STUDIES ON INTERACTIONS OF NOROVIRUS CAPSID PROTEIN WITH FUCOSYLATED GLYCANS AND GALACTOSYLCERAMIDE AS SOLUBLE AND MEMBRANE BOUND LIGANDS

Waqas Nasir

Department of Clinical Chemistry and Transfusion Medicine
Institute of Biomedicine
Sahlgrenska Academy at University of Gothenburg

UNIVERSITY OF GOTHENBURG
Gothenburg, 2014

Cover illustration: Norovirus P dimer in complex with blood group B HBGA trisaccharide.

Studies on interactions of norovirus capsid protein with fucosylated glycans and galactosylceramide as soluble and membrane bound ligands
Institute of Biomedicine
Department of Clinical Chemistry and Transfusion Medicine
Sahlgrenska Academy at the University of Gothenburg, 2014
ISBN (e-pub): 978-91-628-9072-8
ISBN (print): 978-91-628-9071-1
URL: http://hdl.handle.net/2077/35454

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِيْمَ

In the name of Allah, the Gracious, the Merciful.

To His Holiness, Mirza Ghulam Ahmed[as] of Qadian (1835-1908)

The Promised Messiah.

Abstract

Noroviruses (NVs) are among the most common viral pathogens which target the gastrointestinal (GI) tract and cause severe diarrhea, vomiting and episodes of abdominal cramps with fever. Millions of people around the world get infected with NVs annually, of which 200,000 cases are estimated to be fatal. Yet for decades, the failure of propagating the human NVs in a cell-culture model has hampered NV infection research and consequently treatment and vaccine development. Thus, the research has mainly been focused on epidemiology and studies on the interaction of model virus-like particles (VLPs) with potential host receptors or attachment factors to gain understanding of the first steps of the infection in order to successfully pave the way for effective clinical therapy or prophylaxis. Motivated by this theme, the emphasis of the present work is mainly on protein-carbohydrate interactions of host glycans with NV VLPs. About 80 % of NV outbreaks reported to date are caused by GII.4 genocluster of NVs. Therefore, due to its dominating clinical importance GII.4 NV-like particles and capsid protein dimers are used for *in vitro* and *in silico* studies included in the thesis work.

NVs have been shown to recognize host histo-blood group antigens (HBGAs) as viral receptors or attachment factors. To investigate the molecular details of interactions of GII.4 NVs with a repertoire of fucosylated HBGAs, molecular dynamics studies were initially carried out based on the crystal structure of B-trisaccharide HBGA in complex with VA387 GII.4 norovirus P dimer. The results, which were later confirmed by crystallographic studies, could explain, on an atomic level, the binding characteristics from a mutagenesis study carried out earlier on the same NV strain. Along with the modelling studies theoretical binding energies were also estimated for different HBGAs binding to VA387 P dimers. The atomic details of binding modes revealed how a single fucose binding site could exploit two different binding modes of the same glycan. This was supported by a literature review of the occurrence of similar fucose binding sites and modes observed in nature for fucose binding lectins and antibodies.

One of the objectives of the thesis was to understand the dynamics of virus host interactions at the cell surface membrane, as it holds clues to very early steps of virus infection. Therefore, total internal reflection fluorescent microscopy (TIRFM) was developed to study the binding events of glycosphingolipid (GSL)-containing vesicles to single NV like particles bound to a supported lipid bilayer (SLB). The advantage of this single vesicle binding assay is the ability to analyze the attachment-detachment kinetics both in transient and steady state conditions. Therefore, it enabled us, for the first time, to discriminate between compositionally different GSL-containing vesicles based on their detachment activation energy. This relates directly to the binding strength of the virus-vesicle complex thereby providing new insights into the characteristics of binding virus-like particles to various lipid bound glycans. Moreover, the differences in the distribution of detachment energy of activation for different GSL-containing vesicles were also analyzed.

Microdomains or clustered patches of GSLs with or without cholesterol are dynamic integral parts of most of the plasma membranes. Their role has been implicated in virus infection of HIV and influenza virus but not in NVs. However for the first time, NVs were shown to recognize galactosylceramide (GalCer) microdomains in supported lipid bilayers. The atomic details of the binding mode of these interactions are, however, still to be clarified.

In conclusion, the thesis describes details of viral protein - host carbohydrate interactions at the molecular level, of relevance for understanding virus infection and design of novel anti-viral strategies.

Keywords: norovirus, molecular dynamics, molecular docking, total internal reflection fluorescent microscopy, desorption activation energy, virus-host interactions, histo-blood group antigens, fucose

Sammanfattning på svenska

Norovirus (NV) räknas som en av de vanligaste gastrointestinala patogenerna. Norovirus infektioner leder ofta till kortvariga episoder av diarréer, kräkningar och magkramper med feber. Miljontals människor runt om i världen smittas årligen med NVs och i ca 200 000 fall har infektionen en dödlig utgång. En av de största utmaningarna för NV-forskningen är kopplade till svårigheterna att odla det mänskliga viruset i cellkulturer. Detta utgör ett stort hinder för infektionsforskningen och därmed utvecklingen av nya terapeutiska strategier och vacciner. Sålunda har forskningen hittills varit främst inriktad på de epidemiologiska aspekterna av NV infektionen. Det förekommer även en intensiv forskning kring grunderna för adhesion och interaktion med olika värdfaktorer som kan vara kritiska under de första stegen av en infektion. Sådana studier är av stor betydelse för att bana väg för en effektiv klinisk profylax. Motiverad av detta tema ligger tyngdpunkten i detta avhandlingsarbete främst på studier av protein - kolhydrat- interaktioner av fukosylerade värd glykaner med s.k. virus-liknade-partiklar (VLPs). Cirka 80% av NV-utbrotten orsakas av stammar tillhörande GII.4 genoklustret. På grund av dess dominerande kliniska betydelse har även vi använt GII.4 NV liknande virus partiklar och protein dimerer för våra *in vitro* och *in silico* studier.

Flera olika studier har visat att NV kan använda blodgruppsantigener (HBGAs) som receptorer eller adhesionsfaktorer. För att börja undersöka de molekylära detaljerna av dessa interaktioner genomförde vi initialt molekyldynamikstudier av ett GII.4 NVs med en bred repertoar av fukosylerade HBGAs. Studien utgick från kristallstrukturen av en HBGA B-trisackarid i komplex med VA387 GII.4 NV protein dimeren. Våra resultat gav oss insikt i de strukturella grunderna för bindning på atomärnivå och lyckades dessutom förklara resultaten från en mutagenes studie som tidigare genomförts på samma NV–stam. Dessa teoretiska insikter bekräftades senare med en kristallografisk studie från en annan grupp. Efter modellering fortsatte vi med att beräkna de teoretiska bindningsenergierna för interaktionen mellan olika HBGAs och VA387 P dimeren. Vi lyckades nu visa hur ett enda fukos-bindningsställe på virusproteinet kan utnyttja två olika bindningssätt för samma glykan. Detta stöddes av en litteraturgenomgång av förekomsten av liknande fukos-bindningsställen på andra fukosbindande lektiner och antikroppar.

Ett av målen med avhandlingen var att förstå dynamiken av virus-värd interaktioner vid cellmembranet, något som kan ge ledtrådar om de tidiga stegen i en virusinfektion. Därför använde vi en nyutvecklad metod, Total Intern Reflektion Fluorescens-mikroskopi (TIRFM), för att studera bindningen mellan glykosfingolipid (GSL) innehållande vesiklar med NV-liknande partiklar bundna till modellmembraner (supported lipid bilayers). En stor fördel med denna analys var att vi kunde följa bindningskinetiken för enskilda vesiklar under olika experimentella förhållanden. På så sätt kunde vi diskriminera mellan olika GSL-innehållande vesiklar, med olika glykolipid sammansättningar, genom att beräkna aktiveringsenergin för dissociationen av vesiklar från VLP. Denna parameter är direkt relaterad till bindningsstyrkan av virus-vesikel komplexet vilket gett oss nya insikter i bindningsegenskaperna av virus-liknade partiklar till olika lipidbundna glykaner.

Mikrodomäner av GSLs, med eller utan kolesterol, förekommer i de flesta plasmamembran. De har visat sig vara av betydelse för olika virala infektioner såsom HIV och influensa. Om de kan vara viktiga för norovirus är hittills okänt. I detta sammanhang visade vi också för första gången att NV interagerar med galaktosylceramid (GalCer) i mikrodomäner i artificiella lipidmembraner. Ytterligare studier behövs dock för att klargöra mekanismen bakom denna observation.

Sammanfattningsvis beskriver denna avhandling detaljerna bakom virala protein-kolhydrat interaktioner på molekylär nivå, något som är relevant för att förstå de första stegen vid norovirusinfektioner och för att kunna utforma nya antivirala strategier mot dessa.

List of Publications

The following papers are included in this thesis. They will be referred to in the text by their roman numerals.

