virions are still able to bind rgp120 and adhere to host cells. Figure 3-9 similarly confirms that *N*B-DNJ does not disrupt the trimer's ability to undergo conformational change upon CD4 binding. These two figures are summarised in Figure **3-10**, where it can clearly be seen that kifunensine treatment greatly affects 2G12 and PGT145 binding, as previously reported. However, while *N*B-DNJ treatment does appear to induce a change in the binding of antibodies VRC01, 14e, 2G12 and 412d (b) relative to PBS treatment alone, (c) makes it clear that, with the exception of 2G12, none of these changes are significantly different from those induced by kifunensine alone, suggesting that they are merely the result of glycan differences. Where there is a significant difference between the kifunensine and *N*B-DNJ treated binding, it is in the antibodies 2G12 and PGT145 where the terminal mannose is vital to binding.

Covering this with a glucose cap is thus likely to affect binding anyway, thus making any structural changes in the actual protein impossible to separate out.

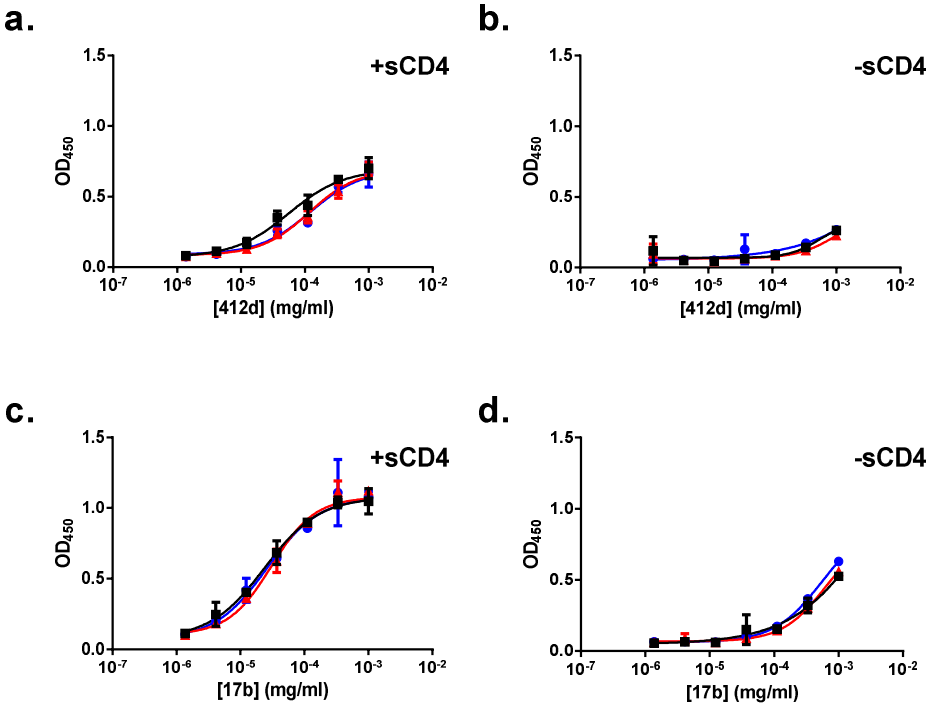



**Figure 3-9** The binding of CD4i monoclonal antibodies to BG505 SOSIP produced in HEK 293T cells treated with PBS (black squares), 2 mM *N*B-DNJ (blue circles) or 20 μM kifunensine (red triangles) in the presence (a and c) or absence (b and d) of sCD4. Bars show standard deviation (n = 3).

### 3.3.6 V1/V2 loop deletion ameliorates, but does not eliminate, the antiviral activity of *N*B-DNJ

Deletion of the V1/V2 loop was long thought to be fatal to the virus but recent work using virus evolution found that the deletion is survivable if various compensatory mutations are included [209,210]. By repeatedly passaging these viruses the infectivity was allowed to improve by viral evolution, thus creating a series of relatively replication competent mutants. These mutants were, like the antibody mapping above, based on the laboratory strain LAI.

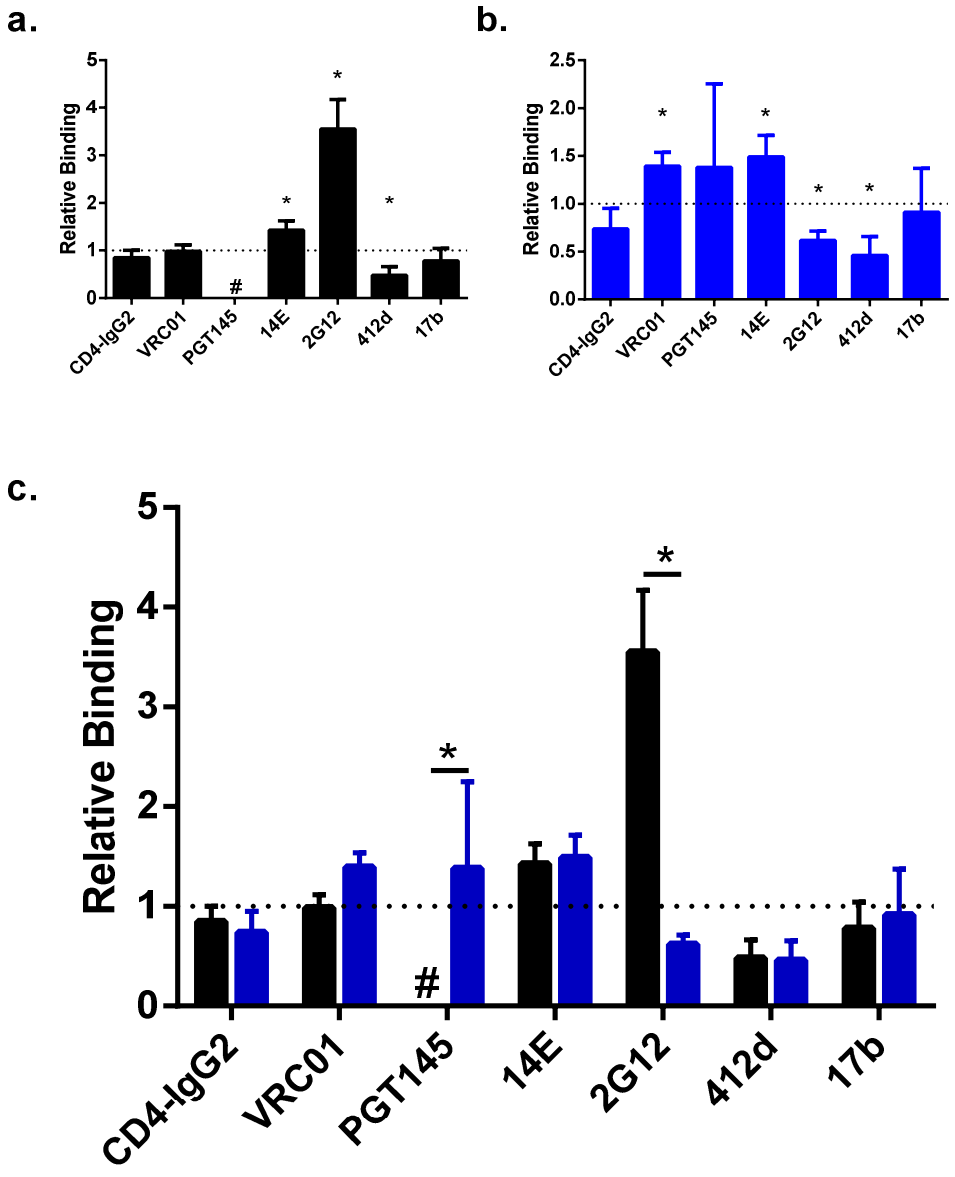



**Figure 3-10** The relative binding of antibodies to BG505 SOSIP treated wither with 20 μM kifunensine versus PBS (a) or 2 mM *N*B-DNJ versus PBS (b). Both data are presented together in c. Data for antibodies 412d and 17b are those obtained only in the presence of sCD4. Bars show standard deviation (n = 3). * in a and b indicate significant difference from 1 (no difference in binding) by a one sample student *t*-test ($p < 0.01$). * in c indicate significant differences between the two pairs of data (ie. *N*B-DNJ treatment has a significantly different effect on antibody binding to kifunensine treatment) as determined by a two-way ANOVA with an alpha of 0.01, using the Sidak correction for multiple comparisons. # indicates that an $EC_{50}$ could not be determined and hence a ratio could not be taken.

If the drop in infectivity induced by *N*B-DNJ is due to misfolding of the V2 loop then one would predict that viruses that have no V1/V2 loop would no longer be susceptible to the drug. In other words, these viruses have already suffered a loss in infectivity from having no V1/V2 and so the drug cannot make it worse. Such an outcome would validate the hypothesis that the site of misfolding lies within the V1/V2 loop.

Four V1/V2 mutants were used in this study, ΔV1/V2.2, ΔV1/V2.4.DNGSEK, ΔV1/V2.9.VE and ΔV1/V2.VK. However, one could hypothesise that any significant mutation could affect susceptibility to *N*B-DNJ as one might expect a mutated protein to need more time and more rounds of the calnexin cycle to fold. To control for this, an unrelated mutant, C119A C205A, was also used. This mutant lacks a disulphide bond at the base of the bridging sheet (which connects the V1/V2 loops to the rest of the protein) and is known to substantially decrease, but not eliminate, LAI infectivity.

The structures of the V1/V2 and bridging sheet regions (covering amino acids 114-210) are shown in Figure 3-11. Briefly, ΔV1/V2.2 expunges the entire V1/V2 loop down to the C126:C196 disulphide bond at its base, replacing these two cysteines with peptide bonded alanines. ΔV1/V2.4.DNGSEK, ΔV1/V2.9.VE and ΔV1/V2.VK all maintain the C126:C196 disulphide and instead add a 3-5 amino acid linker in place of the loop (SAG for ΔV1/V2.4.DNGSEK and VDAGS for the two ΔV1/V2.9 mutants). Additionally the two ΔV1/V2.9 mutants possess a V120E or V120K substitution, respectively- these are substitutions that arose naturally during the viral evolution process and were found to enhance infectivity. The mutations of the control, C119A C205A, are self-explanatory.

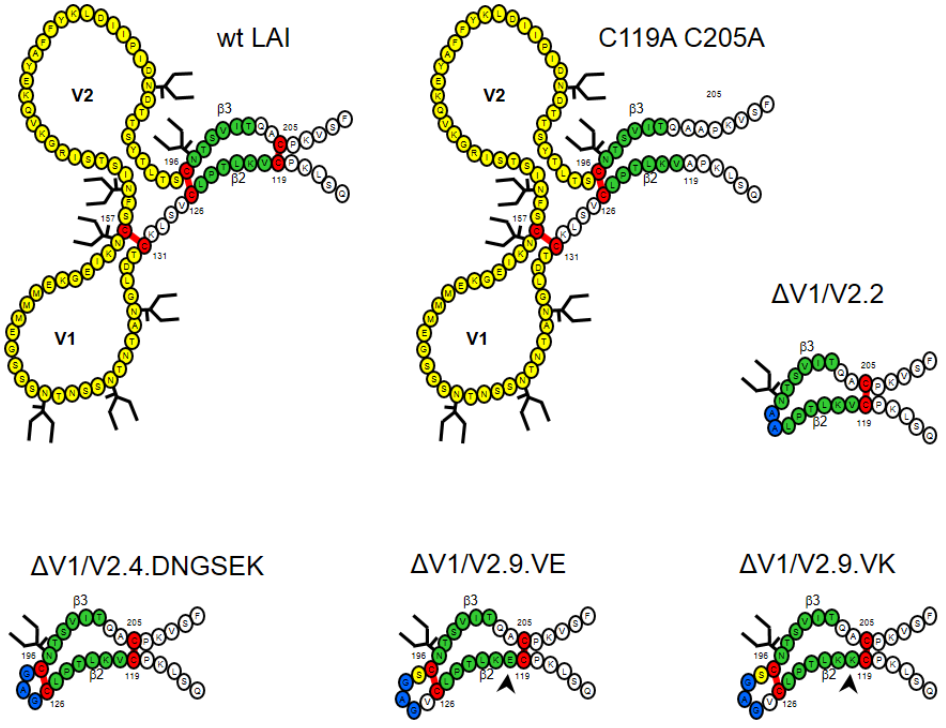



**Figure 3-11** The primary amino acid sequences of the V1/V2 loop and bridging sheet region of HIV LAI and five mutants thereof. Branched lines depict glycans (not to scale); the V1/V2 loop is shown in yellow, the bridging sheet (consisting of the β2 and β3 sheets) in green and the disulphide linked cysteines in red. Amino acids that have been introduced to replace truncated parts of the V1/V2 loop are shown in blue, and arrowheads highlight the V120 residue that is mutated in the two

ΔV1/V2.9 mutants. A short stretch of both the C1 and C2 domains, connecting the illustrated region to the rest of the protein, are shown in white. Figure adapted from *Bontjer et al.* [210].

Firstly, the mutants, presented as full length molecular clones, were transfected into HEK 293T cells. Two days later, the supernatant was removed and assayed for p24, the HIV capsid protein, by ELISA, a common and well-validated correlate of HIV concentration. These data are presented in Figure 3-12 a. This shows that all of the mutants expressed at a substantially higher rate than the wild-type, with the exception of ΔV1/V2.2 which failed to express at all, and thus was excluded from further studies. This may be explained by reasoning that because the envelope protein was shorter, the cells were able to produce it more quickly, though the increase is substantial and the deletions (at most 76 amino acids from an envelope protein of 861 amino acids) small. Also, this would not explain why C119A C205A also expressed at a higher rate. An alternative explanation for this mutant could be that by dispensing of a disulphide bond the chances of a mispairing were reduced, speeding up the time it takes for the protein to reach its native disulphide arrangement. However, the most likely explanation, in my opinion, is that the viruses have evolved to compensate for low virion infectivity by simply producing more of them.

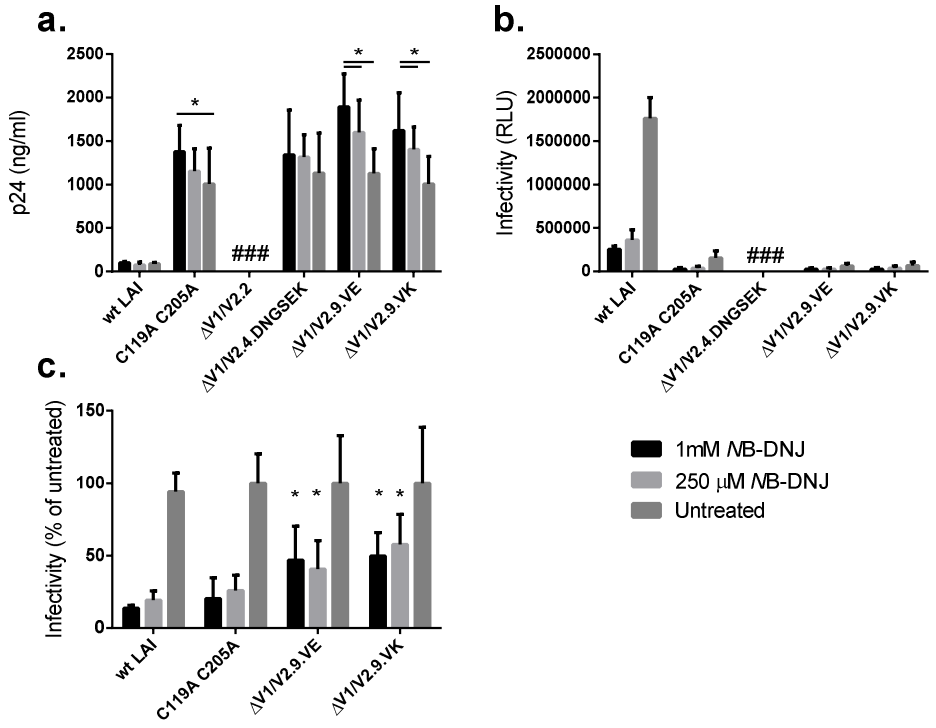



**Figure 3-12** The effects of *N*B-DNJ on the secretion and infectivity of HIV LAI V1/V2 mutants. a shows the levels of secretion of p24, a correlate of HIV concentration. b shows the infectivity of those viruses as determined by the levels of luciferase produced by the reporter TZM-bl cell line when infected with 500pg p24 equivalents. c shows the same data as b but normalised as a % of the infectivity of the untreated control for each mutant (thus a low number indicates that the drug has decreased the infectivity of the mutant). In all cases bars show standard deviation (n = 4; except

wild type, n = 3) and # indicates that recorded levels were indistinguishable from background. In b, * indicates a significant difference between two treatments within each mutant set, with the columns in question highlighted by lines between them while in c, * indicates that the difference in infectivity between the treated and untreated columns is significantly different to that seen in the wild-type (two-way ANOVA with an alpha of 0.01, using the Dunnett correction for multiple comparisons). Abbreviations: wt, wild-type; RLU, relative luminescence units.

For C119A C205 and the two ΔV1/V2.9 mutants *N*B-DNJ causes a dose dependent, statistically significant increase in HIV secretion, something not observed for the wild-type (although not statistically significant ΔV1/V2.4.DNGSEK seems to show a similar trend). This may be because *N*B-DNJ, by indirectly inhibiting mannose trimming from the D3 arm, prevents degradation of some of the gp160 mutants by ERAD, allowing more of them to reach the plasma membrane and thus allowing for slightly greater secretion of virions. To measure the infectivity of the virions independently of this effect on secretion the supernatants were normalised into 500 pg doses, which were used to infect the TZM-bl reporter cell line. TZM-bl is a HeLa cell line transfected with CD4 and the co-receptors CCR5 and CXCR4, making it highly permissive to HIV infection. Upon infection the cells produce luciferase under the control of the HIV promoter, thus the luminescence produced when luciferase substrate is added is directly proportional to the level of HIV infection in the well. To maximise infection DEAE dextran is used, and indinavir, an HIV protease inhibitor, is added to the medium to prevent the cells secreting a second round of virions that could amplify the signal. All of the mutants were less infectious than the wild-type (Figure 3-12 b), especially ΔV1/V2.4.DNGSEK which, contrary to previous reports, showed no infectivity at all. However, we are more interested in whether *N*B-DNJ is able to decrease the infectivity further, so the data were each normalised to be expressed as a percentage of the infectivity recorded for the untreated control of each mutant (c). Here it can be seen that the control, C119A C205A, experienced a drop in infectivity that was statistically indistinguishable from that of the wild type, to 21.35% (99% Confidence Interval [CI] 10.39% to 32.3%) cf. 13.68% ((99% CI 11.88% to 15.48%). However, the two ΔV1/V2.9 mutants did show a significantly smaller drop in infectivity, to 48.77% (99% CI 31.01% to 66.53%) and 49.79% (99% CI 36.06% to 63.52%) for VE and VK, respectively. This means that *N*B-DNJ is less efficacious against these mutants than it is against the wild type, yet the action of the drug cannot solely be due to inducing misfolding within the V1/V2 loop.

# 4 The mechanism of action of *N*B-DNJ as an HIV antiviral

## 4.1 Contributions

Dominic Alonzi performed the quantification of the free oligosaccharides. Nicholas McCaul performed all experiments involving radioactive pulse-chase labelling. Analysis of these data was by the author.

## 4.2 Introduction

In Chapter 3 we were able to confirm the finding of *Fischer et al.* that *N*B-DNJ affects the conformation of the gp120 V1/V2 loop, and also expand upon it. Firstly we refined this result to show that it is the V2 loop which is specifically less immunoreactive and that this is due to an altered conformation as opposed to steric constraints, as evidenced by the maintained reactivity to antibodies recognising linear epitopes. However, we also showed that the drug must affect the protein beyond the V1/V2 loop as well, as HIV clones lacking this region were still partly susceptible to the antiviral effects of *N*B-DNJ.

In this chapter, we will expand upon these findings to answer a number of questions concerning *how* *N*B-DNJ is able to induce this misfold and how such a misfold translates into an antiviral effect. Firstly we will develop a method of determining the relationship between *N*B-DNJ concentration and the proportion of rgp120 that is misfolded and relate that information to the level of inhibition of the glucosidase enzymes. More importantly, we can correlate misfolding to the antiviral effect to try and establish the link between the two.

The other major aim of this chapter will be to determine why α-glucosidase inhibition leads to misfolding by using radioactive pulse-chase labelling to follow a synchronous sample of rgp120 through the folding process in the presence and absence of *N*B-DNJ to identify any differences and their timing. Finally, we will ask why some HIV strains are more susceptible than others to *N*B-DNJ by looking for the genetic basis of glucosidase inhibitor sensitivity. As glucosidase inhibitors affect glycans, this will be achieved by constructing a panel of HIV clones with glycan deletions and examining how this affects their susceptibility to *N*B-DNJ.

## 4.3 Results

### *4.3.1* *N*B-DNJ treatment induces a mixed population of folded and misfolded rgp120, both of which can bind conformation sensitive antibodies

In the previous chapter it was established that 2 mM *N*B-DNJ decreases the ability of conformation sensitive monoclonal antibodies (mAbs) to bind to the gp120 V2 loop, showing that they bound with only ~50% of the $EC_{50}$ to treated protein as to the untreated. However, while this is a qualitative measure that something has changed in V2 the actual numbers are not particularly quantitative, i.e. does a 50% drop in $EC_{50}$ translate to 100% of the gp120 being misfolded, all of which bind with a lower affinity, or does it translate to a mixed population with some proteins binding natively and others binding not at all?

This question can be addressed by looking at the $E_{max}$ reached on the ELISA curves. This is a measure of the total amount of binding that occurs, so if $E_{max}$ remains constant while the concentration of *N*B-DNJ changes then it can be inferred that the sensitive antibodies are able to bind the misfolded gp120 to the same maximal level as to the native protein. These data are presented in Figure 4-1, where the binding of conformation sensitive anti-V2 antibody ARP3075 was compared to that of the insensitive ARP3077. These two antibodies will be used for all further experiments on folding in this chapter.

In Figure 4-1 a we can see that increasing the concentration of *N*B-DNJ ("redder" curves) creates an increasing right shift in ARP3075 binding, demonstrating a dose response of misfolding to drug concentration. More significantly for current purposes, however, all of the curves appear to converge on the same point. The same convergence, but lack of right shift, can be seen with the conformation insensitive antibody ARP3077, which acts as a control for concentration differences as previously described in Chapter 3. GraphPad Prism 6 can be used to calculate the plateau of each curve, our $E_{max}$, which are then plotted against *N*B-DNJ concentrations in c. The gradients of the linear models describing the effect of *N*B-DNJ on the $E_{max}$ of ARP3077 and ARP3075 do not deviate significantly from 0 (i.e. no relation between drug and $E_{max}$) at an alpha of 0.01. However, the line for ARP3075 does show a slight positive gradient which is significantly different from the horizontal at a $p = 0.0125$, suggesting that there is a relationship if one accepts significance at an alpha of 0.05. However, the gradient is only 0.074 so even if it were statistically significant it is unlikely to be a practically significant effect.

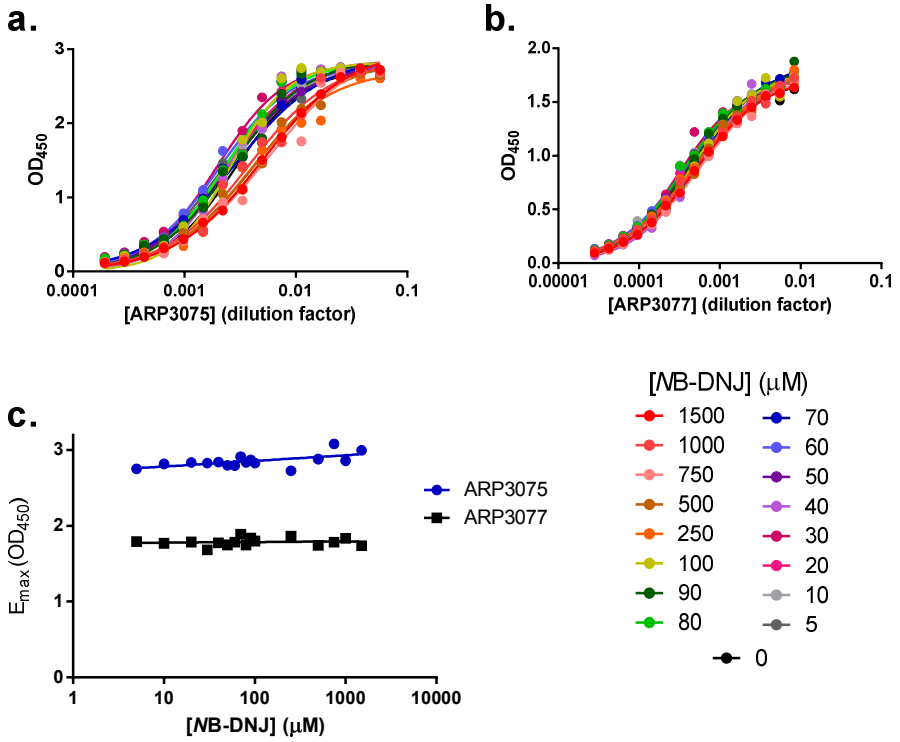



**Figure 4-1** Treatment with increasing concentrations of *N*B-DNJ has no effect on the maximal ELISA reading ($E_{max}$) of $rgp120_{HXB2}$, produced in HEK 293T cells, probed with either the conformation sensitive ARP3075 (a) or the conformation insensitive ARP3077 (b) anti-V2 mAbs. GraphPad Prism 6 was used to find the $E_{max}$ for each curve; these are then plotted as a function of the concentration of *N*B-DNJ with which the cells were treated (c). a and b were analysed using a standard 4-parameter non-linear model while c was fitted with linear model.

We can, therefore, conclude that the antibodies, whether their $EC_{50}$ in binding to rgp120 is or is not affected by the drug, still bind the protein at the same maximal level. Thus we can reject the hypothesis that *N*B-DNJ treatment creates a population of rgp120 that cannot bind ARP3075 at all in favour of the hypothesis that *N*B-DNJ induces the creation of a mixed population of two rgp120 states, native and misfolded, each of which bind the ARP3075 at a different affinity. Thus the right shift that is seen with increasing concentrations of NB-DNJ is due to the ratio of misfolded to folded rgp120 increasing, shifting the population $EC_{50}$ closer to that that would be seen with 100% misfolded protein.

### 4.3.2 Developing a model to calculate the ratio of folded to misfolded rgp120 by ELISA

We have hypothesised that treatment of cells transfected with rgp120 with NB-DNJ will lead to the creation of two populations of rgp120 being secreted into the cell supernatant; the native rgp120 (*gp120*) and the misfolded rgp120, which we will label *gp120**, each of which bind the

sensitive antibody ($Ab$) with a different dissociation constant ($Kd$) to form two distinct antibody/antigen complexes:

(1) $$Ab + gp120 \overset{Kd_1}{\longleftrightarrow} C_1$$

(2) $$Ab + gp120^* \overset{Kd_2}{\longleftrightarrow} C_2$$

(3) $$B = C_1 + C_2$$

An ELISA can only measure the total concentration of the complexes (B) and cannot distinguish the two, hence why the $E_{max}$ remains constant. However, the $EC_{50}$ will be proportional to the ratio of the two $Kds$ and thus we can use this to infer the ratio of $C_1$:$C_2$, which will be the ratio of *gp120:gp120**.

To discern this quantity we must first recognise that in a given ELISA well there will be a mixture of free gp120s, free antibodies, and complexes:

(4) $$gp120_T = gp120 + C_1$$

(5) $$gp120^*_T = gp120^* + C_2$$

(6) $$Ab_T = Ab + C_1 + C_2 \approx Ab$$

In all of the above $_T$ indicates the total quantity of each species (i.e. that of both in and not in a complex). In (6) we approximate that the concentration of free antibody (that not in complex with protein) can be approximated by the total concentration of antibody, on the assumption that the amount of antibody actually involved in binding will be only a small proportion of the whole. This means that the value $Ab$, the amount of antibody not in complex, can be taken to be the same concentration that we add to the ELISA well, thus making it a known quantity. We can rearrange equations (1) and (2) thus:

(7) $$Kd_1C_1 = Abgp120$$ (8) $$Kd_2C_2 = Abgp120^*$$

Then substitute in equations (4) and (5) respectively:

(9) $$Kd_1C_1 = Ab_T(gp120_T - C_1)$$

(10) $$Kd_2C_2 = Ab_T(gp120^*_T - C_2)$$

Rearrange to:

(11) $C_1 = \frac{Ab_T gp120_T}{Kd_1 + Ab_T}$ (12) $C_2 = \frac{Ab_T gp120^*_T}{Kd_2 + Ab_T}$

Because equation (3) already explains the relationship between $C_1$ and $C_2$ we can merge the two equations:

(13) $B = \frac{Ab_T gp120_T}{Kd_1 + Ab_T} + \frac{Ab_T gp120^*_T}{Kd_2 + Ab_T}$

Let us now introduce a term for the total amount of rgp120 (whether misfolded or not) in the system and call it G.

(14) $G = gp120_T + gp120^*_T$

Substitute (14) into (13) to replace $gp120^*_T$

(15) $B = \frac{Ab_T gp120_T}{Kd_1 + Ab_T} + \frac{Ab_T (G - gp120_T)}{Kd_2 + Ab_T}$

Then divide by G

(16) $\frac{B}{G} = \frac{Ab_T (\frac{gp120_T}{G})}{Kd_1 + Ab_T} + \frac{Ab_T (1 - \frac{gp120_T}{G})}{Kd_2 + Ab_T}$

Since what we are interested in is the fraction of total gp120 that is misfolded, let us create at term for that:

(17) $f_1 = \frac{gp120_T}{G}$

We can now substitute $f_1$, the fraction of rgp120 that is correctly folded, into (16):

(18) $\frac{B}{G} = \frac{Ab_T f_1}{Kd_1 + Ab_T} + \frac{Ab_T (1 - f_1)}{Kd_2 + Ab_T}$

Since the ELISA readout is B, it would be useful to eliminate G from the left side of the equation. This can be achieved by normalising the $E_{max}$ to 1 in the ELISA curves, as G, like $E_{max}$ is a measure of the total amount of rgp120 of both species.

As the *x* axis of the ELISA curves is logarithmic, let us convert the equation to reflect that, by first creating a term for the logarithm of $Ab_T$ and substitute it into the normalised (18):

(19) $$\overline{Ab_T} = Log_{10}Ab_T$$

(20) $$B = \frac{10^{\overline{Ab_T}} f_1}{Kd_1 + 10^{\overline{Ab_T}}} + \frac{10^{\overline{Ab_T}}(1-f_1)}{Kd_2 + 10^{\overline{Ab_T}}}$$

This can be simplified by dividing by $10^{\overline{Ab_T}}$:

(21) $$B = \frac{f_1}{\frac{Kd_1}{10^{\overline{Ab_T}}} + 1} + \frac{1-f_1}{\frac{Kd_2}{10^{\overline{Ab_T}}} + 1}$$

And simplified further if we also express the two Kds as base 10 logarithms ($\overline{Kd}$):

(22) $$B = \frac{f_1}{1 + 10^{\overline{kd}_1 - \overline{Ab}_T}} + \frac{1-f_1}{1 + 10^{\overline{kd}_2 - \overline{Ab}_T}}$$

Thus we end up with an equation with the following 5 terms:

| Term | Definition |
|---|---|
| **B** | The total amount of antibody/gp120 complexes, regardless of gp120 folding |
| **$f_1$** | The fraction of gp120 in those complexes that is correctly folded |
| $\overline{Ab_T}$ | The $\log_{10}$ of the concentration (or dilution factor) of the total antibody present |
| $\overline{Kd_1}$ | The $\log_{10}$ of the dissociation constant of the antibody from the correctly folded gp120 |
| $\overline{Kd_2}$ | The $\log_{10}$ of the dissociation constant of the antibody from the incorrectly folded gp120* |

Of these terms, $\overline{Ab}_T$ is known, as it is simply the concentration (or dilution factor) of antibody added, and is the *x* value on the ELISA curve (when expressed logarithmically). B is also known, as it is the optical density value measured by the ELISA, and therefore is just the *y* value on the ELISA curves, after they have been normalised such that $E_{max} = 1$.

This leaves three unquantified terms- the two Kds and the term we are interested in, $f_1$. However, we can assume that if no drug is added to the cells then all the protein should be correctly folded,

ie. $f_1=1$. Therefore, when equation (22) is applied to the control wells it will simplify to the following:

$$B = \frac{1}{1+10^{\overline{kd}_1-\overline{Ab}_T}} \quad (23)$$

In equation (23) there is a single unknown, $Kd_1$, which can therefore be readily interpolated by GraphPad Prism. This value for $Kd_1$ can now be substituted into (22) leaving just two unknown terms for GraphPad to calculate for the treated wells: $Kd_2$ and the elusive $f_1$.

To improve GraphPad's ability to determine the correct values for these terms, the following constraints are enforced:

i. $0 \leq f_1 \leq 1.1$

ii. $\overline{Kd}_1 < \overline{Kd}_2$

iii. $\overline{Kd}_2$ is shared for all data sets

### 4.3.3 The proportion of misfolded rgp120 secreted by HEK 293T cells is increased in a dose-dependent fashion by *N*B-DNJ

As shown above, to calculate the proportion of correctly folded rgp120 in a population we must first know the dissociation constant of each antibody to the native rgp120 ($Kd_1$). This assumes that cells untreated with *N*B-DNJ will only produce perfectly folded protein. This assumption is almost certainly untrue; however, since we are only interested in the effects of the drug compared to the untreated baseline this should not matter, in other words "correctly folded" should be taken to mean "folded as usual". However, the value of $Kd_1$ is key to the equations and getting it even slightly out can have large effects on the calculation of $f_1$.

To ensure that $Kd_1$ was calculated accurately, equation (23) was not only applied to the untreated control but to all of the treated curves as well. Because the treated rgp120 will be less able to bind to ARP3075 it will give an erroneously high value of $Kd_1$. However, if one fits a simple 4 parameter non-linear model to the data then the bottom ($E_{min}$) should converge as the concentration of drug trends to zero. This asymptote will be a better estimate of $Kd_1$ than just taking a single reading of an untreated control. ARP3077 should show a flat line as it cannot distinguish the treated and untreated proteins. Such curves are presented in Figure 4-2, which calculates the $Kd_1$ for ARP3075 as 0.0028 (a 1/354 dilution of the antibody stock) and for ARP3077 as 0.000412 (a 1/2427 dilution).

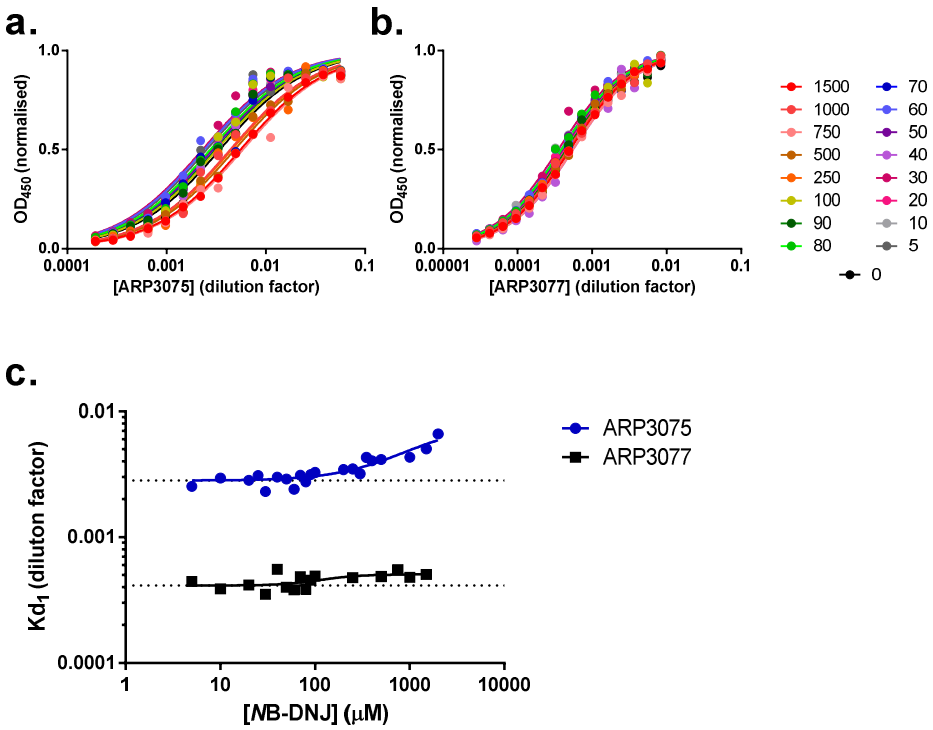



**Figure 4-2** Calculating the dissociation constants of antibodies from native rgp120. The binding of a dilution series of the conformation sensitive anti-V2 antibody ARP3075 (a) and the insensitive antibody ARP3077 (b) to the supernatants of HEK 293T cells producing $rgp120_{HXB2}$ treated with varying concentrations of *N*B-DNJ is determined by ELISA. Curves are fitted to these data using the model described above as equation (23). $Kd_1$ is interpolated from each of these curves by GraphPad Prism 6 and plotted as a function of *N*B-DNJ concentration (c). A 4 parameter non-linear fit is performed and the lower asymptote found (i.e. where $Kd_1$ trends to as the concentration of drug trends to 0); these asymptotes are described by the dotted lines.

These values of $Kd_1$ can now be substituted into equation (22) permitting the calculation of $Kd_2$ and, more importantly, $f_1$, which can trivially be converted from a proportion into a percentage. These data are displayed in Figure 4-3.

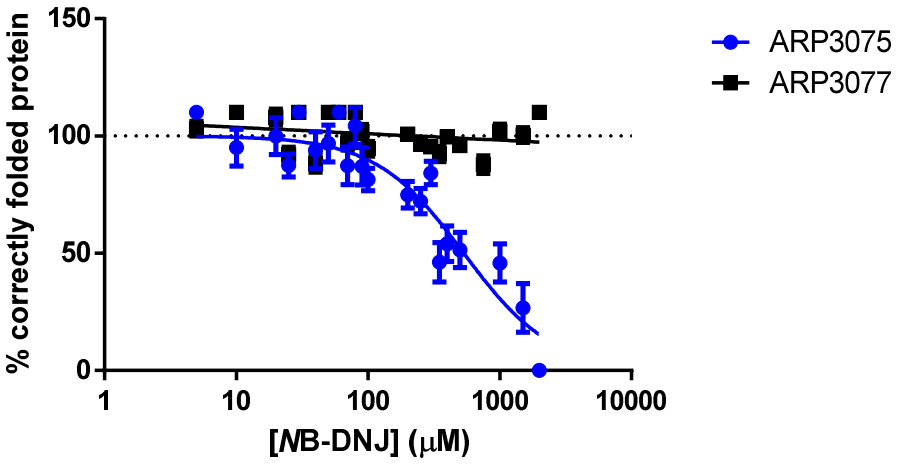



**Figure 4-3** Increasing the concentration of *N*B-DNJ causes the percentage of rgp120 that is correctly folded to decrease in a dose dependent fashion, as measured by the conformationally sensitive

antibody ARP3075. The percentage as measured by the conformation insensitive antibody ARP3077 does not deviate from 100% (marked with a dotted line) as it is unable to distinguish the misfolded from folded protein. Bars show standard error (n is between 1 and 4, generally 3, for ARP3075 and is 1 for ARP3077).