I. Koppisetty CA, **Nasir W**, Strino F, Rydell GE, Larson G, Nyholm PG. 2010. *Computational studies on the interaction of ABO-active saccharides with the norovirus VA387 capsid protein can explain experimental binding data*. Journal of Computer-Aided Molecular Design 24(5):423-431.

II. **Nasir W**, Frank M, Koppisetty CA, Larson G, Nyholm PG. 2012. *Lewis histo-blood group α1,3/α1,4 fucose residues may both mediate binding to GII.4 noroviruses*. Glycobiology 22(9):1163-1172.

III. **Nasir W**, Bally M, Zhdanov VP, Larson G, Höök F. *Interaction of virus-like particles with vesicles containing glycolipids: Kinetics of detachment*. Manuscript.

IV. **Nasir W**, Bally M, Kunze A, Zhdanov VP, Parra F, Peters T, Höök F, Larson G. *Binding and inhibition studies on interactions of GII.4 norovirus-like particles with membrane bound fucosylated histo-blood group antigens*. Manuscript.

V. Bally M, Rydell GE, Zahn R, **Nasir W**, Eggeling C, Breimer ME, Svensson L, Hook F, Larson G. 2012. *Norovirus GII.4 virus-like particles recognize galactosylceramides in domains of planar supported lipid bilayers*. Angew Chem Int Ed Engl 51(48):12020-12024.

Contents

Abbreviations

A-1	A type 1
AFM	Atomic Force Microscopy
ALe^b	A Lewis b
ALe^y	A Lewis y
AMBER	Assisted Model Building with Energy Refinement
B-1	B type 1
BSA	Bovine Serum Albumin
CBA	Chromatogram Binding Assay
CD	Cluster of Differentiation
Cer	Ceramide
CHARMM	Chemistry at HARvard Macromoleuclar Mechanics
Da	Dalton
ELISA	Enzyme-Linked ImmunoSorbent Assay
ER	Endoplasmic Reticulum
Fuc-T	FucosylTransferase enzyme
FUT	FUcosylTransferase gene
Gal	Galactose
GalNAc	N-acetylgalactosamine
GBSA	Generalized Born Surface Area
Glc	Glucose
GlcNAc	N-acetylglucosamine
GROMOS	GROningen MOlecular Simulation
GSL	Glycosphingolipid
H-1	H type 1

HBGA	Histo-Blood Group Antigen
HSA	Human Serum Albumin
ICTV	International Committee on Taxonomy of Viruses
IEL	IntraEpithelial Lymphocyte
IFN-γ	Interferon gamma
IL	Interleukin
LAD	Leukocyte Adhesion Deficiency
Le^a	Lewis a
Le^b	Lewis b
Le^x	Lewis x
Le^y	Lewis y
MD	Molecular Dynamics
MM	Molecular mechanics
NAMD	NAnoscale Molecular Dynamics
Neu5Ac	N-acetyl neuraminic acid
NMR	Nuclear Magnetic Resonance
NV	Norovirus
ORF	Open Reading Frame
PBMC	Peripheral Blood Mononuclear Cell
PBS	Phosphate Buffer Saline
PDB	Protein Data Bank
PRR	Pattern Recognition Receptor
QCM-D	Quartz Crystal Microbalance with Dissipation
RdRp	RNA dependent RNA polymerase
RNA	Ribonucleic acid
SLe^x	Silayl Lewis x

TIRFM	Total Internal Reflection Fluorescent Microscopy
TLR	Toll-Like Receptor
VF1	Virulence Factor 1
VLP	Virus-Like Particle
VP1	Viral Protein 1
VPg	Viral genome protein

1. Introduction

Infectious diarrhea or gastroenteritis is a major global public health concern (Krenzer, 2012). The etiological agents identified are mainly bacteria (Bishop and Ulshen, 1988), viruses (Krenzer, 2012; Steinhoff and John, 1980) and parasites (Guerrant and Bobak, 1991). The viral pathogens infecting human gut, mostly in children, are mainly from four families; Caliciviridae (Norovirus and Sapovirus), Reoviridae (Rotavirus), Adenoviridae (Adenovirus) and Astroviridae (Astrovirus).

The advent of modern molecular biology techniques has attributed the majority of gastroenteritis cases, previously of unknown origin, to noroviruses (NV) as the major etiological agent (Atmar and Estes, 2001). Today, noroviruses are identified to cause about half of all gastroenteritis outbreaks worldwide (Patel *et al.*, 2009). Due to the extremely contagious nature, noroviruses to this day pose great threat to all kinds of community settings including hospitals, day care centers, nursing homes, schools, restaurants and cruise ships (Fankhauser *et al.*, 2002; Lindesmith *et al.*, 2008; Matthews *et al.*, 2012). During the last decade, there has been an increase in the number of reported norovirus outbreaks and a new variant has appeared every 2-3 years causing large epidemics worldwide (Donaldson *et al.*, 2008) with Sydney 2012 being the most recent one (Eden *et al.*, 2014).

There is a well-established body of evidence that human noroviruses recognize histo-blood group antigens (HBGAs) as potential receptors or attachment factors (Shirato, 2011). Although studies showing their role in virus infection and cellular entry are virtually lacking in the absence of cell culture model, *in vitro* binding studies are still helpful in designing much needed antiviral therapeutics. Following this theme, the current thesis aims at revealing novel insights into the norovirus-HBGA interactions to facilitate the understanding of very early steps of virus recognition over the cellular membrane and design of novel antiviral strategies.

1.1. Norovirus

1.1.1. Historical perspective

The phrase “winter vomiting disease” was first coined by Zahorsky in 1929 who described a disease which was previously unknown. The symptoms included vomiting, diarrhea and abdominal cramps which all faded off in only a couple of days and with full recovery (Zahorsky, 1929). During the next few decades, several outbreaks with similar symptomatic description occurred but none could be attributed to the viral etiology because of the failure of classical tissue-culture virology approaches (Kapikian, 2000). It was not before 1972 that the clear serological evidence of virus infection was shown by Kapikian et al., through immune electron microscopy where they identified a 27 nm virus particle from infectious stool specimen (Kapikian *et al.*, 1972). Kapikian and co-workers later demonstrated a clear difference between the quantities of antibody coating of this 27 nm virus particle using pre-challenge and convalescent sera of the same individual (Kapikian *et al.*, 1975). The virus was named

Norwalk virus based on the outbreak in an elementary school in Norwalk, Ohio, 1968 (Adler and Zickl, 1969) from which the samples for virus identification were prepared. In addition to Norwalk virus, other names like "Norwalk-like virus" or "small rounded structure virus" have also been used over the years (Appleton, 1987). However, in 2002 the international committee on taxonomy of viruses (ICTV) standardized the name as "Norovirus".

1.1.2. Clinical features

General symptoms

The general characteristic clinical manifestations of a norovirus infection are sudden onset of vomiting and watery diarrhea. The constitutional symptoms include nausea, abdominal cramps, myalgias and fever (37-45% of cases) (Glass *et al.*, 2009). The incubation period of the virus is 10-51 hours and the illness lasts typically for 2-3 days but could be longer in case of children (< 11 years), elderly or immunocompromised individuals (Lopman *et al.*, 2004a; Lopman *et al.*, 2004b; Rockx *et al.*, 2002). The virus can be identified in stools for 8 weeks post infection (Atmar *et al.*, 2008) and in some cases months or even years (Milbrath *et al.*, 2013; Nilsson *et al.*, 2003). Norovirus induces a self-limiting disease and the symptoms usually disappear without sequelae. In otherwise healthy individuals only rehydration is enough treatment for total restoration. However, for the vulnerable population (children, elderly and immune compromised) complications may occur which include renal insufficiency, electrolyte imbalance and high loss of body fluids (Mattner *et al.*, 2006). Even fatalities have been reported for a number of outbreaks involving immunocompromised and elderly individuals (Desai *et al.*, 2012; Harris *et al.*, 2008; Trivedi *et al.*, 2013; van Asten *et al.*, 2011). Indeed, it has been estimated that more than 200,000 deaths among children (< 5 years) are caused annually by noroviruses in developing countries (Patel *et al.*, 2008).

Unusual symptoms

Several reports have described set of symptoms which are not manifested in common cases of norovirus infections. For e.g. the famous outbreak of norovirus associated gastroenteritis among British soldiers in Afghanistan resulted in symptoms like headache, neck stiffness and photophobia. Even disseminated intravascular coagulation was observed for one of the patients (Centers for Disease and Prevention, 2002). Necrotizing enterocolitis was also observed in patients in a neonatal intensive care unit, half of which were norovirus positive (Turcios-Ruiz *et al.*, 2008). Moreover, norovirus gastroenteritis has also been associated with exacerbations of inflammatory bowel disease (Khan *et al.*, 2009). Apart from otherwise healthy persons, the immunocompromised individuals, particularly the recipients of organ or stem cell transplant have shown extremely complicated syndromes (Nilsson *et al.*, 2003; Roddie *et al.*, 2009; Westhoff *et al.*, 2009).

Subclinical symptoms

The norovirus symptoms may vary from person to person. Some individuals may observe only vomiting or only diarrhea, while others may not present any symptom at all. In one of the challenge studies, around one third of the infected population did not

develop the disease (Graham *et al.*, 1994). Such illness-free carriers of virus do develop a serological response and shed the virions in stools (Hutson *et al.*, 2004a). The asymptomatic infections have been identified in several outbreaks of prototype Norwalk virus and other norovirus strains (Bucardo *et al.*, 2010; Gallimore *et al.*, 2004; Huynen *et al.*, 2013; Ozawa *et al.*, 2007).

1.1.3. Classification

Norovirus together with Lagovirus, Sapovirus, Nebovirus and Vesivirus constitute the known genera of the Caliciviridae family (Green *et al.*, 2000; Kaplon *et al.*, 2011). Following the discovery of novel caliciviruses in rhesus monkey, chicken and swine, new genera have been proposed but not yet officially approved (Farkas *et al.*, 2008; Wolf *et al.*, 2011).

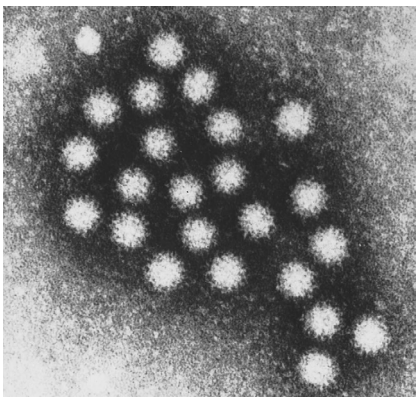

Figure 1.1 Norwalk agent as visualized by electron microscopy for the first time in 1972. *Modified from (Kapikian, 2000).*

The name "calici" (derived from greek work *calyx* meaning cup) comes from the distinct morphology of these viruses showing "cup-like" structures when visualized under the electron microscope (Hansman *et al.*, 2010) (Fig. 1.1).The prominent clinical symptoms associated with animal caliciviruses include upper respiratory tract infections, gastrointestinal infections and severe haemorrhagic syndromes (Hansman *et al.*, 2010) (Fig. 1.2). Noroviruses and sapoviruses infect humans and serve as etiological agents for acute, mild and asymptomatic gastroenteritis. Noroviruses belong to the genetically diverse group of caliciviruses. Since at present a cell culture to propagate all noroviruses is not available, the conventional serotypes based method to cluster the genetically relevant strains is not possible (Duizer *et al.*, 2004). Therefore, several classification schemes have been proposed based on the genetic diversity (Ando *et al.*, 2000; Kroneman *et al.*, 2013; Vinje *et al.*, 2000; Zheng *et al.*, 2006). The most widely used system of classification for noroviruses considers pairwise sequence similarity between whole capsid proteins (Zheng *et al.*, 2006). Using this system, noroviruses were segregated into as many as 5 genogroups and more than 30 genotypes within these genogroups (Patel *et al.*, 2009). More recently, a sixth genogroup (GVI) has been proposed to represent newly discovered dog noroviruses. The most widely used system of classification for noroviruses considers pairwise sequence similarity between whole capsid proteins (Zheng *et al.*, 2006). Using this system, noroviruses were segregated into as many as 5 genogroups and more than 30 genotypes within these genogroups (Patel *et al.*, 2009). More recently, a sixth genogroup (GVI) has been proposed to represent newly discovered dog noroviruses (Mesquita *et al.*, 2010). Human noroviruses are combined in genogroups I, II and IV. Genogroup III mainly infects cattle, while genogroup V includes murine noroviruses (Donaldson *et al.*, 2010; Patel *et al.*, 2009). The cutoff value for the capsid protein sequence similarity is 14.1 % between the variants within single genotype and 15 % between the two nearest genotypes (Zheng *et al.*, 2006).

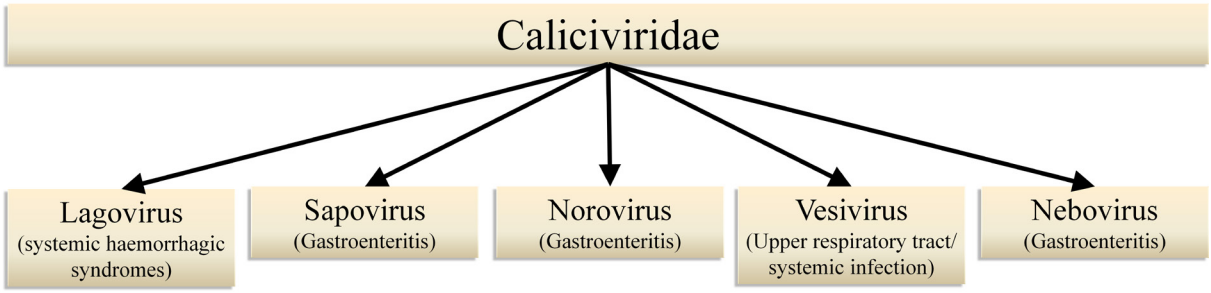

Figure 1.2 Known genera in Caliciviridae family with their clinically defined syndromes.

1.1.4. Epidemiology

Norovirus outbreaks and sporadic cases

Norovirus is estimated to cause half of all gastroenteritis outbreaks worldwide and is the leading cause of acute non-bacterial gastroenteritis in humans (Atmar and Estes, 2006). The unique combination of long periods of virus shedding both in symptomatic and asymptomatic infections, low infectious dose, high degree of environmental contamination, viral stability and extreme antigenic variability enable these pathogens to efficiently dominate the outbreaks in all kinds of community settings (Hansman *et al.*, 2010) mainly including, but not limited to, hospitals (Fretz *et al.*, 2009; Hoffmann *et al.*, 2013; Kanerva *et al.*, 2009; Leuenberger *et al.*, 2007; Ohwaki *et al.*, 2009; Verbelen *et al.*, 2004), day care centers (Lyman *et al.*, 2009), military settings (Grotto *et al.*, 2004; Mayet *et al.*, 2011; Wadl *et al.*, 2010; Yap *et al.*, 2012), cruise ships (Morillo *et al.*, 2012; Wikswo *et al.*, 2011; Vivancos *et al.*, 2010), restaurants (Baker *et al.*, 2011; Centers for Disease and Prevention, 2006; Smith *et al.*, 2012) and schools (Centers for Disease and Prevention, 2008; Gomez, 2008; Morioka *et al.*, 2006; Oogane *et al.*, 2008).

Although, norovirus outbreaks are usually more efficiently reported, sporadic norovirus cases are also common among family settings with a high prevalence among young children and elderly all around the world (Gao *et al.*, 2011; Kele *et al.*, 2009; Muhsen *et al.*, 2013; Park *et al.*, 2012; Patel *et al.*, 2009; Tan *et al.*, 2010; Yoneda *et al.*, 2014). These cases could be complicated if medical care, especially for children and elderly, is not provided in time (Desai *et al.*, 2012). Yoneda *et al.* reported the norovirus detection rate of 39.5 % (in both bacterial and non-bacterial gastroenteritis combined) among infants in a study conducted over the period of 5 years in Japan (Yoneda *et al.*, 2014) while a similar study in Israel found noroviruses in 17.3 % of stool samples collected over 3 years period for children under 5 years (Muhsen *et al.*, 2013), only second to rotaviruses for which a detection rate of 21% was observed. Similar trends have been observed in several other studies. It is clear that the disease burden for norovirus is much higher in developing countries. Patel *et al.* has estimated that in these countries 200,000 deaths among children under five years are caused by noroviruses annually (Patel *et al.*, 2008). Individual deaths for elderly as a consequence of norovirus infection have also been reported (Lopman *et al.*, 2004b; Okada *et al.*, 2006).

Transmission modes

Noroviruses are highly contagious with 50% human infectious dose of as low as 1320 genomic equivalents (Atmar *et al.*, 2013). The fecal-oral route is the preferred mode of transmission which is facilitated by the high loads of norovirus shedding in stools of infected individuals (see section 1.1.2). Person to person contacts have been reported to be associated with a majority (70-80%) of the enteric and norovirus outbreaks (Greig and Lee, 2009; Kroneman *et al.*, 2008b). The control strategies targeting to neutralize person to person transmission routes have proven to be most effective (Friesema *et al.*, 2009). Noroviruses do exploit their ability to sustain the environmental conditions for transmission. An infected individual could potentially contaminate everything in close contact including common households, computer keyboards, door handles and alike (Clay *et al.*, 2006). This is potentially dangerous if not taken care of (Evans *et al.*, 2002; Kuusi *et al.*, 2002) particularly in healthcare facilities and hospitals. Perhaps it also contributes to norovirus being one of the most common reasons to close down the hospital wards (Hansen *et al.*, 2007). Furthermore, projectile vomiting, a common norovirus symptom, has also been shown to contaminate relatively larger areas (Caul, 1995). Food and water-borne outbreaks are also reported but their proportion varies in different countries (Doyle *et al.*, 2009; Hamano *et al.*, 2005; Kroneman *et al.*, 2008a; Lynch *et al.*, 2006). Among food related diseases, noroviruses are estimated to cause 67 % of the cases (Koopmans and Duizer, 2004). Foodborne or water borne norovirus transmission could potentially infect a very wide range of populations from small restaurants (due to an ill food handler) to even across the continents (de Wit *et al.*, 2007; Simmons *et al.*, 2007). Oysters, frozen berries, salads and mussels are well known vehicles for norovirus transmission (Hjertqvist *et al.*, 2006; Makary *et al.*, 2009; Simmons *et al.*, 2007).