From Figure 4-3 we can determine the concentration of *N*B-DNJ that induces 50% of the rgp120 to be misfolded (inhibitory constant 50%, $IC_{50}$) to be 534 μM with a standard error of 262 μM. Part of the reason that the error is so high is because the curve does not reach its bottom asymptote. These data could be obtained by treating the cells with concentrations of *N*B-DNJ higher than 2 mM. *N*B-DNJ can be used to treat HEK 293T cells at concentrations up to 10 mM without seeing any decreased viability as determined by the reduction of 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (MTS) to formazan (data not shown). However, on visual examination the cells are noticeably different shapes at concentrations of 4 mM and above. This discrepancy is unexplained so all experiments were performed using a maximum of 2 mM lest it be a sign of toxicity.

One way to avoid this problem would be to assume that as *N*B-DNJ concentration tends to infinity the percentage of correctly folded protein would trend to zero. With such a constraint the $IC_{50}$ would barely change, going down to 533 μM, but the error would be reduced to 70 μM, a far more acceptable degree of error. However, this is not a trivial assumption as it presupposes that rgp120 is completely incapable of folding without calnexin; it may be that a small population is capable of reaching the native state without it and the cycle only exists to aid this process. Nonetheless, the fact that the course of the curve is imperceptibly different for this constraint, altering the $IC_{50}$ by less than 0.2%, I believe that this constraint is justified in this model.

### 4.3.4 *N*B-DNJ induced misfolding correlates to the level of inhibition of α-glucosidase I & II

*N*B-DNJ is known to inhibit ER α-glucosdiases but it is not known how glucosidase activity correlates to protein folding. α-glucosidase activity can be measured *in vitro* using isolated enzymes, but these assays suffer from a multitude of problems, such as not using human enzyme and using small artificial substrates that may not behave similarly to a large multiply glycosylated protein. They also ignore the effects of cell and ER membranes on excluding the drug from, or concentrating it in, various cellular compartments.

A better assay, therefore, would be to look at the activity of the ER α-glucosidases actually within the HEK 293T cells that are producing rgp120. Such an assay can be performed by looking at the production of free oligosaccharides (FOS). FOS are produced when proteins are degraded as

part of the ER-associated degradation process (ERAD). During ERAD proteins are translocated from the ER to the cytoplasm where their glycans are removed by PNGase F. These glycans can be extracted from cell lysates, labelled with fluorescent markers and identified by HPLC in a manner similar to that used for the glycan profiling in Chapter 2. Thus, if one assumes that ERAD covers a representative sample of all proteins, one can look for glucosylated FOS as a measure of glucosidase inhibition. As a FOS assay uses a cell lysate, while the folding assay uses the cell supernatant, the same cells can be used for both assays, thus ensuring maximum relevance.

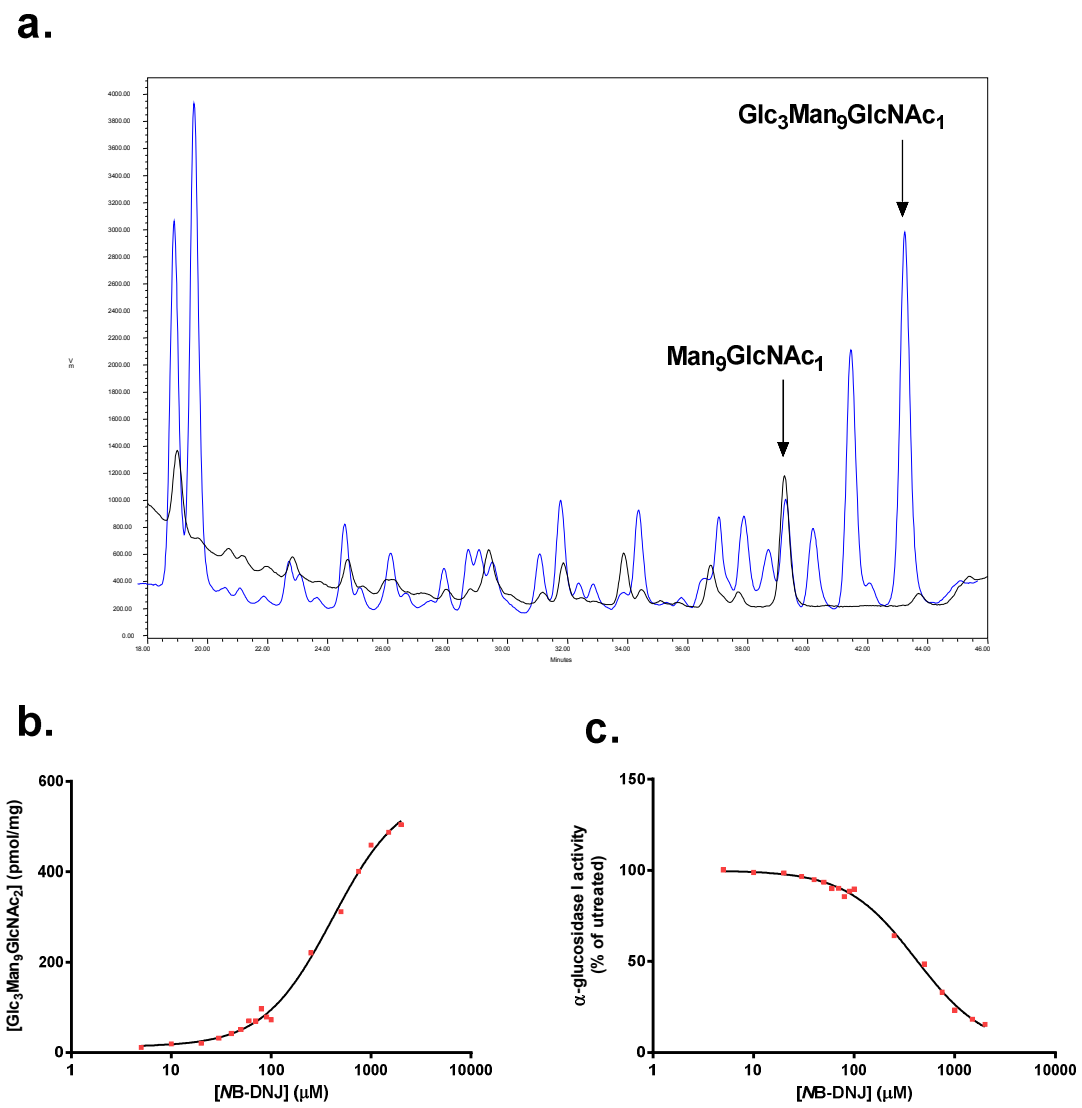



**Figure 4-4** ER α-glucosidase I inhibition as measured by FOS analysis. An HPLC profile for FOS from cells treated with 20 μM kifunensine (black) and the same plus 2 mM *N*B-DNJ (blue) (a). The largest peak in the kifunensine treated cells is that of $GlcNAc_1Man_9$, while upon additional treatment with *N*B-DNJ this shifts to the $GlcNAc_1Man_9Glc_3$ peak, which was absent previously. Measuring the area under this peak to calculate its concentration allows us to show that the presence of $GlcNAc_1Man_9Glc_3$ FOS is dose dependent (b). These data can then be transformed into a measure of α-glucosidase I (c). Curves fitted using unconstrained four-parameter non-linear regression.

As the cells are also treated with 20 μM kifunensine the majority of FOS will usually be $Man_9GlcNAc_1$ with smaller populations of glycans with fewer mannoses as a result of degradation by cytoplasmic mannosidases. These populations can be seen in α-glucosidase inhibition and will show up as a shift in the $Man_9GlcNAc_1$ to $Glc_3Man_9GlcNAc_1$. We can therefore measure the level of α-glucosidase I inhibition by integrating the area under this peak. Kifunensine's inhibition of mannosidases means that it is an ERAD inhibitor, yet its presence does not seem to prevent the detection of FOS. This could be due to non-glycan dependent ERAD pathways taking over.

The FOS profiles for the kifunensine only and 2 mM *N*B-DNJ plus kifunensine cells are presented in Figure 4-4 a. Figure 4-4 b shows that the effect is dose dependent, with rising doses of *N*B-DNJ inducing increasing concentrations of the $Glc_3Man_9GlcNAc_1$ FOS. This can then be translated into a measure of GlcI activity by taking each concentration as a percentage of the maximum possible $Glc_3Man_9GlcNAc_1$ (i.e. that when inhibition is total, or *N*B-DNJ concentration is infinite). This is defined as the $E_{max}$ of Figure 4-4 b. This gives the percentage of inhibition, which can then be subtracted from 100 to get the percentage activity (c).

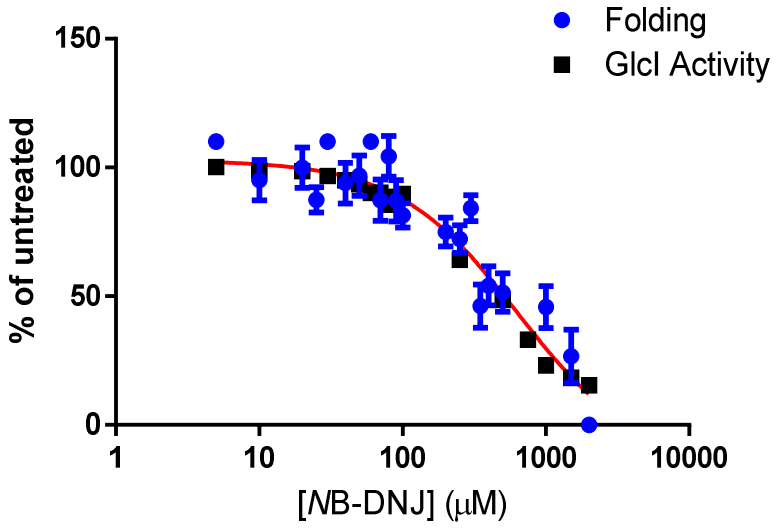



**Figure 4-5** The correlation of rpg120 folding to α-glucosidase I (GlcI) activity in HEK 293T cells. The red line is the curve common to both data sets, plotted using four-parameter non-linear regression.

The effect of *N*B-DNJ on GlcI seems to correlate well with its effects on *N*B-DNJ folding (Figure 4-5). To see if this correlation is perfectly 1:1 we can use an additional-sum-of-squares F test. This asks whether one curve adequately describes both data sets, a value of F greater than 1 with a low $p$ value indicating significance. In this case F is 0.912 and the $p$ is 0.470, showing that it is not necessary to use two curves to describe the data.

This means that rgp120 folding is directly correlated to GlcI function, so when the GlcI is inhibited to 50% activity, 50% of the rgp120 is misfolded. Note that this does not necessarily mean that 50% of the enzymes are inhibited as there will probably be some reserve capacity in the system, so it may be that a certain proportion of the enzymes need to be inhibited before glucosidase activity begins to decrease.

### 4.3.5 The effect of *N*B-DNJ on the folding of rgp120 does not correlate directly with its antiviral activity

To know that the $IC_{50}$ of *N*B-DNJ against rgp120 folding is roughly 500 μM is interesting but would be far more so if this were put in the context of its antiviral activity. The effects of *N*B-DNJ on α-glucosidases is likely to be cell dependent so it is important to assess both criteria in as similar a model as possible. The most realistic model would be to infect primary human peripheral blood mononuclear cells (PBMCs) in the presence of drug with a primary isolate of HIV and measure the infectivity of the secreted virions while simultaneously looking at the folding of their envelope proteins. This experiment is difficult for a number of reasons.

The first reason is that the sensitive antibodies thus far discovered, such as ARP3075, only work against the laboratory strain HIV LAI, thus ruling out using primary isolates. Secondly, it is possible that the infection will spread throughout the cell population; the antiviral nature of the *N*B-DNJ means that it could limit this spread, forcing the collection of virus at a time point less than two viral cycles to prevent this from affecting the concentration of envelope protein between treatment groups. Thirdly, to perform the ELISAs it is necessary to remove the supernatants to a biosafety level 2 (BSL 2) laboratory. This necessitates inactivating the virus, a process which involves the use of heat and detergents, conditions which are unlikely to preserve protein conformation reliably.

An alternative solution, therefore, is to grow HIV in HEK 293T cells, the same line in which the folding data were elucidated. HEK 293T cells are not a natural target of HIV but they are readily transfectable with molecular clones- such a clone of LAI is already a commonly used HIV model. Another advantage of this system is that the virions cannot re-infect the cells, so there is no possibility of confounding from multiple cycles of infection. Unfortunately, the inactivation problem still applies so the folding assay cannot be carried out on the envelope protein produced in the context of the whole virus. Thus it is necessary to compare the effects of *N*B-DNJ on the folding of recombinant gp120 expressed in transfected 293T cells to its effects on the infectivity of HIV virions produced from transfected 293T cells- not identical, but very similar systems. With

that caveat in mind, the effects of *N*B-DNJ on the secretion and infectivity of HIV LAI virions produced in HEK 293T cells is shown in Figure 4-6.

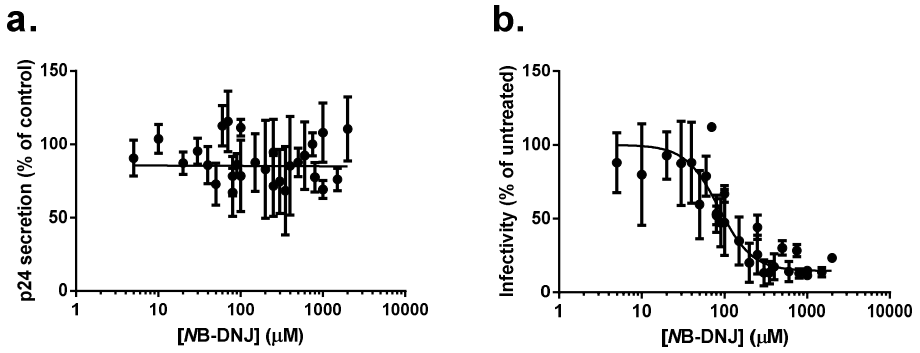



**Figure 4-6** The effects of *N*B-DNJ on HIV LAI produced in HEK 293T cells transfected with the BRU2 molecular clone. Shown are the drug's effects on the level of HIV secretion as determined by p24 ELISA (a) and the infectivity of said virions when normalised to equal concentrations as determined using the TZM-bl reporter cell line (b). Bars show standard deviation, n is the result of three or four replicates, each assayed for p24 with n = 3 or for infectivity with n = 4). Curves are plotted using either linear regression (a) or non-linear regression using a 4-parameter model in which the $E_{max}$ is constrained to 100%.

As can be seen in Figure 4-6 a, *N*B-DNJ does not significantly affect the production of HIV virions, as measured by the concentration of capsid protein p24 by ELISA, with the gradient of the linear model being statistically indistinguishable from the horizontal ($p = 0.94$). That said, the actual variation in the amount of HIV produced within treatment groups is highly variable, as evidence by the large standard deviations visible in Figure 4-6 a. To prevent this variation confounding the effects of the drug (e.g. a higher concentration of virus could be mistaken for less virus with a lower level of infectivity) all virus-containing supernatants were normalised into doses equivalent to 500 pg of p24, which were used to infect TZM-bl reporter cells to measure viral infectivity (b).

This assay also produced data with a large degree of variation, as visible in the error bars. However, fitting a regression line with a 4-parameter non-linear model reports an $IC_{50}$ (the concentration of drug that decreases infectivity to 50% of its untreated level) of 85.5 μM to within a 95% confidence interval of just 76.07 μM to 94.98 μM. This high degree of confidence can be achieved because of the large number of individual data (200 points) used to interpolate the curve. Also worthy of note is that *N*B-DNJ is not capable of reducing infectivity to zero, instead the $E_{min}$ trends towards 16% (95% CI 5.4% to 27.4%). This implies that the calnexin cycle is helpful, but not vital, for HIV infectivity. It is now possible to look at the effects of *N*B-DNJ on rgp120 misfolding and HIV infectivity and attempt to correlate the results (Figure 4-7).

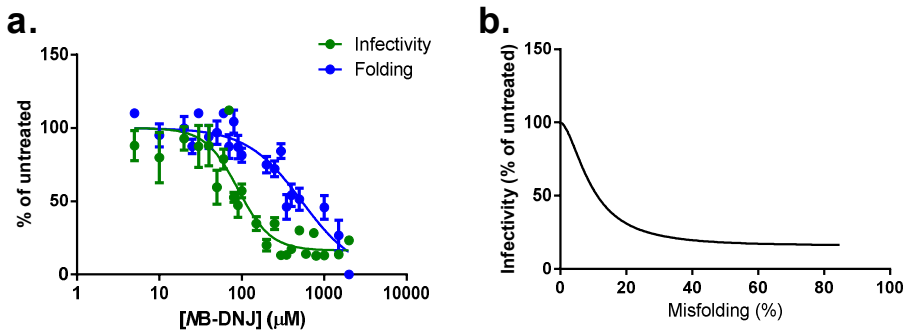



**Figure 4-7** The effects of *N*B-DNJ on the folding of $rgp120_{HXB2}$ (as detailed in Figure 4-3) and on the infectivity of HIV LAI (as detailed in Figure 4-6), both produced in HEK 293T cells (a). b shows these two curves plotted against one another such that one readily see the effects of misfolding (a transform of folding, simply $100(1\text{-}f_1)$, or 100-(% correctly folded)) on the infectivity of the virions. Bars show standard error, n is three or four replicates.

Increasing misfolding of gp120 appears to have a biphasic effect on virion infectivity, initially causing a steep drop (10% misfolding decreases infectivity by 37%) but by the time a fifth of the protein is misfolded the effect becomes severely diminished until beyond 40% misfolding no further effect on infectivity occurs (b). We have already established that *N*B-DNJ cannot reduce infectivity to below ~16% yet it reaches the upper bound of the confidence interval (27% infectivity) when only 24% of the protein is misfolded.

### 4.3.6 *N*B-DNJ increases the rate of rgp120 oxidative folding in the ER

Now we have an understanding of the effects of *N*B-DNJ on the secreted rgp120 it is worth taking a look at the effects of the drug on the actual folding process and ask how it is that glucosidase inhibition results in misfolded protein. It is known that glucosidase activity is vital for calnexin interaction, and that calnexin recruits chaperones that potentially aid the correct folding of proteins. gp120 is a useful model protein for studying folding in the ER because it is a large protein with 9 disulphide bonds. As these bonds form, the protein becomes more and more mobile on SDS-PAGE, presumably because these bonds compact the protein thus reducing its drag. Thus one only needs to look at the changes in gel motility of ER-stage gp120 to see how far oxidative folding has progressed.

Protein production in the ER is a massively parallel but asynchronous process, so at any one time the ER will contain proteins at all stages of folding. To study the effects of *N*B-DNJ on the folding of gp120 it will be necessary to follow a synchronous population. This can be achieved using radioactive pulse-chase labelling. In this technique, cells transfected to produce rgp120 are starved of sulphur containing amino acids and then fed radioactively labelled ones for a five minute period. When the five minutes are up the radioactive amino acids are removed and the cells are flooded

with unlabelled amino acids; this means that only the proteins translated in that five minute window will contain radioactive methionine and cysteine residues. The cells can then be left to continue the normal folding process for as long as required whereupon they can be treated with alkylating agents (to freeze all disulphides in their current oxidation states) and lysed. The rgp120 can then be purified from the lysates by immunoprecipitation and run on SDS-PAGE. Autoradiography of the gel will then show up only the positons of the radioactive rgp120, thus showing you the mobility of just that synchronous population.

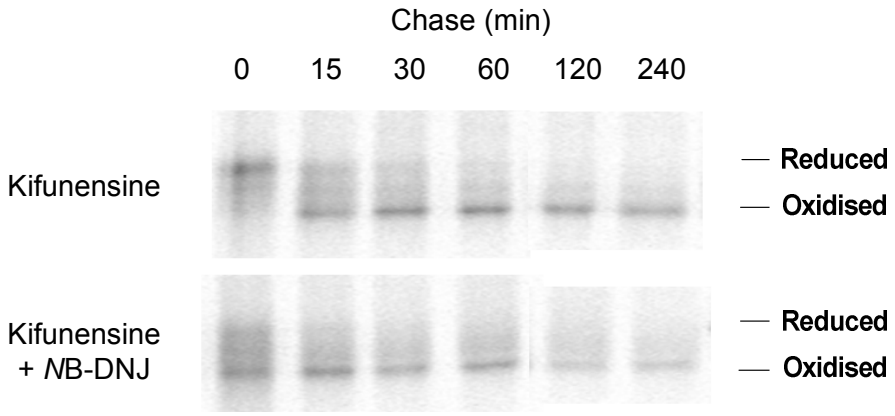



**Figure 4-8** Autoradiograph of non-reducing SDS PAGE of radioactively pulse-labelled $gp120_{LAI}$ expressed in HeLa cells. Each lane represents a separate experiment stopped at the indicated number of minutes after the completion of a ten minute pulse ("chase"). Cells were pre-treated with 20 μM kifunensine ± 2 mM *N*B-DNJ, as indicated, for 45 minutes and these concentrations were then maintained during the pulse and chase periods. The "reduced" annotation marks the minimum motility of the protein and hence represents the reduced population, while "oxidised" represents the most compact, fully oxidised species. While both data for the same timepoints were performed on the same gels, the timepoints 120 and 240 minutes were performed on a separate gel to the other timepoints. The two gels have been digitally composited to appear as one, for clarity. Data are representative of two replicates.

This technique was used to produce a time course of rgp120 folding both in the presence and absence of 2 mM *N*B-DNJ. To maintain consistency with previous work all experiments were performed in the presence of 20 μM kifunensine (Figure 4-8 and Figure 4-9).

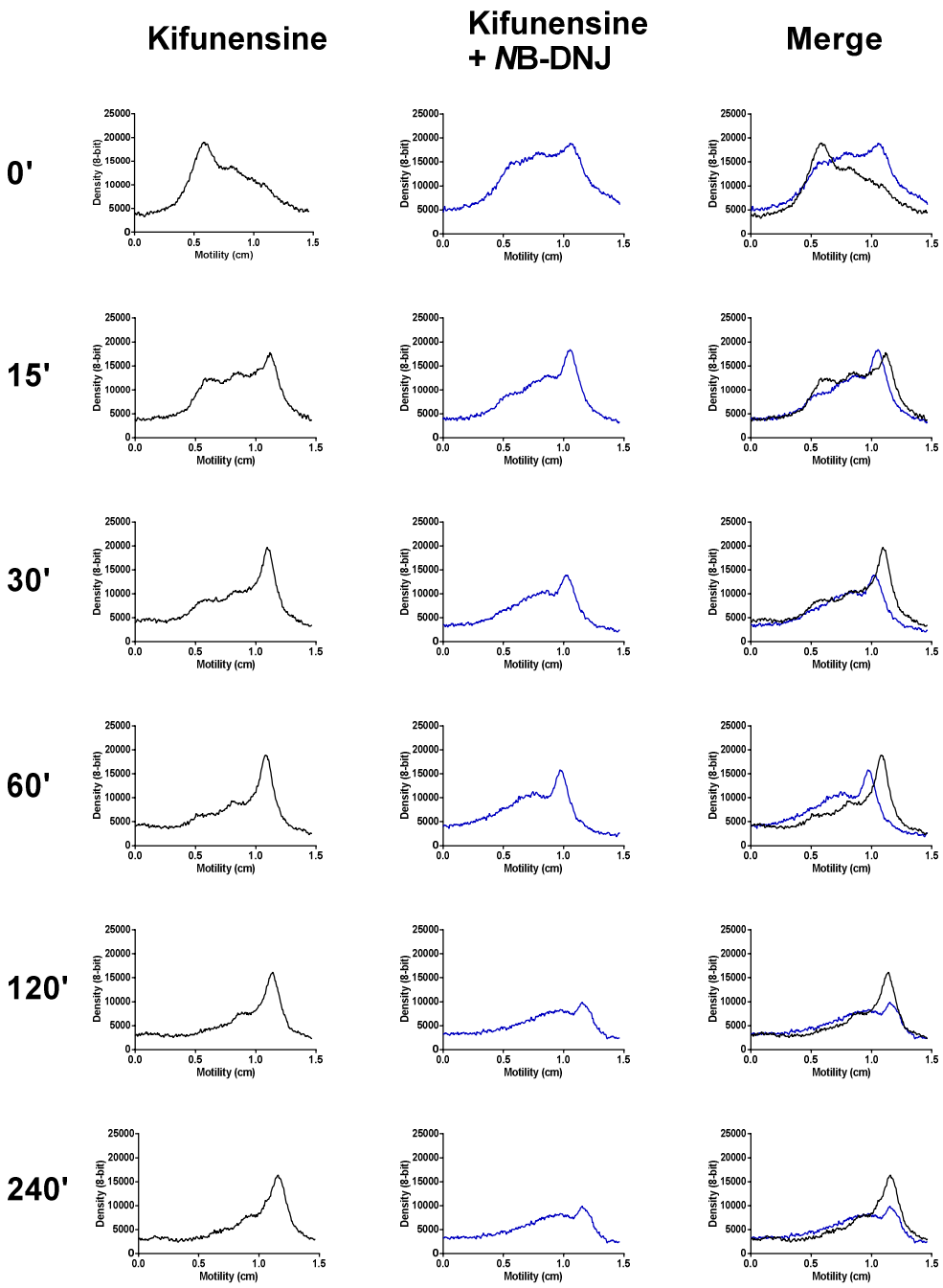



**Figure 4-9** Lane profiles of the autoradiographs presented in Figure 4-8. The profiles show the gel motility (Motility) of rgp120$_{IIIB}$ expressed in HeLa cells at the indicated number of minutes after translation against the concentration of each species as measured by 8-bit densitometry (Density). Motility is expressed as centimetres travelled from an arbitrary, but shared, point on the gel.

Looking at both at Figure 4-8 and Figure 4-9 together, it is first worth understanding how the oxidative folding of rgp120 progresses in the absence of *N*B-DNJ. Immediately following the pulse (0' chase) the major band/peak is relatively poorly mobile, representing protein that is almost completely reduced, then there is a "smear" of more mobile species beyond it. By 15' chase the dominant peak is the most motile band, though the majority of the protein is still either in the

original band or in the smear. This motile band represents fully oxidised rgp120, in which all the disulphide bonds are formed, while the smear represents intermediate forms which are partially oxidised. As the chase progresses the proteins move from the reduced population, through the intermediates, into the oxidised population. This is probably not a one-way transition and it is likely that the disulphides will have to isomerise, in a constant process of redox reactions with protein-disulphide isomerases, moving it in and out of the smear. However, as time progresses the oxidised band becomes dominant and the reduced band disappears.

In the presence of 2 mM *N*B-DNJ the 0' timepoint more closely resembles the 15' than 0' control point, suggesting that the rgp120 oxidises far more quickly. However, later in the process, from 60', the oxidised peak is still not dominant, with most of the protein still being in an intermediate oxidation state. Even by 240' the oxidised peak is still relatively small compared to area of the intermediate "peaks". A possible explanation for this is that *N*B-DNJ induces the protein to oxidise much more quickly than normal. However the disulphide bonds it forms are not native, requiring extensive isomerisation

If this hypothesis is true then one might expect that *N*B-DNJ would cause rgp120 to remain in the ER for a longer period than usual so that it can undergo this isomerisation. Luckily, gp120 has an unusual and useful quirk that means that its folding time can be accurately measured. Whereas most proteins have their signal peptides removed co-translationally, the signal peptide of gp120 is only removed once the protein has finished folding. Thus signal peptide cleavage can be used as marker for the completion of folding. The cleavage of the signal peptide can be measured by reducing the rgp120 with dithiothreitol (DTT) prior to SDS-PAGE. This reduces all of the disulphides forcing all of the proteins into a single, reduced, population at all time points ($R_U$, reduced uncleaved). The removal of the signal peptide will result in a reduction in the mass of the protein which can be visualised as a jump in motility on the autoradiograph ($R_C$, reduced cleaved). Densitometry can then be used to measure the concentration of $R_C$ as a percentage of the whole rgp120 population, thus giving a measure of the proportion of rgp120 that has finished folding (Figure 4-10). Densitometry was restricted to timepoints of 60 minutes or less as the $R_U$ band was so faint beyond this point that it became difficult to accurately distinguish a single peak for it. Both by eye (a) and by densitometry (b) it can be seen that the rgp120 appears to be cleaved sooner when treated with *N*B-DNJ. Taking the results from Figure 4-9 and Figure 4-10 together one can conclude that *N*B-DNJ causes rgp120 to not only oxidise faster but also to reach a conformation recognised by the cell as "folded" sooner. This is in agreement with the work from the previous

chapter that shows that the secreted protein is of a near native conformation, but the accelerated nature of the folding could account for the failure to reach the native V2 loop conformation.

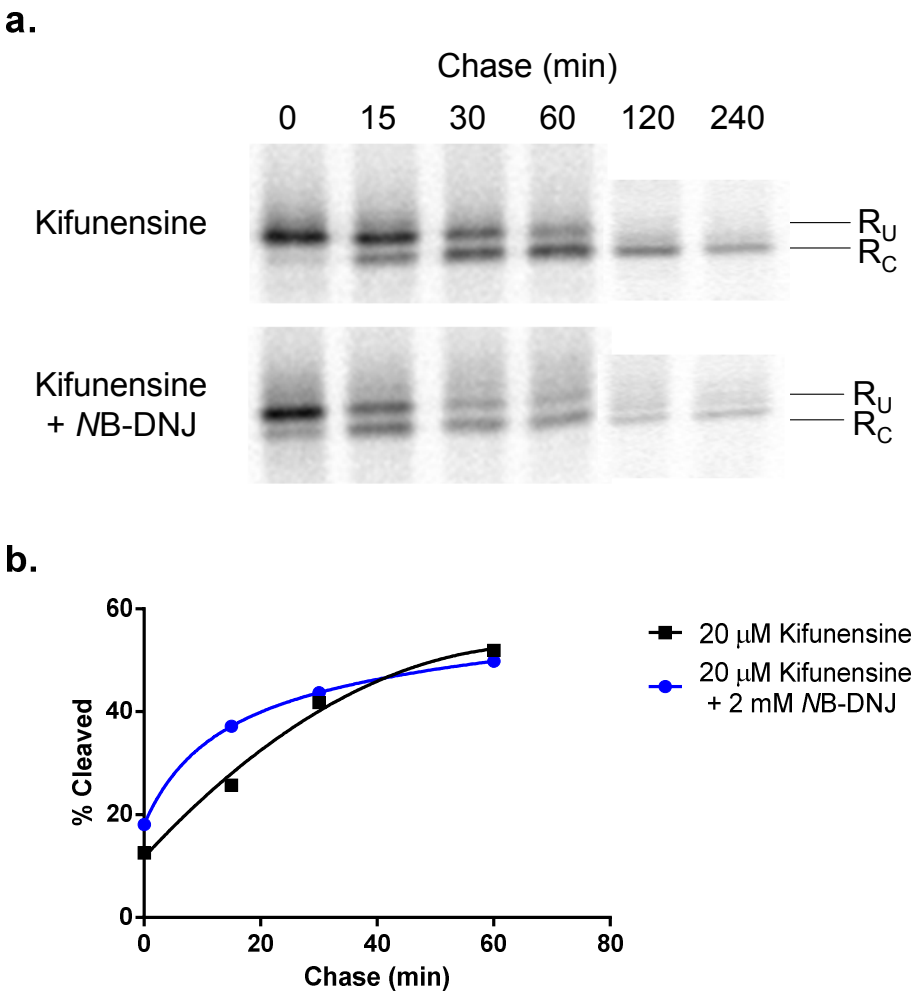



**Figure 4-10** Cleavage of the signal peptide from radiolabelled $rgp120_{LAI}$ divides the protein into two populations visible on reducing SDS-PAGE, reduced uncleaved ($R_U$) and reduced cleaved ($R_C$). These two populations are distinguished by the lower mass, and hence greater gel motility, of $R_C$, which comes to dominate over $R_U$ as folding progresses. This change can be visualised using radioactive pulse chased material (a). Densitometry can be used to measure the relative proportion of rgp120 that is in the cleaved state- this can be used to measure the speed of folding (b). Data are representative of two replicates.

### 4.3.7 $HIV_{LAI}$ mutants lacking potential *N*-glycosylation sites have altered susceptibility to *N*B-DNJ

In the 2008 paper *Pollock et al.* the authors determined that while *N*B-DNJ was effectively antiviral against a range of HIV strains, the $IC_{50}$ was variable [214]. Since *N*B-DNJ targets a cellular enzyme this suggests that the different strains have different requirements of α-glucosidase activity to achieve infectivity. Pollock found that susceptibility to *N*B-DNJ was positively correlated with the number of occupied *N*-glycosylation sites in the C2/V3 region of the gp120; no correlation was

found with glycan occupancy in C1/V1/V2. This appears at odds with the conclusion that it is V2 that is misfolded as a results of *N*B-DNJ activity. To investigate this hypothesis molecular clones of LAI were produced each lacking one of the potential *N*-glycosylation sites (PNGS) across C1-V3, which were then tested for their susceptibility to the drug.

Mutants were designed based on the work of *Wang et al.* who made systematic mutations of PNGS sites across the FE strain of HIV and tested them for effects on infectivity [215]. All PNGS deletions were made by substituting the asparagine residue for an alternative amino acid. While this was usually glutamine in some cases these mutants were not infectious and alternative substitutions were trialled, such as serine or aspartate. In each case, I have used the substitution that resulted in the highest infectivity virions. Where a glycan site exists in LAI that does not exist in FE the substitution of the nearest shared PNGS was used. All the attempted mutations are shown in Table 4-1. PNGS occupancy (whether or not a glycan is likely to be added there) was predicted using the NetNGlyc 1.0 server (www.cbs.dtu.dk/services/NetNGlyc); not all glycan sites were predicted to be occupied but, as this was only a prediction, the mutants were produced anyway.

Unfortunately, only seven of the mutants were successfully produced- N88D, N143cQ, N191D, N197D, N230D, N234D and N241S, representing sites in C1, V1/V2, the bridging sheet and C2. These mutants were expressed in HEK 293T cells using transfections of both 5 μg DNA per well, as with the wild-type, or just 1 μg per well. After 2 days the supernatants were assayed for HIV concentration by p24 ELISA, as previously. All mutants were found to express at least as well as the wild type although, interestingly, the cells looked abnormal and produced less p24 when transfected with the 5 μg cf. 1 μg (Figure 4-11 a). This may be because the mutant proteins are less able to fold than the wild type and accumulate in the ER, triggering stress responses. Using the lower concentration avoids this complication and allows for better survival of the cells, leading to higher overall secretion. For this reason, 1 μg DNA was used for subsequent experiments with the mutant clones, while 5 μg was still used for the wild type.

| Region | Mutation | Occupancy |
|---|---|---|
| C1 | N88D | + |
| V1/V2 | N136Q | + |
| | N141Q | + |
| | N143cQ | + |
| | N156D | - |
| | N160D | - |
| | N186D | + |
| Bridging Sheet | N197D | + |
| C2 | N230D | + |
| | N234S | + |
| | N241S | + |
| | N261D | + |
| | N276D | + |
| | N289D | - |
| V3 | N295Q | + |

**Table 4-1** Mutations introduced to potential N-glycosylation sites in HIV LAI. PNGS occupancy was predicted using the NetNGlyc 1.0 server. + indicates an above threshold probability of PNGS occupancy, - below threshold. All numbering is relative to the HXB2 standard, with lower case letters representing insertions not found in HXB2.

Several of the glycan mutants showed decreased infectivity, with mutants N230D and N234S abolishing infectivity all together (b). This is contrary to the *Wang et al.* study that found no substantial effect on infectivity when the N243S mutation was introduced into FE, perhaps indicating variation in glycan requirements across strains. N143cQ was the only mutation to have no effect on infectivity ($p = 0.273$, one-way ANOVA with Sidak correction for multiple comparisons). LAI is identical to the HXB2 strain, from which the numbering is derived, except for a duplication of the amino acids from 138 to 143 (labelled 143a-e). Since 143c is then a duplication of a glycan just 5 residues away it is not surprising that it can be removed without effect.

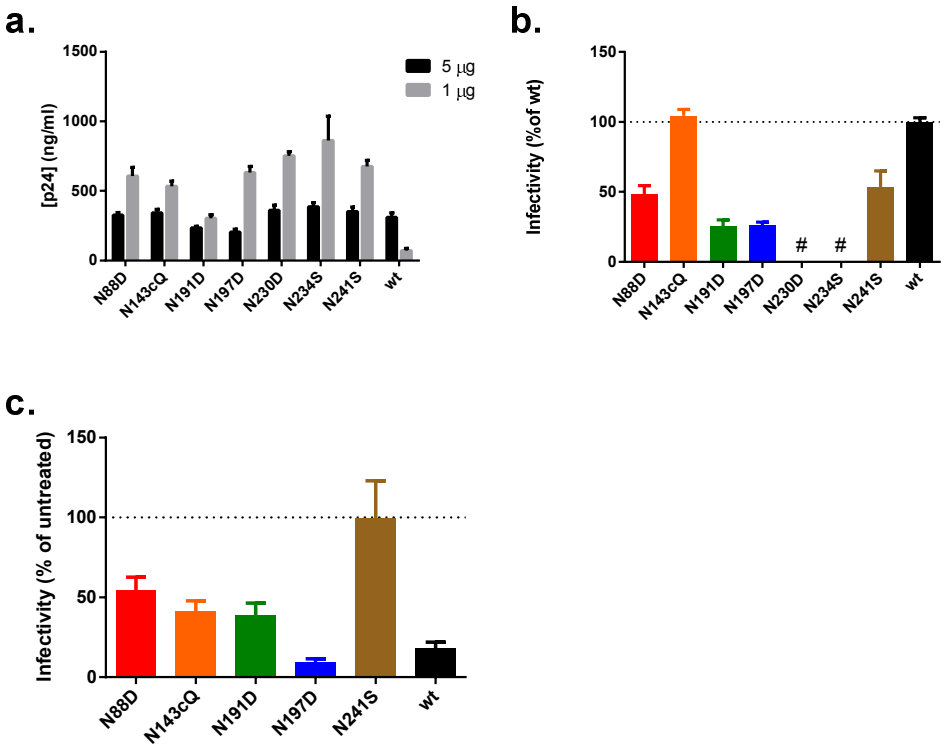



**Figure 4-11** The effects of mutating out potential *N*-glycosylation sites in HIV LAI on the production (a) and infectivity (b) of HIV virions, and the effects of *N*B-DNJ on that infectivity (c). The expression of each mutant was tested by transfecting HEK 293T cells with 1 or 5 µg of plasmid DNA bearing each molecular clone in the wells of 6-well plate and then measuring p24 produced by ELISA 48 hours later (a). Supernatants were normalised to 500 pg p24 equivalent doses which were then used to infect TZM-bl reporter cells, with the level of infection reported by luciferase assay and expressed as a percentage of wild type LAI (b). Simultaneously, identically transfected cells were treated with 250 µM *N*B-DNJ. The ratio was found of the infectivity of virions produced in the treated cells to that of those produced in the untreated cells (c). Bars show standard deviation, ELISAs were conducted in triplicate while luciferase assays were conducted in sextuplicate. # indicates infectivity was indistinguishable from background. RLU, relative light units; wt, wild type.

Figure 4-11 c shows that treatment with *N*B-DNJ causes a reduction in infectivity of all of the mutants except for N241S, which is unaffected by the drug. This finding could explain how *N*B-DNJ induces misfolding of the V2 loop and hence is antiviral. One can hypothesise that this glycan is vital for calnexin dependent V1/V2 loop folding- deletion of the glycan will therefore cause a decrease in infectivity, as seen in Figure 4-11b. The infectivity is not reduced to zero because, as we know from titrating *N*B-DNJ against the wild type, calnexin is only helpful, not vital, to HIV production. Thus we can further hypothesise that, in the wild type, *N*B-DNJ treatment leads to the maintenance of glucose residues on the N241 glycan preventing it from associating with calnexin and thus preventing efficient V1/V2 folding. *N*B-DNJ can, therefore, have no effect on the N241S mutant as there is no glycan there to be affected.