Direct zoonotic transmission has not been shown for noroviruses. However, potential risks cannot be denied as an increased number of animal noroviruses closely related to human strains, particularly in pigs, have been reported in recent past (Mattison *et al.*, 2007; Wang *et al.*, 2005). Furthermore, the GIV human norovirus strains, which are not common in human populations, have been identified in different animals (Martella *et al.*, 2007; Martella *et al.*, 2008). Also, cattle and swine have been shown to be susceptible to human noroviruses (Cheetham *et al.*, 2006; Souza *et al.*, 2008). The most striking finding of human GII.4 norovirus particles in the stools of pig raises several questions (Mattison *et al.*, 2007).

Molecular epidemiology

The advancement in the sequencing techniques with more sensitive and rapid virus detection methods has enabled many laboratories to report genetically relevant data together with the epidemiological information. As mentioned before, GI, GII and GIV norovirus strains cause human infection. Irrespective of outbreaks or sporadic cases, GII.4 noroviruses outnumber other norovirus strains in causing the disease in human populations. Patel *et al.* has reported around 75-100 % of sporadic cases are caused by GII.4 noroviruses (Patel *et al.*, 2009). A recent literature review nicely describes the association of different genogroups in norovirus outbreaks and their impact on outbreak characteristics like modes of transmission, outbreak settings and attack rates (Matthews *et al.*, 2012). The authors found that GII noroviruses dominated the

outbreaks (563/754; 75%) with GII.4 strains observed in every third GII norovirus associated outbreak (Matthews *et al.*, 2012). Interestingly, the authors could also find the correlation between different genogroups and modes of transmission and outbreak settings. For instance, waterborne transmission was more likely to be facilitated by GI noroviruses as compared to GII norovirus strains which were predominant in the foodborne, person to person and environmental outbreaks (Matthews *et al.*, 2012). A high number of water- and foodborne outbreaks were associated with the mix of both GI and GII strains. Similarly, GII.4 noroviruses were more likely to cause outbreaks in healthcare centers than other community settings (Matthews *et al.*, 2012). The norovirus outbreaks have also been reviewed elsewhere with fairly similar findings (Fankhauser *et al.*, 2002; Kroneman *et al.*, 2008b; Lopman *et al.*, 2003a).

Seasonality

Norovirus cases are diagnosed throughout the year but usually peak during the winter and adjacent seasons (Ahmed *et al.*, 2013) with some exceptions (Boga *et al.*, 2004; Lopman *et al.*, 2003b). Interestingly, in a recent study this peak was shown to be associated with the rainy season in Cameroon (Ayukekbong *et al.*, 2013), where the minimum temperatures are usually at least as high as in the Scandinavian summer. The reason behind the norovirus seasonality is poorly understood. However, it is likely that the virus transmission and survival are facilitated by the change in human behavior and environment during this time of the year resulting in more closed settings and increased person to person contacts (Rohayem, 2009).

1.1.5. Genetic diversity

Norovirus genome

Noroviruses are small non-enveloped single stranded positive-sense RNA viruses (Donaldson *et al.*, 2010; Thorne and Goodfellow, 2014). The norovirus genome is divided in to three open reading frames (ORFs) with the exception of murine noroviruses in GV genocluster which has an alternative fourth ORF (Thorne and Goodfellow, 2014). The genome is covalently linked to a viral genome protein (VPg) at the 5’ end and is polyadenylated at the 3’ end (Thorne and Goodfellow, 2014). The non-translated region of the human norovirus genome at the 5’ end is typically 48 nucleotides long (Gutiérrez-Escolano *et al.*, 2000; Pletneva *et al.*, 2001) and represents evolutionary conserved RNA secondary structures which are repeated throughout the genome. These secondary structures are implicated to have important roles in virus replication, translation and pathogenesis (Bailey *et al.*, 2010; Simmonds *et al.*, 2008).

The ORF1 (~5 kb) of the genomic RNA covers around 60% of the genome and encodes a ~200 kD polyprotein. At least 6 non-structural proteins are produced by cleavage of this polyprotein by a 3C-like protease, also encoded by the virus (Sosnovtsev *et al.*, 2006). Starting from 5’ end these proteins include: p48 (~48 kDa), which is important in replication complex formation; nucleoside triphosphatase; p22 (~22 kDa), 3A-like protein which is also implicated in replication; VPg, suggested to be involved in translation and replication (Daughenbaugh *et al.*, 2003); 3C-like protease (viral protease) and RNA-dependent RNA polymerase (Donaldson *et al.*, 2010; Hardy, 2005; Hyde and Mackenzie, 2010; Thorne and Goodfellow, 2014). The

major capsid protein, viral protein 1 (VP1, 57 kDa), is encoded by ORF2 whereas, ORF3 encodes a minor structural protein (VP2) which has been shown recently to associate with VP1 (Vongpunsawad *et al.*, 2013). The ORF4 in murine norovirus encodes virulence factor 1 (VF1) which has been shown to have effects on the innate immune response in infected mice (McFadden *et al.*, 2011).

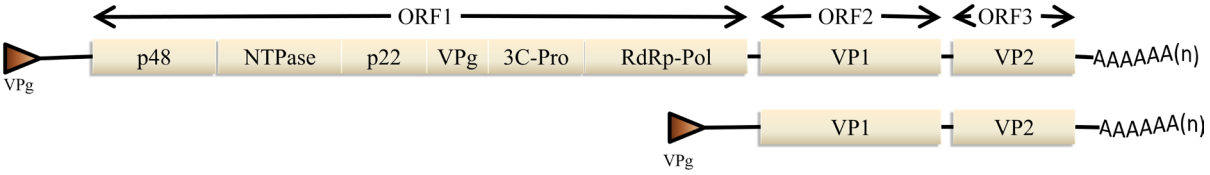

Figure 1.3 Organization of genomic and subgenomic RNA of noroviruses. *The norovirus genome is positive sense single stranded non-segmented RNA that is covalently attached to VPg (viral- protein genome) at the 5' end and is polyadenylated at the 3' end. ORF1 encodes for non-structural viral proteins that are important mainly in replication and translation. The major capsid protein (viral protein 1, VP1) is encoded by ORF2 and ORF3 codes for the minor structural protein VP2.*

Human norovirus recombination

Intra- and inter-genogroup recombination

The first study showing the natural recombination in noroviruses was reported in 1997 when the capsid region of GII snow mountain virus was shown to be 94% identical to that of the Melksham strain, another GII norovirus, with 79% identity in the RdRp region (Hardy *et al.*, 1997). The study demonstrated that snow mountain virus was a naturally occurring recombinant through phylogenetic analysis. The recombination breakpoints in noroviruses are commonly located at the junction of ORF1 and ORF2 resulting in strains which contain capsid structure (ORF2) from one strain while RdRp (ORF1) from another. Among the human noroviruses (GI, GII and GIV) recombinant strains have been found mostly in GII and less frequently in GI noroviruses, whereas GIV norovirus recombinants have not been identified so far. The GI recombinant strain, WUG1/01/JP was first found in Japan in 2001 which contained GI.2 RdRp and GI.6 capsid regions (Katayama *et al.*, 2002). Later on, between 2001 and 2004 several other WUG1/01/JP-like recombinant strains were isolated from Japan and the US (Bull *et al.*, 2007). For GII noroviruses, more than 40 recombinant strains have been isolated which are a mix of RdRp from one of at least 12 different genotypes and the capsid region from one of at least 17 distinct genotypes. Apart from these intra-genogroup recombinants, there is one report on inter-genogroup recombination as well (Nayak *et al.*, 2008). The recombinant strain was found in Kolkata, India and contained GI.3 RdRp and GII.4 capsid region with the breakpoint at ORF1/ORF2 overlap.

GII.4 intra-genotype recombination

Over the last 15 years, GII.4 noroviruses have seen a critical rise both in the number of reported outbreaks and the number of intra-genotype recombinant strains. This is mainly attributed to the number of circulating GII.4 noroviruses since 2002 which is doubled as compared to the early 1990s (Eden *et al.*, 2013). Higher number of circulating strains increases the probability of co-infections and hence recombinations. A recent study has analyzed the GII.4 intra-genotype recombinants from 1974 to 2012. The temporal analysis of 10 different recombinant strains demonstrated the expansion

in genetic variability of these human noroviruses with time (Eden *et al.*, 2013). The authors have also shown that since 2002, after the Farmington Hills strain was isolated, these recombinations have influenced the newly emerging strains either as being recombinant or giving rise to the recombinant circulating strains. Interestingly, the last two pandemic GII.4 strains, i.e. New Orleans 2009 and Sydney 2012, are also actually intra-genotype GII.4 recombinant strains (Eden *et al.*, 2013).