Figure 4-11c shows that treatment with *N*B-DNJ causes a reduction in infectivity of all of the mutants except for N241S, which is unaffected by the drug. This finding could explain how *N*B-DNJ

induces misfolding of the V2 loop and hence is antiviral. One can hypothesise that this glycan is vital for calnexin dependent V1/V2 loop folding- deletion of the glycan will therefore cause a decrease in infectivity, as seen in Figure 4-11 b. The infectivity is not reduced to zero because, as we know from titrating *N*B-DNJ against the wild type, calnexin is only helpful, not vital, to HIV production. Thus we can further hypothesise that, in the wild type, *N*B-DNJ treatment leads to the maintenance of glucose residues on the N241 glycan preventing it from associating with calnexin and thus preventing efficient V1/V2 folding. *N*B-DNJ can, therefore, have no effect on the N241S mutant as there is no glycan there to be affected.

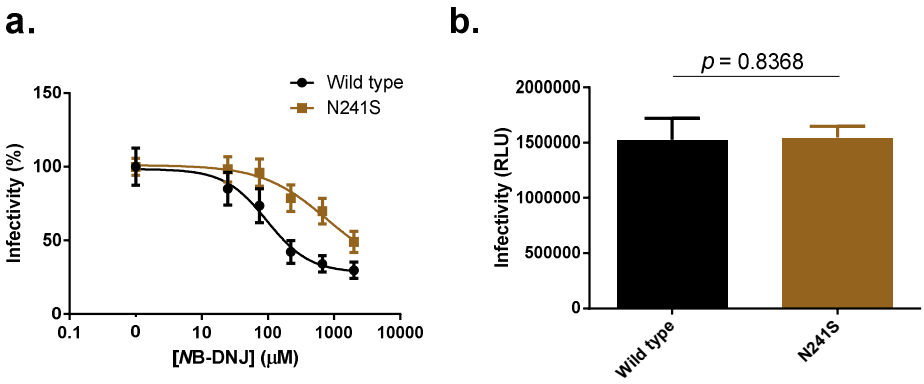



**Figure 4-12** The effect of *N*B-DNJ on the infectivity of the N241S mutant, which lacks a single *N*-glycosylation site in the C2 region. Either wild-type or N241S HIV LAI viruses were grown in HEK 293T cells treated with varying concentrations of *N*B-DNJ. The supernatants of these cells were then adjusted into 500 pg doses and used to infect TZM-bl reporter cells to gain a measure of their infectivity (a). The infectivity of untreated wild-type and N241S HIV LAI is shown in (b). Bars show 95% confidence intervals for two (wild-type) or three (N241S) biological replicates each with seven replicates of the TZM-bl assay. Data are normalised to the infectivity of the untreated control and curves are fitted using four component non-linear regression constrained to an upper bound of 100%. The $p$ value in (b) is generated using a two-tailed Student's $t$ test. Abbreviation: RLU, relative light units.

To investigate this hypothesis further *N*B-DNJ was titrated against the N241S mutant to confirm these results and see if the drug became effective at higher concentrations (Figure 4-12). This showed that *N*B-DNJ is able to inhibit the infectivity the N241S mutant (a), but the curve is significantly (F test, $p < 0.0001$) right shifted, thus demonstrating increased but not total resistance to the drug. Interestingly, when more replicates are included there is no significant difference between the infectivity of the untreated mutant versus that of the wild-type, in contrast to that observed in the pilot study above (Figure 4-11c).

These data suggest that the original hypothesis cannot be correct- the N241 glycan cannot be vital for V1/V2 folding as the virus appears to be completely competent without it. However, deletion of this site does then give the virus significant and substantial resistance to *N*B-DNJ treatment.

# 5 Mannosidase Inhibitors as Antiretroviral Drugs

## 5.1 Contributions

Glycan profiling was performed by Laura Pritchard and the endomannosidase digest was done by Dominic Alonzi. Nicholas McCaul performed the radioactive pulse-chase labelling experiment.

## 5.2 Results

### 5.2.1 Kifunensine is antiviral against molecular clones of HIV LAI

As much of the previous work with *N*B-DNJ against rgp120 (see Chapters 3 and 4) was performed in the presence of kifunensine, to control for glycosylation differences, it was important to check that kifunensine did not have any effects on the gp120 population beyond restricting it to $Man_9GlcNAc_2$. The easiest way to see whether it had any significant effects of its own was to observe whether it, either alone or in combination with *N*B-DNJ, had any antiviral activity. HIV LAI was expressed from a molecular clone in HEK 293T cells as previously, in the presence of a titration of *N*B-DNJ and either 20 μM kifunensine or the equivalent volume of phosphate buffered saline (PBS) (Figure 5-1). 20 μM kifunensine was chosen as this is the concentration used in previous studies on rgp120. At all concentrations of *N*B-DNJ, including 0 μM (untreated), 20 μM kifunensine was antiviral (a), reducing infectivity by up to 85% (b). Looking only at the untreated population, kifunensine also reduced the secretion of HIV p24 by 37.5% (c).

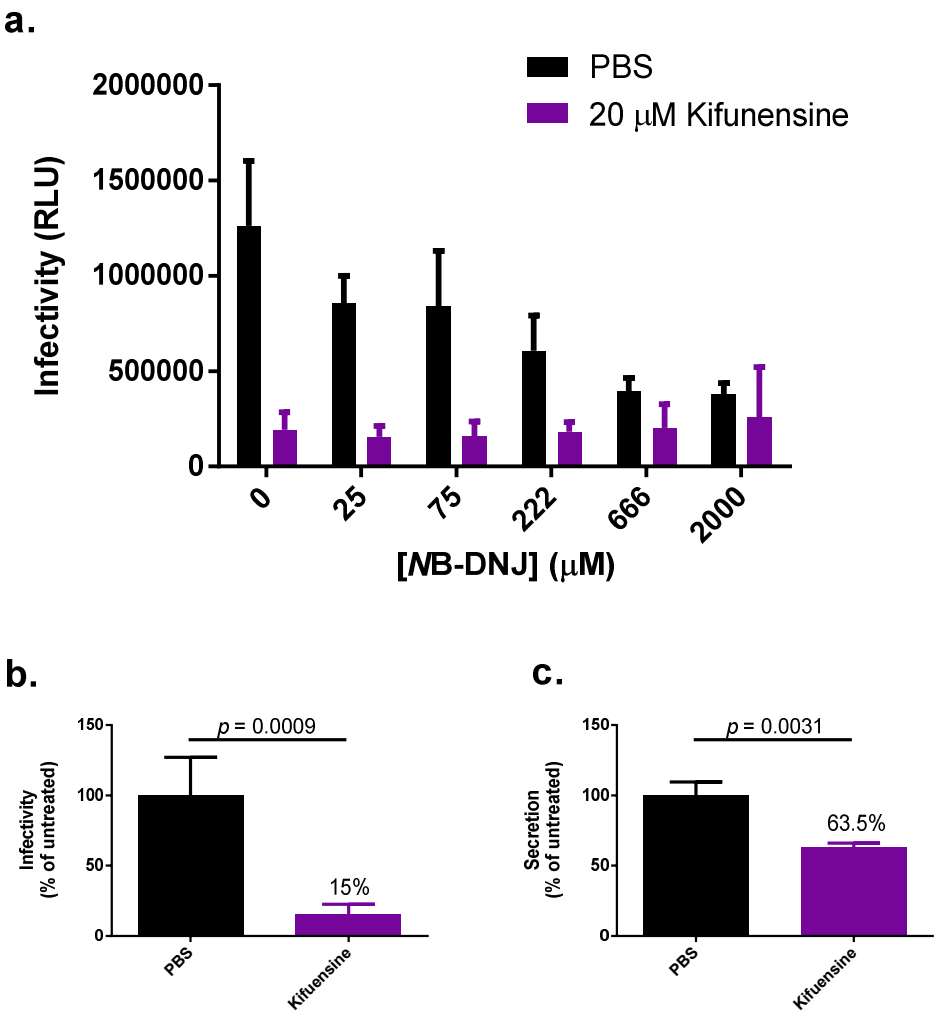



**Figure 5-1** Kifunensine was investigated for antiviral activity either alone or in synergy with *N*B-DNJ. HIV LAI was expressed in HEK 93T cells, as previously, against a titration of *N*B-DNJ and either 20 μM kifunensine or a PBS control. The supernatants were harvested 2 days later, normalised into 500 pg p24 equivalent doses and used to infect TZM-bl reporter cells, from which the level of HIV infection could be calculated by luciferase assay (a). Looking only at those cells untreated with *N*B-DNJ (b) it is clear that 20 μM kifunensine is antiviral even in the absence of *N*B-DNJ. It also has an effect on the secretion of HIV particles as measured by p24 ELISA (c). Bars show standard deviations, n = 4 for infectivity assays and 3 for secretion assays. Analyses in b and c are by two-tailed student's *t* test. RLU, relative light units.

To confirm the veracity of these data, transfected 293T cells were incubated with a titration of kifunensine, in the absence of any other treatments, and again assayed for HIV p24 secretion and virion infectivity Figure 5-2.

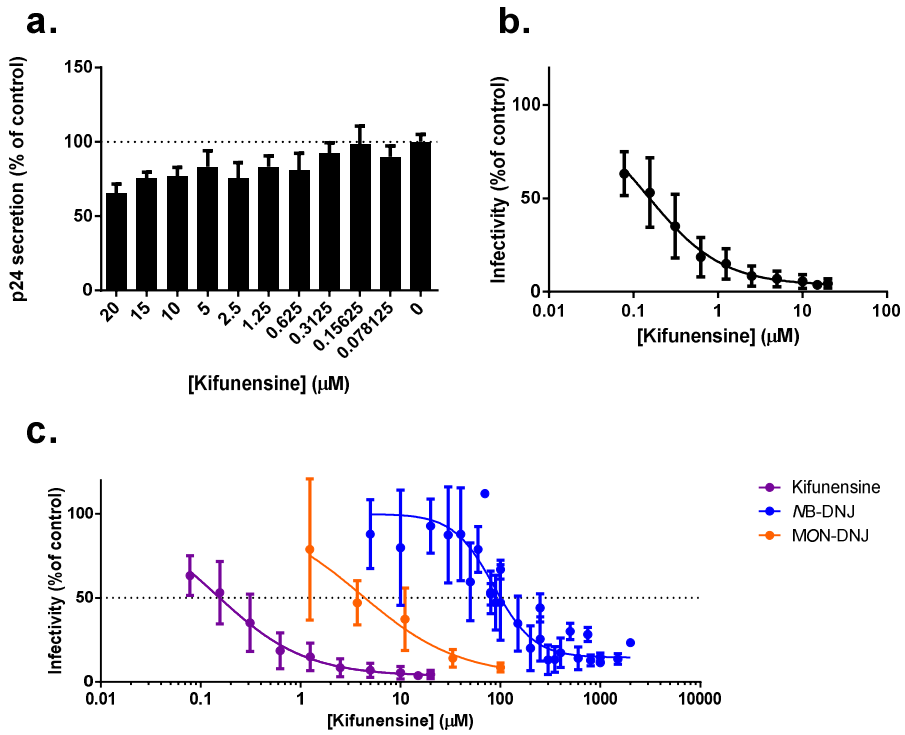



**Figure 5-2** Kifunensine causes a dose-dependent decrease in HIV LAI secretion, as measured by p24 ELISA (a), and in the infectivity of those virions, as determined by luciferase activity in TZM-bl reporter cells infected with 500 pg of p24 equivalents (b), both shown as a percentage of that produced from untreated cells. The effects of kifunensine on HIV LAI infectivity are compared with the effects of two other iminosugars, *N*B-DNJ and M*O*N-DNJ. Bars show standard deviation; a and b have an n of 3 and 4 respectively, both performed in triplicate, while c has an n of 4 performed in triplicate, singlicate and quadruplicate for kifunensine, M*O*N-DNJ and *N*B-DNJ respectively. The dotted line in c indicates the point on each curve where infectivity has been reduced by 50% (the $IC_{50}$).

These data show that kifunensine has a dose dependent effect on virion infectivity and, unlike other iminosugars, on secretion. Treatment with kifunensine causes a decrease in p24 secretion of up to 35% ($p = 0.0092$, one way ANOVA with Sidak correction for multiple comparisons), agreeing favourable with the data obtained previously (Figure 5-2a and Figure 5-1). Even when this secretion difference is controlled for by normalising HIV concentrations the virions are still less infectious than those left untreated (b). This effect is two orders of magnitude greater than that of *N*B-DNJ, and one order of magnitude greater than that of M*O*N-DNJ, an iminosugar currently undergoing phase I clinical trials for use in humans against dengue virus

| Iminosugar | IC50 (μM) | 95% Confidence Interval (μM) |
|---|---|---|
| Kifunensine | 0.1445 | 0.1205 to 0.1685 |
| M*O*N-DNJ | 4.073 | 0 to 8.291 |
| *N*B-DNJ | 85.52 | 76.07 to 94.98 |

**Table 5-1** The relative potencies of three iminosugars against the infectivity molecular clones of HIV LAI expressed from HEK 293T cells.

### 5.2.2 The antiviral effects of kifunensine are not due to ERAD inhibition

Kifunensine has two known biological effects, each of which may be the source of its antiviral activity. Both of these effects stem from its ability to inhibit mannosidases- general mannosidase inhibition prevents processing of *N*-linked glycans beyond $Man_{8/9}GlcNAc_2$ as explained in Section 3.3.1 . However, kifunensine can also inhibit the ER degradation enhancing α-mannosidase like protein (EDEM). This protein is responsible for the escape of misfolded proteins from the calnexin cycle to ER associated degradation (ERAD). Proteins that persist in the ER for too long, as a result of repeated reglucosylation by UGGT1, have their D3 mannose removed in a time dependent fashion by EDEM2. This makes the glycan a ligand for a $Man_8GlcNAc_2$ binding lectin OS-9 that targets the protein for retrotranslocation to the cytoplasm where it is degraded by the proteasome.

It is possible to hypothesise that ERAD inhibition may have antiviral properties. Knowing that *N*B-DNJ is antiviral by forcing the incorporation of poorly folded proteins into the virion envelope it is possible that, by inhibiting ERAD, envelope protein that would otherwise fail quality control could end up in the virion. This could have a similar effect to *N*B-DNJ, wherein very few poorly folded proteins could have an amplificative effect resulting in the inability of the virus to fuse with the host cell.

To test this hypothesis cells transfected with the LAI molecular clone were treated with eeyarestatin I (EeyI), an ERAD inhibitor that works independently of mannosidase inhibition, instead inhibiting p97, an enzyme associated with the retrotranslocation complex that removes misfolded proteins from the ER. By inhibiting this pore complex misfolded proteins will persist in the ER, and potentially become incorporated into the virion, without their glycans being affected. Thus if EeyI is able to reduce the infectivity of virions then it supports the hypothesis that ERAD inhibition alone could be the antiviral mechanism of kifunensine.

EeyI is reported to be active at concentrations around 10 μM, but unfortunately this concentration proved highly toxic if HEK 293T cells were treated for a whole 48 hours as with kifunensine (Figure 5-3).

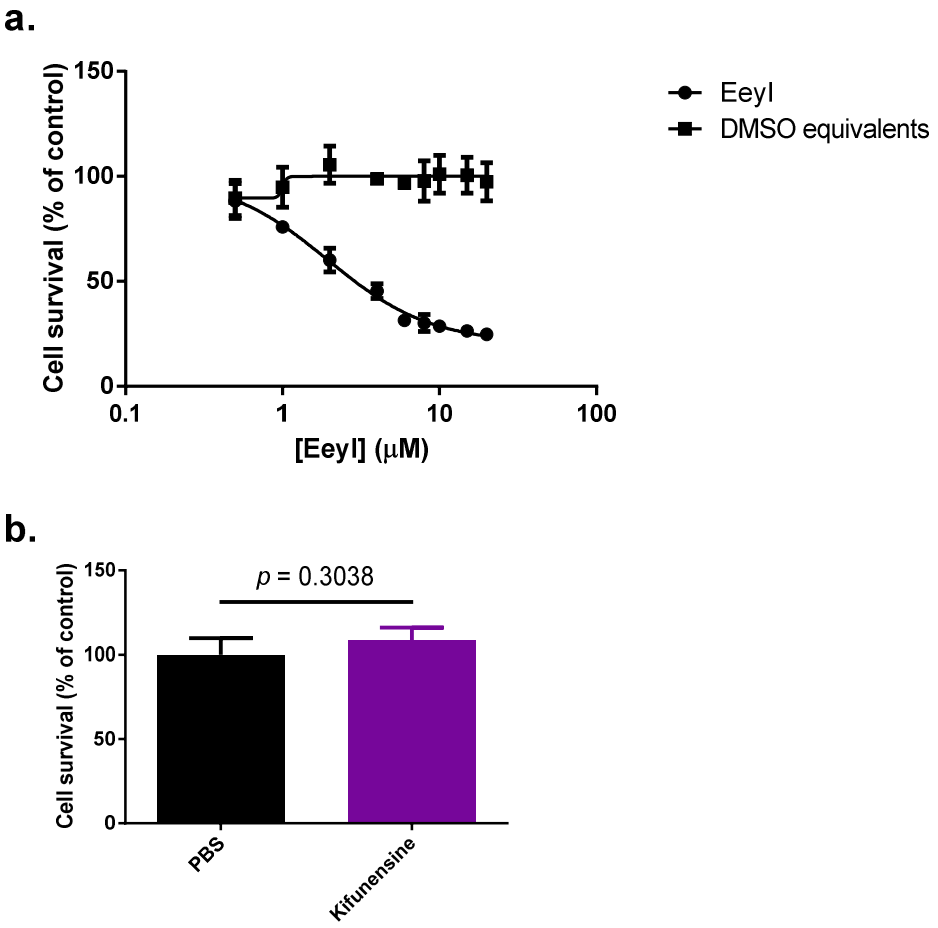



**Figure 5-3** Treatment with Eeyarestatin (EeyI) for 48 hours is toxic. HEK 293T cells were treated with a titration of EeyI, or an equivalent volume of DMSO, for 48 hours (a). 20 μM kifunensine, by comparison, had no effect on viability over the same time period (b). Viability was then measured using the AqueousOne Cell Proliferation Assay and expressed as a percentage of the viability of the untreated cells, minus background. Bars show standard deviation, n = 3, *p* value obtained using two-tailed Student's *t*-test. .

48 hour treatment was obviously not possible, but treatment of 293T cells for 12 hours with 10 μM EeyI is reported in the literature. Thus it was decided to treat the transfected cells for 8 hours, a timepoint determined as a reasonable length of time for adequate virus production while leaving as little time as possible for toxicity. Since toxicity was still possible the cell viability would be determined after the removal of the infected supernatant at 8 hours, thus allowing a direct measure of toxicity on the cells in question. The treated and untreated cell supernatants could then be normalised to the same p24 concentration, thus accounting for any decreased secretion due to cell toxicity. Thus, if toxicity were detected then it would be difficult to be conclusive on the effects of ERAD inhibition on virion secretion, but it may still be possible to detect effects on the infectivity of those particles (Figure 5-4).

EeyI is still substantially toxic after just 8 hours of treatment (Figure 5-4 a), a finding which casts a degree of uncertainty onto all following data. For example, the finding that p24 secretion is decreased by EeyI (b) is thoroughly equivocal as this is far more likely to be due to toxicity than ERAD inhibition. The effect of EeyI on infectivity is potentially more interesting because these experiments were performed using viral material that was normalised to equivalent p24 concentrations (c). Thus it may be that EeyI is actually capable of decreasing virion infectivity. However, there are still caveats to this finding.

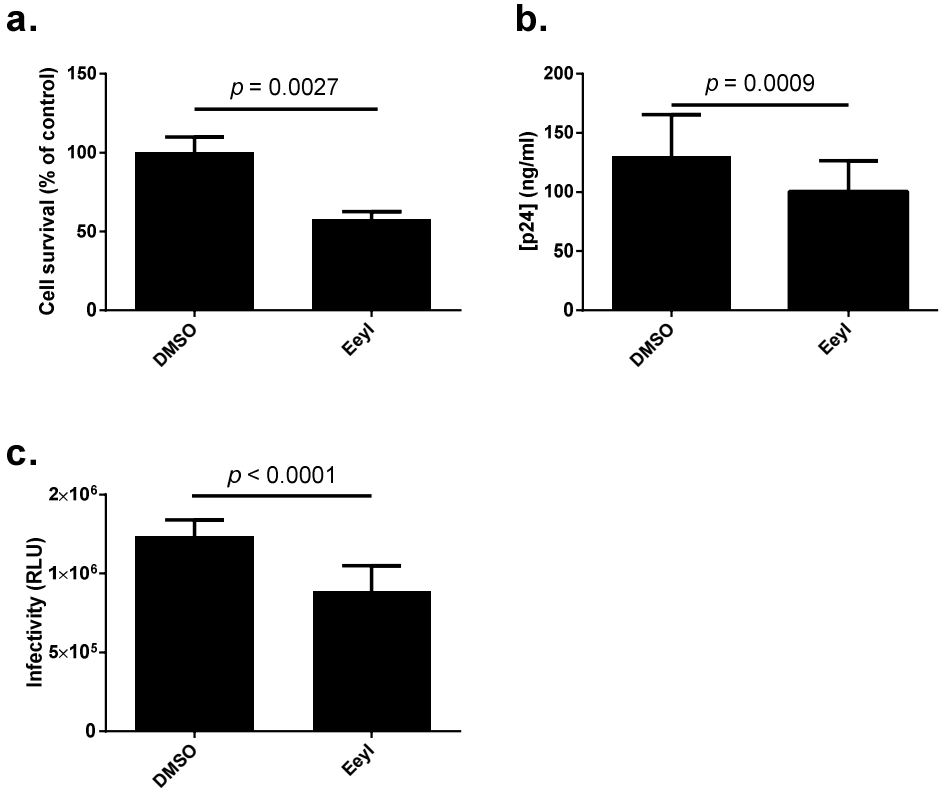



**Figure 5-4** Treatment of HEK 293T cells with 10 μM EeyI for even just 8 hours is toxic, causing a decrease in cell viability of 42% (a). However, in that time virion production is relatively unaffected, and EeyI treated cells produce 77% as much virus as DMSO treated cells (b). If the supernatants are adjusted such that 500 pg of p24 equivalents are used to infect TZM-bl reporter cells then the virions produced in EeyI treated cells are 28% less infectious than those from DMSO treated cells (c). All bars show standard deviations; the experiment was repeated three times with cell viability assays performed in singlicate, p24 ELISAs in nonuplicate and infectivity assays in septuplicate. *p* values for a and b were obtained by two-tailed Student's *t*-test and for c by two-way ANOVA.

Firstly, it is possible that the changes in the virions are due to the fact that the cells are dying or otherwise stressed. Such changes are likely to interfere with the correct packaging of the virus as well as production of its key proteins, so even though virion number is the same virion quality is not solely ERAD dependent. A second problem is that the supernatant used to infect the TZM-bl reporter cells still contains EeyI by necessity, as it cannot be separated from the virus. Assuming its concentration has not been at all affected by the HEK 293T cells (such as by degradation or it

being sequestered in cell compartments or by cell proteins) then the TZM-bl cells were exposed to a maximum concentration of 0.25 μM EeyI, one fortieth the concentration that the 293T cells were exposed to. Extrapolating from Figure 5-3, this concentration is not predicted to be toxic in 293T cells but its effects on TZM-bl cells and on luciferase production are unknown. A potential solution to this would be to treat the control TZM-bl cells with EeyI as well, though this still would not account for the problems from general cell toxicity. For these reasons, the role of ERAD inhibition as an antiviral mechanism of kifunensine remains unresolved.

### 5.2.3 The antiviral effect of kifunensine correlates to its effects on gp120 glycosylation

The alternative possibility is that kifunensine is antiviral because HIV requires the presence of complex, or at least trimmed high-mannose, glycans for the correct functioning of the envelope glycoprotein; the purely $Man_{8/9}GlcNAc_2$ phenotype induced by kifunensine is somehow substantially suboptimal. As set out in the introduction, this hypothesis is unlikely for two main reasons. Firstly, other mannosidase inhibitors have been tested as antiretroviral agents and have generally been found to be ineffective. Secondly, kifunensine is not reported to have any effect on gp120 EndoH sensitivity (an endoglycosidase that only cleaves off high mannose glycans) below 500 nm, and not to have a substantial effect below 5 μM. Similarly, the binding of 2G12, an antibody that is specific for high mannose glycans on gp120, is not affected below 5 μM [150]. Since the $IC_{50}$ of kifunensine against LAI clones in 293T cells is nearer 145 nM and the effect is maximal by 5 μM it seems unlikely that the antiviral effect stems from the change in glycosylation.

To test whether this is the case, HEK 293T cells were transfected with $His_6$ tagged $rgp120_{BaL}$ and treated with a titration of kifunensine. The rgp120 was purified by nickel column affinity and then had all glycans removed and labelled. These glycans were subsequently identified by HPLC. Although BaL was not the strain used in the previous experiments, its gp120 can be expressed in $His_6$ tagged recombinant form at far greater levels than that of LAI, thus making it better suited to this purpose when large quantities of material are needed. The glycan profiles are presented in Figure 5-5.

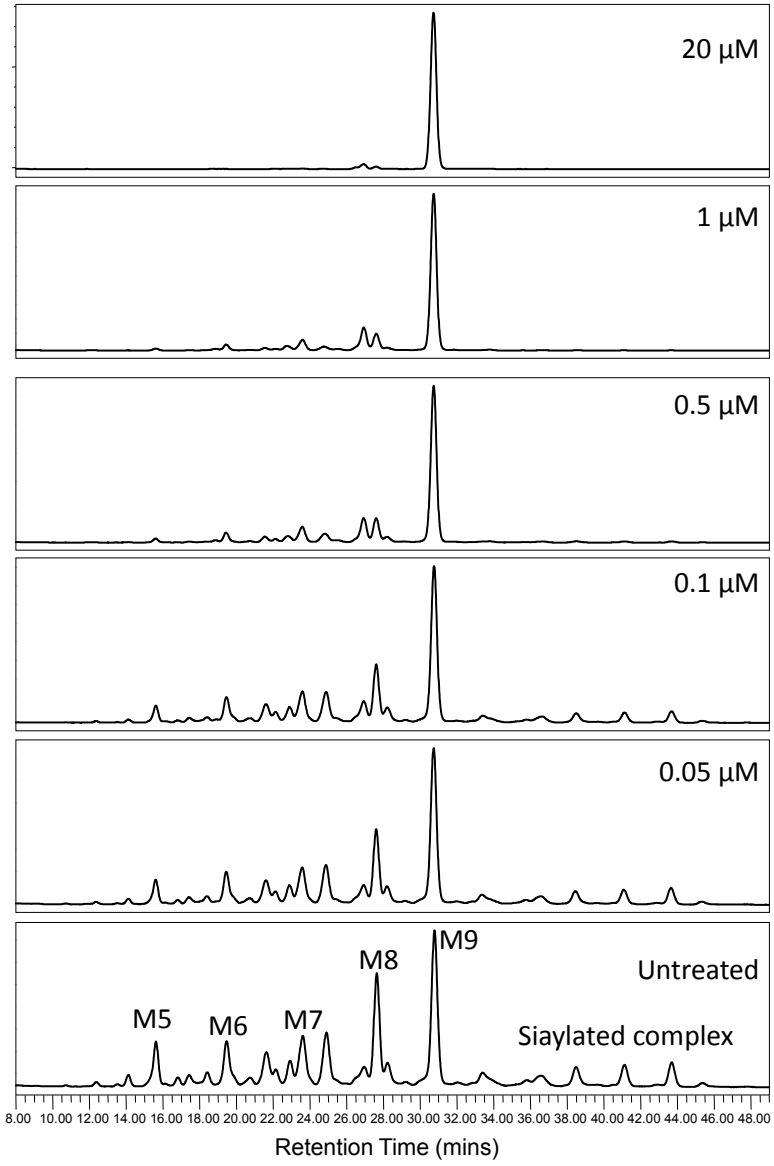



**Figure 5-5** Glycan profiles of $rgp120_{BaL}$ produced in HEK 293T cells treated with the indicated concentration of kifunensine. The *y* axes are arbitrary units representing fluorescence of the labelled glycans and are not all to the same scale. Abbreviations: M5, $Man_5GlcNAc_2$, M6, $Man_6GlcNAc_2$ *et cetera*.

These data show a clear dose-dependent effect of kifunensine on the glycans of $rgp120_{BaL}$, from 50 nM treatment giving a near native profile while 20 µM gives an exclusively $Man_9GlcNAc_2$ phenotype. As the 20 µM treatment results in this one single glycan we can use its presence as a measure of the magnitude of kifunensine's effects. This can be achieved by expressing the area under this peak as percentage of the area under all peaks. Since $Man_9GlcNAc_2$ glycans are still present in the untreated control this will be used as the baseline, so measurements will be taken as the increase in percentage $Man_9GlcNAc_2$ above the normal.

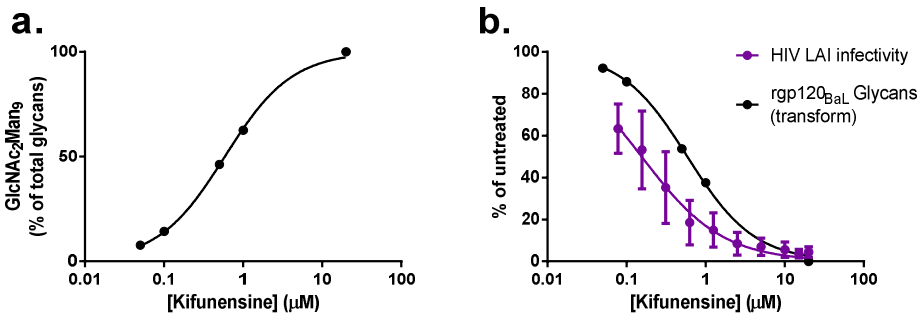



**Figure 5-6** Kifunensine causes a dose dependent increase in the percentage of $Man_9GlcNAc_2$ glycans (above those normally present) on $rgp120_{BaL}$ expressed in HEK 293T cells (a). For ease of comparing these data with those on kifunensine's antiviral effects on molecular clones of HIV LAI, the same data have been expressed as a transform of 100 - $x$, where $x$ is the percentage of total glycans that are $Man_9GlcNAc_2$. The glycan profiles were not repeated, data and bars for the antiviral effect are as in Figure 5-2.

These data are at odds with those of *Scanlan et al.* [150] and suggest that kifunensine starts affecting glycans at much lower concentrations (a) than previously reported, with effects being seen with doses as low as 50 nM and an $EC_{50}$ of 586 nM (95% confidence interval 505 to 666 nM). This is significantly different from the 145 nM $IC_{50}$ for kifunensine's antiviral effects (b), as determined by the extra-sum-of-squares F test ($p < 0.0001$, $F = 20.38$). However, by showing that kifunensine does have effects on the glycans at lower concentrations than previously thought, on the same order of magnitude as the antiviral effect, this lends credence to the possibility that it is the shift to a solely $Man_9GlcNAc_2$ phenotype that is the source of the antiviral effect.

### 5.2.4 Kifunensine does not affect gp120 oxidative folding

My current hypothesis is that *N*B-DNJ affects the oxidative folding of rgp120 by inhibiting glucose trimming and hence association with calnexin and ERp57. As mannose trimming occurs after calnexin association it has always been assumed that mannosidase inhibitors cannot affect oxidative folding. However, one can think of hypotheses as to how they might. For example, mannose trimming of the D2 and D3 branches by ER ManI occurs during the "off" phases of the calnexin cycle. This demannosylation can change the affinity of calnexin cycle partners to the glycoprotein; for example, calnexin and GlcII associate strongly with $Glc_1Man_9GlcNAc_2$ glycans but significantly less so with $Glc_1Man_7GlcNAc_2$ and UGGT1 is better able to reglucosylate $Man_9GlcNAc_2$ than $Man_7GlcNAc_2$. This effect serves to limit the amount of time a protein can stay in the calnexin cycle and thus prevents the accumulation of terminally misfolded proteins in the ER. Thus one can hypothesise that kifunensine, as an inhibitor of ER ManI, could prevent this process causing longer calnexin association times. Such longer times seem like they should be

pro-viral, as they could allow for a higher proportion of correctly folded glycoprotein, but they could be antiviral, by inducing saturation of chaperones and foldases with inconsequential proteins creating competitive inhibition.

To investigate whether kifunensine can, in fact, affect the oxidative folding of rgp120 radio active pulse chase was employed using the exact same techniques and conditions described in Section 4.3.6. Looking first at the nature of the oxidative folding, it appears that rgp120 produced in the presence of 20 μM kifunensine folds according to the same pattern as the untreated control. Initially the majority of the protein is reduced but this slowly becomes fully oxidised over the course of four hours (Figure 5-7). Unlike when *N*B-DNJ is used (see previous section) folding appears to be an orderly progression towards increasing compactness.

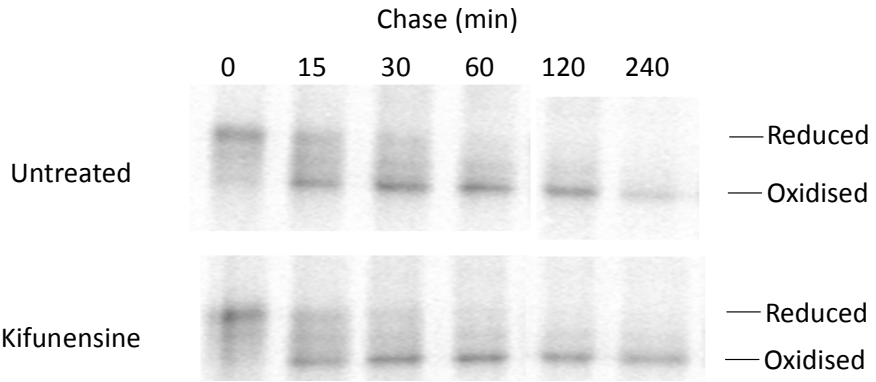



**Figure 5-7** Autoradiograph of non-reducing SDS PAGE of radioactively pulse-labelled $gp120_{LAI}$ expressed in HeLa cells. Each lane represents a separate experiment stopped at the indicated number of minutes after the completion of a ten minute pulse ("chase"). Cells were pre-treated with 20 μM kifunensine or left untreated, as indicated, for 45 minutes and this concentrations was then maintained during the pulse and chase periods. The "reduced" annotation marks the minimum motility of the protein and hence represents the reduced population, while "oxidised" represents the most compact, fully oxidised species. While both data for the same timepoints were performed on the same gels, the timepoints 120 and 240 minutes were performed on a separate gel to the other timepoints. The two gels have been digitally composited to appear as one, for clarity. Data are representative of two replicates.

This shows that the hypothesis that kifunensine could indirectly block calnexin association by saturating the cycle with terminally misfolded proteins is incorrect- kifunensine does not affect the pattern of oxidative folding of rgp120. However, it is still possible that kifunensine could affect the rate of folding. We can test this by looking at the timing of signal peptide cleavage in the presence and absence of kifunensine. Signal peptide cleavage only occurs when protein folding is complete and thus can be used as a measure of the rate of folding. Cleavage can be observed by looking for an increase in electrophoretic mobility of reduced, pulse-chased rgp120 (Figure 5-8).

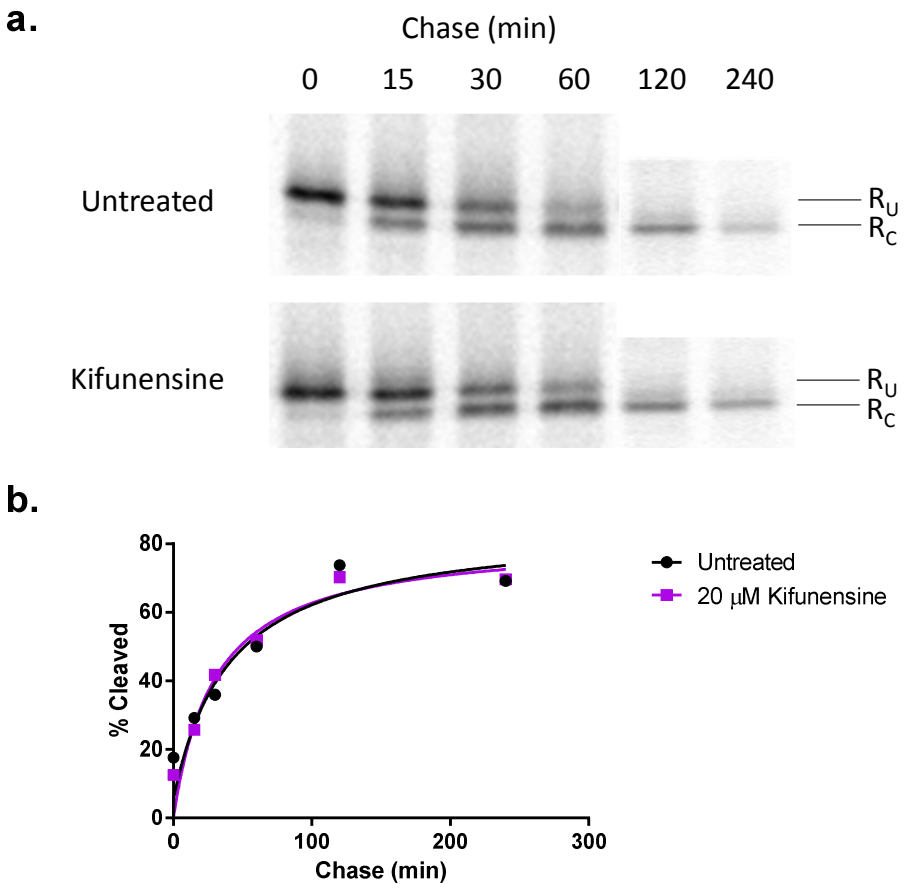



**Figure 5-8** Cleavage of the signal peptide from radiolabelled $rgp120_{LAI}$ divides the protein into two populations visible on reducing SDS-PAGE, reduced uncleaved ($R_U$) and reduced cleaved ($R_C$). These two populations are distinguished by the lower mass, and hence greater gel motility, of $R_C$, which comes to dominate over $R_U$ as folding progresses. This change can be visualised using radioactive pulse chased material (a). Densitometry can be used to measure the relative proportion of rgp120 that is in the cleaved state- this can be used to measure the rate of folding (b). Data are representative of two replicates.

It is thus clear that kifunensine does not affect the rate of oxidative folding. However, it does appear to affect the amount of gp120 secreted- Figure 5-7 and Figure 5-8 both show an increased ER concentration of gp120 in the presence of kifunensine (especially at later timepoints) while autoradiographs of the supernatant (thus measuring secreted gp120) show decreased secretion (Nicholas McCaul, personal communication). This implies that kifunensine causes gp120 to be retained in the ER rather than secreted to the plasma membrane- an observation that explains the decreased secretion of HIV virions shown in Figure 5-2a.

That an ERAD inhibitor would cause a decrease in protein secretion is counterintuitive as one would expect the opposite- indeed, ERAD inhibition by kifunensine *increases* secretion of HCV envelope protein and virions [216]. The answer may lie in the specificity of the L-type lectins that act as gatekeepers for ER exit (see Section 1.3.3). One of these lectins, VIP36, has been known to actively sort high-mannose glycoproteins out of the Golgi and back towards the

ER [217]. It may be, therefore, that kifunensine disrupts the balance between ER-Golgi secretion and ER-cytoplasm retrotranslocation (ERAD), potentially in either direction depending on the specific glycoprotein.

### 5.2.5 The effects of indirect mannosidase inhibition by *N*B-DNJ

*N*B-DNJ is a glucosidase inhibitor and is not known to have any activity against mannosidases. However, since removal of mannose residues is subsequent to the removal of glucose residues in glycan processing inhibition of the glucosidase leads to a *de facto* inhibition of exomannosidase activity as the enzymes cannot access their substrates. Since only one the three glycan branches possesses glucoses the inhibition is limited to that branch and the other two may be processed normally. Thus one would expect *N*B-DNJ to induce a population of mixed high-mannose glycans and potentially complex glycans if endomannosidase activity is significant.

Although it has been shown in the previous chapters that *N*B-DNJ has an effect on the folding of rgp120 which is independent of any effect on glycosylation, it is still not clear whether this one effect adequately accounts for all of *N*B-DNJ's antiretroviral activity, as the amplification hypothesis remains untested. Now that kifunensine has raised the spectre of glycan changes being antiviral it is worth investigating the magnitude of *N*B-DNJ's glycan effects at antiviral concentrations in a similar fashion to the above. Therefore, $rgp120_{BaL}$ was produced in the presence of a titration of *N*B-DNJ and the glycans profiled exactly as previously (Figure 5-9).

The changes seen on treatment with *N*B-DNJ are somewhat harder to quantify that the changes seen with kifunensine, as *N*B-DNJ does not favour a single species. The most prominent changes as concentration increases are the increasingly significant peaks eluted after 32 minutes, thus making them more massive than $Man_9GlcNAc_2$. These peaks in the untreated cells are almost certainly sialylated complex glycans, while those at higher concentrations of drug are most likely glucosylated high mannose obscuring or replacing the relatively small complex peaks. Eluting last of these, on treatment with 2 mM *N*B-DNJ, will be exclusively $Glc_3Man_9GlcNAc_2$ but the smaller peaks are harder to label. Since mannose and glucose are epimers it is almost impossible to distinguish, for example, $Glc_1Man_9GlcNAc_2$ from the identically massive but structurally dissimilar $Man_9GlcNAc_2$. The two would elute at slightly different times due to their different structures, but these two peaks would be so close together as to be indistinguishable at this resolution. For this reason it is difficult, if not impossible, to accurately assign all of the peaks based on the dextran ladder alone. Instead it is necessary to use enzymes to selectively remove the glucose residues and then observe the changes in the resulting profile.

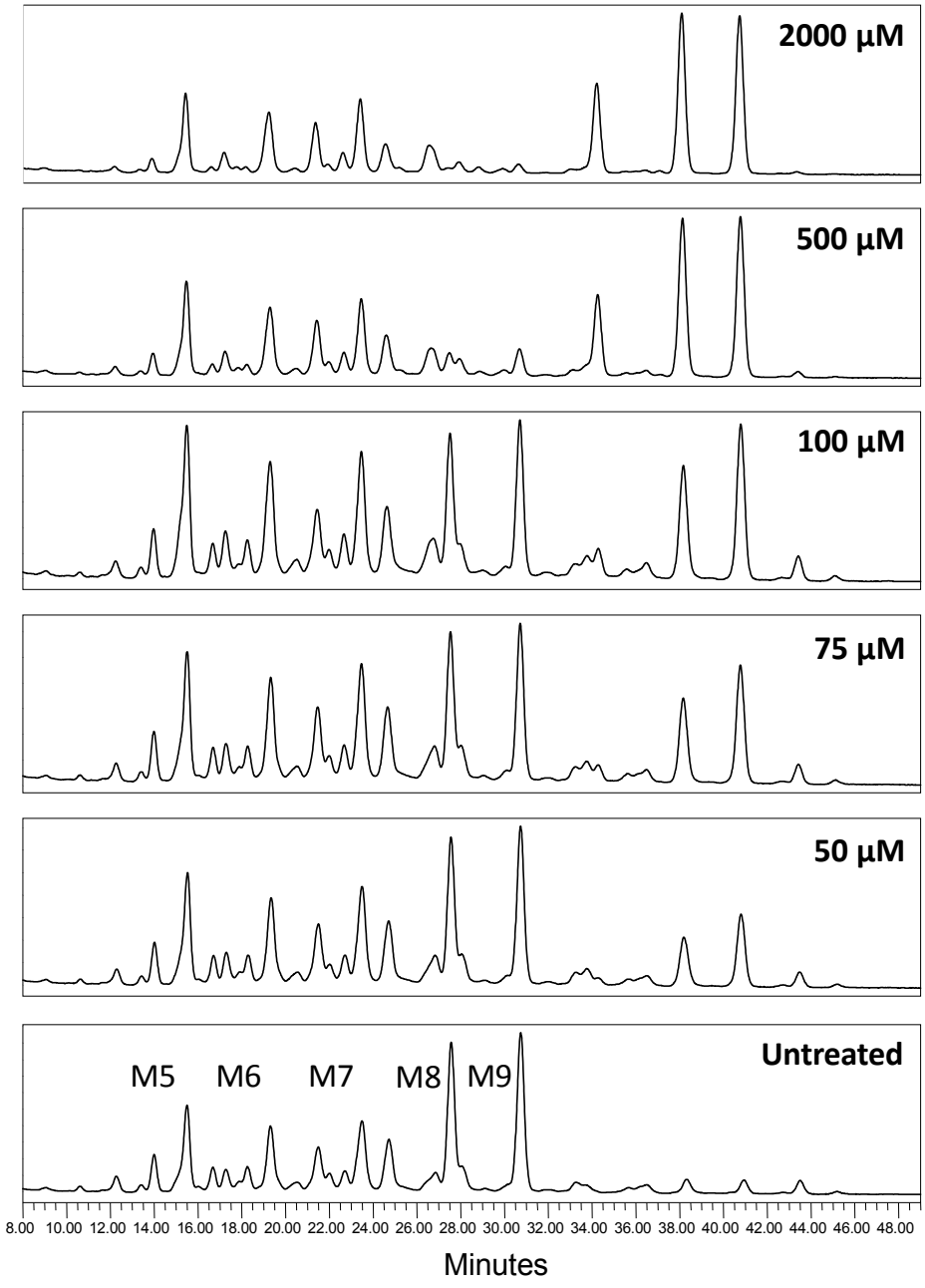



**Figure 5-9** Glycan profiles of $rgp120_{BaL}$ produced in HEK 293T cells treated with the indicated concentration of NB-DNJ. The y axes are arbitrary units representing fluorescence of the labelled glycans and are not all to the same scale. Abbreviations: M5, $Man_5GlcNAc_2$, M6, $Man_6GlcNAc_2$ *etcetera*.

These shifts can be seen in Figure 5-10 a, where the gp120 produced in the presence of 2 mM *N*B-DNJ has been digested with endomannosidase. The three largest peaks have disappeared, leaving small stumps representing complex glycans- this is very similar to the profile shown by the untreated gp120 (Figure 5-9) proving that there is some natural endomannosidase activity in the HEK 293T cells. While these peaks shrink, the peaks at 19.5, 23.5 and 26.5 minutes, representing $Man_{6\text{-}8}GlcNAc_2$ glycans all grow, while the $Man_9GlcNAc_2$ peak disappears altogether. If one uses integration to find the relative areas of each peak we can determine to where each peak shifted; for

example, the area of the $Man_8$ peak is the rough sum of the former $Man_9$ peak plus that of the peak at 40.5 minutes, showing that this peak must have been $Glc_3Man_9$. Such integrations indicate that the peaks at 34 and 38 minutes are $Glc_3Man_8$ and $Glc_3Man_7$ respectively. This shows that at 2 mM *N*B-DNJ GlcI inhibition is total and no glucose residues are removed at all, except by the endomannosidase.

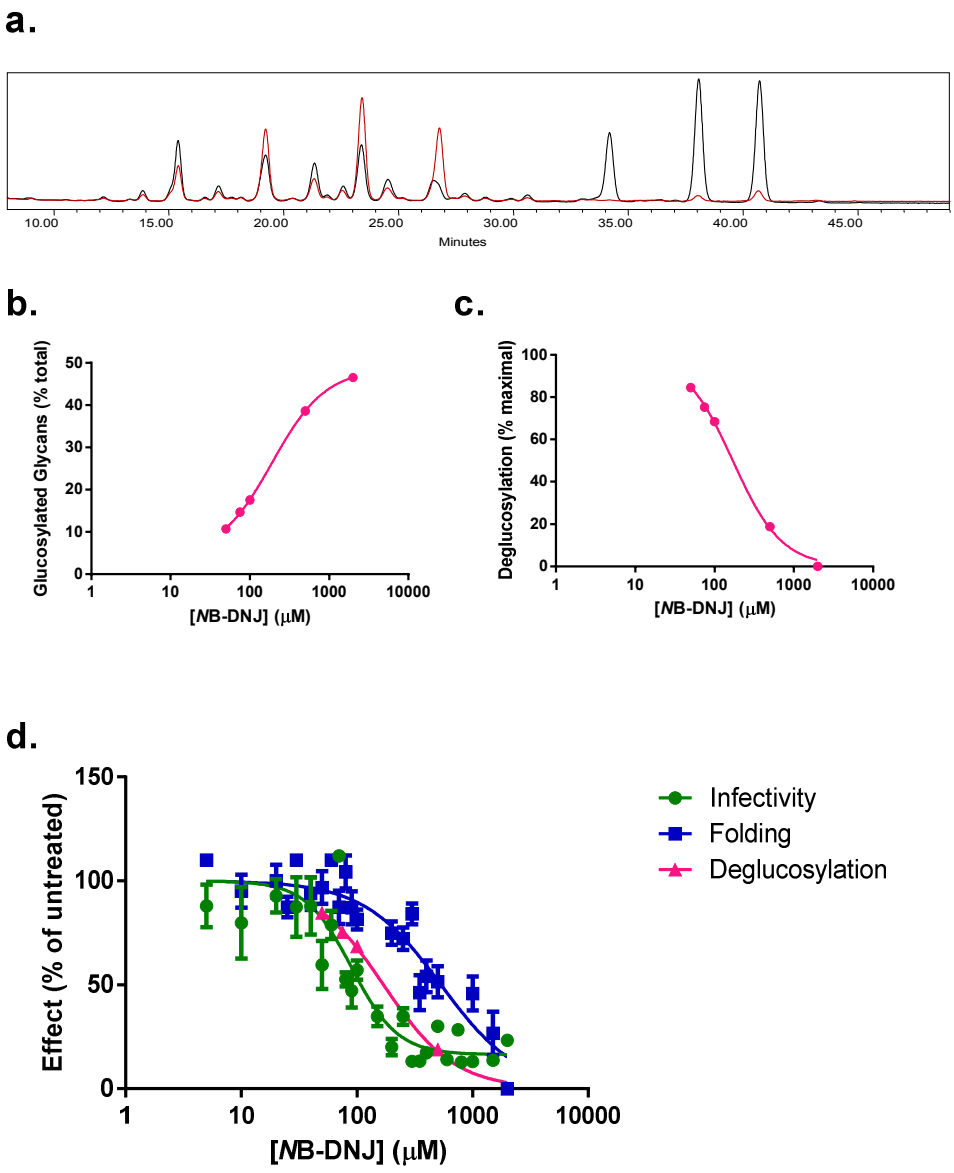



**Figure 5-10** Endmannosidase digestions can be used to assign peaks to $gp120_{BaL}$ glycans. The glycan profile of $gp120_{BaL}$ produced in the presence of 2 mM *N*B-DNJ is shown in black, and the profile of the same glycans after endomannosidase digestion is overlain in red (a). These assignments can be applied to data from Figure 5-9 to show the effect of *N*B-DNJ on the percentage of glucosylated glycans in $gp120_{BaL}$ (b). These data can then be transformed to show the percentage of deglucosylated glycans on the secreted gp120 (c). Note that this is *not* the same as measure of deglucosylation activity as that measure by FOS as this is only looking at the glycans on the secreted protein, not on the total population. This measure is then compared with *N*B-DNJ's effects on gp120 folding and HIV LAI infectivity (d), using data from the previous chapter.

Now that the peaks are assigned, we can quantify the percentage all gp120 glycans that are glucosylated at each concentration of *N*B-DNJ and compare the effect of the drug on the glycans to the effects on viral infectivity (Figure 5-10 b-d). The effect of the drug on the retention of glucose residues is dose dependent in a sigmoidal fashion, as one would expect, and tracks relatively closely to *N*B-DNJ's antiviral activity- far more so than the effect on folding does.

### 5.2.6 The mannosidase inhibitors 1-deoxymannojirimycin and swainsonine also show antiviral activity

Previous investigations of the antiretroviral capabilities of mannosidase inhibitors have used either the ManI inhibitor 1-deoxymannojirimycin (DMJ) or the Golgi ManII inhibitor swainsonine. These studies have been divided on the antiviral capabilities of DMJ, with finding it have activity and others not, while all studies have concurred that swainsonine is not antiretroviral (for full details and references see Sections 1.5.1 and 1.6.5). These compounds antiviral effects against molecular clones of HIV LAI were compared to those of kifunensine (Figure 5-11).

These experiments confirmed that none of the compounds were toxic at the concentrations used (Figure 5-11 a) and that both DMJ and swainsonine treatment resulted in a small decrease in the amount of HIV p24 secreted compared to the untreated control (b). This experiment confirmed that kifunensine can supress the infectivity of LAI virions and also showed that DMJ was capable of a similar, though lesser, effect (c). DMJ and kifunensine are both ER ManI inhibitors so this effect is not surprising. Kifunensine has antiviral activity at a far lower concentrations than DMJ, but this is consistent with the fact that it is a less potent ManI inhibitor.

The more surprising finding was that 20 μM swainsonine was able to suppress infectivity by a mean of 60% (compared to kifunensine's 85%). Swainsonine blocks Golgi ManII and thus permits substantial mannose trimming and the formation of hybrid-type glycans but prevents the final stages of mannose trimming that permit the production of complex glycans (see Section 1.5.1).

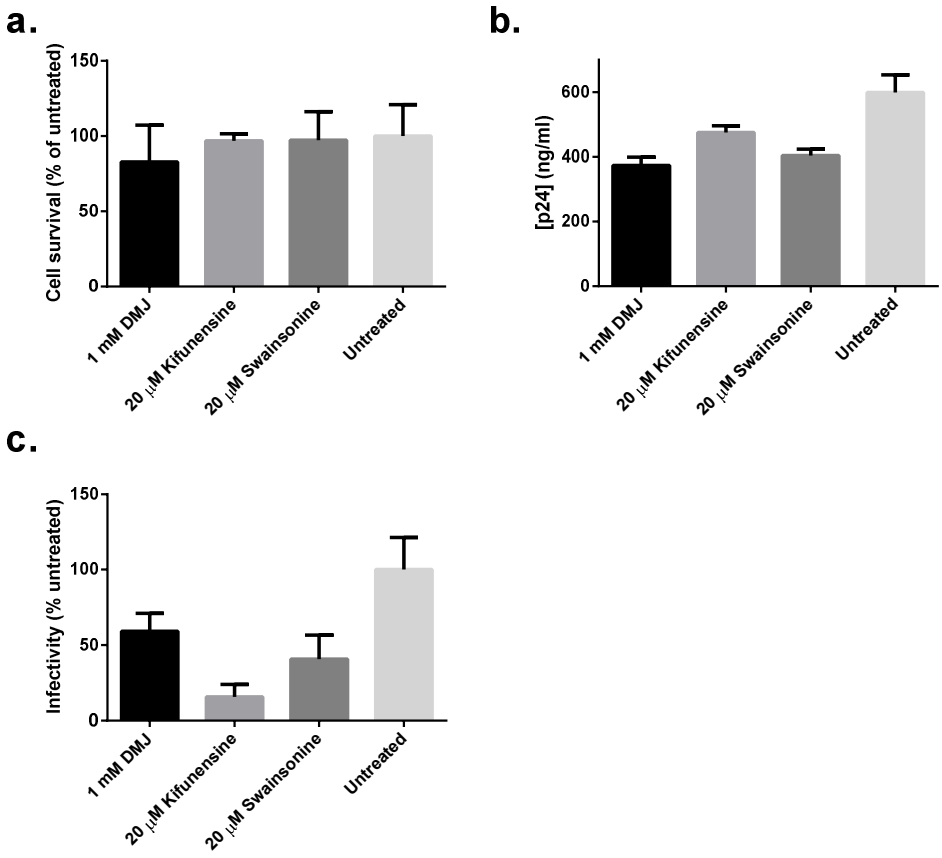



**Figure 5-11** The antiretroviral effects of 1 mM DMJ, 20 μM kifunensine and 20 μM swainsonine were all compared to that of a PBS control using the methods described previously in Section 5.2.1. First the compounds were determined to be non-toxic against transfected HEK 293T cells by MTS assay (a). HIV LAI was grown in 293T cells treated with the compounds for 2 days and the concentration of p24 produced determined by ELISA as previously. This showed that, like kifunensine, DMJ and swainsonine treatment resulted in decreased p24 secretion (b). Doses equivalent to 500 pg p24 were prepared and used to infect TZM-bl reporter cells, whereupon a luciferase assay was used to determine the relative infectivity of each viral stock. This showed that both DMJ and swainsonine can lower viral infectivity though not as much as by kifunensine (c). All bars show standard deviations and are the result of three biological replicates combined with different numbers of technical replicates: the MTS assays were performed in singlicate, the p24 ELISA in triplicate and the infectivity assays in septuplicate. Comparisons were made using a one-way ANOVA with an alpha of 0.05 and using the Tukey correction for multiple comparisons.

# 6 Interlude

## 6.1 Iminosugars as HIV pharmaceuticals

In the previous three chapters we have looked at the antiviral effects of iminosugars on HIV. However, I do not think that iminosugars will ever be a viable treatment for HIV infection. In clinical trials it was not possible to reach antiviral concentrations of *N*B-DNJ in patients' serum, despite dosing with *grams* of the drug, due to the very rapid initial elimination half-life [184,185]. Even the lowest doses used in that trial resulted in the unpleasant and antisocial side effects of diarrhoea and flatulence. While new derivatives of DNJ such as M*O*N-DNJ appear to show greater potency against HIV there is evidence that increasing the potency against ER α-glucosidases correlates with increased activity against the intestinal carbohydrate digestive enzymes such as sucrose-isomaltase (Atsushi Kato, personal communication). This implies that while lower doses may be able to achieve adequate plasma concentrations the side effects will not be diminished.

Perhaps even more significantly, iminosugars are not an economically viable HIV therapy. When HIV/iminosugar research began in 1987 only one drug, azidothymidine (AZT), had been licensed by the US Food and Drug Administration (FDA) for the treatment of the disease and HIV infection was perceived as uniformly fatal. Three decades later, over 32 anti-HIV drugs have been licensed by the American Food and Drug Administration (FDA) [218]. This has led to a dramatic improvement in both the quality of life and life expectancy of those lucky enough to have access these treatments [219,220].

It is true that iminosugars may offer some advantages that these other drugs cannot; for example, unlike iminosugars, existing HIV medications target viral enzymes and proteins, making them vulnerable to the evolution of resistant strains. For example, AZT resistant virions were first identified just two years after FDA approval [221], and were found to be spreading to newly infected patients by 1993 [222]. Meanwhile, HIV has been exposed to *N*B-DNJ *in vitro* for 6 months without the emergence of any resistance [214]. However, the threat of antiviral resistance is already effectively controlled using combination of existing therapies to make the emergence of escape mutants statistically prohibitive [223].

Also, iminosugars cannot hope to eradicate the proviral reservoir that prevents HIV cure [39]. This means that, like all the other existing therapies, *N*B-DNJ treatment would have to be for life. The last trials of *N*B-DNJ for HIV were performed in patients in the advanced stages of the disease when there were few other drugs available to help them. It is hard to imagine HIV patients now

tolerating the gastrointestinal side effects of *N*B-DNJ for months at a time when milder, more potent drugs are available.

So, does this make iminosugars a dead end line of research? I strongly disagree. Currently, antiviral drugs tend to be specific for individual families of viruses but have no clinical use beyond that, for example despite the large number of antiviral therapies for HIV none have been FDA approved for any other infection outside the retrovirus family. In addition to this, new viral diseases are constantly emerging into human populations with potentially devastating results- such as the emergence of Severe Acute Respiratory Syndrome (SARS) coronavirus in the early 21st Century which killed almost 800 people across two continents. α-glucosidase inhibitors' broad spectra of activity against a wide range of viruses combined with their low cytotoxicity profile could make them vital front-line drugs in the war on emerging diseases and a defence against bioterror attacks.

There are a number of reasons why I believe that *N*B-DNJ's failure against HIV clinically should not preclude its use in treating other viral diseases. Firstly, although diarrhoea was a significant side effect in the clinical trial I believe that it could be avoided or mitigated in this instance. Firstly, α-glucosidase inhibitor treatment could be limited to acute, severe infections. Because they are acute, the treatment will only have to be for a short period, making the side effects more tolerable and higher doses easier to justify. Secondly, if one is dealing with severe infections then one can assume that such drugs will only be administered to in-patients. This facilitates the use of parenteral dosing, thus bypassing the intestinal enzymes responsible for side effects. Also, parenteral dosing can allow for the administration of higher concentrations of drug; for example, intravenous infusions could be used to maintain antiviral concentrations in the plasma. To investigate the use of α-glucosidase inhibitors against a severe, acute viral infection *in vivo* we used a guinea pig model of ebolavirus disease (EVD).

## 6.2 Ebolavirus and EVD

### 6.2.1 Ebolavirus disease

EVD is a potentially fatal disease characterised by sudden onset fever, myalgia, arthralgia, anorexia, headache, abdominal pain, nausea, sore throat and prostration ("flu-like symptoms"). These generic symptoms can become complicated with symptoms of coagulopathy and haemorrhage such as melena, hematemesis, gingival and vaginal bleeding, epistaxis and bleeding at injection sites; symptoms associated with increased mortality [224]. Ebolavirus disease takes its name from the Ebola River in the former Zaire, close to which the first cases were reported in 1976. Since then,

there have been a number of localised outbreaks of the disease across Africa and potentially many unreported, sporadic cases which do not progress [225]. EVD is a zoonosis, with fruit bats thought to be the natural, asymptomatic reservoir; humans become infected by eating bushmeat, either infected bats or other mammals that have become infected with the virus [226]. The virus then spreads within human populations via close personal contact, with a significant nosocomial component [227].

EVD is caused by four separate viruses of the *Filovirus* family: Zaire ebolavirus (EBOV), Bundibugyo ebolavirus (BDBV), Sudan ebolavirus (SUDV) and Tai Forest ebolavirus (TAFV). A fifth virus, Reston ebolavirus (RESTV), causes EVD in non-human primates (NHPs) but appears asymptomatic in humans. It is also the only ebolavirus not originating in Africa, instead coming from East Asia. EBOV, BDBV and SUDV have all caused significant outbreaks across the African continent while TAFV (formerly Cote d'Ivoire (sic) ebolavirus) has only been confirmed in a single case [228]. EVD is particularly feared because of the very high mortality rates associated with some epidemics; for example, outbreaks of the Zaire strain of the disease has caused mortalities of up to 90% (see Table 6-1).

| Dates | Location | Cases | Fatalities | Mortality |
|---|---|---|---|---|
| 2014- present | Guinea, Liberia, Sierra Leone, Nigeria | >6263[1] | >2917[1] | 47%[1] |
| 2008 | DRC | 32 | 14 | 44% |
| 2007 | DRC | 264 | 187 | 71% |
| 2005 | Republic of Congo | 12 | 10 | 83% |
| 2003 (Nov-Dec) | Republic of Congo | 35 | 29 | 83% |
| 2003 (Jan- Apr) | Republic of Congo | 143 | 128 | 90% |
| 2001-2002 | Republic of Congo | 59 | 44 | 75% |
| 2001-2002 | Gabon | 65 | 53 | 82% |

**Table 6-1** Confirmed outbreaks of Zaire ebolavirus since the year 2000. Data courtesy of the World Health Organisation (WHO) and Centre for Disease Control (CDC). Abbreviation: DRC, Democratic Republic of Congo. [1] As of 29th September 2014.

Aside on nomenclature: The International Committee for the Taxonomy of Viruses (ICTV) dictates that the causative agent of EBV should be described as "ebolavirus" while many people continue to refer to it as "Ebola virus", a term which in many contexts only applies to the Zaire strain. In this thesis, the term "ebolavirus" will be used to refer to members of the family in general

while “Ebola virus” or “EBOV” will be used to specifically refer to the Zaire strain, in accordance with the system established by *Kuhn et al.* [229]

### 6.2.2 Ebolavirus structure

Ebolaviruses are negative-sense RNA viruses. The 19 kb genome encodes 7 open reading frames- from 3’ to 5’ these are a nucleoprotein (NP), the polymerase cofactor (VP35), a matrix protein (VP40), the envelope glycoprotein (GP), a viral transcription factor (VP30), another matrix protein (VP24) and an RNA-dependent-RNA-polymerase (L). VP35, VP30, L and the NP bound vRNA form an infectious ribonucleoprotein complex; this sits embedded in a matrix of VP40 and VP24 bound by a cell-derived lipid envelope decorated with GP trimers (Figure 6-1) [230]. Virions possess an unusual filamentous morphology and can be micrometres long, though all particles have a uniform 80 nm cross-sectional diameter [230,231].

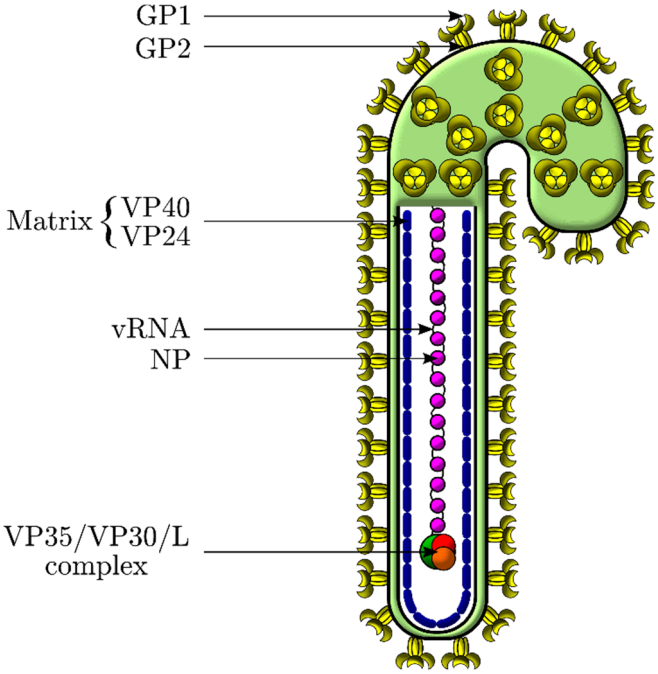



**Figure 6-1** Structure of the ebolavirus virion.

### 6.2.3 The ebolavirus glycoprotein and membrane fusion

As the putative target for any iminosugar mediated antiviral effect, the envelope glycoprotein of ebolavirus deserves closer inspection. Ebolavirus’s trimers are certainly reminiscent of those of HIV but there are many significant differences. Like HIV gp160, GP is cleaved by furin in the trans-Golgi into two proteins GP1 and GP2; the former being responsible for cell-attachment and the later possessing the transmembrane domain and fusion peptide- thus making the pair roughly analogous to gp120 and gp41 respectively. Unlike gp120/gp41, however, GP1 and GP2 remain

covalently associated via an interchain disulphide bond [230,231]. Also unlike gp160, this cleavage event is not vital for infectivity either in cell culture or in NHPs [232,233]. GP1/GP2 trimers adopt a chalice-like conformation with the putative receptor binding site at the base of the bowl.

GP1/GP2 is also heavily glycosylated, possessing at least 17 *N*-linked glycans and an unknown number of *O*-linked glycans [234,235]. The edges and rim of the chalice are heavily decorated with glycans; four highly conserved *N*-linked glycans form the so-called "glycan cap" atop the chalice and large mucin-like domains with both *N* and *O*-linked glycans act to further deepen the bowl. These domains almost certainly work to protect the receptor binding site from neutralisation by antibodies [234], although, interestingly, their removal does not affect the immunoreactivity of GP1 to human serum. This could suggest that GP1 glycans are evolved to shield the virus from neutralisation by the antibodies of bats, in which the infection is chronic [236].

Only 20% of GP transcripts result in the expression of structural GP1/GP2. This is because to encode the *C*-terminal, including the transmembrane domain, GP transcripts must undergo transcriptional RNA editing, consisting of a non-template coded adenosine insertion in a poly-A region, in order to provoke a frame shift that evolves the transmembrane domain. 80% of transcripts do not undergo this edit, resulting in the expression and secretion of a soluble glycoprotein, sGP [237]. Furin cleavage of sGP results in the so-called Δ peptide; this short peptide is of unknown function and significance but does appear to compete with the virus for cellular receptors, despite not including the receptor binding domain [238].

The cellular receptor for ebolaviruses, and hence GP1, is unknown: a number of cell surface molecules have been proposed but no one of them is individually able to account for virion targeting. These include DC-SIGN, β1 integrins, the Axl receptor tyrosine kinase and folate receptor α [230,239]. Ebolaviruses are also able to enter cells via interactions with the TIM-1 receptor via phosphatidyl-serine lipids in their envelope, implying that GP1 may not be required for cell attachment, only in regulating fusion [240]. Following attachment the virus is internalised by either macropinocytosis or clathrin dependent endocytosis [239].

Virion/membrane fusion occurs in the endosome and is preceded by a decrease in pH [241]. However, low pH is necessary but not sufficient to trigger fusion. The mucin domain of GP1 is proteolytically removed in the endosome by cathepsins B and L, inhibition of which substantially diminishes infection indicating that cleavage is a necessary precursor to fusion [235].

A GP2 crystal structure revealed a similar structure to HIV gp41, consisting of two antiparallel helices in a hairpin conformation, holding the fusion peptide close to the virion membrane in the pre-fusogenic state [242]. A crystal structure of the GP1/GP2 complex then revealed that GP1 holds

a helix of GP2 in a "clamp", lending further credence to the notion that cleavage could promote GP2 extension and fusion peptide insertion, followed by six-helical bundle formation and membrane fusion [234]. While cleavage is necessary, it is not sufficient to explain GP2 release; there must be a second step following cathepsin cleavage. This step was identified as the binding of GP1 to Niemann-Pick C 1 (NPC1) receptor; the residues that bind NPC1 are associated with the residues that clamp GP2, thus this binding weakens GP1's hold allowing extension of GP2 [243].

Thus GP mediated fusion is a multi-step process analogous to, but distinct from, HIV envelope mediated fusion. First GP1 (or lipids in the virion envelope) binds to a cellular receptor, probably a C-type lectin, and the virus is internalised into an endosome. Endosome acidification triggers cathepsins to cleave off a large portion of GP1 leaving a "primed" glycoprotein ready to initiate entry but awaiting a second signal, similarly to how CD4 binding to gp120 induces conformational changes to expose the co-receptor site. GP1 then binds NPC1 freeing the GP2 to extend and insert its fusion peptide into the endosome membrane, similarly to how co-receptor binding leads to exposure of gp41. In both systems fusion is then mediated by six-helical bundle formation, probably involving multiple glycoproteins working together to form a fusion pore.

## 6.3 Iminosugars and ebolaviruses

There are currently no licensed direct acting antivirals for EVD- treatment is mostly supportive and aimed at controlling symptoms. A number of experimental treatments are in development, many of which are based on small interfering RNA (siRNA) or monoclonal antibody technologies, though some small molecule inhibitors are in the early phases of development [244]. Small molecule inhibitors would be advantageous over antibodies and RNAs due to the greater ease of production, storage and distribution; factors which are of particular importance when dealing with a disease like EVD, which often occurs in remote areas with poor access to services.

With this in mind, we propose investigating the α-glucosidase inhibitors *N*B-DNJ and M*O*N-DNJ for antiviral effects against Ebola virus. As *N*B-DNJ is already a licensed drug (Miglustat) and M*O*N-DNJ is currently in phase I clinical trials for the treatment of dengue fever we anticipate that it would facilitate approval of these compounds for the treatment of EVD compared to a completely new molecule. Such is the nature of EVD that even experimental compounds that have never been used in humans have been employed clinically in the most recent outbreak, when the mAb based ZMAPP treatment was given to several Western aid workers infected with EBOV in West Africa [245].

There are only two precedents for the use of α-glucosidase inhibitors as therapeutics for the treatment of EVD. The first is a study by *Chang et al.* which investigated the antiviral effects of three novel DNJ derivatives. These derivatives were all able to inhibit the infectivity of EBOV GP pseudotyped lentiviruses and were also able to protect mice from challenge with an adapted EBOV strain [246]. *N*B-DNJ and M*O*N-DNJ have been patented for the treatment of filoviral infections by United Therapeutics, the same company which is currently carrying out the phase I trials of M*O*N-DNJ. In this patent they describe both drugs as being able to inhibit EBOV *in vitro* and M*O*N-DNJ as effectively reducing mortality in a mouse model of EVD (*N*B-DNJ was not investigated *in vivo*) [247].

We have expanded on these previous studies by carrying out an investigation into the safety and efficacy of intravenous *N*B-DNJ and M*O*N-DNJ in the treatment of adapted EBOV infection in guinea pigs.

# 7 Iminosugars as treatments for Ebola virus infection in a guinea pig model

## 7.1 Contributions

This chapter represents the results of collaborative experiments between the Oxford Institute of Glycobiology and Public Health England (PHE). Due to the requirement of specialist training and facilities (BSL-4) for the handling of ebolaviruses the majority of experiments were performed by specialist PHE staff. Dominic Alonzi performed the FOS analysis. The author, in close collaboration with Joanna Miller, was responsible for study design and data interpretation, and was personally responsible for the necropsy of uninfected animals and the histopathological examination of tissues.

## 7.2 Study design

### 7.2.1 Animal model

Of all the ebolaviruses, Ebola virus (EBOV, formerly Zaire ebolavirus) is the most pathogenic, having been responsible for the greatest mortality in the family, both as a proportion of all cases and in total. For this reason, we chose Ebola virus as the subject of this study. There are three commonly used animal models for Ebola; non-human primates (NHPs), guinea pigs and mice. NHPs are the gold standard model as, once infected with EBOV, they closely recapitulate the pathogenesis and symptoms of human EVD. However, there are considerable ethical implications involved in working with NHPs as well as high costs. As this was to be a proof-of-concept experiment it was decided to not use NHPs unless efficacy could be shown in a rodent model.

The most commonly used rodent model for EVD is the mouse. Adult mice cannot be naturally infected with EBOV. Instead, the virus must be serially passaged through neonatant mice until it is adequately adapted, whereupon it can infect adults but only via intraperitoneal inoculation [248]. Guinea pigs, in contrast, can be naturally infected as adults but serial passage is necessary to increase the virulence of the virus. Guinea pig disease more closely replicates the pathology and symptoms of human disease than the mouse model, as guinea pigs experience fever and clotting disorders (pathognomonic for "haemorrhagic fevers") while mice do not. However, mice do experience the key pathological sign of bystander apoptosis, in which uninfected lymphocytes apoptose, while guinea pigs do not [248].

A significant factor in the choice of a rodent model is their size. Guinea pigs are roughly 10 times the mass of mice. This makes mice appealing for trials of drugs as doses are much smaller.

However, as we are testing iminosugars for intravenous administration the small size of mice is a disadvantage. Intravenous (iv) injection of mice is very difficult and is usually substituted with intraperitoneal (ip) injection. Therefore, mice would need to be given several dozen ip injections over the course of the study. This represents a significant risk to the operator as it requires the use of sharps in a BSL-4 context. It is also highly stressful for the animals and carries a risk of organ damage from misplaced needles. Guinea pigs are large enough that they can be cannulated, allowing for administration of drugs directly into central veins. This does not eliminate the need for needles but does substantially lower the risk to the operator and minimises suffering and stress for the animal. For all of these reasons, a cannulated-guinea pig model was selected.

### 7.2.2 Treatments and dosing

The iminosugars chosen for investigation were *N*B-DNJ and M*O*N-DNJ. *N*B-DNJ was selected because it is already licensed for use in humans for the treatment of Gaucher's Disease, thus making it possible for it to be used in the treatment of EVD off-label if shown to be effective in animals. M*O*N-DNJ was selected as it is currently in phase I clinical trials for the treatment of dengue fever.

As this is a proof-of-concept, the desire was to give the maximum safe dose of each compound to see if any antiviral effect was achievable; if successful, future studies could identify the minimum therapeutic dose. Previous work on *N*B-DNJ had highlighted its short half-life in plasma, thus making continuous intravenous infusion the preferred route of administration. As this is impossible in guinea pigs without substantially compromising welfare (they would have to be immobilised to prevent them chewing through tubing) large doses could help maintain plasma concentrations above effective antiviral levels for longer periods.

Therefore, it was decided to perform a preliminary toxicity study in which guinea pigs were treated with escalating concentrations of each of the drugs or a carrier control. These animals were monitored for weight gain and temperature as well being regularly observed for clinical signs of illness. After 14 days of treatment the animals were euthanized and then necropsied. To assess whether the drugs were reaching effective concentrations, post-mortem liver samples were assayed for the presence of free oligosaccharides (FOS), a correlate of ER α-glucosidase inhibition.

## 7.3 Dose escalation safety study

Nine juvenile, female guinea pigs with cannulae surgically implanted into the right jugular vein were obtained from Charles River Laboratories. These guinea pigs were allowed to recuperate from the stress of transportation for 24 hours before the study began. *N*B-DNJ is a relatively well studied

drug that has proved safe at all concentrations (up to 2400 mg/kg/day *per os*) previously tested in rodents. For this reason, guinea pigs were dosed with just one intermediate concentration, 433 mg/kg thrice daily (tid), for two days after which doses were maintained at 617 mg/kg tid for a further 14 days. As there are no published studies specifically looking at the tolerability of M*O*N-DNJ in rodents, a slower dose escalation was performed with this drug. Guinea pigs still received thrice daily dosing, but on the first day received sequential doses of 5, 10 and 20 mg/kg M*O*N-DNJ. On the second day they received three doses of 30 mg/kg and then from day three they received 40 mg/kg for the remaining 14 days. Should any adverse reactions be observed the dose would be returned to the previous, safe dose for the remainder of the experiment. Compounds were prepared daily, based on weights from the previous day, from stocks of 460 mg/ml *N*B-DNJ (2.1 M), pH 6.8, 30 mg/ml M*O*N-DNJ (94 mM), pH 3.0 or water, pH 3.0 (placebo).

### 7.3.1 Clinical observations

One guinea pig in the M*O*N-DNJ group suffered a substantial loss of weight over the course of the first 24 hours of the experiment, losing 9% of its body weight over 24 hours (Figure 7-1 a). However, this loss of weight was not accompanied by any other signs of illness or distress so dosing was maintained as planned. From day 2, this guinea pig resumed weight gain and continued to gain weight at the same rate as the other animals in the group. This weight loss remains unexplained but appears to be unrelated to the drug and may be due to anorexia brought on by the stress of the start of the experiment. This anomaly makes the average weight gain of the M*O*N-DNJ group appear lower than that of the placebo group, but if this single datum is discounted then the two groups are indistinguishable (data not shown).

All of the animals gained weight over the course of the study but the weight gain of the *N*B-DNJ group was slower than that of the other groups and plateaued at day 13 (Figure 7-1). This is similar to the effects seen in mice treated with *N*B-DNJ [200]. The decreased weight gain could be explained with reference to inhibition of digestive enzymes by the drug, but no increase in the volume or consistency of faeces was observed, as in the previous mouse study. The authors of that study hypothesise that the drug may function as an appetite suppressant or affect metabolism [200].