The overall genetic diversity of human noroviruses, and GII.4 noroviruses in particular, is the result of recombination events and mutations under evolutionary pressure due to host factors (section 1.2.5). The recombinant strains have also been reported for animal noroviruses (Bull *et al.*, 2007). However, despite the finding of human noroviruses in the pig stools (Mattison *et al.*, 2007), no human/animal recombinant norovirus strain has been reported in the literature so far. It is clear that a mix between human and animal norovirus will have potentially greater impact in terms of pathogenesis and virulence on the human population (de Jong *et al.*, 1997; Nguyen-Van-Tam and Sellwood, 2007; Stincarelli *et al.*, 2013). The understanding of the origin and impact of genetic variability due to recombination may lead to the prediction of newly emerging pandemic strains and to improved strategies to neutralize such norovirus perils (Hansman *et al.*, 2010).

1.1.6. Cell culture and animal models

Norovirus propagation in cell culture

The research on human norovirus pathogenesis and consequent prophylaxis has long been hampered by the lack of an *in vitro* cell culture system to cultivate the virus, despite of extensive efforts by several laboratories (Duizer *et al.*, 2004). The virus was shown to bind to cultured cells (White *et al.*, 1996) but later on was lost upon several passages suggesting that the cause of failure was not related to the initial binding step (Duizer *et al.*, 2004). Furthermore, a later study has shown replication of viral RNA following transfection and thereafter subsequent release with virus particles from human hepatoma cells (Guix *et al.*, 2007). This suggests that the viral RNA is also infectious *in vitro* and that the block of replication might be related to cellular entry or uncoating. Moreover, the virus inoculum used from the infectious stools was demonstrated to be sufficient to observe replication of other human enteric viruses (e.g. adenovirus, echovirus) present in the same inoculum suggesting that the lack of replication was specific to noroviruses despite of high titers in the virus inocula (Duizer *et al.*, 2004).

Apart from developing conventional cell culture systems, attempts have been made to replicate human noroviruses in 3D-cell culture systems with limited success (Straub *et al.*, 2007). The assay using human colon carcinoma cell line (HT-29) and small intestinal enterocyte-like cell line (caco-2) did not show evidence of any successful replication. However, conflicting results have been reported for 3D-organoid intestinal cell culture model using differentiated human embryonic small intestinal cell line INT-407 underlining the complexity and challenges to cultivate human noroviruses *in vitro* (Papafragkou *et al.*, 2013; Straub *et al.*, 2007).

Animal models

Noroviruses have been identified in bovine, porcine, canine and murine species (Mesquita *et al.*, 2010; Scipioni *et al.*, 2008). In the absence of a human norovirus propagating cell culture system, several animal models have been developed in an attempt to increase the understanding of biology and pathogenesis of human noroviruses.

Murine norovirus discovery (Karst *et al.*, 2003) and subsequent cell culture propagation (Wobus *et al.*, 2004) were important breakthroughs for norovirus research (Wobus *et al.*, 2006). To date murine norovirus remains the only norovirus to be cultivated successfully in a conventional cell culture system. The virus was shown to produce a systemic infection in immunocompromised hosts which initiated in the small intestine and subsequently reached liver, lungs, lymph nodes and spleen (Wobus *et al.*, 2006). The symptoms including encephalitis, vasculitis in cerebral vessels, pneumonia and hepatitis were observed in immunocompromised mice. The wild type mice despite of serological response did not show any clinical symptoms except mild diarrhea (Liu *et al.*, 2009). The virus could be identified in stools and had a tropism for macrophages and dendritic cells *in vitro* (Wobus *et al.*, 2004) contrasting to human norovirus tropism which is thought to be distinct from any immune cells (Lay *et al.*, 2010).

Gnotobiotic pigs and calves have been shown to be experimentally sensitive to human GII.4 noroviruses and develop diarrhea, virus shedding in stools, intestinal lesions and viremia (Cheetham *et al.*, 2007; Souza *et al.*, 2008). The pigs share with humans the expression of H and A HBGAs on mucosal surfaces which are implicated in both human and porcine norovirus infections (Shirato, 2011; Tian *et al.*, 2007). Likewise, norovirus challenge studies have also been conducted on non-human primates including chimpanzees (Bok *et al.*, 2011; Wyatt *et al.*, 1978) and macaques (Rockx *et al.*, 2005b; Subekti *et al.*, 2002). In these studies immunocompromised and newborn monkeys were shown to develop serological response to the infection along with moderate diarrhea while immunocompetent hosts remained mostly asymptomatic. Virus shedding was also observed in stools and in some cases the duration was comparable to that in humans (Bok *et al.*, 2011). In one of the studies, chimpanzees were also shown to be protected from homologous norovirus infection for longer periods of time but not from heterologous infection following immunization using virus-like particles (Bok *et al.*, 2011). Similar study on mice did not demonstrate long term protection even from homologous infections (Liu *et al.*, 2009).

1.1.7. The viral capsid

In Norovirus genomic RNA (see section 1.1.5) is surrounded by a capsid protein which is mainly composed of 60 kDa polypeptide called VP1 (Bertolotti-Ciarlet *et al.*, 2002; Thorne and Goodfellow, 2014). In early 90s, it was shown that, when expressed in recombinant baculovirus transfected insect cells, copies of VP1 can self-assemble to form non-infectious “virus like particle” (VLP) lacking the nucleic acids (Jiang *et al.*, 1992). The VLPs were shown to be immunogenic and morphologically similar to the native virions. The only difference is the presence of VP2, a minor structural protein

which was recently shown to be associated with VP1 (Vongpunsawad *et al.*, 2013), and VPg-linked viral genome (section 1.1.5) in the native virions. However, some labs have produced norovirus VLPs with intact VP2 and 3' untranslated region as well, for instance see methods section in Larsson *et al.* (2006).