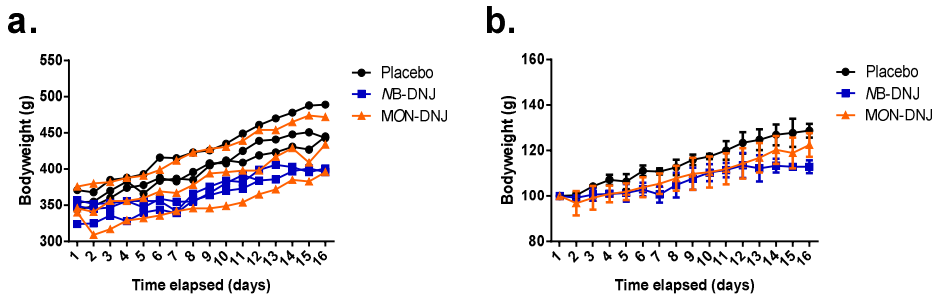



***Figure* 7-1** The growth rates of guinea pigs treated iv with *N*B-DNJ, M*O*N-DNJ or carrier placebo, as detailed in the text. (a) shows the weight of each individual animal over the course of the experiment while (b) shows the mean mass of each group as a percentage of the day 1 mass of each animal. Bars show standard deviation, n = 3. *Nota bene:* there are only 2 animals in the *N*B-DNJ group from day 13 onwards.

The most significant finding of the study occurred on day 11 when one of the *N*B-DNJ treated guinea pigs went into severe respiratory distress during dosing and was euthanized to prevent further suffering. Due to a failure of communication the cadaver was disposed of before a thorough post-mortem examination could be performed. Therefore, the cause of the respiratory distress could not be determined. No other animals showed any signs of illness or adverse reactions to dosing at any point and all animals' temperatures were within normal limits throughout.

Bacterial pneumonia is the most common cause of death in domestic guinea pigs but is unlikely to be responsible for this fatality as there were no signs of pneumonia prior to the incident and the animal was gaining weight and had no fever. A more likely possibility is viral pneumonia from guinea pig adenovirus (GAV). Guinea pigs can carry this virus with asymptomatically and the disease is then triggered by stress. Disease manifests as sudden death without prior clinical signs [249,250]. GAV is moderately prevalent in laboratory guinea pigs but Charles River Laboratories routinely test for it and so seems unlikely to be the source of this mortality.

The major concern, of course, is that this fatality was associated with the drug. If so, there is a question as to why it only affected one of the three guinea pigs in the *N*B-DNJ treatment group and why it should have affected that animal so suddenly. The fact that the symptoms began at dosing is unlikely to be coincidental. The drug was injected directly into the jugular vein whereupon it would quickly enter the right side of the heart and into the respiratory circulation. The drug is in a hyperosmotic solution so I was concerned that this may have caused some sort of shock by transiently affecting pulmonary fluid balance. However, human patients are regularly given 50 ml 50% glucose boluses, which is comparable to the concentration and relative volumes here, into central veins without concern.

The most likely differentials, in my clinical opinion, are a pulmonary embolism or hypersensitivity. The latter implies that the guinea pig may have become sensitized to the drug

and went into anaphylactic shock on administration. There are two likely causes of embolism in these animals; either thrombus formation on the cannula which was displaced by injection of the drug, or an air embolism caused by bubbles in the preparation. The risk of thromboembolism was minimised as the cannulae were flushed with heparinised saline before each dose, but it is not uncommon on an indwelling cannula. Air embolism would be a result of operator error when preparing the doses. I am reluctant to ascribe this, as the staff responsible are highly trained and experienced, but I cannot rule it out.

All other animals survived until day 16, with no signs of any respiratory or other problems, whereupon they were humanly euthanized using a rising concentrations of carbon dioxide ($CO_2$), in accordance with Schedule 1 of the Welfare of Animals (Scientific Procedures) Act 1986. Reduced spleen size is a known effect of *N*B-DNJ in mice and we were concerned that euthanasia with pentobarbital, which is known to cause splenomegaly, would interfere with such measurements, despite being (in my opinion) a more humane method of euthanasia. The animals were then necropsied and the livers and spleens weighed. One lobe of each liver (with the exception of that of the fatality) was frozen at -80 °C for FOS analysis.

Samples of the liver, spleen, lungs and (where possible to dissect) the right jugular vein were fixed in formalin, mounted, stained with haematoxylin and eosin (H&E) and examined for histopathology. The necropsy and histopathology results are detailed below. For ease of reference, the placebo, *N*B-DNJ and M*O*N-DNJ groups are labelled groups 1 to 3 respectively, with individual animals within those groups being numbered 1 to 3 (e.g. 1:2 refers to the second animal in the placebo group). Guinea pig 2:3 was the fatality and is not included in this report.

### 7.3.2 Pathology report

#### GROSS DESCRIPTION

EXTERNAL EXAMINATION

The animals are good body condition. Each animal has a catheter inserted into the right jugular vein, which is protruding from the cranio-ventral thorax. A red liquid (blood) is visible bilaterally in the external nares, which was not present prior to euthanasia.

INTERNAL EXAMINATION

**Subcutaneous tissue:** There is a mild quantity of cream coloured subcutaneous fat.

**Peritoneal Cavity:** The peritoneum is smooth and shiny. The peritoneal cavity contains a mild quantity of a clear fluid (not quantified).

**Cardiovascular system:** The pericardium contains a mild quantity of clear fluid. The endocardial and epicardial surfaces are pale pink and uniformly smooth and shiny. The chambers contain a black, gelatinous substance (clotted blood). The great vessels are similar. The jugular veins were examined as they ran subcutaneously across the cranio-ventral thorax in animal 2:2. The catheterised vein was within normal limits.

**Alimentary system**: The liver is uniformly brown, smooth and shiny. The gall bladder is full. The alimentary canal is full and shows post-mortem peristaltic movements. The lumen of the gut was not examined.

**Respiratory System:** The lungs are dark pink with numerous focal dark red/black circular patches across the surface and throughout the parenchyma measuring 1-3 mm in diameter across all lobes. The pleural surfaces are smooth and shiny.

**Urinary System:** The kidneys are dark brown/red and smooth. In longitudinal section the cortex and medulla are clearly defined with no signs of pathology. The bladder and reproductive tract were not examined.

**Lymphoid System:** The mediastinal lymph nodes are pale pink and 1-2 mm diameter. The spleens are within normal limits.

The **musculoskeletal**, **endocrine** and **nervous** systems were not examined.

GROSS DIAGNOSES

Lungs: Haemorrhage, acute, diffuse, moderate to marked

Nares: Epistaxis, acute, mild, secondary to the above

COMMENT

The guinea pigs appear to have been healthy with no grossly visible signs of pathology that could be associated with toxicity from the treatments; indeed there were no observed differences between the animals either within or between the treatment groups. The epistaxis and pulmonary haemorrhage can be explained as injury sustained agonally as a consequence of $CO_2$ asphyxiation. To confirm this diagnosis histology was performed on the affected lungs (see below).

### Histopathological Description

As all 9 guinea pigs were grossly identical on post mortem only organs from one animal in each group were prepared for histological examination. The lungs, spleens and hearts from animals 1:3, 2:2 and 3:1 were fixed in formaldehyde, mounted on microscope slides and stained with H&E for histological examination, as was the catheterised right jugular vein of guinea pig 2:2.

**Lungs:** The lungs show diffuse congestion in the interstitial vasculature with erythrocytes. There are numerous haemorrhagic foci in which erythrocytes are visible within the alveoli, though the septa are still intact and are clearly defined. In some areas of congestion the alveoli are filled with pink-staining fluid (proteinaceous fluid). Leukocytes are visible only rarely and are well within normal limits. The visceral pleura is within normal limits.

**Heart**: The right atria were examined for signs of damage from the acidic (all groups, especially group 3) and hyperosmolar (group 3 and especially 2) solutions that they were exposed to on each infusion. The endocardium was intact and showed no signs of inflammation. The myocardium was within normal limits. On guinea pig 3:1 the tricuspid valve was examined and was within normal limits; it was not in the plane of section and hence not examined on the other animals. The atrial and ventricular walls appeared to be relatively thick (3-4mm in the left ventricles, ˜1mm in the right atria) but this is hard to judge as I do not have experience of the normal limits of the guinea pig myocardial thickness.

**Spleen:** Within normal limits.

**Jugular vein (GP 2:2 only)**: The endothelium is intact and smooth. The tunica media contains spindle shaped cells (smooth myocytes and fibroblasts) but no leukocytes and there are no signs of inflammation, hypertrophy, hyperplasia, oedema or fibrosis.

### Diagnosis

Lungs: Perimortem congestion, acute, diffuse, moderate to marked with focal haemorrhages, acute, diffuse, moderate to marked.

### Comment

The absence of inflammatory cells in the lungs confirms the diagnosis from the post mortem that the pathology in the lungs is an agonal change and hence is most likely a result of the method of euthanasia rather than the treatments. No pathology is visible either grossly or histologically that could be attributed to the treatments, leading to the tentative conclusion that, at the doses

given, the treatments are not toxic in guinea pigs. The lack of the damage to the jugular vein or right atria suggest that the formulations are not causing any mechanical damage as they enter the body due to low pH or osmotic shock. Indeed, the lack of any fibrosis or phlebitis in the jugular vein of animal 2:2 is surprising, as it will have been potentially irritated by the catheter for over 18 days as well as by the drug for 16 days.

### 7.3.3 Effects on guinea pig organ masses

As mentioned above, a previous tolerability study of *N*B-DNJ in mice showed that the animals all suffered a considerable reduction in spleen mass compared to the untreated controls (30% lower when on 1800 mg/kg/day) [200]. Flow cytometry showed that these spleens also had a higher proportion of T-cells than the untreated (31% $CD4^+$ cells in the treated mice versus 20.5% in the untreated and 16% versus 9% for $CD8^+$ cells). Importantly, the spleens returned to the normal size within two weeks of cessation of therapy. To see whether the spleens were similarly affected in guinea pigs the spleens and livers (as control) were removed and weighed in their entirety at post mortem. All large vessels and omentum were dissected away but the gall bladders were left in place; the weights of the organs as a percentage of the mass of each guinea pig are presented in Figure 7-2.

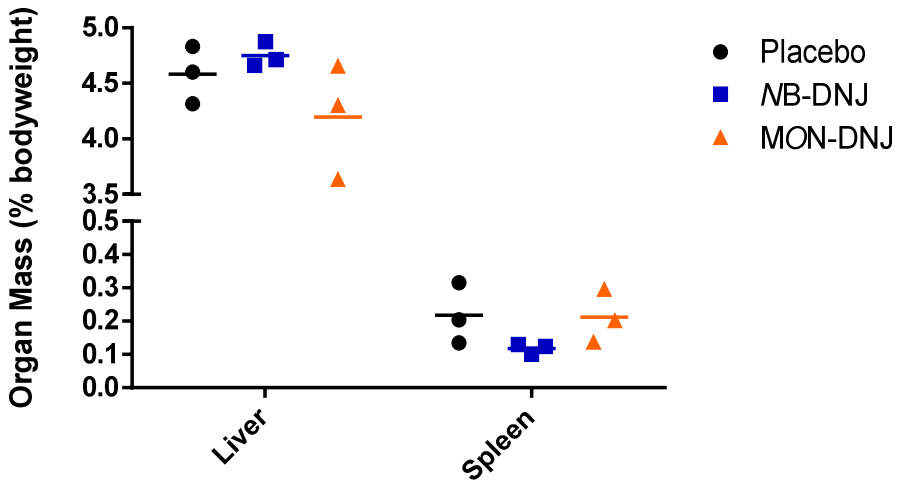



**Figure 7-2** Masses of the liver and spleen of each animal as a percentage of bodyweight. Groups were analysed using two-way ANOVA at an alpha of 0.05; no significant differences were observed between treatment groups.

As Figure 7-2 shows, there *N*B-DNJ does not have a statistically significant effect on the weights of either the spleen or the liver. The spleens of the *N*B-DNJ group are, on average, smaller than those of the other two groups but this effect is not significant. As there are relatively few animals in each group it is worth asking whether this effect is real but the study simply does not have the power to resolve this question. Therefore, instead of comparing the three animals in the *N*B-DNJ

group solely to the placebo group we would be better off looking at the average spleen/body mass ratio for normal guinea pigs.

Surprisingly, someone has gone to the effort of weighing almost 500 guinea pig spleens to come to a definitive answer as to what normal is. In this, frankly genocidal, study the author determined that 400-500 g female guinea pigs (as these were) should have spleens that are about 0.15 % of their total body mass, with a 25% co-efficient of variation [251]. The *N*B-DNJ treated guinea pigs' spleens were 0.10, 0.12 and 0.13 %, putting them inside these margins for normality.

### 7.3.4 Generation of free oligosaccharides in guinea pig livers

*N*B-DNJ and M*O*N-DNJ show two-phase elimination kinetics when administered intravenously, with an initially rapid redistribution phase resulting in a precipitous drop in drug plasma concentration, followed by an exponential elimination phase with a half-life of 0.68 hours (United Therapeutics, personal communication). This caused concern that effective concentrations might not be reached in tissues. To test this, a portion of liver tissue from each animal (except the fatality) was examined for the presence of certain free oligosaccharides. Free oligosaccharides (FOS) are cytoplasmic glycans that are free from proteins; they are generated as part of the ERAD pathway and thus are representative of the glycans present on proteins that have recently been in the ER. As the ER α-glucosidase enzymes usually very rapidly deglucosylate *N*-linked glycans, glucosylated FOS are rarely detected in normal tissues. However, if the glucosidases are inhibited then unique FOS begin to appear: $Glc_3Man_5GlcNAc_1$ being indicative of GlcI inhibition and $Glc_1Man_4GlcNAc_1$ representing GlcII inhibition. Typically GlcII is inhibited at lower concentrations of drug than GlcI and GlcI inhibition masks inhibition of GlcII (i.e. GlcI inhibition depletes the substrate for GlcII so less GlcII FOS will be generated in the presence of GlcI inhibitors). The FOS detected in the treated animals are shown in Figure 7-3.

*N*B-DNJ caused the generation of both $Glc_3Man_5GlcNAc_1$ and $Glc_1Man_4GlcNAc_1$ FOS, while M*O*N-DNJ only induced GlcII inhibition. M*O*N-DNJ was given at a >15-fold lower dose than *N*B-DNJ which may explain this distinction. M*O*N-DNJ appears to be a greater inhibitor of GlcII than *N*B-DNJ, but this is almost certainly due to the masking effect from GlcI inhibition described above. Importantly, both drugs are reaching adequate tissue concentrations to inhibit at least α-glucosidase II. GlcII is thought to be the significant enzyme for the antiviral effect; with dengue virus, another haemorrhagic fever virus, maximum inhibitory effect correlates with the maximum GlcII FOS (Andrew Sayce, personal communication). This makes intuitive sense as calnexin is

insignificantly more able to bind biglucosylated as triglucosylated glycans, so adding GlcI inhibition on top of maximal GlcII should not have any additive effect.

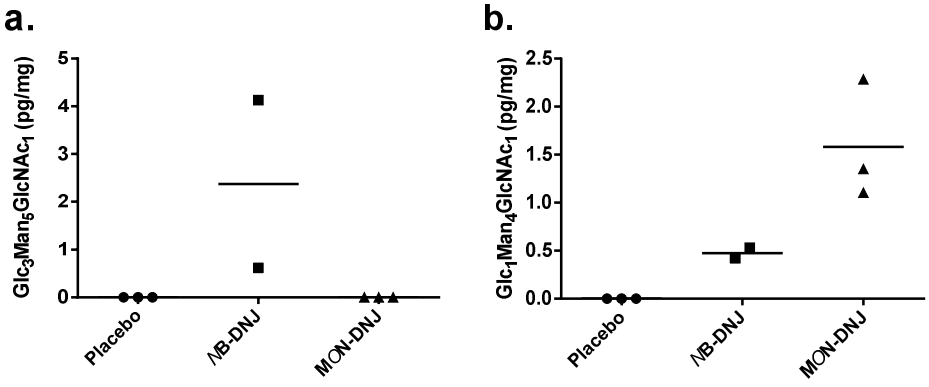



**Figure 7-3** Free oligosaccharides present in the liver tissue of iminosugar treated guinea pigs at time of sacrifice.

### 7.3.5 Conclusions

The safety study confirms that M*O*N-DNJ is non-toxic to guinea pigs when given for two weeks at 120 mg/kg/day iv, therefore making this the concentration that will be used in efficacy studies against Ebola. The interpretation of the *N*B-DNJ data is more equivocal; one guinea pig did die during treatment but due to the lack of a necropsy it is hard to determine the cause of death. As *N*B-DNJ is safe at high concentrations in mice and this fatality occurred suddenly, late in the experiment, my preferred hypothesis is that this fatality was the result of a pulmonary embolism (most likely a thromboembolism) and not the result of the drug.

This hypothesis cannot be confirmed directly, but with it in mind I dissected out the catheterisation sites of the remaining two guinea pigs in the *N*B-DNJ group and examined them both grossly and one histologically. No inflammation, sclerosis or thrombosis was evident in either vein. Unfortunately, while the presence of these would have helped confirm the hypothesis their absence dose not disprove it, as these animals did not suffer similar symptoms. To allay fears the experiment was repeated on a further two guinea pigs, treated with *N*B-DNJ only (data not shown). Neither of these animals showed any adverse effects, meaning that, in total, 1 in 5 guinea pigs died during treatment with *N*B-DNJ. The fatality was ascribed to idiopathic causes and the decision was made to maintain the current dose of 1850 mg/kg/day in the efficacy studies.

As previously observed with mice, guinea pigs on *N*B-DNJ showed a slower rate of growth than their untreated compatriots. Guinea pigs treated with M*O*N-DNJ showed normal growth rates; as, like *N*B-DNJ, M*O*N-DNJ is a known inhibitor of digestive enzymes this is probably reflective of the lower doses used. Both drugs reached inhibitory concentrations in liver tissue, suggesting that

the doses were sufficient to see an effect, though *N*B-DNJ had a greater effect on the enzymes than M*O*N-DNJ.

## 7.4 Efficacy study

To test the efficacy of iminosugars against Ebola virus 3 groups of 8 guinea pigs were planned to be infected with the virus then ascribed to receive either 120 mg/kg/day M*O*N-DNJ, 1850 mg/kg/day *N*B-DNJ or a pH-matched water placebo. However, due to the challenges of working with animals in a BSL-4 environment the study was split into two consecutive studies, each with 3 groups of 4 animals. Should the first study show no protection at all then the second study would be abandoned on welfare grounds, but otherwise these should be thought of as a single study. For this study we were able to obtain the chloride salt of M*O*N-DNJ, which is soluble at a higher pH than that used previously. Therefore, both drugs and the placebo had a pH in the range of 6.3 to 6.9.

Cannulated animals were obtained as previously and allowed to recover for two days. At this point they were anaesthetised with isoflurane and $10^3$ TCID (tissue culture infectious dose, 50%) of guinea pig adapted Ebola virus was inoculated subcutaneously into the right caudal dorsum of each animal, in a volume of 0.2 ml. This has been determined to be a 100% lethal dose. Tid iv dosing began immediately upon recovery. The animals' weights, temperatures and clinical signs were then assessed for 14 days. Humane end points were defined as any animal losing 20% of its body weight, losing 10% while also showing clinical signs of disease or if showing signs of distress at the discretion of the named animal care and welfare officer (NACWO).

### 7.4.1 Results

Survival curves for the three groups are presented in Figure 7-4. These curves were analysed using the Mantel-Cox log-rank test and the two treatment arms were found to be insignificantly different from the placebo arm, with $p = 0.4595$ for *N*B-DNJ and 0.9096 for M*O*N-DNJ. Nonetheless, looking at the curves one can see a qualitative difference for the *N*B-DNJ group; one animal survived until the end of the study. Although it does not show up as significant due to the low $n$ number this is very interesting as the dose of Ebola administered is known to be 100% lethal; in all the previous studies performed using this stock of virus no animal in a placebo wing has ever survived (more than 50 animals, PHE, personal communication), indicating that survival is unlikely to be due to chance.

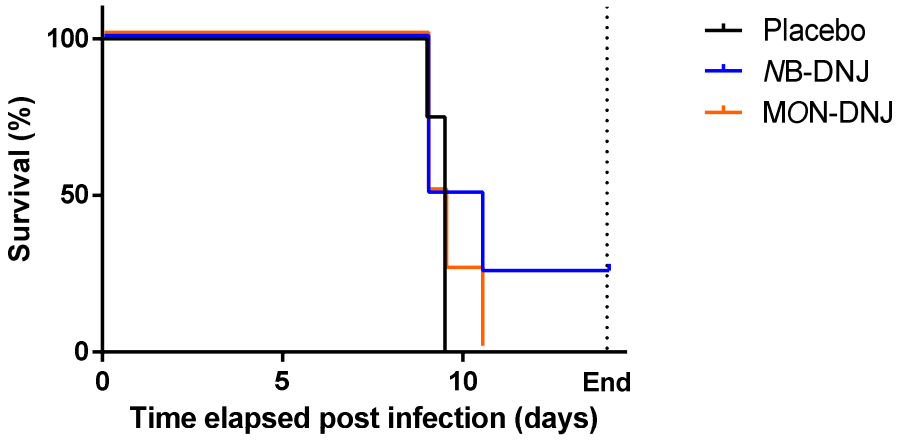



**Figure 7-4** Survival curves for guinea pigs infected with Ebola virus then treated with iv iminosugars (120 mg/kg/day MON-DNJ, 1850 mg/kg/day NB-DNJ, acidified water placebo). n = 4 for all groups. "End" represents the end of the experiment, 14 days after infection.

If one looks at the clinical history of the surviving animal (Figure 7-5) one can see that it suffered from fever and weight loss on a similar time scale to the placebo group, indicating that it was successfully infected and did develop EVD. However, although the guinea pig did not gain weight as quickly as the controls at the start of the experiment (as expected), it was much more able to maintain weight in the presence of the disease. Also, while the placebo group are all dead by day 9.5 the surviving animal in the *N*B-DNJ group appears to actively recover from this point, with its temperature normalising and its body mass increasing. These data suggest that, in at least one animal, *N*B-DNJ was able to limit the severity of EVD sufficiently that the animal was able to make a full recovery from the usually lethal infection.

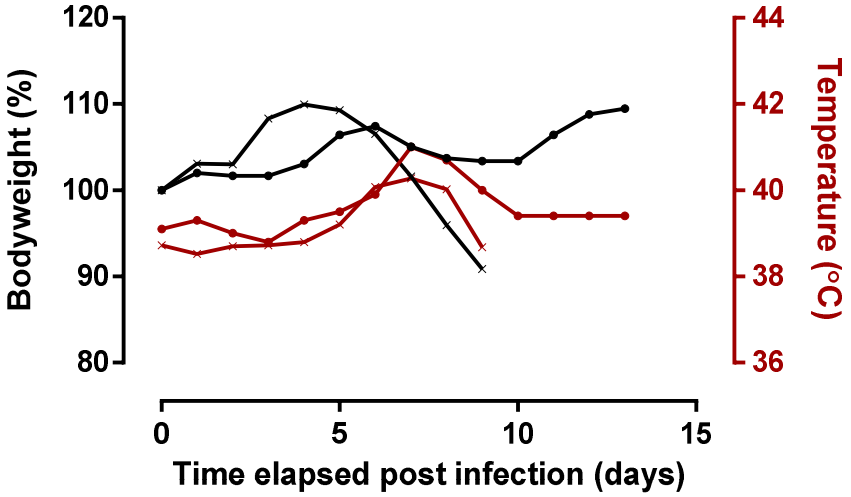



**Figure 7-5** Clinical history of the surviving animal in the *N*B-DNJ arm (•) versus the mean scores of the animals in the placebo arm (×, n = 4).

Looking at the clinical data for all 12 guinea pigs a few trends are noticeable, though not necessarily statistically significant (Figure 7-6). For example, *N*B-DNJ treatment appears to cause decreased weight gain relative to that in the placebo treated guinea pigs, as previously observed. Two M*O*N-DNJ treated animals gain weight substantially faster than the controls, though they begin losing this weight and die at the same rate as the placebo group suggesting that this may be incidental rather than as a result of the drug (a). All of the animals show a rising body temperature over the course of the infection which then falls immediately prior to death (b). While the normothermic range for guinea pigs extends up to 40 °C, temperatures in this experiment were recorded from subcutaneous implants and thus are not directly recording the core body temperature, thus underreporting the temperature of each animal. Thus it is reasonable to assume that the temperature increases constitute fever in all cases. The decline in temperature in the terminal stages of the disease is likely due to failure of thermoregulation due to organ failure.

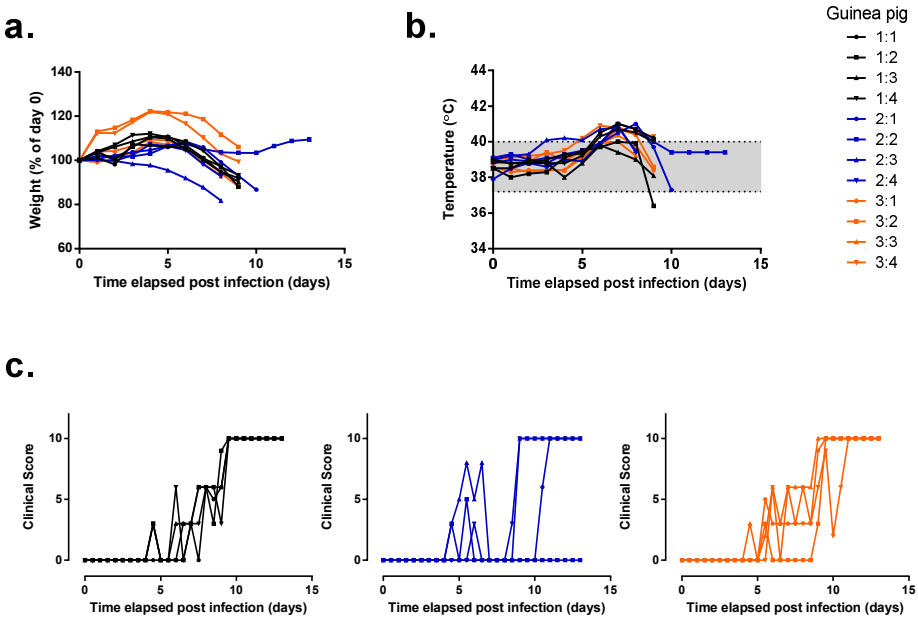



**Figure 7-6** Clinical information for each guinea pig following infection with EBOV on day 0. In all graphs, guinea pigs in the placebo group (1:1 to 1:4) are coloured black, the *N*B-DNJ group (2:1 to 2:4) blue and the M*O*N-DNJ group (3:1 to 3:4) orange, with the shape of the symbol consistently identifying each animal. Graphs show guinea pig bodyweight as a percentage of that at the time of infection (a) and subcutaneous temperature (b, the grey box indicates the guinea pig normothermic range). Clinical scores were recorded twice daily according to an arbitrary scale where 0 is no clinical signs and 10 is death or euthanasia (c).

The clinical scores are more interesting- the placebo and M*O*N-DNJ treated groups' scores (c, groups 1 and 3, black and orange respectively) progress steadily, generally trending towards worsening symptoms, the *N*B-DNJ treated group's (2, blue) symptoms fluctuate, with some animals reporting illness but then apparently recovering before relapsing. A complication of this study was

that the BSL-4 staff were not able to give treatments outside of working hours; this meant that the tid dosing was not at 8 hours intervals but was instead distributed over the working day whereupon the animals went untreated overnight, meaning that plasma levels of the drug were likely to fluctuate substantially over each 24 hour period. Thus it could be that, in the case of *N*B-DNJ, the virus was able to replicate in these intervals but was then suppressed when dosing restarted, thus resulting in paroxysms of symptoms.

### 7.4.2 Conclusions

*N*B-DNJ was able to induce what appeared to be a complete recovery in a single guinea pig out of the four treated with the drug. This is a 100% lethal model, so this strongly suggests that, in some circumstances, the drug is able to protect against EVD. This effect is not statistically significant ($p = 0.4595$, Mantel-Cox log-rank test) but this may be because this test, like most survival curve analyses, was designed for use in large clinical trials where the event rate (death, in this case) was not expected to be 100% in the placebo wing. For example, PHE Porton Down have used this guinea pig model previously; in total there have been more than 50 guinea pigs in the placebo arms of various trials, all of which have died. Thus if we were to count these animals as part of our placebo arm and assume they all died on day 9 then the $p$ value would counterintuitively rise to 0.7827! This highlights the inadequacy of such statistical tests, as any reasonable person would assume that 1 in 4 animals surviving a disease that has a survival rate below 1 in 50 would be unlikely to occur by chance alone.

Thus the major question should be why only one animal survived and not all or none? One possible explanation for this is that the guinea pigs are a relatively outbred population (compared to other laboratory rodents) and it may be that the animals responded heterogeneously to both the virus and the drug. However, the more likely explanation, in my opinion, is that fault lay in the dosing intervals. The large overnight gaps between doses appear to have allowed the disease to rebound repeatedly. This may have resulted in a "tug-of-war" between the virus's ability to damage versus the body's ability to repair, with the outcome for each guinea pig depending on which side the equilibrium fell. This is also consistent with the findings from the toxicity study that levels of α-glucosidase inhibition varied between animals- this may also be a result of inconsistent plasma levels of the inhibitors and may be linked to survival. FOS analysis of tissues from the infected guinea pigs is currently pending, but I predict that the surviving animal will show higher levels of α-glucosidase inhibition than those that did not survive.

This study was split into two separate studies, each with 4 animals in each group, for practical reasons. However, with the results of the first half of the study as they are it is hard to justify performing the rest of the study as planned. To do so without addressing the problem of the dosing intervals would be detrimental to the welfare of the guinea pigs without substantially advancing our understanding of whether the drug is or is not effective. The obvious solution would be to change the dosing intervals so that they are uniformly 8 hourly; however, as only 12 animals can be studied at a time this would necessitate small group sizes of only 4 per group, as there is no practical possibility of a third experiment. Therefore, it has been decided that, since M*O*N-DNJ did not provide for any survival or decrease in clinical signs, it will be dropped from the subsequent trial. Instead, 8 animals will receive *N*B-DNJ and 4 the placebo, all at 8 hour intervals. The increased time between doses during the day will also be beneficial for the animals as it will allow them more time to recuperate in between each, decreasing stress and thus diminishing the negative effects of cortisol on immune function.

## 7.5 Addendum

After this thesis was submitted but prior to its publishing the second efficacy study was completed. Disappointingly, no significant differences were observed at all between the treatment and placebo groups with all guinea pigs requiring sacrifice due to severe symptoms at similar times. There are two possibilities as to the failure of this study in the face of the (qualified) promise of the pilot. Firstly, it may be that the original survivor was an outlier. It may be that the inoculum of virus it received was lower than that of its compatriots, such as due to misadministration or partial inactivation of that vial either before or after freezing or, alternatively, that animal may have had a particularly strong genetic resistance.

The alternative hypothesis is that in the original experiment animals were receiving high doses during the day (due to accumulation) but then had long periods overnight in which plasma concentrations of *N*B-DNJ are likely to have fallen below antiviral minima. I originally hypothesised that these low concentration periods were why the drug failed in most cases, but it may be that the reverse was true, and it was actually the high dose periods, when the drug had accumulated, that were responsible for saving the animal. As these periods were eliminated by changing the dosing interval the drug may have never been able to reach the necessary plasma concentrations.

Whatever the reason for the failure of this study, I believe that the safest and most reasonable conclusion to these studies is that *N*B-DNJ is not antiviral against Ebola virus in guinea pigs, even

at very high, intravenous doses and that it would be inappropriate to pursue any further *in vivo* studies of this compound for this indication.

# 8 Discussion

## 8.1 Iminosugars as broad-spectrum antiviral drugs

If you read a novel on the beach this summer about terrorists threatening genocide, zombies taking over the world or even about a homosexual lawyer fighting for acceptance in the 1980s, there will be one theme in common: killer viruses. Just 60 years ago novels were being published in which the world is decimated by bacterial plagues (e.g. *I Am Legend*, by Richard Matheson) yet today you would struggle to find any contemporary fiction in which bacteria are treated as a credible threat. The state of the public consciousness is clear: viruses are an unstoppable, unknowable enemy while bacteria are a solved, old-fashioned problem that does not pose a credible threat.

The reason for this sea change is obvious- antibiotics have completely changed not only the way we view extant bacteria, but also emerging bacterial infections. Even though new bacterial diseases have emerged in the antibiotic era and caused some localised concern, such as Legionnaires' and Lyme Disease, existing antibiotics have generally proved useful allowing for rapid development of successful treatment protocols. Conversely, when new viral diseases emerge, such as Ebola, SARS or Hepatitis C, specific therapies take decades to be licensed, if at all.

DNJ derived iminosugars have been known for decades to be effective against a broad spectrum of enveloped viruses and viral families. Could they be the solution to the problem of viral disease in the same way penicillin heralded the turning point in the war on bacteria? Biotechnology companies such as United Therapeutics (UT) believe they could be; UT's lead molecule M*O*N-DNJ (also known as UV-4) is currently in phase I clinical trials which, if successful, will lead to phase II trials for dengue fever and possibly influenza and EVD.

This thesis has aimed to aid the development of iminosugars as antivirals by coming to a better understanding of their mechanism of action. By understanding which of the effects of iminosugars are vital for their antiviral activity I hope to influence the design of future drugs, while by explaining the drugs' selectivity for viruses I hope to aid licensing by assuaging fears of side effects on the patient. Although I do not seriously believe that iminosugars will have much of a role to play in the treatment of HIV (see Interlude), I chose to use HIV as a model virus for this study due to the huge body of knowledge already in existence and the wide variety of reagents in place. Equally *N*B-DNJ was chosen as a model iminosugar despite the existence of newer derivatives as it was more freely available and had already been the subject of significant research. In this section I will

review and discuss the findings of this thesis and how they advance both the development of the drugs but also our general understanding of HIV biology.

## 8.2 Iminosugars induce misfolding of the gp120 V2 loop

The existing synthesis for the mechanism of action of *N*B-DNJ against HIV was that it inhibited glucose trimming from gp120, preventing its association with calnexin and thus preventing correct V1/V2 folding leading to the production of virions with suboptimal capacity to infect new cells [177]. This synthesis supplanted previous work suggesting that V3 was misfolded [173,174]. In this study, I was able to confirm the findings of *Fischer et al.* that the V1/V2 loop was in an aberrant conformation. In addition, I was able to further refine the localisation to the V2 loop and show that this was specifically a conformational change and not just the result of steric hindrance. In addition, I was the first to perform these experiments while controlling for the "indirect mannosidase inhibitor" effects of *N*B-DNJ by using kifunensine to control for changes in glycosylation induced by the drug. In this way, I was able to be confident that the antibodies showed differential binding due to changes in the protein's tertiary structure and not just because of hindrance by or interaction with the changed glycans. In general, the variable loops possess complex glycans while the constant regions have high mannose, so if one were just observing changes in glycosylation one would expect regions like V1/V2, with its 6 complex glycans in HXB2, to be affected [48].

The antibodies that were able to distinguish *N*B-DNJ treated and untreated rgp120 were ARP3074, ARP3075, ARP3218 and ARP3211. The first three of these all target conformation dependent epitopes in the crown and *C*-terminal half of V2, between residues 176 and 192 [252]. The exact epitope of ARP3211 is unknown, only that it is conformational and within V2. Meanwhile, the insensitive antibody ARP3077 targets a linear region in the *N*-terminus of V2, between residues 162 and 172, meaning that it does not overlap with the region detected by the conformation sensitive antibodies [252]. This could have two interpretations- either the change in conformation is limited to the *C*-terminal half of V2 or the misfolding is present over all of V2 but is only detectable by conformation-sensitive antibodies. The latter seems more likely but, without more antibodies one cannot say for certain, therefore it makes most sense to consider the V2 loop as a whole rather than try to localise the exact portions involved more closely. One concern is that the epitopes of the conformation sensitive mAbs are centred on a complex glycan, N176, so it is possible that the nature of this glycan could have substantial effects on the access of antibodies to the epitopes beneath. I have tried to control for this as much as possible by using kifunensine to prevent processing of this glycan beyond $Man_9GlcNAc_2$ in the untreated and $Glc_3Man_9GlcNAc_2$ in the

*N*B-DNJ treated samples. This is not so great a difference as that between complex and high mannose glycans but it is possible that those three glucose residues may be enough to hinder access of the antibodies. The best way to confirm this would be to digest the glucose residues away with α-glucosidase I and II, but these enzymes are not readily available. By moving the rgp120 expression to HEK 293T cells I was able to expose the proteins to the actions of endomannosidase, the first time this has been done with *N*B-DNJ treated rgp120. Endomannosidase is able to cleave both $Glc_3Man_9GlcNAc_2$ and $Man_9GlcNAc_2$ to $Man_8GlcNAc_2$, thus ensuring that both treated and untreated protein have identical glycosylation by the time they are secreted. The sensitive antibodies were still able to distinguish treated and untreated rgp120 produced in this way, theoretically confirming that the differential binding was not due to retained glucosylation. However, the endomannosidase is far from efficient and glycan profiling showed that glucosylated glycans still remained.

It was interesting that no changes could be observed of the binding of antibodies to recombinant gp140 "SOSIP" trimers. This was almost certainly due to the fact that BG505 SOSIP is designed to induce broadly neutralising antibodies and so these tend to be the reagent available for it. Recombinant $gp120_{IIIB}$ has a multiplicity of antibodies available for it that each target well defined epitopes, but this is as a result of relatively misguided work in the early days of HIV vaccinology, when laboratory strains were of major interest. Now that we are aware of bnAbs there is no great value in developing a large number of defined antibodies to just one specific protein. By definition, bnAbs have to be able to target broadly conserved regions of the protein and be able to contend with largely dissimilar conformations around those regions. This is likely to make them particularly tolerant of the potentially quite subtle changes in the V2 loop, if present, in the SOSIP molecule. These experiments confirmed that treated proteins were still able to bind to CD4 and then undergo the conformational rearrangements necessary to expose the co-receptor binding site, so it is worth considering how misfolding in the V2 region might affect the fusion capabilities of the protein.

## 8.3 Significance of V2 misfolding

The structure of the V1/V2 loop has been expanded greatly by recent cryo-EM and crystallographic studies. These studies have either looked at the region on the BG505 SOSIP [59,61] or at a scaffolded V1/V2 in complex with neutralising antibodies [253,254]. These studies have shown that V1 and V2 each contribute two β-sheets to form a 4-sheet Greek key motif. These sheets are relatively well conserved thus allowing the region to conserve conformation, and presumably function, while still permitting diversity of sequence [253]. From between these sheets the actual

"loops" of V1 and V2 emerge; these appear to be unstructured and hence have not been visible in structural studies.

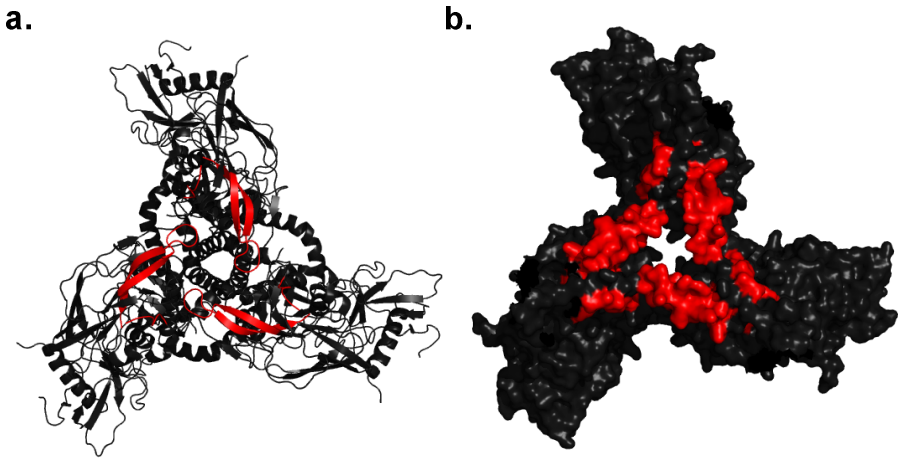



**Figure 8-1** The V2 region highlighted on the BG505.664 SOSIP trimer crystal structure (PDB: 4NCO). The V2 loop is highlighted in red on cartoon (a) and surface (b) models. Glycans are omitted for clarity. The V2 is only partially reconstructed on this model, with 10 residues (of 40 in the whole loop) between 181 and 191 omitted due to inconsistent or absent electron density. The trimer is viewed from above, looking down towards the virion membrane along the axis of symmetry.