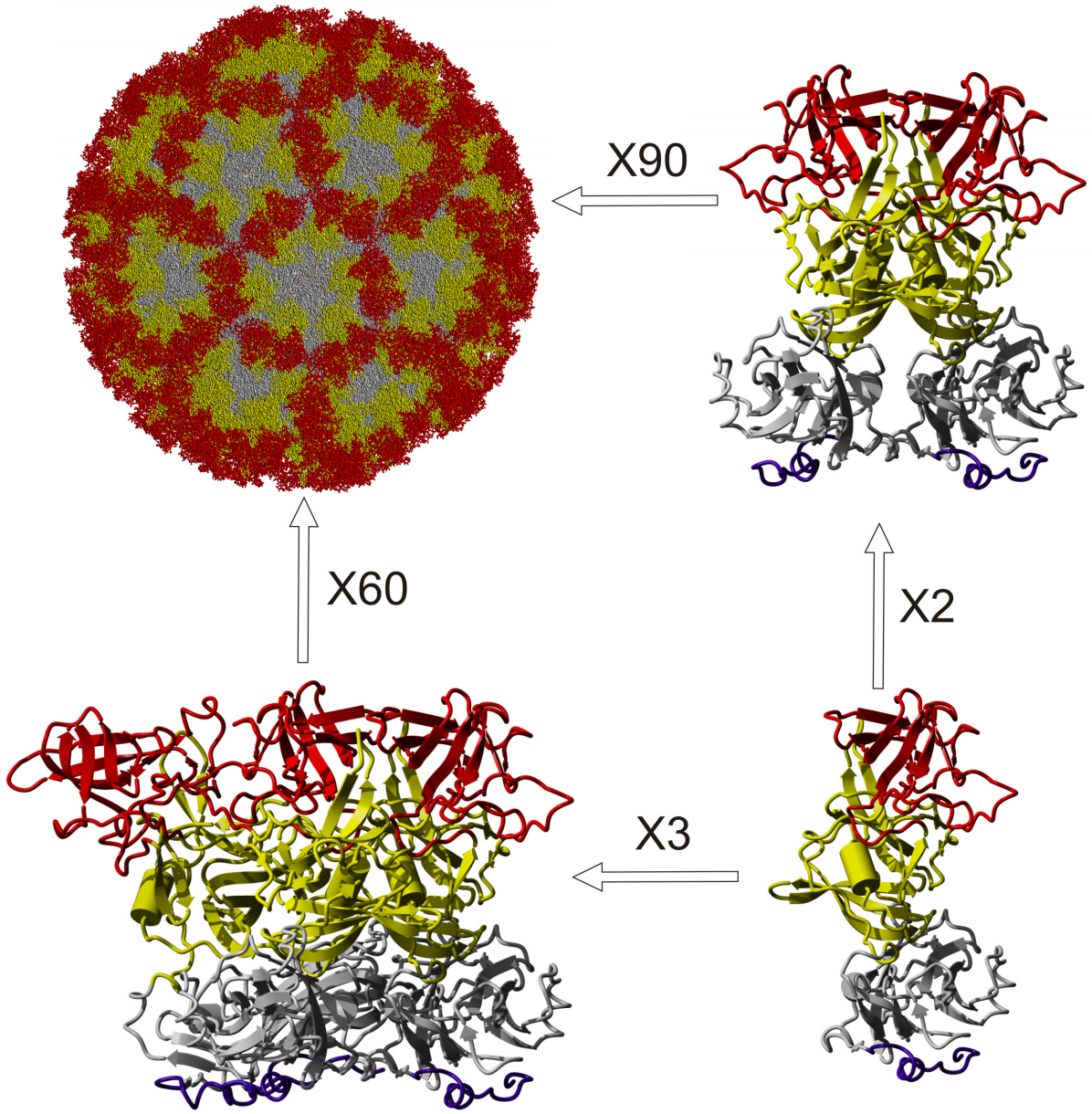

Figure 1.4 The GI.1 Norwalk VLP capsid reconstructed from crystallographic data. *The capsid is composed of an N-terminal arm (blue), S (grey) and P (P1: yellow; P2: red) domains, where P2 subdomain is surface exposed and contains the antigenic and glycan recognition sites. The capsid displays T=3 icosahedral symmetry and could be reconstructed by 60 equilateral triangles of 3 monomers of VP1 protein (single subunit), 90 copies of VP1 homodimers or 180 copies of VP1 monomeric units.*

A major breakthrough in the field of norovirus research was the crystallization and reconstruction of human GI.1 Norwalk VLP at 3.4 Å resolution (Prasad *et al.*, 1999). Prasad and co-workers demonstrated that this VLP is composed of 180 copies of VP1 and exhibits a T=3 icosahedral symmetry which essentially means that the VLP can be geometrically reconstructed using 60 identical equilateral triangles each consisting of 3 copies of VP1 (Prasad *et al.*, 1999). The VLP crystal structure demonstrated that VP1

contains an N-terminal arm (residues 10-49 in Norwalk virus) followed by a shell (S) domain (residue 50-225 in Norwalk virus) which forms the interior of the capsid and provides stability to the VLP through a β-sandwich motif (Prasad *et al.*, 1999). The protruding (P) domain forms the rest of the capsid structure. The P domain is further divided into two subdomains P1 and P2. The P1 subdomain is connected to the S domain with a flexible stretch of amino acids while the P2 subdomain forms the outermost part of the capsid. Sequentially, the P2 subdomain (residues 279-405) is an insertion between the N and C termini of the P1 subdomain. The degree of sequential and structural variation is lowest in the S domain and highest in the P2 subdomain (Chakravarty *et al.*, 2005). The S domain is involved in the highest number of inter subunit contacts in 3-fold, 5-fold and 6-fold rotational axes of symmetry whereas the P1 or P2 sub domains in one subunit interact with only a single adjacent subunit at 2-fold rotational axis of symmetry (Hansman *et al.*, 2010; Prasad *et al.*, 1999). The P2 subdomain is particularly solvent exposed and contains immunogenic and ligand binding sites which later have been pinpointed through several structural determination methods and mutagenesis studies (for further discussion, see section 1.2.5) (Bu *et al.*, 2008; Cao *et al.*, 2007; Tan *et al.*, 2008b).

Cryo-electron microscopy reconstructions of VLPs from norovirus GII Grimsby strain along with the VLPs and native virions of several other caliciviruses demonstrate a common structural arrangement of VP1 (Barcena *et al.*, 2004; Bhella *et al.*, 2008; Chen *et al.*, 2004; Prasad *et al.*, 1996). However, the crystal structure of native murine norovirus has demonstrated that the P domains are projected away from the spherical S domain organization in this virion (Katpally *et al.*, 2008). This structural feature is unique to murine noroviruses. All other caliciviruses are observed to have P domain connected to S domain by a short stretch of amino acids and folds directly above the S domain sphere in the virions (Hansman *et al.*, 2010).

1.1.8. Histopathology

In the absence of virus propagation cell culture systems, the knowledge of human norovirus tropism for a specific target cell is limited. However, several intestinal biopsies from volunteer challenge studies involving both GI (Norwalk) and GII (Hawaii) human noroviruses have demonstrated histological lesions (Agus *et al.*, 1973; Schreiber *et al.*, 1973, 1974). The early studies showed alterations in the jejunal villi cells with intact mucosa. The jejunal lesions were also observed for infected individuals who did not develop the clinical illness (Schreiber *et al.*, 1973, 1974). These findings were further strengthened later on by another challenge study (Troeger *et al.*, 2009) where reduction in the height of duodenal villus but not in the length of the crypt was observed along with an increase in CD8+ lymphocytes belonging to a unique T-cell population, i.e. intraepithelial lymphocytes (IELs). This was in line with the previous findings of an increase in mononuclear cells infiltrating *lamina propria* of small intestine in biopsies (Schreiber *et al.*, 1973, 1974). Although the duodenal villi damage was mainly attributed as a consequence of direct viral infection the increased level of a cytotoxic molecule, perforin, in IELs suggested that IELs might also contribute to the observed blunting of villi via epithelial apoptosis (Troeger *et al.*, 2009).

Despite the histological lesions observed in the reports summarized above, no virus particles could be visualized in the tissues. However, there are reports on binding of virus-like particles and virions to histological sections of human duodenal tissues. Norwalk (GI) VLPs were shown to bind exclusively to gastroduodenal epithelial cells (Marionneau *et al.*, 2002). In contrast, a recent *in vitro* whole virus binding study performed on human adult duodenal tissues revealed that the virus (stool filtrate of norovirus GII.4 Sakai variant) attached mainly to the cells in lamina propria and not to the duodenal epithelial cells (Chan *et al.*, 2011). The apparent discrepancy may be explained by the use of different norovirus strains (GI vs. GII) having distinctly different binding characteristics (Donaldson *et al.*, 2010). This is particularly interesting in the light of a recent finding from the chimpanzee norovirus model where tissue tropism to the cells of small intestinal lamina propria was demonstrated in chimpanzees infected with human GII.4 norovirus (Bok *et al.*, 2011).

It is not surprising that different strains of human noroviruses could potentially have different tissue tropism for small intestinal cells because of the variety of binding specificities to HBGAs exhibited by different strains (Shirato, 2011). The expression of these histo-blood group antigens on different cells in the small intestine may direct the tissue tropism of different norovirus strains. This could potentially explain the differences in the observed clinical symptoms as well (see section 1.1.2).

1.1.9. Immunity

Natural Immunity

In the absence of a cell culture system to propagate noroviruses *in vivo* the markers for natural immunity are challenging to investigate and explain the scarcity of available data on the subject. In particular, there are extremely few studies characterizing the role of various pattern recognition receptors (PRRs), e.g. TLRs, in human norovirus infections. However, one vaccination study has recently demonstrated that TLR7 and TLR9 agonists when used as adjuvants with VLPs elicit a stronger immune response indicating their potential role in norovirus infection (Hjelm *et al.*, 2013). Another preliminary study has suggested that retinoic acid reducible gene 1 (RIG-1) pathway is not important in inhibiting human Norwalk RNA replication in human hepatoma Huh-7 cells (Guix *et al.*, 2007). Most of the studies on human immune responses to norovirus infections have mainly relied on volunteer challenges following infection or vaccination (Debbink *et al.*, 2012b; Lindesmith *et al.*, 2005; Lindesmith *et al.*, 2010; Tacket *et al.*, 2003). In norovirus infection challenge studies, the analysis of pre- and post-challenge peripheral blood mononuclear cells (PBMCs) from volunteers have shown that norovirus infection induces CD4+ Th1 responses leading to IFN-γ production and IL-2, but not IL-6, up regulation. Similar results have been reported in an *ex vivo* study using human PBMCs which were stimulated with VLPs (Ponterio *et al.*, 2013).

Acquired immunity

The first indication of short term immunity came from early challenge studies (Dolin *et al.*, 1972; Parrino *et al.*, 1977; Wyatt *et al.*, 1974). In one cross challenge study this immunity was shown to be genogroup specific using GI (Norwalk) and GII (Hawaii)

viral agents for challenges and re-challenges (Wyatt *et al.*, 1974). Similar lack of cross protection was reported quite recently in a chimpanzee norovirus model after infection with GI and GII norovirus agents (Bok *et al.*, 2011). A recent vaccination study demonstrated at least 3 weeks protection from clinical Norwalk virus infection using the corresponding VLP as a vaccine (Atmar *et al.*, 2011).