It is tempting to presume that *N*B-DNJ might somehow disrupt this Greek key motif, since its conservation suggests import. However, it seems unlikely that one could disrupt the V2 portion of the region without affecting the V1 component, which appeared to fold normally in the presence of *N*B-DNJ. Indeed, the epitope of the sensitive antibodies is not in the ordered region but in the disordered true "loop" region of V2. It seems somewhat paradoxical that there should be conformation-sensitive antibodies to a disordered region so my hypothesis is that it is not the "loop" conformation on which they are dependent but actually the Greek key beneath that may, for example, hold the two ends of the loop close enough together for the antibody to bind both of them. When the protein is denatured the Greek key is lost and the two ends drift apart, drastically lowering antibody affinity. This hypothesis is supported by the fact that the residues at either end of the "loop" are vital for all of the sensitive antibodies to bind [252] and these are the only residues of the loop visible in the crystal structures. This may suggest that the misfolding really does occur within the ordered, β-sheet region of V2.

If one looks at the location of V2 in the trimer structure (Figure 8-1) one can see that the V2 loop sits right across the interfaces of the three protomers; V1/V2 and V3 are known to be vital for maintaining the stability of the inter-protomer interactions [59,61]. Hence one could imagine that even just small changes in V2 as measured on the rgp120 monomer be far more disruptive to the quaternary structure of the trimer as whole. However, experiments on *N*B-DNJ treated BG505.664

SOSIP failed to show any difference in the binding of quaternary sensitive bnAbs. It may be that the mutations in SOSIP designed to stabilise the trimer help counter the destabilising effects of the drug. To test this hypothesis I tested the antiviral activity of *N*B-DNJ against BG505.664 pseudotyped virions. This meant that the virions had envelope proteins identical to those used in the SOSIP trimer but for the extra disulphide bond, which would have rendered them non-functional due to an inability to expose gp41. Against these pseudotypes, *N*B-DNJ had an $IC_{50}$ of 91 μM (95% CI 40 to 142 μM), making it insignificantly different from the sensitivity of LAI ($IC_{50}$ 86 μM, 95% CI 76 to 95 μM). Thus the failure to see any differential binding of antibodies to treated SOSIP is most likely either due to the fact that very few bnAbs were used (only 8, versus over 40 tried against the monomer) or that bnAbs, by their very nature, are more forgiving of slight conformation variations than other antibodies.

Importantly, the fact that the V1/V2 deleted HIV LAI mutants were still susceptible to inhibition by *N*B-DNJ proves that misfolding cannot be limited to that region alone. It is possible that V2 misfolding is the usual mechanism of action and that the drug has a different mechanism against the deletion mutants, but all the evidence is that V1/V2 folding is independent of the rest of the gp160 protein [209] and hence it is unlikely that compensated V1/V2 deletion such as these would greatly change the structure of the protein as a whole. However, while the structure remains constant it is likely that the interaction of the protein with calnexin is significantly affected by the loss of the roughly a third of gp120's glycans. Hence it is not unreasonable to assume that the ΔV1/V2 gp160 mutants associate with calnexin and foldases differently to the wild type, potentially opening up novel avenues for inhibition.

The alternative interpretation is that, as well as V1/V2 misfolding, other regions of gp160 also misfold, such as regions of the inner domain or gp41. These regions' response to *N*B-DNJ has not been studied, either by myself or others, due to the scarcity of antibodies targeting these regions. None the less, ΔV1/V2 mutants do have some resistance to the drug, at least confirming that V1/V2 misfolding is at least partially responsible for the mechanism of action.

## 8.4 The proportion of misfolded gp120 is dose dependent, but does not directly correlate with the antiviral effect

All previous studies on *N*B-DNJ induced misfolding have looked at a single, high concentration of the drug. This is sensible as one wants the greatest possible effect when one is screening as it will be easier to detect. However, since these doses (usually 1 or 2 mM) are way above antiviral

concentrations the state of the proteins at high doses may not be an accurate model for what happens at lower, more relevant doses. To investigate this hypothesis I, with substantial help from Dr Omer Dushek, developed a mathematical model that could be used to infer the proportion of any population of proteins that was misfolded using solely ELISA data.

By applying this model to rgp120 I was able to show that, in HEK 293T cells, the proportion of misfolded protein is dose dependent, increasing with the concentration of drug. Interestingly, this increase in misfolded protein perfectly correlated with the increase in $Glc_3Man_9GlcNAc_1$ FOS. This FOS is representative of the level of *functional* GlcI inhibition- that is to say it may be necessary to inhibit 90% of the enzymes before one actually sees a decrease in deglucosylation due to reserve capacity. FOS is a functional measure and so when FOS is at 50% of its maximum level we can infer that 50% of glycans going through the ER are retaining their glucoses. This correlation, therefore, shows that V2 folding is directly related to glucose removal and thus FOS can be used as a surrogate measure of folding. FOS is regularly used as an iminosugar screening tool to select for the most effective glucosidase inhibitors; with these data one can now say that FOS is representative of the level of protein misfolding occurring.

A significant observation is that when one compares the levels of misfolding with the antiviral effects of the drug one sees infectivity of virions falling sharply as the proportion of misfolded proteins approaches 20%, decreasing more slowly between 20 and 40% and then any increase in misfolded protein above 40% does not have any additional effect on infectivity (Figure 8-2 a). These data potentially hold the key to solving the age old paradox of why iminosugars are selectively damaging to viruses without harming cellular functions.

These data show that only a tiny proportion of proteins need to be misfolded for a substantial antiviral effect- only 11% misfolding is enough to cause a 50% decrease in virion infectivity. I believe that the reason for this amplificative effect lies in the nature of the fusion process. gp120/gp41 functions as a trimer, so immediately we can hypothesise that it would only take one misfolded protomer in that trimer to abrogate function completely. To understand the effects of such amplification it is worth constructing functions to show how such a requirement could impact on virion infectivity (Figure 8-2 b and c).

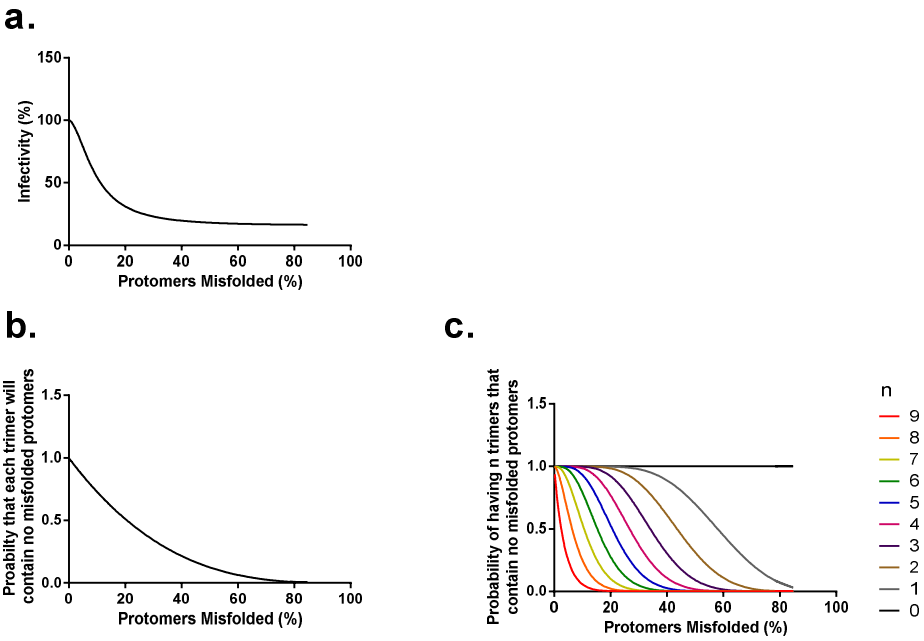



**Figure 8-2** The relationship between gp120 protomer misfolding and the infectivity of HIV LAI virions (a) as interpolated through treatment with varying concentrations of *N*B-DNJ (see Section 4.3.5 for full details). For comparison are two mathematical functions, the first demonstrating the probability of producing a trimer comprising only correctly-folded protomers as the proportion of misfolded protomers increases (b). The second function expands on the former by showing the probabilities that a virion with 9 trimers in total would possess *at least* n trimers containing no misfolded protomers at all.

The simplest function asks what the probability is of selecting three correctly folded protomers from a pool of folded and misfolded protomers (Figure 8-2 b). As the proportion of misfolded protomers grows the probability of selecting a perfect triplet falls. However, HIV virions do not possess just one single trimer on their surface, they possess several, with recent estimates placing the average number at around 7-8 spikes [42]. These spikes have to function together to form a fusion pore as no one spike would be able to overcome the energy barrier on its own. Thus we can hypothesise not only that any one misfolded protomer is enough to render a trimer inactive, but also that any one inactive trimer is enough to render a cluster inactive, potentially explaining the amplification effect.

To model this I imagined a virion with 9 spikes (ergo 27 protomers) and calculated the probabilities that, under these conditions, the virion would have n functional trimers (Figure 8-2c). In this model, if the virus needed only one functional trimer to be infectious then the chances of it having one are still good even when a large proportion of protomers are misfolded; there is a 57% chance even when 50% of protomers are misfolded. However, current estimates are that a virion with 9 spikes would need at least 5 of them to cooperate in order to overcome the fusion energy barrier [73]. Looking at this model (the blue line in Figure 8-2c) we can see that the odds go against

the virus having at least five perfect trimers when just 20% of the protomers are misfolded. This could be extrapolated to infectivity- if one considers a population of viruses budding from a cell, picking up a random sample of trimers as they bud, then one would predict that the probability of any one virus getting 5 correct trimers would translate directly into the infectivity of population as a whole. For example, when 20% of the protomers are misfolded 50% of the virions could expect to have at least 5 working trimers, therefore infectivity would be at 50% of the normal level when all the virions would have 9 correct trimers.

These mathematical models are very simple and obviously do not explain the observed data perfectly. One of the biggest discrepancies is that infectivity of the virion never reaches 0 however high the drug concentration (or its correlate, misfolded proteins) goes. This could suggest that the misfolded proteins are not completely non-functional, merely less functional, thus once they are in the majority infectivity cannot be further diminished, but is not at zero.

Another problem with such simplistic models is that it is unlikely that, even in the presence of no drug, all the protomers are correctly folded. It is known that only around 10% of HIV virions are infectious anyway and it is possible that this is partly due to misfolded gp120- ergo, *N*B-DNJ would merely be magnifying a problem that is already in effect. However, although these simplistic models are not necessarily predictive they can be thought of as illustrative of the problems HIV virions have assembling their fusion machinery and how increasing the proportion of misfolded protomers could affect the virion as a whole.

Thus we have seen how only a small perturbation of protomer folding can have a massive effect on viral infectivity through this amplificative process induced by these large, multi-component complexes. However, such complexes are not unique to viruses- why do human cells not suffer similar effects on their $F_1/F_0$ ATP-synthase complexes, for example? I believe that the answer lies in the lack of redundancy in viral systems. Viruses are simple pathogens that have small genomes and are under constant attack from the immune system. This means that single genes (such as *env*) have to perform multiple functions and that evolution will favour variants that are increasingly novel (and hence able to escape neutralisation by the immune system) rather than increasingly functional.

This means that viruses are always on the edge of failure, relying on their small generation times and massive "litters" (thousands of daughter viruses being produced from each parent genome) to compensate for the fact that they are incredibly inefficient. Thus it only takes a small perturbation, such as the insertion of a few misfolded protomers in a larger complex, to tip the balance. Conversely, humans have very long generation times and small litters and hence evolution

has favoured redundancy and error correction pathways designed to make our cellular processes much less sensitive to external perturbation. For example, it may be that human proteins are able to utilise alternative folding pathways in the absence of calnexin, such as via BiP, or just fold on their own, albeit more slowly. gp160, however, on the edge of adequacy, has greater need of calnexin's help and the tiny number of spikes actually incorporated into each virion leave little redundancy.

A counter argument to this is that *N*B-DNJ has previously been shown to induce misfolding of the E2 envelope protein of bovine viral diarrhoea virus (BVDV), as measured by decreased binding of conformationally sensitive antibodies. Importantly, this effect was directly correlated to a decrease in virion infectivity [192], refuting the notion of any amplificative effect. It may be that gp120 is indeed less sensitive to *N*B-DNJ induced misfolding than BVDV but the amplification effect makes up for this. This highlights the need to examine the effects of α-glucosidase inhibition on a range of cellular and viral proteins to understand the link between misfolding and decreased activity in more general terms. However, the difficulty of assessing correct folding makes such large scale studies prohibitive. An alternative explanation for the better correlation in BVDV may be that there is an alternative mechanism of action of the drug, which is more significant than α-glucosidase inhibition, that could be accounting for the antiviral effect independently of the E2 misfolding. The obvious candidate would be inhibition of the BVDV viroporin p7 as the related HCV p7 is known to be inhibited by long alkyl chain iminosugars such as *N*N-DNJ [193] (see Section 1.7.2). However, it is unlikely that *N*B-DNJ could have this effect on BVDV p7 as it is has been shown to have no activity against HCV p7 and *N*B-DGJ, an enantiomer of *N*B-DNJ which has no anti-α-glucosidase activity, is not antiviral [192,193].

## 8.5 *N*B-DNJ affects the oxidative folding of gp120

The pulse chase studies of gp120 oxidative folding in the presence of *N*B-DNJ showed that the drug causes gp120 to reach its final conformation more quickly, as evidenced by the increased rate of signal peptide cleavage. The trigger for signal peptide cleavage is not known for certain, but is believed to be the result of contacts between C1 and C5, the two domains that are most distant in the primary sequence (Ineke Braakman, personal communication). This supports the findings from the antibody and functional studies that the core gp120 folds normally in the presence of *N*B-DNJ, hence why these signal peptide cleavage and secretion are permitted. Our data suggest that the V1/V2 loop has a greater reliance on calnexin than the core (as it is the only region to be

measurably misfolded) so it may be that the folding of this region is the rate limiting step; misfolded V1/V2 is recognised by UGGT1 and the protein is reglucosylated, forcing it to linger in the ER and have more chances to fold. When *N*B-DNJ prevents calnexin association the only limiting step is the cleavage of the signal peptide, thus the protein leaves the ER when the core is folded but not necessarily V1/V2. This hypothesis is supported by the fact that deletion of any of the three disulphide bonds in V1/V2 permits gp160 secretion while deletion of any (bar one) of the other seven disulphides prevents it [129].

The other finding of the pulse-chase experiments was that gp120 formed disulphide bonds more quickly when treated with *N*B-DNJ, with fully reduced species becoming undetectable at 5 minutes after the beginning of the pulse labelling. However, these are almost certainly aberrant disulphides rather than the correct arrangements. An analogy might be the idea of trying to tie complex knot in a piece of sticky string being threaded through a hole. If the knot tyer grasps the string as it emerges then he or she can tie the knot relatively easily. If however the tyer waits for the whole string to emerge it is likely to become tangled and stuck to itself, meaning that they must first untangle it before it can be correctly tied. Calnexin, with its accompanying disulphide isomerase ERp57 can function as this knot tyer, with ERp57 forming mixed disulphides with, and hence protecting, the emerging cysteines, thus accounting for the reduced appearance on the pulse-labelling gels. Thus I predict that while disulphide isomerisation is still possible in the absence of calnexin (by, for example, PDI) it is now harder and it is possible that the protein leaves the ER with mispaired disulphide bonds, and the majority of the disulphides that are dispensable for gp160 secretion are in V1/V2. Therefore, these data provide a potential explanation for why *N*B-DNJ induced misfolding is limited to V1/V2.

These oxidative folding data largely support previous data that show that α-glucosidase inhibition can affect disulphide bond formation of glycoproteins. In the presence of *N*B-DNJ, the host enzyme tyrosinase also exits the ER faster and also appears to fold faster, reaching a compact state almost instantly instead of after several hours normally. Treated tyrosinase is also non-functional, being unable to recruit the copper ions which are vital to its activity [255]. The study questions the assumption that *N*B-DNJ does not affect host proteins. Tyrosinase does not form any sort of multi-protein complex, nor is it likely to be as sparsely available in the cell as gp120 is on the virion surface, yet it is acutely sensitive to calnexin cycle inhibition.

Tyrosinase is vital to the process of melanogenesis and B16 mouse melanoma cells treated with *N*B-DNJ go from being dark to pale due to its inhibition. However, children suffering from CDG-IIb (the genetic lack of α-glucosidase I, see Section 1.7.3) do not have any disorders of pigmentation at

all (Sergio Rosenzweig, personal communication), though their cells are resistant to viral infection [202]. Thus it may be that B16 cells have unreported deficiencies either in the melanogenesis or protein folding pathways which makes their tyrosinase especially sensitive to *N*B-DNJ mediated inhibition. For example, there may be deficiencies in another chaperone or foldase that would normally compensate for calnexin inhibition; this could lead to misfolding of a variety of cellular proteins in the presence on *N*B-DNJ but tyrosinase inhibition stands out due the black-and-white nature of its phenotype.

## 8.6 Deletion of glycan N241 can imbue partial resistance to *N*B-DNJ

One of the most interesting findings of this study was that of a single glycan in C2, N241, whose deletion can decrease the antiviral potency of *N*B-DNJ by almost a whole order of magnitude, increasing the $IC_{50}$ to 767 μM (95% CI 378 to 1156 μM). The confidence interval on this figure is large because the virus was so resistant that it was not possible to reach maximum inhibition with safe concentrations of *N*B-DNJ. Nonetheless, in a survey of the susceptibility of over 30 HIV strains to *N*B-DNJ none showed an $IC_{50}$ greater than 400 μM [214], making this the most resistant strain yet discovered.

In one experiment the N241S mutants was also less infectious to TZM-bl cells, even in the absence of any drug, but this effect was not observed on repetition of the experiment, for unknown reasons. However, N241S is a highly conserved site and other studies have identified a substantial (though not total) loss of infectivity when it is deleted [215]. The evidence, therefore, suggests that this glycan is helpful but not vital for virion infectivity, and that its deletion makes *N*B-DNJ less able to suppress infectivity further. I believe that this is evidence that folding of individual protein domains may be dependent on specific glycans and, in this case, the glycan for V1/V2 is N241.

One might assume that calnexin-mediated folding of specific domains would be dependent on glycans located within that domain. However, folding must require flexibility so it would be more likely that calnexin would hold the protein at a glycan distal to the site that needs to fold. If one looks at the location of N241 on the crystal structure of BG505.664 SOSIP one can see that while it is close to the V2 loop in the primary sequence (it is 40 amino acids away), in the tertiary structure it is located on the opposite side of the protein, sitting underneath the bulk of gp120 near the virion membrane. Opposite as it is, calnexin could bind the protein here thus leaving the opposite face unobstructed for foldase access. This is consistent with the crystal structure of calnexin, which shows the ERp57 binding domain removed from the lectin domain by a long 140 Å

arm [256], which may make it easier for ERp57 to bind domains further from the calnexin linked glycan than those close. It is also consistent with *Papandreou et al.* which showed that calnexin binding to gp160 occludes access of antibodies to the V1/V2 and C2 domains [257].

Thus we can construct the following model for *N*B-DNJ mechanism of action (Figure 8-3): under normal circumstances gp160's glycan N241 would be trimmed to $Glc_1Man_9GlcNAc_2$ whereupon it would bind to calnexin. This interaction facilitates the folding of the V1/V2 loop. This loop is slower to fold than the core of the protein so UGGT prevents the protein from dissociating from calnexin until it is correctly folded, then the protein trimerises and is secreted to the Golgi. When *N*B-DNJ is present, N241 remains triglucosylated and the protein does not bind calnexin. The core of the protein is able to fold without it but V1/V2 is not and without calnexin to retain it in the ER the gp160 is limited only by signal peptide cleavage and is able to leave the ER early, before V1/V2 can fold. In the N241S mutant the V1/V2 loop cannot use calnexin to fold anyway, and the drug can only work by preventing other domains from utilising calnexin. As these domains are far less reliant on the chaperone than V1/V2 it takes much higher concentrations for any effect to be seen.

## 8.7 Failure to remove mannose residues from gp160 can severely diminish infectivity of HIV LAI virions

One of the most surprising findings of this study was that the iminosugar kifunensine, a powerful inhibitor of ER mannosidase I, was itself antiviral. This was concerning as kifunensine had been used in the previous antibody binding and pulse-chase studies as a control for the effects of *N*B-DNJ on glycan processing beyond the calnexin cycle. As kifunensine was used in both the *N*B-DNJ treated and untreated controls this should not cause any confounding, however, as the only variable was the presence of *N*B-DNJ so any changes in the protein can only be ascribed to that drug.

Nonetheless, it was important to see whether kifunensine could affect the gp160 structure on its own. Such an effect would be a significant finding as crystallographers often use kifunensine to make glycoproteins completely EndoH sensitive for easier crystallisation- if the drug did affect protein conformation it could call these structures into question! There were two possible reasons why this may occur: firstly, mannose trimming by ER ManI can serve to decrease a protein's ability to be reglucosylated by UGGT1, providing an escape route from the calnexin cycle for terminally misfolded proteins. Kifunensine could, therefore, cause misfolded proteins to stay in the calnexin cycle indefinitely and thus act as competitive inhibitors of gp160 folding. This hypothesis was

disproved using radioactive pulse-chase experiments, which showed that gp120 folded with the normal kinetics in the presence of kifunensine.

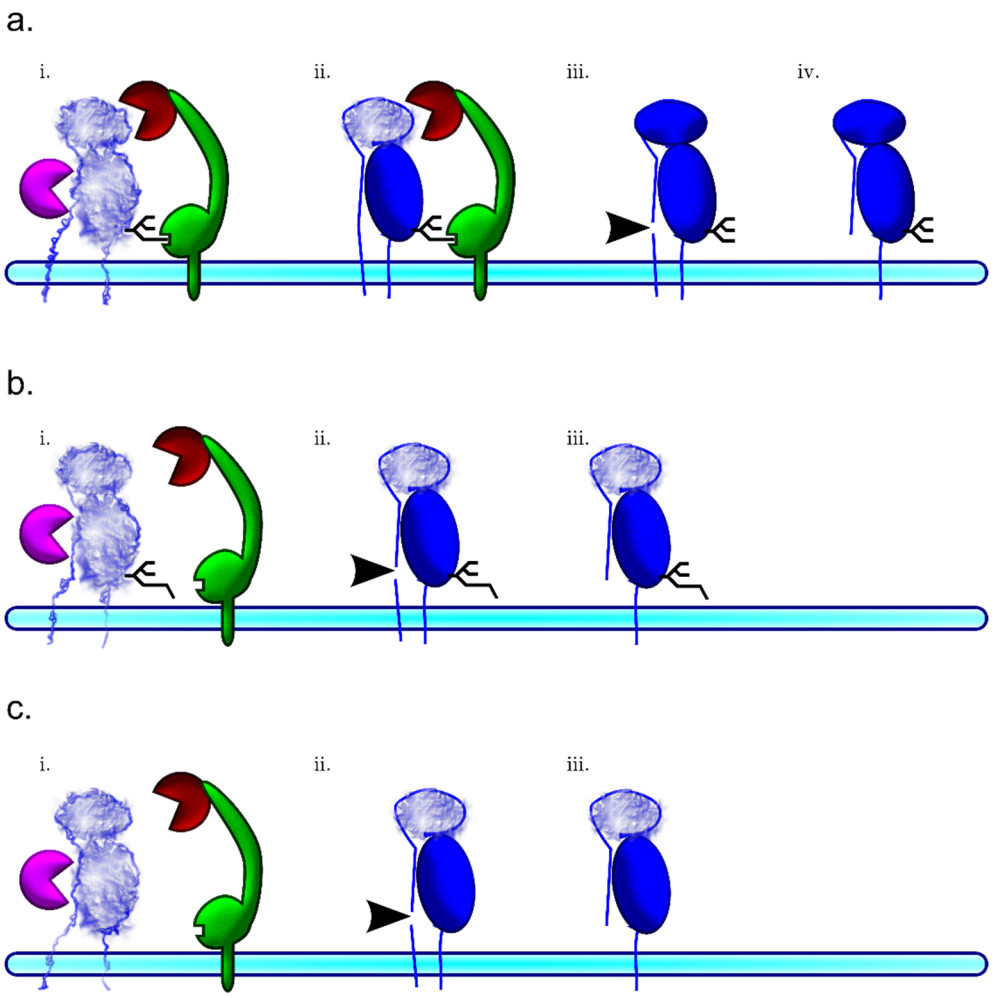



**Figure 8-3** Model for the calnexin-dependent folding of gp160 on the luminal face of the ER membrane (a) and how it is affected by *N*B-DNJ (b) or glycan N241 deletion (c). gp160 is shown in blue and comprises a core (membrane proximal), V1/V2 domain (membrane distal) and two transmembrane domains- that of gp41 (right) and the signal peptide (left); the N241 glycan is shown as a branched structure. Calnexin is shown in green with the associated ERp57 in red, while the purple pac-man represents alternative chaperones and foldases such as PDI and BiP; the signal peptidase is shown as a black arrowhead. Under normal circumstances (a) gp160 associates with calnexin via N241 (i) and is acted upon by calnexin dependent (ERp57) and non-dependent foldases. The core folds first, allowing C1 and C5 to come together (ii) but signal peptide cleavage cannot occur until V1/V2 is folded and the protein is released from the calnexin cycle by GlcII (iii); the protein is now fully folded and can be exported to the Golgi. In the presence of *N*B-DNJ (b) calnexin cannot bind N241 as it remains triglucosylated (i). The core folds normally but without calnexin to prevent it the signal peptide is removed prior to V1/V2 folding (ii) permitting early export of the immature protein (iii). N241 deletion recapitulates this process in the absence of *N*B-DNJ (c) as calnexin has no ligand by which to bind. Folding of the core is likely partially dependent on calnexin binding alternative glycans as *N*B-DNJ still has some inhibitory ability, but is mostly dependent on alternative pathways.

The alternative hypothesis was that, because kifunensine is also an inhibitor of the mannosidases involved in the ERAD pathway, it caused misfolded proteins that would normally be degraded to instead be secreted and incorporated into the virions, whereupon they would induce the same amplificative inhibition of fusion as hypothesised for *N*B-DNJ. I attempted to test this hypothesis by using an alternative ERAD inhibitor, eeyarestatin I, which would prevent ERAD without

affecting the glycan structures. This inhibitor was not antiviral, suggesting that it was not the ERAD inhibiting properties of kifunensine that were the root of its antiviral effects. However, eeyarestatin I proved to be highly toxic to the cells and so only very short, 8 hour incubation periods could be used. Eight hours is enough time for virions to be produced and secreted but the population is likely to be contaminated with numerous virions whose production began before the inhibitor was added, potentially confounding the experiment. An alternative reason why ERAD inhibition was unlikely to be the cause of the antiviral effect is that gp160 production is reported to be very efficient, with an undetectable level of degradation [128]. Conversely, inhibition of ERAD by kifunensine is known to be proviral in HCV, apparently by increasing levels of the envelope glycoprotein E2 [216]. Antibody binding experiments were not appropriate for this as the dramatic glycan differences induced by the drug would have caused considerable confounding.

This leaves changes in the glycans themselves as the potential source of the effect; $Man_9GlcNAc_2$ glycans are large and could feasibly cause steric hindrances that affect gp120/gp41 function. Although kifunensine was previously reported to only have significant effects on gp120 glycosylation at relatively high concentrations (>5 μM [150]) this study confirmed that the glycans are still being affected even at sub micromolar concentrations, meaning that the antiviral effect does correlate with the effect on glycans. Kifunensine has been observed to cause small decreases in infectivity in gp160 pseudotyped viruses in the past (Katie Doores, personal communication) but this has not been published. However, DMJ, which is a less potent ER ManI inhibitor, has been variously described as being both antiviral and not, depending on the study (see Section 1.6.5 for full details); this study confirmed that it was antiviral, though it was considerably less potent than kifunensine or *N*B-DNJ.

What is most surprising is that swainsonine also showed antiviral activity, when all previous studies have shown that viruses produced in the presence of swainsonine are fully infectious. This could be put down to differences in methodology or the virus used, as I have been using molecular clones to look specifically at envelope biogenesis and thus excluding many of the stages of the viral lifecycle that would have been seen if I had used primary cells and isolates. However, HIV LAI has been produced from the exact same clone in HEK 293 cells deficient in the $GnTI^{-/-}$ enzyme (thus preventing the formation of hybrid or complex glycans) with no decrease in infectivity [180].

Inhibition of complex glycan formation in mice by deletion of Golgi ManII has been shown to induce severe autoimmune disease [258], yet neither patients on *N*B-DNJ nor the children with ER GlcI deficiency have had any such immunopathy. This highlights the importance of the endomannosidase as an escape route for glycoproteins to undergo complex glycosylation. As there

is no such escape route for mannosidase inhibition the side effects are much more severe. This means that even if kifunensine is antiviral it is unlikely to make a good pharmaceutical. Instead, the significance of this finding is in our understanding of the mechanism of action of ER α-glucosidase inhibitors as antivirals. If we accept that changes in glycosylation have antiviral capabilities on their own, we must now ask what proportion of the antiviral effects of α-glucosidase inhibitors is due to their effects on glycans versus their effects on protein folding via calnexin. The best way to answer this question would be to produce HIV virions in calnexin$^{-/-}$ and calreticulin$^{-/-}$ cells, to allow for normal glycan trimming while still inhibiting the folding pathways.

## 8.8 *N*B-DNJ has limited activity against Ebola virus in a guinea pig model

This study has shown the *N*B-DNJ has been able to save 1/4 guinea pigs from EVD mortality and reduced the severity of clinical signs in the other 3, although time of death was not affected. This may not seem clinically significant but as the model is usually 100% lethal it is cause for further investigation. This study was also able to show that both *N*B-DNJ and M*O*N-DNJ are safe and well tolerated, even at high doses, in guinea pigs. Both drugs were able to inhibit ER α-glucosidase II *in vivo,* although only *N*B-DNJ inhibited α-glucosidase I. This effect is probably due to the 15x higher dose of *N*B-DNJ that was administered (M*O*N-DNJ is only twice as active against Ebola i*n vitro* [247]), and may explain why *N*B-DNJ was effective yet M*O*N-DNJ was not.

Interestingly, while this study was in progress a paper was published showing that if all of the potential *N*-glycosylation sites (PNGS) in the Ebola glycoprotein GP1 were mutated out then not only could the glycoprotein still fold but the resulting pseudotyped virions were actually more infectious [236]! This study concluded that the glycans on GP1 are there to protect the protein from neutralisation by antibodies or digestion by proteases, and are not vital for protein folding or receptor binding. In fact, they decrease receptor binding but this is an acceptable sacrifice for the increased serum resistance.

This fascinating paper has substantial implications for my study as if GP1 can fold even without its glycans, when calnexin binding would be impossible, then it is hard to argue that inhibiting this step should lead to misfolding. So how is it that α-glucosidase inhibitors can inhibit the infectivity of these viruses [246,247]? There are several possible explanations for this discrepancy. Firstly, there is growing evidence that α-glucosidase inhibition can lead to modification of the cytokine profile induced by another haemorrhagic fever virus, dengue virus, with potentially protective effects (Andrew Sayce, unpublished data). However, while this effect could explain how iminosugars have

been used to protect animals from EVD it is unlikely to explain the *in vitro* data since the cells used were not cells of the immune system.

A more likely explanation is that while all of the glycans were removed from GP1, the two glycans on GP2 were not mutated out. As the calnexin cycle occurs in the ER, when GP1 and GP2 have not yet been cleaved from the precursor GP protein, it is possible that these two glycans are sufficient to permit calnexin dependent folding. I would be very interested in obtaining these mutant glycoproteins and seeing whether they immunoprecipitate with calnexin and whether *N*B-DNJ still has antiviral activity against them.

A final explanation could be that it is the inhibition of processing to complex glycans, rather than the effects on the calnexin cycle, which underlie the antiviral effects of iminosugars in this virus. *Lennemann et al.* suggests that the glycans of GP1 limit virion infectivity as the price for improved serum resistance; therefore, it is possible that more massive high-mannose glycans induced by the drug could disrupt this balance and make the virus even less infectious. If this hypothesis is true, then one would expect kifunensine to be antiviral against Ebola as well, although, as detailed above, it is unlikely to be as suitable for clinical use as *N*B-DNJ.

Indeed, it is the fact that *N*B-DNJ is already licensed for clinical use in the treatment of type I Gaucher's Disease that is its best hope for being licensed for use against Ebola as well. The 2014 West Africa outbreak, which is continuing as I write, has resulted in a drive to accelerate clinical Ebola research and get treatments into the clinic as soon as possible. *N*B-DNJ has been in people for over a decade, with no serious adverse effects, and the protocols for manufacturing it are well established. These factors combined could make it a valuable addition to the antiviral pharmacopeia in the face of this terrible tragedy.

## 8.9 Conclusion and future directions

This thesis has expanded the field of antiviral iminosugar research by demonstrating exactly how ER α-glucosidase inhibition results in less infectious viruses without compromising the functions of the host. The majority of this work was carried out using HIV as a model virus, so the obvious next step would be to see if the same principles can be extended to other virus families, particularly those of immediate clinical importance like dengue and influenza viruses. One of the problems with extending this work is the cumbersome and inefficient method of using antibody binding to measure misfolding. Since the misfolding might be highly localised, as it is with gp120, one would need a wide variety of well characterised antibodies to hope to find one that is affected

by the drugs. If one had a better, faster method of assessing protein folding then one could not only identify susceptible viruses but also susceptible host proteins that could be the source of side effects.

One such method could be to use phage display libraries to quickly generate lists of the epitopes exposed on both treated and untreated proteins; one could then clone any $F_{ab}$ (antigen binding fragment) that showed differential binding and determine its specificity, thus identifying regions that were changed. However, even this method is technically demanding and requires the operator to know which protein they are interested in to begin with, thus making it less useful for screening for affected host proteins.

This study also does not address existing problems with iminosugar therapy; those of low serum concentrations and unpleasant gastrointestinal side effects. This is a matter for the medicinal chemists to contemplate, but I would like to propose a potential way to exploit the same pathway as iminosugars but with fewer side effects. I believe that UGGT1 could be a target for broad spectrum antiviral activity. UGGT1 is responsible for returning incompletely folded glycoproteins back into the calnexin cycle for further chances at folding. If UGGT1, rather than ER α-glucosidases, were inhibited then all proteins would have one chance to bind calnexin rather than zero. In lower organisms, such as trypanosomes, UGGT1 is dispensable for cell function until the cells become stressed [259], and UGGT1 can be safely deleted in mouse embryonic fibroblasts [260]. I suspect that UGGT1 would become more important during viral infections, when large quantities of glycoproteins are being synthesised simultaneously, or for particularly slow folding glycoproteins like gp160. Thus UGGT1 inhibition could be seen as a more selective form of calnexin cycle inhibition than that of iminosugars, as it could disproportionately affect viral glycoproteins. What is more, as they would not seek to inhibit glycosidases, such drugs would not have the intestinal side effects of iminosugars.

In conclusion, broad spectrum antivirals would be a paradigm shift in the history of medicine. Iminosugars are not yet that magic bullet, but this thesis has shown how glycoproteins misfolding can be exploited to have a disproportionate effect on viruses over the host cell. Even if iminosugars are not the optimal way to exploit this weakness, they are a useful experimental tool to better understand viral glycoprotein folding. Through better understanding of these pathways we can hope to design better targeted, more specific drugs that really could be the clinically effective broad spectrum antivirals the world needs.

# Appendix: Iminosugars against an unenveloped virus

## A.1 Results

Despite being tested against a wide range of enveloped viruses, iminosugars have never before been tested against an unenveloped virus. This is because it has always been presumed that as these viruses, by definition, do not have envelope glycoproteins, the iminosugars' inhibition of the calnexin cycle could not affect them. I chose to investigate this hypothesis to see whether *N*B-DNJ could affect the replication of an unenveloped virus. I chose to work with an adenovirus called AdHu5-GFP, a deletion mutant of the human pathogen human adenovirus 5. In AdHu5-GFP the E1 gene has been replaced by the green fluorescent protein (GFP), causing infected cells to fluoresce. I used this unique feature of AdHu5-GFP to develop a novel assay for screening antiviral drugs against adenoviruses. The first protocol I developed involved infecting HEK 293T cells at a low multiplicity of infection (MOI) and allowing the virus to complete approximately three replication cycles in the presence or absence of *N*B-DNJ. The cells were then harvested, stained with propidium iodide (a vital stain) and counted using a flow cytometer. This differentiated the proportion of cells that were dead, and the proportion of cells expressing GFP. The data are presented in Figure A1-1.

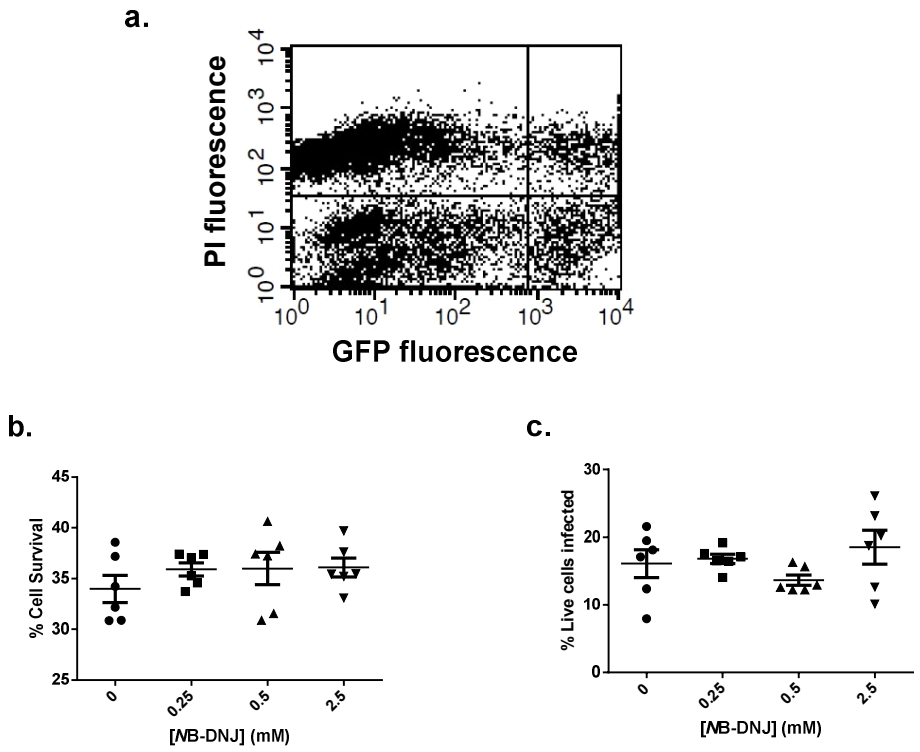


**Figure A1** HEK 293T cells were infected with AdHu5-GFP at an MOI of 1 and then treated with *N*B-DNJ for 3 days. A typical flow cytometry plot is shown in (a). This is used to divide cells into live/dead and infected/uninfected categories. Cell survival was calculated by the ratio of propidium iodide (PI) stained (dead) cells to unstained (live) cells (b). Of those cells found to be alive, those with GFP fluorescence were considered infected (c). 50,000 events were captured in (a); in (b) and (c) error bars represent standard error, n=6. There are no significant differences (1-way ANOVA, alpha of 0.05) between any of the treatment groups.