While there is some evidence for short term immunity, the long term immunity against human noroviruses seems to be more complicated. In one of the reports all of the individuals who were symptomatically infected with norovirus fell ill again upon challenge with the same virus 2-4 years later (Parrino *et al.*, 1977). In contrast, another study showed that almost half of genetically susceptible individuals were protected against noroviruses upon re-infection suggesting long-term acquired immunity (Lindesmith *et al.*, 2003). These contradictory results might be indicative of the fact that protective immunity might be confounded by the pre-exposure to a norovirus strain in circulation and the frequency of infection before challenge. Earlier challenge studies suggested that individuals with high pre-challenge antibody titers to noroviruses were more prone to getting infected than those with lower antibody titers before challenge, suggestive of herd immunity (Baron *et al.*, 1984; Gray *et al.*, 1994; Okhuysen *et al.*, 1995). Moreover, since noroviruses are extremely prevalent in all populations almost everyone has been exposed to circulating norovirus strain during his lifetime (Donaldson *et al.*, 2010). This exposure history severely complicates the interpretation of human norovirus immunity studies.

Antibody cross-reactivity has been observed within the GI and GII genogroups but not in between (Cannon *et al.*, 2009; Lindesmith *et al.*, 2010; Lindesmith *et al.*, 2011; Lindesmith *et al.*, 2008; Rockx *et al.*, 2005a). In general the GI antibody response has been shown to be extended to heterologous strains within the GI genogroup (Lindesmith *et al.*, 2010) however, GII antibodies have more restricted homologous cross reactivity (Lindesmith *et al.*, 2011).

1.2. Glycobiology and pathogenesis

Glycans, for the most part of 20th century, were in the shadow of great advances in the research on proteins and nucleic acids as principle biomolecules of life. It was only recently that glycans were recognized as major biomolecules through considerable developments in the still growing and expanding field of Glycobiology (Varki and Sharon, 2009). Over the last few decades the physiological roles of glycans, soluble or membrane bound on the surface of cells, have been established in a wide range of biological functions including protein maturation, protein folding, cell signaling, adhesion and recognition (Hakomori, 1990; Helenius and Aebi, 2001; Lichtenstein and Rabinovich, 2013; Molinari, 2007; Sakuma *et al.*, 2003; Sarter *et al.*, 2007; Sharon and Lis, 1989, 1993; Zanetta *et al.*, 1992). The biological and functional roles of glycans are manifested through (i) protein-carbohydrate interactions (Collins and Paulson, 2004; van Kooyk and Rabinovich, 2008) (ii) carbohydrate-carbohydrate interactions (Hakomori, 2004) (iii) aggregation of lipid-linked carbohydrates in microdomains or lipid rafts in cellular membranes (Hakomori, 2000). These features, particularly protein-carbohydrate interactions are specifically exploited by several pathogens to

adhere and subsequently invade the host cells (Karlsson, 1995; Skehel and Wiley, 2000; Smith and Helenius, 2004). Moreover, several reports have linked impairment in glycan expression and biosynthesis to severe health disorders including atherosclerosis (Shrikhande *et al.*, 2010), Alzheimer's disease (Schedin-Weiss *et al.*, 2014), muscle dystrophies (Kanagawa *et al.*, 2013; Saito *et al.*, 2005), lysosomal storage diseases (Futerman and van Meer, 2004) and a growing list of congenital disorders of glycosylation (Cylwik *et al.*, 2013a; Cylwik *et al.*, 2013b).

1.2.1. Glycan diversity

Glycans, carbohydrates or sugars are "hydrates of carbon" with the empirical formula $C_x(H_2O)_n$ where x is typically between 3 and 9. All living organisms synthesize and metabolize glycans which play a multitude of physiological roles and serve as energy sources for different cellular functions. The hallmark of Glycobiology is the diversity of carbohydrate structures found in living systems. This heterogeneity which is unique to sugars comes from several characteristics of glycans including their capacity to form glycosidic bonds with different types and numbers of sugars, the differences in their sequence, configurations and positions of anomeric bonds and their ability to form branched structures. To appreciate the complexity in glycan structures one can compare them with proteins. For example two identical amino acid residues can produce only a single peptide while two identical hexose monosaccharides, like glucose, can produce more than 12 unique sugar molecules. This combinatorial diversity increases exponentially with the number of monosaccharides. For e.g. four different amino acids can produce 24 different peptides while 4 different hexose monosaccharides can theoretically give rise to at least 35,560 unique glycans. However, in living cells only a subset of these possible structures have been found to date but it gives some idea of the potentially enormous information coding capacity of glycans and the challenges of structural Glycobiology (Bertozzi and Rabuka, 2009; Ghazarian *et al.*, 2011; Taylor and Drickamer, 2011).

1.2.2. Glycan biosynthesis

In contrast to proteins, glycans are not primary gene products. The biosynthesis of glycans is instead carried out by the *sequential* action of glycosyltransferases which catalyze the glycosylation reactions of the substrate nucleotide donors or transporters, which are essentially for delivering the activated form of monosaccharides in the right place and time, and of substrate acceptors which commonly include to oligosaccharides, proteins and ceramides. In eukaryotes, most of the biosynthetic pathways characterized for the assembly of the main classes of glycans are confined to the cellular compartments of the Golgi apparatus and the endoplasmic reticulum (ER) with the exception of some cytoplasmic and nuclear glycans which assemble in cytoplasm or on plasma membranes. In analogy to proteases, glycans are also cleaved at specific glycosidic linkages by glycosidases for recycling purposes or to form intermediates for other glycan processing enzymes. The regulation of glycan expression and diversity is governed primarily by the specificity, availability and competition between glycosyltransferases and other glycan processing enzymes. The level of gene expression, post translational protein modifications and the subcellular

localization of expressed glycosyltransfeases and glycosidases together with sugar donors and substrate availability thus controls glycan biosynthesis (Freeze and Elbein, 2009; Rini *et al.*, 2009; Taylor and Drickamer, 2011; Varki *et al.*, 2009a).

1.2.3. Glycoconjugates

Most of the biomolecules in eukaryotic cells are transported through Golgi apparatus after originating in ER. During this passage lipids and amino acids are often covalently modified or *glycosylated* by a selection of glycosyltransferases to form *glycoconjugates*. The cellular membrane of a eukaryotic cell is densely covered by these glycoconjugates which include glycolipids, glycoproteins and proteoglycans. The latter two are abundant in body fluids including saliva, serum and cerebrospinal fluid (Halim *et al.*, 2013; Ramachandran *et al.*, 2008; Wu *et al.*, 2014). Nuclear and cytoplasmic glycoconjugates are also common. However, their biosynthesis is distinct from the secretory and cell surface glycoconjugates (Varki *et al.*, 2009a).

Glycoproteins

More than 50% of eukaryotic proteins have been estimated to be glycosylated (Apweiler *et al.*, 1999). The sugars could be covalently attached to proteins through the amide group of asparagine (N-linked) or hydroxyl group of serine, threonine or tyrosine (O-linked). The asparagine in N-linked glycosylation is usually present in the sequence motif Asn-X-Ser/Thr. Both N- and O-linked glycosylations have distinct pathways and functions (Brockhausen *et al.*, 2009; Hart and Akimoto, 2009; Stanley *et al.*, 2009). N-glycosylations are important in protein folding and quality control (Helenius and Aebi, 2001) but has also fundamental biological roles in cell signaling, adhesion and migration. O-linked glycans are most commonly found as mucin glycoproteins both in membrane bound and soluble secretory form (Hattrup and Gendler, 2008). These proteins are heavily glycosylated at multiple sites and contribute in hydrating and protecting the underlying epithelium from bacteria, viruses and other particles. Another major class of O-linked glycoproteins consists of proteoglycans which are also heavily glycosylated and contain extremely long sulfated sugar chains (Esko *et al.*, 2009). Their glycan moiety is characterized by linearly arranged repeated disaccharide units which are named after their histological origin. Examples include hyaluronic acid, chondroitin sulfate, dermatan sulfate and keratan sulfate (Taylor and Drickamer, 2011). Some proteoglycans are typically found both on membranes and in the extracellular matrix and are characterized by extremely compact molecular arrangements that help resist compression under pressure, for example in case of cartilage. Defects in the biosynthesis of N- and O- glycans have been assembled in the group of inherited diseases termed *congenital disorders of glycosylation* which appear with severe but varying clinical symptoms due to the loss of normal glycosylation (Cylwik *et al.*, 2013a; Cylwik *et al.*, 2013b).