The flow cytometry data show no effect of the drug on the proportion of live cells that are infected with AdHu5-GFP or on their survival. However, the high proportion of cell death, the fact that infected dead cells may no longer fluoresce and the lack of clear quartiles in the dot plot made me concerned that flow may not be accurate enough. I devised an alternative experiment in which HEK 293T cells were infected with a very high MOI of 10 and pre and post treated with 1 mM *N*B-DNJ. The supernatant from these cells was collected at different timepoints and titrated onto naïve HEK 293Ts, using fluorescence microscopy to count the number of fluorescent (i.e. infected) cells, allowing calculation of infectious particles per ml (ip/ml). These data are presented in Figure A1-2. These data are consistent with the hypothesis that *N*B-DNJ is ineffective against AdHu5.

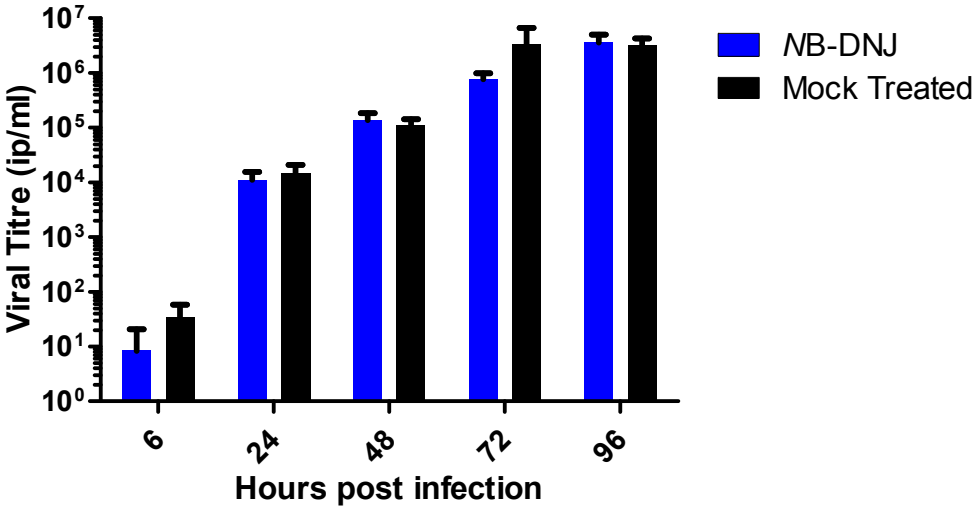



**Figure A1-2** HEK 293T Cells were treated with either 1 mM *N*B-DNJ or an equivalent volume of PBS for the indicated period of time. The supernatant was then titrated onto naïve 293T cells in quadruplicate. 48 hours later the number of fluorescent (infected) cells were counted at the lowest dilution at which there were a countable number of cells. Each column represents the mean of 3 biological replicates, each with 4 titration replicates, thus producing error bars showing 95% confidence intervals with n = 12. There are no significant (Student's $t$ test, $p < 0.05$) differences between treated and untreated cells at any point.

## A.2 Discussion

While the AdHu5 experiments failed to show any efficacy of *N*B-DNJ as an anti-adenoviral, the data are valuable for two reasons. Firstly, these data are consistent with the hypothesis that iminosugars are antiviral by means of interference with the *N*-linked glycosylation of virion envelope proteins, as has been shown for HIV [176-178]. As adenovirus appears to lack *N*-linked glycosylation of its proteins [261,262] any antiviral activity of *N*B-DNJ would have proved that there must be at least one alternative mechanism of action of the drug. Secondly, I was able to demonstrate the effectiveness of a novel assay for adenovirus treatments. Existing assays generally rely on qPCR [263] or plaque assays [264]. Plaques can take several days to form and though the use of immunohistochemical or immunofluorescent dyes can speed up the process [265] it also increases the expense. HEK 293 lineage cells are also extremely weakly adherent, so the washing and fixing steps required can result in significant damage to the monolayer. qPCR is faster, but also more expensive and requires specialist equipment.

By using a fluorescent virus I have developed an assay that is capable of detecting and quantifying infectious adenovirus particles in just 48 hrs without requiring the purchase of any specialist equipment or reagents except for the virus itself. Subsequently to developing this assay I found that a similar technique has been used to quantify adenovirus particles before [266], though this is, to my knowledge, the first time it has been used to test an antiviral drug.

## A.3 Materials and methods

The adenovirus strain AdHu5-GFP was a kind gift from the Jenner Institute, Oxford, UK. HEK 293T cells were maintained as described previously. For the flow cytometry experiments 1 x $10^6$ HEK 293Ts were infected with AdHu5-GFP at an MOI of 1 for 2 hours. The inoculum was then removed and replaced with fresh medium containing varying concentrations of *N*B-DNJ. These cells were incubated for 3 days then scraped and resuspended in their own medium. This medium was then centrifuged at 1200 rpm for 5 minutes thus pelleting both the adherent cells and any dead cells which had detached. This pellet was briefly treated with trypsin to break down cell clumps then resuspended in flow cytometry buffer containing propidium iodide. It was then subjected to analysis using a FACScan flow cytometer (Becton and Dickenson, Oxford, UK).

For the follow-up adenovirus experiments 1.25x$10^5$ cells were treated with *N*B-DNJ for 90 minutes then infected at an MOI of 10 for a further 90 minutes. The inoculum was then removed, the cells washed gently with PBS, and then replaced with medium containing drug or the PBS

equivalent. Experiments were harvested at set timepoints between 6 and 96 hours, in which the supernatant would be removed and frozen at -80°C and the cells discarded. These samples were then titrated onto naive HEK 293T cells to quantify the concentration of infectious particles. 3 x $10^4$ HEK 293Ts were added to each well of a 96-well plate and were infected with a 1-in-2 dilution series of each sample plus fresh medium, in quadruplicate. After 48 hours the plates were examined with a fluorescence microscope and the number of infected cells was quantified at the lowest countable dilution. These figures were then used to calculate the titre of the original samples in infectious particles per millilitre (ip/ml).

# References


1 Ribatti, D. Sidney Farber and the treatment of childhood acute lymphoblastic leukemia with a chemotherapeutic agent. *Pediatr Hematol Oncol* **29**, 299-302, (2012).

2 Hutchings, B. L., Mowat, J. H., Oleson, J. J., Stokstad, E. L. R., Boothe, J. H., Waller, C. W., Angier, R. B., Semb, J. & SubbaRow, Y. Pteroylaspartic acid, an antagonist for pteroylglutamic acid. **170**, 19, (1947).

3 Farber, S., Diamond, L. K., Mercer, R. D., Sylvester, R. F. & Wolff, J. A. Temporary Remissions in Acute Leukemia in Children Produced by Folic Acid Antagonist, 4-Aminopteroyl-Glutamic Acid (Aminopterin). *New England Journal of Medicine* **238**, 787-793, (1948).

4 Farber, S. Some observations on the effect of folic acid antagonists on acute leukemia and other forms of incurable cancer. *Blood* **4**, 160-167, (1949).

5 Claeyssens, M. & Bruyne, C. K. d-Xylose-derivatives with sulfur or nitrogen in the ring: Powerful inhibitors of glycosidase-activities. *Naturwissenschaften.* **52**, 515-515, (1965).

6 Paulsen, H. Carbohydrates Containing Nitrogen or Sulfur in the "Hemiacetal" Ring. *Angew.Chem.Int.Ed.Engl.* **5**, 495-510, (1966).

7 Nishikawa, T. & Ishida, N. A new antibiotic R-468 active against drug resistant *Shigella. J Antibiot (Tokyo)* **18**, 132-133, (1965).

8 Ishida, N., Kumagai, K., Niida, T., Tsuruoka, T. & Yumoto, H. Nojirimycin, a new antibiotic. II. Isolation, characterization and biological activity. *Journal of Antibiotics* **20**, 66-71, (1967).

9 Niwa, T., Inouye, S., Tsuruoka, T., Koaze, Y. & Niida, T. " Nojirimycin" as a Potent Inhibitor of Glucosidase. *Agric.Biol.Chem.* **34**, 966-968, (1970).

10 Argoudelis, A. D., Reusser, F. & Mizsak andBaczynskyj, S. A. L. Antibiotics produced by Streptomyces ficellus. II. Feldamycin and nojirimycin. *Journal of Antibiotics* **29**, 1007-1014, (1976).

11 Schmidt, D. D., Frommer, W., Müller, L. & Truscheit, E. Glucosidase-Inhibitoren aus Bazillen. *Naturwissenschaften.* **66**, 584-585, (1979).

12 Yagi, M., Kouno, T., Aoyagi, Y. & Murai, H. *The structure of moranoline, a piperidine alkaloid from Morus species.*, 1976).

13 Saunier, B., Kilker, R. D. & Tkacz Jr, J. S. Inhibition of N-linked complex oligosaccharide formation by 1-deoxynojirimycin, an inhibitor of processing glucosidases. *Journal of Biological Chemistry* **257**, 14155-14161, (1982).

14 Romero, P. A., Saunier, B. & Herscovics, A. Comparison between 1-deoxynojirimycin and N-methyl-1-deoxynojirimycin as inhibitors of oligosaccharide processing in intestinal epithelial cells. *Biochemical Journal* **226**, 733-740, (1985).

15 Gruters, R. A., Neefjes, J. J., Tersmette, M., De Goede, R. E. Y., Tulp, A., Huisman, H. G., Miedema, F. & Ploegh, H. L. Interference with HIV-induced syncytium formation and viral infectivity by inhibitors of trimming glucosidase. *Nature.* **330**, 74-77, (1987).

16 Tyms, A. S., Berrie, E. M., Ryder, T. A., Nash, R. J., Hegarty, M. P., Taylor, D. L., Mobberley, M. A., Davis, J. M., Bell, E. A., Jeffries, D. J., Taylor-Robinson, D. & Fellows, L. E. Castanospermine and other plant alkaloid inhibitors of glucosidase activity block the growth of HIV. *Lancet* **2**, 1025-1026, (1987).

17 Fleet, G. W. J., Karpas, A., Dwek, R. A., Fellows, L. E., Tyms, A. S., Petursson, S., Namgoong, S. K., Ramsden, N. G., Smith, P. W., Son, J. C., Wilson, F., Witty, D. R., Jacob, G. S. & Rademacher, T. W. Inhibition of HIV replication by amino-sugar derivatives. *FEBS.Lett.* **237**, 128-132, (1988).

18 Walker, B. D., Kowalski, M., Goh, W. C., Kozarsky, K., Krieger, M., Rosen, C., Rohrschneider, L., Haseltine, W. A. & Sodroski, J. Inhibition of human immunodeficieny virus syncytium formation and virus replication by castanospermine. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 8120-8124, (1987).

19 Gallo, R., Salahuddin, S., Popovic, M., Shearer, G., Kaplan, M., Haynes, B., Palker, T., Redfield, R., Oleske, J., Safai, B. & et, a. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* **224**, 500-503, (1984).

20 Popovic, M., Sarngadharan, M., Read, E. & Gallo, R. Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. *Science* **224**, 497-500, (1984).

21 Sarngadharan, M., Popovic, M., Bruch, L., Schupbach, J. & Gallo, R. Antibodies reactive with human T-lymphotropic retroviruses (HTLV-III) in the serum of patients with AIDS. *Science* **224**, 506-508, (1984).

22 Schupbach, J., Popovic, M., Gilden, R., Gonda, M., Sarngadharan, M. & Gallo, R. Serological analysis of a subgroup of human T-lymphotropic retroviruses (HTLV-III) associated with AIDS. *Science* **224**, 503-505, (1984).

23 Barre-Sinoussi, F., Chermann, J., Rey, F., Nugeyre, M., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W. & Montagnier, L. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* **220**, 868-871, (1983).

24 Clavel, F., Guetard, D., Brun-Vezinet, F., Chamaret, S., Rey, M. A., Santos-Ferreira, M. O., Laurent, A. G., Dauguet, C., Katlama, C., Rouzioux, C. & et al. Isolation of a new human retrovirus from West African patients with AIDS. *Science* **233**, 343-346, (1986).

25 Clavel, F., Guyader, M., Guetard, D., Sallé, M., Montagnier, L. & Alizon, M. Molecular cloning and polymorphism of the human immune deficiency virus type 2. *Nature.* **324**, 691-695, (1986).

26 Siegal, F. P., Lopez, C., Hammer, G. S., Brown, A. E., Kornfeld, S. J., Gold, J., Hassett, J., Hirschman, S. Z., Cunningham-Rundles, C. & Adelsberg, B. R. Severe acquired immunodeficiency in male homosexuals, manifested by chronic perianal ulcerative herpes simplex lesions. *New England Journal of Medicine* **305**, 1439-1444, (1981).

27 Masur, H., Michelis, M. A., Greene, J. B., Onorato, I., Stouwe, R. A., Holzman, R. S., Wormser, G., Brettman, L., Lange, M., Murray, H. W. & Cunningham-Rundles, S. An outbreak of community-acquired Pneumocystis carinii pneumonia. Initial manifestation of cellular immune dysfunction. *New England Journal of Medicine* **305**, 1431-1438, (1981).

28 Gottlieb, M. S., Schroff, R., Schanker, H. M., Weisman, J. D., Fan, P. T., Wolf, R. A. & Saxon, A. Pneumocystis carinii pneumonia and mucosal candidiasis in previously healthy homosexual men. Evidence of a new acquired cellular immunodeficiency. *New England Journal of Medicine* **305**, 1425-1431, (1981).

29 Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men--New York City and California. *MMWR. Morbidity and mortality weekly report* **30**, 305-308, (1981).

30 UNAIDS. *Report on the Global AIDS Epidemic*. (2010).

31 UNAIDS. http://www.unaids.org/en/dataanalysis/datatools/aidsinfo/. (2013).

32 Turner, B. G. & Summers, M. F. Structural biology of HIV. *Journal of Molecular Biology* **285**, 1-32, (1999).

33 Frankel, A. D. & Young, J. A. T. Vol. 67 1-25 (1998).

34 Li, H., Dou, J., Ding, L. & Spearman, P. Myristoylation is required for human immunodeficiency virus type 1 Gag-Gag multimerization in mammalian cells. *Journal of Virology* **81**, 12899-12910, (2007).

35 Perelson, A. S., Neumann, A. U., Markowitz, M., Leonard, J. M. & Ho, D. D. HIV-1 dynamics in vivo: Virion clearance rate, infected cell life-span, and viral generation time. *Science* **271**, 1582-1586, (1996).

36 De Boer, R. J., Ribeiro, R. M. & Perelson, A. S. Current Estimates for HIV-1 Production Imply Rapid Viral Clearance in Lymphoid Tissues. *PLoS Comput Biol* **6**, e1000906, (2010).

37 Levy, J. A. Pathogenesis of human immunodeficiency virus infection. *Microbiol Rev* **57**, 183-289, (1993).

38 Okoye, A. A. & Picker, L. J. CD4(+) T-cell depletion in HIV infection: mechanisms of immunological failure. *Immunol Rev* **254**, 54-64, (2013).

39 Richman, D. D., Margolis, D. M., Delaney, M., Greene, W. C., Hazuda, D. & Pomerantz, R. J. The challenge of finding a cure for HIV infection. *Science* **323**, 1304-1307, (2009).

40 Stein, B. S. & Engleman, E. G. Intracellular processing of the gp160 HIV-1 envelope precursor. Endoproteolytic cleavage occurs in a cis or medial compartment of the Golgi complex. *Journal of Biological Chemistry* **265**, 2640-2649, (1990).

41 Zhu, P., Liu, J., Bess Jr, J., Chertova, E., Lifson, J. D., Grisé, H., Ofek, G. A., Taylor, K. A. & Roux, K. H. Distribution and three-dimensional structure of AIDS virus envelope spikes. *Nature.* **441**, 847-852, (2006).

42 Chojnacki, J., Staudt, T., Glass, B., Bingen, P., Engelhardt, J., Anders, M., Schneider, J., Müller, B., Hell, S. W. & Kräusslich, H. G. Maturation-dependent HIV-1 surface protein redistribution revealed by fluorescence nanoscopy. *Science* **338**, 524-528, (2012).

43 Montagnier, L., Clavel, F., Krust, B., Chamaret, S., Rey, F., Barré-Sinoussi, F. & Chermann, J. C. Identification and antigenicity of the major envelope glycoprotein of lymphadenopathy-associated virus. *Virology* **144**, 283-289, (1985).

44 Korber, B., Gaschen, B., Yusim, K., Thakallapally, R., Kesmir, C. & Detours, V. Evolutionary and immunological implications of contemporary HIV-1 variation. *Br. Med. Bull.* **58**, 19-42, (2001).

45 Zhang, M., Gaschen, B., Blay, W., Foley, B., Haigwood, N., Kuiken, C. & Korber, B. Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin. *Glycobiology.* **14**, 1229-1246, (2004).

46 Modrow, S., Hahn, B. H., Shaw, G. M., Gallo, R. C., Wong-Staal, F. & Wolf, H. Computer-assisted analysis of envelope protein sequences of seven human immunodeficiency virus isolates: prediction of antigenic epitopes in conserved and variable regions. *Journal of Virology* **61**, 570-578, (1987).

47 Willey, R. L., Rutledge, R. A., Dias, S., Folks, T., Theodore, T., Buckler, C. E. & Martin, M. A. Identification of conserved and divergent domains within the envelope gene of the acquired immunodeficiency syndrome retrovirus. *Proceedings of the National Academy of Sciences of the United States of America* **83**, 5038-5042, (1986).

48 Leonard, C. K., Spellman, M. W., Riddle, L., Harris, R. J., Thomas, J. N. & Gregory, T. J. Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells. *Journal of Biological Chemistry* **265**, 10373-10382, (1990).

49 Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J. & Hendrickson, W. A. Structure of an HIV gp 120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature.* **393**, 648-659, (1998).

50 Chan, D. C., Fass, D., Berger, J. M. & Kim, P. S. Core structure of gp41 from the HIV envelope glycoprotein. *Cell* **89**, 263-273, (1997).

51 Freed, E. O., Myers, D. J. & Risser, R. Characterization of the fusion domain of the human immunodeficiency virus type 1 envelope glycoprotein gp41. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 4650-4654, (1990).

52 Rusche, J. R., Lynn, D. L., Robert-Guroff, M., Langlois, A. J., Lyerly, H. K., Carson, H., Krohn, K., Ranki, A., Gallo, R. C. & Bolognesi, D. P. Humoral immune response to the entire human immunodeficiency virus envelope glycoprotein made in insect cells. *Proceedings of the National Academy of Sciences* **84**, 6924-6928, (1987).

53 Haigwood, N. L., Nara, P., Brooks, E., Van Nest, G., Ott, G., Higgins, K., Dunlop, N., Scandella, C., Eichberg, J. & Steimer, K. Native but not denatured recombinant human immunodeficiency virus type 1 gp120 generates broad-spectrum neutralizing antibodies in baboons. *Journal of virology* **66**, 172-182, (1992).

54 Nara, P., Smit, L., Dunlop, N., Hatch, W., Merges, M., Waters, D., Kelliher, J., Gallo, R., Fischinger, P. & Goudsmit, J. Emergence of viruses resistant to neutralization by V3-specific antibodies in experimental human immunodeficiency virus type 1 IIIB infection of chimpanzees. *Journal of virology* **64**, 3779-3791, (1990).

55 Binley, J. M., Sanders, R. W., Clas, B., Schuelke, N., Master, A., Guo, Y., Kajumo, F., Anselma, D. J., Maddon, P. J. & Olson, W. C. A recombinant human immunodeficiency virus type 1 envelope glycoprotein complex stabilized by an intermolecular disulfide bond between the gp120 and gp41 subunits is an antigenic mimic of the trimeric virion-associated structure. *Journal of virology* **74**, 627-643, (2000).

56 Wyatt, R. & Sodroski, J. The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. *Science* **280**, 1884-1888, (1998).

57 Sanders, R. W., Vesanen, M., Schuelke, N., Master, A., Schiffner, L., Kalyanaraman, R., Paluch, M., Berkhout, B., Maddon, P. J., Olson, W. C., Lu, M. & Moore, J. P. Stabilization of the soluble, cleaved, trimeric form of the envelope glycoprotein complex of human immunodeficiency virus type 1. *Journal of Virology* **76**, 8875-8889, (2002).

58 Harris, A., Borgnia, M. J., Shi, D., Bartesaghi, A., He, H., Pejchal, R., Kang, Y., Depetris, R., Marozsan, A. J., Sanders, R. W., Klasse, P. J., Milne, J. L. S., Wilson, I. A., Olson, W. C., Moore, J. P. & Subramaniam, S. Trimeric HIV-1 glycoprotein gp140 immunogens and native HIV-1 envelope glycoproteins display the same closed and open quaternary molecular architectures. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 11440-11445, (2011).

59 Julien, J. P., Cupo, A., Sok, D., Stanfield, R. L., Lyumkis, D., Deller, M. C., Klasse, P. J., Burton, D. R., Sanders, R. W., Moore, J. P., Ward, A. B. & Wilson, I. A. Crystal structure of a soluble cleaved HIV-1 envelope trimer. *Science* **342**, 1477-1483, (2013).

60 Sanders, R. W., Derking, R., Cupo, A., Julien, J. P., Yasmeen, A., de Val, N., Kim, H. J., Blattner, C., de la Peña, A. T., Korzun, J., Golabek, M., de los Reyes, K., Ketas, T. J., van Gils, M. J., King, C. R., Wilson, I. A., Ward, A. B., Klasse, P. J. & Moore, J. P. A Next-Generation Cleaved, Soluble HIV-1 Env Trimer, BG505 SOSIP.664 gp140, Expresses Multiple Epitopes for Broadly Neutralizing but Not Non-Neutralizing Antibodies. *PLoS pathogens* **9**, (2013).

61 Lyumkis, D., Julien, J. P., De Val, N., Cupo, A., Potter, C. S., Klasse, P. J., Burton, D. R., Sanders, R. W., Moore, J. P., Carragher, B., Wilson, I. A. & Ward, A. B. Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer. *Science* **342**, 1484-1490, (2013).

62 Geijtenbeek, T. B., Kwon, D. S., Torensma, R., van Vliet, S. J., van Duijnhoven, G. C., Middel, J., Cornelissen, I. L., Nottet, H. S., KewalRamani, V. N. & Littman, D. R. DC-SIGN, a Dendritic Cell–Specific HIV-1-Binding Protein that Enhances¡ i¿ trans¡/i¿-Infection of T Cells. *cell* **100**, 587-597, (2000).

63 Klasse, P. J. The molecular basis of HIV entry. *Cellular Microbiology* **14**, 1183-1192, (2012).

64 Myszka, D. G., Sweet, R. W., Hensley, P., Brigham-Burke, M., Kwong, P. D., Hendrickson, W. A., Wyatt, R., Sodroski, J. & Doyle, M. L. Energetics of the HIV gp120-CD4 binding reaction. *Proc. Natl. Acad. Sci. U.S.A.* **97**, 9026-9031, (2000).

65 Wu, L., Gerard, N. P., Wyatt, R., Choe, H., Parolin, C., Ruffing, N., Borsetti, A., Cardoso, A. A., Desjardin, E., Newman, W., Gerard, C. & Sodroski, J. CD4-induced interaction of primary HIV-1 gp120 glycoproteins with the chemokine receptor CCR-5. *Nature.* **384**, 179-183, (1996).

66 Thali, M., Moore, J. P., Furman, C., Charles, M., Ho, D. D., Robinson, J. & Sodroski, J. Characterization of conserved human immunodeficiency virus type 1 gp120 neutralization epitopes exposed upon gp120-CD4 binding. *J. Virol.* **67**, 3978-3988, (1993).

67 Eckert, D. M. & Kim, P. S. Mechanisms of viral membrane fusion and its inhibition. *Annual review of biochemistry* **70**, 777-810, (2001).

68 Berkhout, B. & Sanders, R. W. Molecular strategies to design an escape-proof antiviral therapy. *Antiviral Research* **92**, 7-14, (2011).

69 Weissenhorn, W. D., Harrison, S. C., Skehel, J. J. & Wiley, D. C. Atomic structure of the ectodomain from HIV-1 gp41. **387**, 426-430, (1997).

70 Hart, T. K., Kirsh, R., Ellens, H., Sweet, R. W., Lambert, D. M., Petteway, S. R., Jr., Leary, J. & Bugelski, P. J. Binding of soluble CD4 proteins to human immunodeficiency virus type 1 and infected cells induces release of envelope glycoprotein gp120. *Proc Natl Acad Sci U S A* **88**, 2189-2193, (1991).

71 Moore, J., McKeating, J., Weiss, R. & Sattentau, Q. Dissociation of gp120 from HIV-1 virions induced by soluble CD4. (1990).

72 Thali, M., Furman, C., Helseth, E., Repke, H. & Sodroski, J. Lack of correlation between soluble CD4-induced shedding of the human immunodeficiency virus type 1 exterior envelope glycoprotein and subsequent membrane fusion events. *Journal of Virology* **66**, 5516-5524, (1992).

73 Klasse, P. J. Modeling how many envelope glycoprotein trimers per virion participate in human immunodeficiency virus infectivity and its neutralization by antibody. *Virology* **369**, 245-262, (2007).

74 Eichler, J. Extreme sweetness: protein glycosylation in archaea. *Nat Rev Micro* **11**, 151-156, (2013).

75 Schwarz, F. & Aebi, M. Mechanisms and principles of N-linked protein glycosylation. *Current Opinion in Structural Biology* **21**, 576-582, (2011).

76 Mellquist, J. L., Kasturi, L., Spitalnik, S. L. & Shakin-Eshleman, S. H. The amino acid following an asn-X-Ser/Thr sequon is an important determinant of N-linked core glycosylation efficiency. *Biochemistry* **37**, 6833-6837, (1998).

77 Petrescu, A. J., Milac, A. L., Petrescu, S. M., Dwek, R. A. & Wormald, M. R. Statistical analysis of the protein environment of N-glycosylation sites: implications for occupancy, structure, and folding. *Glycobiology.* **14**, 103-114, (2004).

78 Breitling, J. & Aebi, M. N-Linked Protein Glycosylation in the Endoplasmic Reticulum. *Cold Spring Harbor Perspectives in Biology* **5**, (2013).

79 Li, E., Tabas, I. & Kornfeld, S. The synthesis of complex-type oligosaccharides. I. Structure of the lipid-linked oligosaccharide precursor of the complex-type oligosaccharides of the vesicular stomatitis virus G protein. *Journal of Biological Chemistry* **253**, 7762-7770, (1978).

80 Harvey, D. J., Merry, A. H., Royle, L., P. Campbell, M., Dwek, R. A. & Rudd, P. M. Proposal for a standard system for drawing structural diagrams of N- and O-linked carbohydrates and related compounds. *PROTEOMICS* **9**, 3796-3801, (2009).

81 Ou, W.-J., Cameron, P. H., Thomas, D. Y. & Bergeron, J. J. M. Association of folding intermediates of glycoproteins with calnexin during protein maturation. *Nature.* **364**, 771-776, (1993).

82 Nauseef, W. M., McCormick, S. J. & Clark, R. A. Calreticulin Functions as a Molecular Chaperone in the Biosynthesis of Myeloperoxidase. *Journal of Biological Chemistry* **270**, 4741-4747, (1995).

83 Hammond, C., Braakman, I. & Helenius, A. Role of N-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality control. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 913-917, (1994).

84 Kornfeld, S., Li, E. & Tabas, I. The synthesis of complex-type oligosaccharides. II. Characterization of the processing intermediates in the synthesis of the complex oligosaccharide units of the vesicular stomatitis virus G protein. *Journal of Biological Chemistry* **253**, 7771-7778, (1978).

85 Aebi, M., Bernasconi, R., Clerc, S. & Molinari, M. N-glycan structures: recognition and processing in the ER. *Trends in Biochemical Sciences* **35**, 74-82, (2010).

86 Schallus, T., Jaeckh, C., Fehér, K., Palma, A. S., Liu, Y., Simpson, J. C., Mackeen, M., Stier, G., Gibson, T. J., Feizi, T., Pieler, T. & Muhle-Goll, C. Malectin: A novel carbohydrate-binding protein of the endoplasmic reticulum and a candidate player in the early steps of protein N-glycosylation. *Molecular Biology of the Cell* **19**, 3404-3414, (2008).

87 Deprez, P., Gautschi, M. & Helenius, A. More than one glycan is needed for ER glucosidase II to allow entry of glycoproteins into the calnexin/calreticulin cycle. *Molecular Cell* **19**, 183-195, (2005).

88 Oliver, J. D., van der Wal, F. J., Bulleid, N. J. & High, S. Interaction of the thiol-dependent reductase ERp57 with nascent glycoproteins. *Science* **275**, 86-88, (1997).

89 Solda, T., Garbi, N., Hämmerling, G. J. & Molinari, M. Consequences of ERp57 deletion on oxidative folding of obligate and facultative clients of the calnexin cycle. *Journal of Biological Chemistry* **281**, 6219-6226, (2006).

90 Lynes, E. M., Raturi, A., Shenkman, M., Ortiz Sandoval, C., Yap, M. C., Wu, J., Janowicz, A., Myhill, N., Benson, M. D., Campbell, R. E., Berthiaume, L. G., Lederkremer, G. Z. & Simmen, T. Palmitoylation is the switch that assigns calnexin to quality control or ER Ca2+ signaling. *Journal of cell science* **126**, 3893-3903, (2013).

91 Sousa, M. & Parodi, A. J. The molecular basis for the recognition of misfolded glycoproteins by the UDP-Glc:glycoprotein glucosyltransferase. *EMBO J.* **14**, 4196-4203, (1995).

92 Takeda, Y., Seko, A., Hachisu, M., Daikoku, S., Izumi, M., Koizumi, A., Fujikawa, K., Kajihara, Y. & Ito, Y. Both isoforms of human UDP-glucose:glycoprotein glucosyltransferase are enzymatically active. *Glycobiology.* **24**, 344-350, (2014).

93 Caramelo, J. J. & Parodi, A. J. Getting in and out from calnexin/calreticulin cycles. *Journal of Biological Chemistry* **283**, 10221-10225, (2008).

94 Su, K., Stoller, T., Rocco, J., Zemsky, J. & Green, R. Pre-Golgi degradation of yeast prepro- $\alpha$ -factor expressed in a mammalian cell: Influence of cell type-specific oligosaccharide processing on intracellular fate. *Journal of Biological Chemistry* **268**, 14301-14309, (1993).

95 Jakob, C. A., Burda, P., Roth, J. & Aebi, M. Degradation of misfolded endoplasmic reticulum glycoproteins in saccharomyces cerevisiae is determined by a specific oligosaccharide structure. *Journal of Cell Biology* **142**, 1223-1233, (1998).

96 Hosokawa, N., Wada, I., Hasegawa, K., Yorihuzi, T., Tremblay, L. O., Herscovics, A. & Nagata, K. A novel ER $\alpha$-mannosidase-like protein accelerates ER-associated degradation. *EMBO Rep.* **2**, 415-422, (2001).

97 Frenkel, Z., Gregory, W., Kornfeld, S. & Lederkremer, G. Z. Endoplasmic Reticulum-associated Degradation of Mammalian Glycoproteins Involves Sugar Chain Trimming to Man6-5GlcNAc2. *Journal of Biological Chemistry* **278**, 34119-34124, (2003).

98 Avezov, E., Frenkel, Z., Ehrlich, M., Herscovics, A. & Lederkremer, G. Z. Endoplasmic reticulum (ER) mannosidase I is compartmentalized and required for N-glycan trimming to man5-6GlcNAc2 in glycoprotein ER-associated degradation. *Molecular Biology of the Cell* **19**, 216-225, (2008).

99 Hosokawa, N., Kamiya, Y., Kamiya, D., Kato, K. & Nagata, K. Human OS-9, a lectin required for glycoprotein endoplasmic reticulum-associated degradation, recognizes mannose-trimmed N-glycans. *Journal of Biological Chemistry* **284**, 17061-17068, (2009).

100 Hirao, K., Natsuka, Y., Tamura, T., Wada, I., Morito, D., Natsuka, S., Romero, P., Sleno, B., Tremblay, L. O., Herscovics, A., Nagata, K. & Hosokawa, N. EDEM3, a soluble EDEM homolog, enhances glycoprotein endoplasmic reticulum-associated degradation and mannose trimming. *Journal of Biological Chemistry* **281**, 9650-9658, (2006).

101 Clerc, S., Hirsch, C., Oggier, D. M., Deprez, P., Jakob, C., Sommer, T. & Aebi, M. Htm1 protein generates the N-glycan signal for glycoprotein degradation in the endoplasmic reticulum. *Journal of Cell Biology* **184**, 159-172, (2009).

102 Ninagawa, S., Okada, T., Sumitomo, Y., Kamiya, Y., Kato, K., Horimoto, S., Ishikawa, T., Takeda, S., Sakuma, T., Yamamoto, T. & Mori, K. EDEM2 initiates mammalian glycoprotein ERAD by catalyzing the first mannose trimming step. *The Journal of Cell Biology* **206**, 347-356, (2014).

103 Hosokawa, N., You, Z., Tremblay, L. O., Nagata, K. & Herscovics, A. Stimulation of ERAD of misfolded null Hong Kong $\alpha$1-antitrypsin by Golgi $\alpha$1,2-mannosidases. *Biochemical and Biophysical Research Communications* **362**, 626-632, (2007).

104 Kamiya, Y., Kamiya, D., Yamamoto, K., Nyfeler, B., Hauri, H. P. & Kato, K. Molecular basis of sugar recognition by the human L-type lectins ERGIC-53, VIPL, and VIP36. *Journal of Biological Chemistry* **283**, 1857-1861, (2008).

105 Helenius, A. & Aebi, M. Intracellular functions of N-linked glycans. *Science* **291**, 2364-2369, (2001).

106 Christis, C., Lubsen, N. H. & Braakman, I. Protein folding includes oligomerization - Examples from the endoplasmic reticulum and cytosol. *FEBS Journal* **275**, 4700-4727, (2008).

107 De Lorenzo, F., Goldberger, R. F., Steers Jr, E., Givol, D. & Anfinsen, B. Purification and properties of an enzyme from beef liver which catalyzes sulfhydryl-disulfide interchange in proteins. *Journal of Biological Chemistry* **241**, 1562-1567, (1966).

108 Fuchs, S., De Lorenzo, F. & Anfinsen, C. B. Studies on the mechanism of the enzymic catalysis of disulfide interchange in proteins. *Journal of Biological Chemistry* **242**, 398-402, (1967).

109 Molinari, M. & Helenius, A. Glycoproteins form mixed disulphides with oxidoreductases during folding in living cells. *Nature.* **402**, 90-93, (1999).

110 Puig, A. & Gilbert, H. F. Protein disulfide isomerase exhibits chaperone and anti-chaperone activity in the oxidative refolding of lysozyme. *Journal of Biological Chemistry* **269**, 7764-7771, (1994).

111 Schwaller, M., Wilkinson, B. & Gilbert, H. F. Reduction-reoxidation cycles contribute to catalysis of disulfide isomerization by protein-disulfide isomerase. *Journal of Biological Chemistry* **278**, 7154-7159, (2003).

112 Kersteen, E. A., Barrows, S. R. & Raines, R. T. Catalysis of protein disulfide bond isomerization in a homogeneous substrate. *Biochemistry* **44**, 12168-12178, (2005).

113 Munro, S. & Pelham, H. R. An Hsp70-like protein in the ER: identity with the 78 kd glucose-regulated protein and immunoglobulin heavy chain binding protein. *Cell* **46**, 291-300, (1986).

114 Mayer, M., Reinstein, J. & Buchner, J. Modulation of the ATPase cycle of BiP by peptides and proteins. *Journal of molecular biology* **330**, 137-144, (2003).

115 Meunier, L., Usherwood, Y. K., Tae Chung, K. & Hendershot, L. M. A subset of chaperones and folding enzymes form multiprotein complexes in endoplasmic reticulum to bind nascent proteins. *Molecular Biology of the Cell* **13**, 4456-4469, (2002).

116 Määttänen, P., Gehring, K., Bergeron, J. J. M. & Thomas, D. Y. Protein quality control in the ER: The recognition of misfolded proteins. *Seminars in Cell & Developmental Biology* **21**, 500-511, (2010).

117 Dollins, D. E., Warren, J. J., Immormino, R. M. & Gewirth, D. T. Structures of GRP94-Nucleotide Complexes Reveal Mechanistic Differences between the hsp90 Chaperones. *Molecular Cell* **28**, 41-56, (2007).

118 Argon, Y. & Simen, B. B. GRP94, an ER chaperone with protein and peptide binding properties. *Seminars in Cell and Developmental Biology* **10**, 495-505, (1999).

119 Nielsen, H., Engelbrecht, J., Brunak, S. & von Heijne, G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engineering* **10**, 1-6, (1997).

120 Martoglio, B. & Dobberstein, B. Signal sequences: More than just greasy peptides. *Trends in Cell Biology* **8**, 410-415, (1998).

121 Li, Y., Luo, L., Thomas, D. Y. & Kang, C. Y. The HIV-1 Env protein signal sequence retards its cleavage and down- regulates the glycoprotein folding. *Virology* **272**, 417-428, (2000).

122 Land, A. & Braakman, I. Folding of the human immunodeficiency virus type 1 envelope glycoprotein in the endoplasmic reticulum. *Biochimie* **83**, 783-790, (2001).

123 Doores, K. J., Bonomelli, C., Harvey, D. J., Vasiljevic, S., Dwek, R. A., Burton, D. R., Crispin, M. & Scanlan, C. N. Envelope glycans of immunodeficiency virions are almost entirely

oligomannose antigens. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 13800-13805, (2010).

124 Bernstein, H. B., Tucker, S. P., Hunter, E., Schutzbach, J. S. & Compans, R. W. Human immunodeficiency virus type 1 envelope glycoprotein is modified by O-linked oligosaccharides. *Journal of Virology* **68**, 463-468, (1994).

125 Li, Y., Luo, L., Rasool, N. & Kang, C. Y. Glycosylation is necessary for the correct folding of human immunodeficiency virus gp120 in CD4 binding. *Journal of Virology* **67**, 584-588, (1993).

126 Otteken, A. & Moss, B. Calreticulin interacts with newly synthesized human immunodeficiency virus type 1 envelope glycoprotein, suggesting a chaperone function similar to that of calnexin. *Journal of Biological Chemistry* **271**, 97-103, (1996).

127 Go, E. P., Zhang, Y., Menon, S. & Desaire, H. Analysis of the disulfide bond arrangement of the HIV-1 envelope protein CON-S gp140 DeltaCFI shows variability in the V1 and V2 regions. *J Proteome Res* **10**, 578-591, (2011).

128 Land, A., Zonneveld, D. & Braakman, I. Folding of HIV-1 envelope glycoprotein involves extensive isomerization of disulfide bonds and conformation-dependent leader peptide cleavage. *FASEB.J.* **17**, 1058-1067, (2003).

129 van Anken, E., Sanders, R. W., Liscaljet, I. M., Land, A., Bontjer, I., Tillemans, S., Nabatov, A. A., Paxton, W. A., Berkhout, B. & Braakman, I. Only Five of 10 Strictly Conserved Disulfide Bonds Are Essential for Folding and Eight for Function of the HIV-1 Envelope Glycoprotein. *Molecular Biology of the Cell* **19**, 4298-4309, (2008).

130 Earl, P. L., Moss, B. & Doms, R. W. Folding, interaction with GRP78-BiP, assembly, and transport of the human immunodeficiency virus type 1 envelope protein. *Journal of Virology* **65**, 2047-2055, (1991).

131 Knarr, G., Modrow, S., Todd, A., Gething, M. J. & Buchner, J. BiP-binding sequences in HIV gp160. Implications for the binding specificity of BiP. *Journal of Biological Chemistry* **274**, 29850-29857, (1999).

132 Otteken, A., Earl, P. L. & Moss, B. Folding, assembly, and intracellular trafficking of the human immunodeficiency virus type 1 envelope glycoprotein analyzed with monoclonal antibodies recognizing maturational intermediates. *Journal of Virology* **70**, 3407-3415, (1996).

133 Kantanen, M. L., Leinikki, P. & Kuismanen, E. Endoproteolytic cleavage of HIV-1 gp160 envelope precursor occurs after exit from the trans-Golgi network (TGN). *Archives of Virology* **140**, 1441-1449, (1995).

134 Kowalski, M., Potz, J., Basiripour, L., Dorfman, T., Goh, W. C., Terwilliger, E., Dayton, A., Rosen, C., Haseltine, W. & Sodroski, J. Functional regions of the envelope glycoprotein of human immunodeficiency virus type 1. *Science* **237**, 1351-1355, (1987).

135 McCune, J. M., Rabin, L. B., Feinberg, M. B., Lieberman, M., Kosek, J. C., Reyes, G. R. & Weissman, I. L. Endoproteolytic cleavage of gp160 is required for the activation of human immunodeficiency virus. *Cell* **53**, 55-67, (1988).

136 Moulard, M. & Decroly, E. Maturation of HIV envelope glycoprotein precursors by cellular endoproteases. *Biochimica et Biophysica Acta - Reviews on Biomembranes* **1469**, 121-132, (2000).

137 Decroly, E., Wouters, S., Bello, C. D., Lazure, C., Ruysschaert, J. M. & Seidah, N. G. Identification of the paired basic convertases implicated in HIV gp160 processing based on in vitro assays and expression in CD4+ cell lines. *Journal of Biological Chemistry* **271**, 30442-30450, (1996).

138 Croom Jr, W. J., Hagler Jr, W. M., Froetschel, M. A. & Johnson, A. D. The involvement of slaframine and swainsonine in slobbers syndrome: a review. *J Anim Sci* **73**, 1499-1508, (1995).

139 Dorling, P. R., Huxtable, C. R. & Colegate, S. M. Inhibition of lysosomal alpha-mannosidase by swainsonine, an indolizidine alkaloid isolated from Swainsona canescens. *Biochemical Journal* **191**, 649-651, (1980).

140 Elbein, A. D., Dorling, P. R., Vosbeck, K. & Horisberger, M. Swainsonine prevents the processing of the oligosaccharide chains of influenza virus hemagglutinin. *Journal of Biological Chemistry* **257**, 1573-1576, (1982).

141 Elbein, A. D., Solf, R., Dorling, P. R. & Vosbeck, K. Swainsonine: An inhibitor of glycoprotein processing. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 7393-7397, (1981).

142 Tulsiani, D. R. P., Harris, T. M. & Touster, O. Swainsonine inhibits the biosynthesis of complex glycoproteins by inhibition of Golgi mannosidase II. *Journal of Biological Chemistry* **257**, 7936-7939, (1982).

143 Abraham, D. J., Sidebothom, R., Winchester, B. G., Dorling, P. R. & Dell, A. Swainsonine affects the processing of glycoproteins in vivo. *FEBS.Lett.* **163**, 110-113, (1983).

144 Fuhrmann, U., Bause, E., Legler, G. & Ploegh, H. Novel mannosidase inhibitor blocking conversion of high mannose to complex oligosaccharides. *Nature.* **307**, 755-758, (1984).

145 Bischoff, J. & Kornfeld, R. The effect of 1-deoxymannojirimycin on rat liver α-mannosidases. *Biochemical and Biophysical Research Communications* **125**, 324-331, (1984).

146 Romero, P. A., Friedlander, P., Fellows, L., Evans, S. V. & Herscovies, A. Effects of manno-1-deoxynojirimycin and 2,5-dihydroxymethyl-3,4-dihydroxypyrrolidine on N-linked oligosaccharide processing in intestinal epithelial cells. *FEBS.Lett.* **184**, 197-201, (1985).

147 Elbein, A. D., Tropea, J. E., Mitchell, M. & Kaushal, G. P. Kifunensine, a potent inhibitor of the glycoprotein processing mannosidase I. *Journal of Biological Chemistry* **265**, 15599-15605, (1990).

148 Weng, S. & Spiro, R. G. Demonstration that a kifunensine-resistant α-mannosidase with a unique processing action on N-linked oligosaccharides occurs in rat liver endoplasmic reticulum and various cultured cells. *Journal of Biological Chemistry* **268**, 25656-25663, (1993).

149 Moore, S. E. & Spiro, R. G. Characterization of the endomannosidase pathway for the processing of N-linked oligosaccharides in glucosidase II-deficient and parent mouse lymphoma cells. *J Biol Chem* **267**, 8443-8451, (1992).

150 Scanlan, C. N., Ritchie, G. E., Baruah, K., Crispin, M., Harvey, D. J., Singer, B. B., Lucka, L., Wormald, M. R., Wentworth Jr, P., Zitzmann, N., Rudd, P. M., Burton, D. R. & Dwek, R. A. Inhibition of Mammalian Glycan Biosynthesis Produces Non-self Antigens for a Broadly Neutralising, HIV-1 Specific Antibody. *Journal of Molecular Biology* **372**, 16-22, (2007).

151 Chang, V. T., Crispin, M., Aricescu, A. R., Harvey, D. J., Nettleship, J. E., Fennelly, J. A., Yu, C., Boles, K. S., Evans, E. J., Stuart, D. I., Dwek, R. A., Jones, E. Y., Owens, R. J. & Davis, S. J. Glycoprotein Structural Genomics: Solving the Glycosylation Problem. *Structure* **15**, 267-273, (2007).

152 Elbein, A. D. Glycosidase inhibitors: Inhibitors of N-linked oligosaccharide processing. *FASEB.J.* **5**, 3055-3063, (1991).

153 Datema, R., Romero, P. A., Legler, G. & Schwarz, R. T. Inhibition of formation of complex oligosaccharides by the glucosidase inhibitor bromoconduritol. *Proc Natl Acad Sci U S A* **79**, 6787-6791, (1982).

154 Jacob, G., Scudder, P., Butters, T., Jones, I. & Tiemeier, D. in *Natural Products as Antiviral Agents* (eds ChungK Chu & HoraceG Cutler) Ch. 7, 137-152 (Springer US, 1992).

155 Scofield, A. M., Fellows, L. E., Nash, R. J. & Fleet, G. W. J. Inhibition of mammalian digestive disaccharidases by polyhydroxy alkaloids. *Life Sciences* **39**, 645-650, (1986).

156 Platt, F. M., Neises, G. R., Karlsson, G. B., Dwek, R. A. & Butters, T. D. N-butyldeoxygalactonojirimycin inhibits glycolipid biosynthesis but does not affect N-linked oligosaccharide processing. *Journal of Biological Chemistry* **269**, 27108-27114, (1994).

157 Platt, F. M., Neises, G. R., Dwek, R. A. & Butters, T. D. N-butyldeoxynojirimycin is a novel inhibitor of glycolipid biosynthesis. *Journal of Biological Chemistry* **269**, 8362-8365, (1994).

158 Lubas, W. A. & Spiro, R. G. Golgi endo-alpha-D-mannosidase from rat liver, a novel N-linked carbohydrate unit processing enzyme. *Journal of Biological Chemistry* **262**, 3775-3781, (1987).

159 Lubas, W. A. & Spiro, R. G. Evaluation of the role of rat liver golgi endo- α -D-mannosidase in processing N-linked oligosaccharides. *Journal of Biological Chemistry* **263**, 3990-3998, (1988).

160 Moore, S. E. H. & Spiro, R. G. Demonstration that Golgi endo- α -D-mannosidase provides a glucosidase-independent pathway for the formation of complex N-linked oligosaccharides of glycoproteins. *Journal of Biological Chemistry* **265**, 13104-13112, (1990).

161 Karaivanova, V. K., Luan, P. & Spiro, R. G. Processing of viral envelope glycoprotein by the endomannosidase pathway: Evaluation of host cell specificity. *Glycobiology.* **8**, 725-730, (1998).

162 Roth, J., Ziak, M. & Zuber, C. The role of glucosidase II and endomannosidase in glucose trimming of asparagine-linked oligosaccharides. *Biochimie* **85**, 287-294, (2003).

163 Bosch, F. X., Orlich, M., Legler, G., Schwarz, R. T. & Rott, R. Effect of inhibitors of glycosylation on proteolytic activation of avian influenza virus hemagglutinins: Discrimination between tryptic cleavage and elimination of the connecting peptide. *Virology* **132**, 199-204, (1984).

164 Pan, Y. T., Hori, H., Saul, R., Sanford, B. A., Molyneux, R. J. & Elbein, A. D. Castanospermine inhibits the processing of the oligosaccharide portion of the influenza viral hemagglutinin. *Biochemistry* **22**, 3975-3984, (1983).

165 Romero, P. A., Datema, R. & Schwarz, R. T. N-methyl-1-deoxynojirimycin, a novel inhibitor of glycoprotein processing, and its effect on fowl plague virus maturation. *Virology* **130**, 238-242, (1983).

166 Burke, B., Matlin, K., Bause, E., Legler, G., Peyrieras, N. & Ploegh, H. Inhibition of N-linked oligosaccharide trimming does not interfere with surface expression of certain integral membrane proteins. *EMBO J.* **3**, 551-556, (1984).

167 Schlesinger, S., Malfer, C. & Schlesinger, M. J. The formation of vesicular stomatitis virus (San Juan strain) becomes temperature-sensitive when glucose residues are retained on the oligosaccharides of the glycoprotein. *Journal of Biological Chemistry* **259**, 7597-7601, (1984).

168 Schlesinger, S., Koyama, A. H., Malfer, C., Gee, S. L. & Schlesinger, M. J. The effects of inhibitors of glucosidase I on the formation of Sindbis virus. *Virus Research* **2**, 139-149, (1985).

169 Repp, R., Tamura, T., Boschek, C. B., Wege, H., Schwarz, R. T. & Niemann, H. The effects of processing inhibitors of N-linked oligosaccharides on the intracellular migration of glycoprotein E2 of mouse hepatitis virus and the maturation of coronavirus particles. *Journal of Biological Chemistry* **260**, 15873-15879, (1985).

170 Pinter, A., Honnen, W. J. & Soo Li, J. Studies with inhibitors of oligosaccharide processing indicate a functional role for complex sugars in the transport and proteolysis of Friend mink cell focus-inducing murine leukemia virus envelope proteins. *Virology* **136**, 196-210, (1984).

171 Montefiori, D. C., Robinson Jr, W. E. & Mitchell, W. M. Role of protein N-glycosylation in pathogenesis of human immunodeficieny virus type 1. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 9248-9252, (1988).

172 Karpas, A., Fleet, G. W. J., Dwek, R. A., Petursson, S., Namgoong, S. K., Ramsden, N. G., Jacob, G. S. & Rademacher, T. W. Aminosugar derivatives as potential anti-human immunodeficiency virus agents. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 9229-9233, (1988).

173 Fenouillet, E. & Gluckman, J. C. Effect of a glucosidase inhibitor on the bioactivity and immunoreactivity of human immunodeficiency virus type 1 envelope glycoprotein. *Journal of General Virology* **72**, 1919-1926, (1991).

174 Jones, I. M. & Jacob, G. S. Anti-HIV drug mechanism. *Nature.* **352**, 198, (1991).

175 Ratner, L. & Vander Heyden, N. Mechanism of action of N-butyl deoxynojirimycin in inhibiting HIV-1 infection and activity in combination with nucleoside analogs. *AIDS Research and Human Retroviruses* **9**, 291-297, (1993).

176 Fischer, P. B., Collin, M., Karlsson, G. B., James, W., Butters, T. D., Davis, S. J., Gordon, S., Dwek, R. A. & Platt, F. M. The α -glucosidase inhibitor N-butyldeoxynojirimycin inhibits

human immunodeficiency virus entry at the level of post-CD4 binding. *Journal of Virology* **69**, 5791-5797, (1995).

177 Fischer, P. B., Karlsson, G. B., Butters, T. D., Dwek, R. A. & Platt, F. M. N-butyldeoxynojirimycin-mediated inhibition of human immunodeficiency virus entry correlates with changes in antibody recognition of the V1/V2 region of gp120. *Journal of Virology* **70**, 7143-7152, (1996).

178 Fischer, P. B., Karlsson, G. B., Dwek, R. A. & Platt, F. M. N-butyldeoxynojirimycin-mediated inhibition of human immunodeficiency virus entry correlates with impaired gp120 shedding and gp41 exposure. *Journal of Virology* **70**, 7153-7160, (1996).

179 Pal, R., Hoke, G. M. & Sarngadharan, M. G. Role of oligosaccharides in the processing and maturation of envelope glycoproteins of human immunodeficiency virus type 1. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 3384-3388, (1989).

180 Eggink, D., Melchers, M., Wuhrer, M., van Montfort, T., Dey, A. K., Naaijkens, B. A., David, K. B., Le Douce, V., Deelder, A. M., Kang, K., Olson, W. C., Berkhout, B., Hokke, C. H., Moore, J. P. & Sanders, R. W. Lack of complex N-glycans on HIV-1 envelope glycoproteins preserves protein conformation and entry function. *Virology* **401**, 236-247, (2010).

181 Joubert, P. H., Venter, C. P., Joubert, H. F. & Hillebrand, I. The effect of a 1-deoxynojirimycin derivative on post-prandial blood glucose and insulin levels in healthy black and white volunteers. *EUR. J. CLIN. PHARMACOL.* **28**, 705-708, (1985).

182 Inzucchi, S. E. Oral antihyperglycemic therapy for type 2 diabetes: Scientific review. *JAMA.* **287**, 360-372, (2002).

183 Standards of medical care in diabetes - 2012. *Diabetes Care* **35**, S11-S63, (2012).

184 Tierney, M., Pottage, J., Kessler, H., Fischl, M., Richman, D., Merigan, T., Powderly, W., Smith, S., Karim, A., Sherman, J. & Hirsch, M. The tolerability and pharmacokinetics of N-butyl-deoxynojirimycin in patients with advanced HIV disease (ACTG 100). *J. Acquir. Immune Defic. Syndr.Hum.Retrovirol.* **10**, 549-553, (1995).

185 Fischl, M. A., Resnick, L., Coombs, R., Kremer, A. B., Pottage Jr, J. C., Fass, R. J., Fife, K. H., Powderly, W. G., Collier, A. C., Aspinall, R. L., Smith, S. L., Kowalski, K. G. & Wallemark, C. B. The safety and efficacy of combination N-butyl-deoxynojirimycin (SC- 48334) and zidovudine in patients with HIV-1 infection and 200-500 CD4 cells/mm3. *Journal of Acquired Immune Deficiency Syndromes* **7**, 139-147, (1994).

186 Bryant M, M. R., Partis R, Stolzenbach J, Wyand M, Tiemeier D, Marr J. in *International Conference on AIDS.*

187 Smith S, P. A., Smith M, Wallemark C, Karim A, Sherman J. in *National Conference on Human Retroviruses and Related Infections.*

188 Butters, T. D. Pharmacotherapeutic strategies using small molecules for the treatment of glycolipid lysosomal storage disorders. *Expert Opin. Pharmacother.* **8**, 427-435, (2007).

189 Butters, T. D., Dwek, R. A. & Platt, F. M. Imino sugar inhibitors for treating the lysosomal glycosphingolipidoses. *Glycobiology.* **15**, 43R-52R, (2005).

190 Cox, T., Lachmann, R., Hollak, C., Aerts, J., Van Weely, S., Hrebícek, M., Platt, F., Butters, T., Dwek, R., Moyses, C., Gow, I., Elstein, D. & Zimran, A. Novel oral treatment of Gaucher's disease with N-butyldeoxynojirimycin (OGT 918) to decrease substrate biosynthesis. *Lancet* **355**, 1481-1485, (2000).

191 Zitzmann, N., Mehta, A. S., Carrouée, S., Butters, T. D., Platt, F. M., McCauley, J., Blumberg, B. S., Dwek, R. A. & Block, T. M. Imino sugars inhibit the formation and secretion of bovine viral diarrhea virus, a pestivirus model of hepatitis C virus: Implications for the development of broad spectrum anti-hepatitis virus agents. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 11878-11882, (1999).

192 Branza-Nichita, N., Durantel, D., Carrouée-Durantel, S., Dwek, R. A. & Zitzmann, N. Antiviral Effect ofN-Butyldeoxynojirimycin against Bovine Viral Diarrhea Virus Correlates with Misfolding of E2 Envelope Proteins and Impairment of Their Association into E1-E2 Heterodimers. *Journal of Virology* **75**, 3527-3536, (2001).

193 Pavlovic, D., Neville, D. C. A., Argaud, O., Blumberg, B., Dwek, R. A., Fischer, W. B. & Zitzmann, N. The hepatitis C virus p7 protein forms an ion channel that is inhibited by long-alkyl-chain iminosugar derivatives. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 6104-6108, (2003).

194 Mellor, H. R., Nolan, J., Pickering, L., Wormald, M. R., Platt, F. M., Dwek, R. A., Fleet, G. W. J. & Butters, T. D. Preparation, biochemical characterization and biological properties of radiolabelled N-alkylated deoxynojirimycins. *Biochemical Journal* **366**, 225-233, (2002).

195 Courageot, M. P., Frenkiel, M. P., Duarte Dos Santos, C., Deubel, V. & Desprès, P. $\alpha$-Glucosidase inhibitors reduce dengue virus production by affecting the initial steps of virion morphogenesis in the endoplasmic reticulum. *Journal of Virology* **74**, 564-572, (2000).

196 Whitby, K., Pierson, T. C., Geiss, B., Lane, K., Engle, M., Zhou, Y., Doms, R. W. & Diamond, M. S. Castanospermine, a potent inhibitor of dengue virus infection in vitro and in vivo. *Journal of Virology* **79**, 8698-8706, (2005).

197 Low, J. G., Sung, C., Wijaya, L., Wei, Y., Rathore, A. P. S., Watanabe, S., Tan, B. H., Toh, L., Chua, L. T., Hou, Y. a., Chow, A., Howe, S., Chan, W. K., Tan, K. H., Chung, J. S., Cherng, B. P., Lye, D. C., Tambayah, P. A., Ng, L. C., Connolly, J., Hibberd, M. L., Leo, Y. S., Cheung, Y. B., Ooi, E. E. & Vasudevan, S. G. Efficacy and safety of celgosivir in patients with dengue fever (CELADEN): a phase 1b, randomised, double-blind, placebo-controlled, proof-of-concept trial. *The Lancet Infectious Diseases* **14**, 706-715, (2014).

198 Miller, J. L., Lachica, R., Sayce, A. C., Williams, J. P., Bapat, M., Dwek, R., Beatty, P. R., Harris, E. & Zitzmann, N. Liposome-mediated delivery of iminosugars enhances efficacy against dengue virus in vivo. *Antimicrobial Agents and Chemotherapy* **56**, 6379-6386, (2012).

199 Perry, S. T., Buck, M. D., Plummer, E. M., Penmasta, R. A., Batra, H., Stavale, E. J., Warfield, K. L., Dwek, R. A., Butters, T. D., Alonzi, D. S., Lada, S. M., King, K., Klose, B., Ramstedt, U. & Shresta, S. An iminosugar with potent inhibition of dengue virus infection in vivo. *Antiviral Research* **98**, 35-43, (2013).

200 Platt, F. M., Reinkensmeier, G., Dwek, R. A. & Butters, T. D. Extensive glycosphingolipid depletion in the liver and lymphoid organs of mice treated with N-butyldeoxynojirimycin. *Journal of Biological Chemistry* **272**, 19365-19372, (1997).

201 De Praeter, C. M., Gerwig, G. J., Bause, E., Nuytinck, L. K., Vliegenthart, J. F. G., Breuer, W., Kamerling, J. P., Espeel, M. F., Martin, J. J. R., De Paepe, A. M., Chan, N. W. C., Dacremont, G. A. & Van Coster, R. N. A novel disorder caused by defective biosynthesis of N-linked oligosaccharides due to glucosidase I deficiency. *Am. J. Hum. Genet.* **66**, 1744-1756, (2000).

202 Sadat, M. A., Moir, S., Chun, T. W., Lusso, P., Kaplan, G., Wolfe, L., Memoli, M. J., He, M., Vega, H., Kim, L. J., Huang, Y., Hussein, N., Nievas, E., Mitchell, R., Garofalo, M., Louie, A., Ireland, D. C., Grunes, C., Cimbro, R., Patel, V., Holzapfel, G., Salahuddin, D., Bristol, T., Adams, D., Marciano, B. E., Hegde, M., Li, Y., Calvo, K. R., Stoddard, J., Justement, J. S., Jacques, J., Priel, D. A., Murray, D., Sun, P., Kuhns, D. B., Boerkoel, C. F., Chiorini, J. A., Di Pasquale, G., Verthelyi, D. & Rosenzweig, S. D. Glycosylation, Hypogammaglobulinemia, and Resistance to Viral Infections. *N Engl J Med*, (2014).

203 Davis, S. J., Ward, H. A., Puklavec, M. J., Willis, A. C., Williams, A. F. & Barclay, A. N. High level expression in Chinese hamster ovary cells of soluble forms of CD4 T lymphocyte glycoprotein including glycosylation variants. *Journal of Biological Chemistry* **265**, 10410-10418, (1990).

204 Aricescu, A. R., Lu, W. & Jones, E. Y. A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallographica Section D: Biological Crystallography* **62**, 1243-1250, (2006).

205 Kirschner, M., Monrose, V., Paluch, M., Techodamrongsin, N., Rethwilm, A. & Moore, J. P. The production of cleaved, trimeric human immunodeficiency virus type 1 (HIV-1) envelope glycoprotein vaccine antigens and infectious pseudoviruses using linear polyethylenimine as a transfection reagent. *Protein Expression and Purification* **48**, 61-68, (2006).

206 Kukushkin, N. V., Easthope, I. S., Alonzi, D. S. & Butters, T. D. Restricted processing of glycans by endomannosidase in mammalian cells. *Glycobiology.* **22**, 1282-1288, (2012).

207 Alonzi, D. S., Neville, D. C. A., Lachmann, R. H., Dwek, R. A. & Butters, T. D. Glucosylated free oligosaccharides are biomarkers of endoplasmic-reticulum $\alpha$-glucosidase inhibition. *Biochemical Journal* **409**, 571-580, (2008).

208 Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S. & Alizon, M. Nucleotide sequence of the AIDS virus, LAV. *Cell* **40**, 9-17, (1985).

209 Bontjer, I., Land, A., Eggink, D., Verkade, E., Tuin, K., Baldwin, C., Pollakis, G., Paxton, W. A., Braakman, I., Berkhout, B. & Sanders, R. W. Optimization of human immunodeficiency virus type 1 envelope glycoproteins with V1/V2 deleted, using virus evolution. *Journal of Virology* **83**, 368-383, (2009).

210 Bontjer, I., Melchers, M., Eggink, D., David, K., Moore, J. P., Berkhout, B. & Sanders, R. W. Stabilized HIV-1 envelope glycoprotein trimers lacking the V1V2 domain, obtained by virus evolution. *J Biol Chem* **285**, 36456-36470, (2010).

211 Connolly, B. M., Steele, K. E., Davis, K. J., Geisbert, T. W., Kell, W. M., Jaax, N. K. & Jahrling, P. B. Pathogenesis of experimental Ebola virus infection in guinea pigs. *Journal of Infectious Diseases* **179**, S203-S217, (1999).

212 Dowall, S., Taylor, I., Yeates, P., Smith, L., Rule, A., Easterbrook, L., Bruce, C., Cook, N., Corbin-Lickfett, K., Empig, C., Schlunegger, K., Graham, V., Dennis, M. & Hewson, R. Catheterized guinea pigs infected with Ebola Zaire virus allows safer sequential sampling to determine the pharmacokinetic profile of a phosphatidylserine-targeting monoclonal antibody. *Antiviral Research* **97**, 108-111, (2013).

213 Papandreou, M. J. & Fenouillet, E. Effect of changes in the glycosylation of the human immunodeficiency virus type 1 envelope on the immunoreactivity and sensitivity to thrombin of its third variable domain. *Virology* **241**, 163-167, (1998).

214 Pollock, S., Dwek, R. A., Burton, D. R. & Zitzmann, N. N-Butyldeoxynojirimycin is a broadly effective anti-HIV therapy significantly enhanced by targeted liposome delivery. *AIDS.* **22**, 1961-1969, (2008).

215 Wang, W., Nie, J., Prochnow, C., Truong, C., Jia, Z., Wang, S., Chen, X. & Wang, Y. A systematic study of the N-glycosylation sites of HIV-1 envelope protein on infectivity and antibody-mediated neutralization. *Retrovirology* **10**, 14, (2013).

216 Saeed, M., Suzuki, R., Watanabe, N., Masaki, T., Tomonaga, M., Muhammad, A., Kato, T., Matsuura, Y., Watanabe, H., Wakita, T. & Suzuki, T. Role of the Endoplasmic Reticulum-associated Degradation (ERAD) pathway in degradation of hepatitis C virus envelope proteins and production of virus particles. *Journal of Biological Chemistry* **286**, 37264-37273, (2011).

217 Benyair, R., Ogen-Shtern, N. & Lederkremer, G. Z. Glycan regulation of ER-associated degradation through compartmentalization. *Seminars in Cell & Developmental Biology*.

218 Hughes, V. The outlook for a cure. *Nature.* **466**, S11-S13, (2010).

219 Palella Jr, F. J., Delaney, K. M., Moorman, A. C., Loveless, M. O., Fuhrer, J., Satten, G. A., Aschman, D. J. & Holmberg, S. D. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. *New England Journal of Medicine* **338**, 853-860, (1998).

220 Mocroft, A., Vella, S., Benfield, T. L., Chiesi, A., Miller, V., Gargalianos, P., D'Arminio Monforte, A., Yust, I., Bruun, J. N., Phillips, A. N. & Lundgren, J. D. Changing patterns of mortality across Europe in patients infected with HIV-1. *Lancet* **352**, 1725-1730, (1998).

221 Larder, B. A., Darby, G. & Richman, D. D. HIV with reduced sensitivity to zidovudine (AZT) isolated during prolonged therapy. *Science* **243**, 1731-1734, (1989).

222 Erice, A., Mayers, D. L., Strike, D. G., Sannerud, K. J., McCutchan, F. E., Henry, K. & Balfour, H. H. Primary Infection with Zidovudine-Resistant Human Immunodeficiency Virus Type 1. *New England Journal of Medicine* **328**, 1163-1165, (1993).

223 Clavel, F. & Hance, A. J. HIV drug resistance. *N Engl J Med* **350**, 1023-1035, (2004).

224 Sureau, P. H. Firsthand Clinical Observations of Hemorrhagic Manifestations in Ebola Hemorrhagic Fever in Zaire. *Rev Infect Dis* **11**, S790-S793, (1989).

225 Jezek, Z., Szczeniowski, M. Y., Muyembe-Tamfum, J. J., McCormick, J. B. & Heymann, D. L. Ebola between outbreaks: intensified Ebola hemorrhagic fever surveillance in the Democratic Republic of the Congo, 1981-1985. *The Journal of infectious diseases* **179 Suppl 1**, S60-64, (1999).

226 Olival, K. J. & Hayman, D. T. S. Filoviruses in bats: Current knowledge and future directions. *Viruses* **6**, 1759-1788, (2014).

227 Paessler, S. & Walker, D. H. in *Annual Review of Pathology: Mechanisms of Disease* Vol. 8 411-440 (2013).

228 Formenty, P., Hatz, C., Le Guenno, B., Stoll, A., Rogenmoser, P. & Widmer, A. Human infection due to Ebola virus, subtype Cote d'Ivoire: clinical and biologic presentation. *The Journal of infectious diseases* **179 Suppl 1**, S48-53, (1999).

229 Kuhn, J. H., Becker, S., Ebihara, H., Geisbert, T. W., Johnson, K. M., Kawaoka, Y., Lipkin, W. I., Negredo, A. I., Netesov, S. V., Nichol, S. T., Palacios, G., Peters, C. J., Tenorio, A., Volchkov, V. E. & Jahrling, P. B. Proposal for a revised taxonomy of the family Filoviridae: Classification, names of taxa and viruses, and virus abbreviations. *Archives of Virology* **155**, 2083-2103, (2010).

230 Dolnik, O., Kolesnikova, L. & Becker, S. Filoviruses: Interactions with the host cell. *Cell. Mol. Life Sci.* **65**, 756-776, (2008).

231 Hoenen, T., Groseth, A., Falzarano, D. & Feldmann, H. Ebola virus: unravelling pathogenesis to combat a deadly disease. *Trends Mol. Med.* **12**, 206-215, (2006).

232 Neumann, G., Feldmann, H., Watanabe, S., Lukashevich, I. & Kawaoka, Y. Reverse genetics demonstrates that proteolytic processing of the Ebola virus glycoprotein is not essential for replication in cell culture. *Journal of Virology* **76**, 406-410, (2002).

233 Neumann, G., Geisbert, T. W., Ebihara, H., Geisbert, J. B., Daddario-DiCaprio, K. M., Feldmann, H. & Kawaoka, Y. Proteolytic processing of the Ebola virus glycoprotein is not critical for Ebola virus replication in nonhuman primates. *Journal of Virology* **81**, 2995-2998, (2007).

234 Lee, J. E., Fusco, M. L., Hessell, A. J., Oswald, W. B., Burton, D. R. & Saphire, E. O. Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor. *Nature.* **454**, 177-182, (2008).

235 Schornberg, K., Matsuyama, S., Kabsch, K., Delos, S., Bouton, A. & White, J. Role of endosomal cathepsins in entry mediated by the Ebola virus glycoprotein. *Journal of Virology* **80**, 4174-4178, (2006).

236 Lennemann, N. J., Rhein, B. A., Ndungo, E., Chandran, K., Qiu, X. & Maury, W. Comprehensive functional analysis of N-linked glycans on ebola virus GP1. *mBio* **5**, (2014).

237 Volchkov, V. E., Volchkova, V. A., Mühlberger, E., Kolesnikova, L. V., Weik, M., Dolnik, O. & Klenk, H. D. Recovery of infectious Ebola virus from complementary DNA: RNA editing of the GP gene and viral cytotoxicity. *Science* **291**, 1965-1969, (2001).

238 Radoshitzky, S. R., Warfield, K. L., Chi, X., Dong, L., Kota, K., Bradfute, S. B., Gearhart, J. D., Retterer, C., Kranzusch, P. J., Misasi, J. N., Hogenbirk, M. A., Wahl-Jensen, V., Volchkov, V. E., Cunningham, J. M., Jahrling, P. B., Aman, M. J., Bavari, S., Farzan, M. & Kuhn, J. H. Ebolavirus Δ-peptide immunoadhesins inhibit marburgvirus and ebolavirus cell entry. *Journal of Virology* **85**, 8502-8513, (2011).

239 Aleksandrowicz, P., Marzi, A., Biedenkopf, N., Beimforde, N., Becker, S., Hoenen, T., Feldmann, H. & Schnittler, H. J. Ebola virus enters host cells by macropinocytosis and clathrin-mediated endocytosis. *Journal of Infectious Diseases* **204**, S957-S967, (2011).

240 Jemielity, S., Wang, J. J., Chan, Y. K., Ahmed, A. A., Li, W., Monahan, S., Bu, X., Farzan, M., Freeman, G. J., Umetsu, D. T., DeKruyff, R. H. & Choe, H. TIM-family Proteins Promote Infection of Multiple Enveloped Viruses through Virion-associated Phosphatidylserine. *PLoS Pathog* **9**, e1003232, (2013).

241 Chan, S. Y., Speck, R. F., Ma, M. C. & Goldsmith, M. A. Distinct mechanisms of entry by envelope glycoproteins of Marburg and Ebola (Zaire) viruses. *Journal of Virology* **74**, 4933-4937, (2000).

242 Malashkevich, V. N., Schneider, B. J., McNally, M. L., Milhollen, M. A., Pang, J. X. & Kim, P. S. Core structure of the envelope glycoprotein GP2 from Ebola virus at 1.9- Å resolution. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 2662-2667, (1999).

243 Côté, M., Misasi, J., Ren, T., Bruchez, A., Lee, K., Filone, C. M., Hensley, L., Li, Q., Ory, D., Chandran, K. & Cunningham, J. Small molecule inhibitors reveal Niemann-Pick C1 is essential for Ebola virus infection. *Nature.* **477**, 344-348, (2011).

244 Hirschberg, R., Ward, L. A., Kilgore, N., Kurnat, R., Schiltz, H., Albrecht, M. T., Christopher, G. W. & Nuzum, E. Challenges, progress, and opportunities: Proceedings of the filovirus medical countermeasures workshop. *Viruses* **6**, 2673-2697, (2014).

245 CDC. *Questions and Answers on experimental treatments and vaccines for Ebola*, http://www.cdc.gov/vhf/ebola/outbreaks/guinea/qa-experimental-treatments.html (2014).

246 Chang, J., Warren, T. K., Zhao, X., Gill, T., Guo, F., Wang, L., Comunale, M. A., Du, Y., Alonzi, D. S., Yu, W., Ye, H., Liu, F., Guo, J. T., Mehta, A., Cuconati, A., Butters, T. D., Bavari, S., Xu, X. & Block, T. M. Small molecule inhibitors of ER α-glucosidases are active against multiple hemorrhagic fever viruses. *Antiviral Research* **98**, 432-440, (2013).

247 Butters, T. D., Dwek, R. A., Klose, B., Ramstedt, U. & Zitzmann, N. Iminosugars and methods of treating filoviral diseases. United States patent US 2011/0065754 A1 (2011).

248 Nakayama, E. & Saijo, M. Animal models for Ebola and Marburg virus infections. *Front. Microbiol.* **4**, (2013).

249 Shankar, B. P. Adenovirus infection in guinea pig - A case study. *Vet. World* **1**, 280-280, (2008).

250 Finnie, J. W., Noonan, D. E. & Swift, J. G. Adenovirus pneumonia in guinea pigs. *Aust. Vet. J.* **77**, 191-192, (1999).

251 Garcia-Carrillo, C. Relationship between bodyweight and spleen size in guinea-pigs. *Laboratory Animals* **11**, 175-180, (1977).

252 Shotton, C., Arnold, C., Sattentau, Q., Sodroski, J. & McKeating, J. A. Identification and characterization of monoclonal antibodies specific for polymorphic antigenic determinants within the V2 region of the human immunodeficiency virus type 1 envelope glycoprotein. *Journal of Virology* **69**, 222-230, (1995).

253 McLellan, J. S., Pancera, M., Carrico, C., Gorman, J., Julien, J.-P., Khayat, R., Louder, R., Pejchal, R., Sastry, M., Dai, K., O'Dell, S., Patel, N., Shahzad-ul-Hussan, S., Yang, Y., Zhang, B., Zhou, T., Zhu, J., Boyington, J. C., Chuang, G.-Y., Diwanji, D., Georgiev, I., Do Kwon, Y., Lee, D., Louder, M. K., Moquin, S., Schmidt, S. D., Yang, Z.-Y., Bonsignori, M., Crump, J. A., Kapiga, S. H., Sam, N. E., Haynes, B. F., Burton, D. R., Koff, W. C., Walker, L. M., Phogat, S., Wyatt, R., Orwenyo, J., Wang, L.-X., Arthos, J., Bewley, C. A., Mascola, J. R., Nabel, G. J., Schief, W. R., Ward, A. B., Wilson, I. A. & Kwong, P. D. Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature.* **480**, 336-343, (2011).

254 Pancera, M., Shahzad-ul-Hussan, S., Doria-Rose, N. A., McLellan, J. S., Bailer, R. T., Dai, K., Loesgen, S., Louder, M. K., Staupe, R. P., Yang, Y., Zhang, B., Parks, R., Eudailey, J., Lloyd, K. E., Blinn, J., Alam, S. M., Haynes, B. F., Amin, M. N., Wang, L.-X., Burton, D. R., Koff, W. C., Nabel, G. J., Mascola, J. R., Bewley, C. A. & Kwong, P. D. Structural basis for diverse N-glycan recognition by HIV-1–neutralizing V1–V2–directed antibody PG16. *Nat Struct Mol Biol* **20**, 804-813, (2013).

255 Branza-Nichita, N., Petrescu, A. J., Dwek, R. A., Wormald, M. R., Platt, F. M. & Petrescu, S. M. Tyrosinase folding and copper loading in vivo: A crucial role for calnexin and α-glucosidase II. *Biochemical and Biophysical Research Communications* **261**, 720-725, (1999).

256 Schrag, J. D., Bergeron, J. J., Li, Y., Borisova, S., Hahn, M., Thomas, D. Y. & Cygler, M. The Structure of calnexin, an ER chaperone involved in quality control of protein folding. *Mol Cell* **8**, 633-644, (2001).

257 Papandréou, M. J., Barbouche, R., Guieu, R., Rivera, S., Fantini, J., Khrestchatisky, M., Jones, I. M. & Fenouillet, E. Mapping of domains on HIV envelope protein mediating association with calnexin and protein-disulfide isomerase. *Journal of Biological Chemistry* **285**, 13788-13796, (2010).

258 Chui, D., Sellakumar, G., Green, R., Sutton-Smith, M., McQuistan, T., Marek, K., Morris, H., Dell, A. & Marth, J. Genetic remodeling of protein glycosylation in vivo induces autoimmune disease. *Proc Natl Acad Sci U S A* **98**, 1142-1147, (2001).

259 Lederkremer, G. Z. Glycoprotein folding, quality control and ER-associated degradation. *Current Opinion in Structural Biology* **19**, 515-523, (2009).

260 Molinari, M., Galli, C., Vanoni, O., Arnold, S. M. & Kaufman, R. J. Persistent glycoprotein misfolding activates the glucosidase II/UGT1-driven calnexin cycle to delay aggregation and loss of folding competence. *Molecular Cell* **20**, 503-512, (2005).

261 Russell, W. C. Adenoviruses: Update on structure and function. *Journal of General Virology* **90**, 1-20, (2009).

262 Cauet, G., Strub, J. M., Leize, E., Wagner, E., Van Dorsselaer, A. & Lusky, M. Identification of the glycosylation site of the adenovirus type 5 fiber protein. *Biochemistry* **44**, 5453-5460, (2005).

263 Gainotti, R., Ricarte, C., Ebekian, B., Videla, C., Carballal, G., Damonte, E. B. & Echavarría, M. Real time PCR for rapid determination of susceptibility of adenovirus to antiviral drugs. *Journal of Virological Methods* **164**, 30-34, (2010).

264 Waye, M. M. Y. & Sing, C. W. Anti-viral drugs for human adenoviruses. *Pharmaceuticals* **3**, 3343-3354, (2010).

265 Philipson, L. Adenovirus assay by the fluorescent cell-counting procedure. *Virology* **15**, 263-268, (1961).

266 Zolotukhin, S., Potter, M., Zolotukhin, I., Sakai, Y., Loiler, S., Fraites Jr, T. J., Chiodo, V. A., Phillipsberg, T., Muzyczka, N., Hauswirth, W. W., Flotte, T. R., Byrne, B. J. & Snyder, R. O. Production and purification of serotype 1, 2, and 5 recombinant adeno-associated viral vectors. *Methods* **28**, 158-167, (2002).