Glycosphingolipids (GSLs)

GSLs are based on ceramides, composed of a sphingolipid and a fatty acid core, which together constitute the lipid character of such molecules. The length of glycans in GSLs can vary from one to as much as 60 monosaccharide residues in complex polyglycosylceramides (Miller-Podraza *et al.*, 1993; Podraza *et al.*, 1997). Based on

the chemical composition, more than 300 known species of glycosphingolipids have been characterized (Lingwood, 2011). Almost 90 % of all GSL species are based on glucosyl-ceramide whereas, galactosyl-ceramide serves as precursor for the rest of the GSLs. The major core glycan structures that define the GSL series are lacto-, neo-lacto, globo- and ganglio-series (Hakomori, 2003; Lingwood, 2011). The heterogeneity in GSL ceramides comes from the level of hydroxylation, unsaturation and the length of the fatty acid and sphingosine derivatives. The functional aspects of this heterogeneity are yet unknown but its role in membrane organization and GSL receptor functionality are implicated (Hakomori, 2003). Several studies have established the role of glycosphingolipids as receptors, facilitators of cellular adhesion and signal transduction (Hakomori, 2003). The multitude of biological functions performed by GSLs seem to be crucial for development as mice lacking glucosylceramide synthase die during embryonic development (Yamashita *et al.*, 1999). GSLs have also been shown to be important for intracellular trafficking (Pasquini *et al.*, 1989).

1.2.4. The ABO(H) and Lewis histo-blood group systems

The ABO blood group antigens on red blood cells were discovered in the early 20th century by Karl Landsteiner and colleagues (Landsteiner, 1900). Later, it was realized that these antigens are not specific to erythrocytes but are also found in epithelial linings of gastrointestinal, respiratory and urogenital tracts along with biological secretions such as saliva. Therefore these terminal sugar epitopes, which are produced as result of tissue specific glycosylation of glycolipids and glycoproteins (Clausen and Hakomori, 1989) were termed as *histo*-blood group antigens (HBGAs). The structure of these antigens is determined by the expression of genes that encode different glycosyltransferases. Subsequently, any individual would present HBGAs depending upon the total activities of corresponding glycosyltransferases.

Biosynthesis

The HBGAs are synthesized by the *sequential* action of sets of different glycosyltransferases corresponding to the respective blood group systems (Lowe, 1993; Stanley and Cummings, 2009). Both ABH and the closely related Lewis HBGAs are terminal carbohydrate structures present on the sugar chains of oligosaccharides, glycolipids and glycoproteins. The most common pre-cursors which are glycosylated to form HBGAs are type 1 (Galβ1,3GlcNAcβ1-R), type 2 (Galβ1,4GlcNAcβ1-R), type 3 (Galβ1,3GalNAcα1-R) and type 4 (Galβ1,3GalNAcβ1-R) sugar chains. The type 1-3 chains are common for glycolipids and glycoproteins whereas type 4 chains are typically found only in glycolipids. The biosynthesis of ABO HBGAs utilizes all 4 precursors discussed here while Lewis antigens are synthesized only on the type 1 and type 2 chains.

The human ABH antigens

The biosynthesis of ABO blood group system begins with the action of one of two α1,2-fucosyltransferases which adds α1,2-linked fucose to one of the precursor chains. This produces the **H antigens** (Fig. 1.5). There are two active α1,2-fucosyltransferases encoded by human genome at *FUT1* and *FUT2* loci denoted by FucT-I and FucT-II,

respectively (Kelly *et al.*, 1995; Le Pendu *et al.*, 1985; Rajan *et al.*, 1989). The FucT-I is expressed in red cell precursors and has a somewhat strict preference for type 2 and type 4 chain glycans. The FucT-II is mainly expressed in the epithelial lining of gastrointestinal, respiratory and urinary tracts along with salivary glands and catalyzes the addition of α1,2-linked fucose preferentially to type 1 and type 3 chain precursors, although type 2 chain activity has also been reported (Lowe, 1993; Ravn and Dabelsteen, 2000).

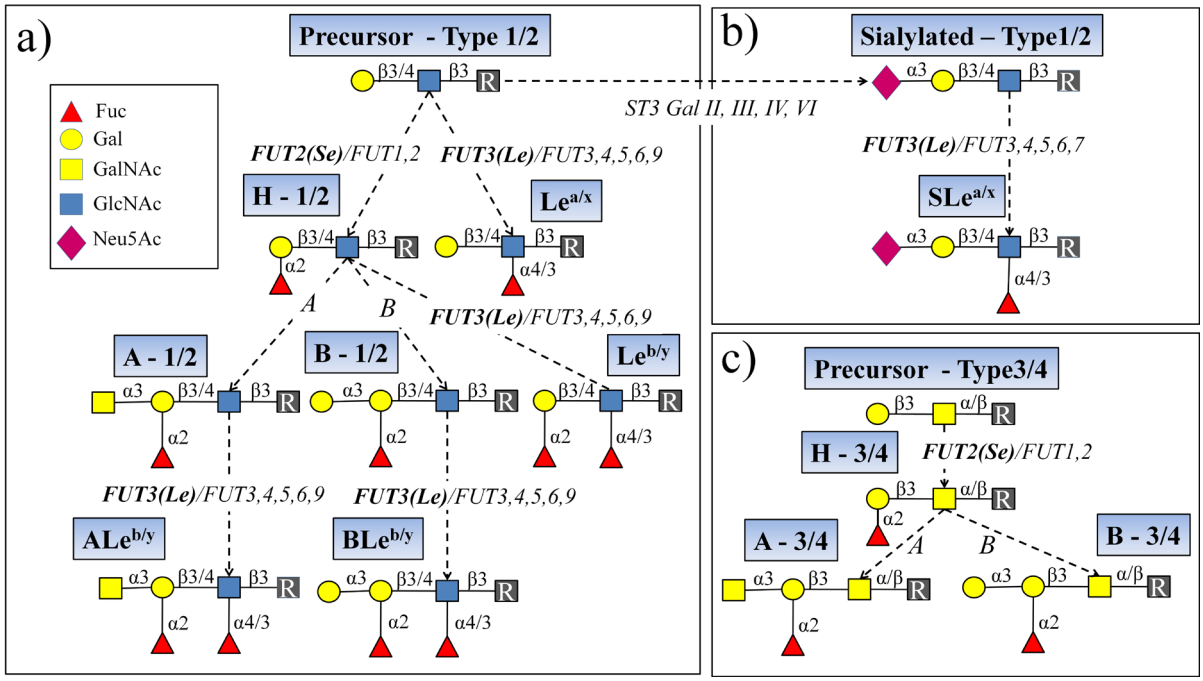

Figure 1.5 Biosynthetic pathways of HBGAs. *The precursors from (a) type1/2 and (c) type 3/4 chain are considered. The antigen names (blue boxes), genetic loci of glycosyltransferases (italic) and the glycosidic linkages (solid lines) are separated by forward slash (/) to specify the pathway for different chains. Biosynthetic pathways for Sialyl Lewis a/x are shown separately in panel b. The FUT2 and the FUT3 genetic loci driving the biosynthesis of the secretor (Se) and the Lewis (Le) antigens, respectively, are also highlighted (bold).*

A and B epitopes are subsequently formed by the action of glycosyltransferases encoded by the *ABO* gene. The A allele at *ABO* locus encodes an α1,3-GalNAc transferase which adds α1,3-linked GalNAc to the Galβ of the H epitopes thereby producing the **A antigens**. Similar action by an α1,3-Gal transferase encoded by the B allele gives rise to the **B antigens** (Fig.1.5). The O allele encodes for an inactivated form of the A or B glycosyltransferase. Blood group A individuals express at least one A allele at the ABO locus (genotype: AA or OA) whereas, those of blood group B express at least one B allele (genotype: BB or BO). Similarly, individuals who are heterozygous for both A and B alleles are classified as blood group AB (genotype: AB). Blood group O individuals are homozygous for inactive O alleles at the ABO locus and consequently express neither A nor B antigens. Both A and B blood group determinants require the type 1, type 2, type 3 or type 4 H antigen precursors as acceptors for the corresponding A and B glycosyltransferases (Watkins, 1980). Therefore, the biosynthesis of A/B epitopes on type 1 chain glycans primarily depend upon the expression of *FUT2* gene whereas, for type 2 chain structures the biosynthesis depends upon the cellular expression of either *FUT1* or *FUT2* genetic loci

(Lowe, 1993; Marionneau *et al.*, 2001; Ravn and Dabelsteen, 2000; Stanley and Cummings, 2009).

Individuals who do not express active *FUT2* gene consequently lack ABH antigens in the epithelial cells or biological secretions such as saliva (Marionneau *et al.*, 2001). Since the soluble or *secreted* form of these antigens cannot be detected in the saliva of these individuals they are denoted as *non-secretors*. On the other hand, individuals with at least one functional form of *FUT2* allele are termed *secretors*. Almost 20% of the Caucasian population comprises of non-secretors (Gaensslen *et al.*, 1987) with both the *FUT2* alleles inactive (Marionneau *et al.*, 2001; Stanley and Cummings, 2009; Watkins, 1980). The most common inactivating mutation for the *FUT2* gene in this population is G428A which codes for a pre-mature stop codon (Kelly *et al.*, 1995). Several other mutations found at *FUT2*, and at other related genetic loci, have been reported from various ethnic populations and are available in the blood group antigen gene mutation database (dbRBC) (Blumenfeld and Patnaik, 2004).

Mutations in *FUT1* are relatively rare but do exist. Since *FUT1* and *FUT2* genes both lie close to each other on chromosome 19, several inactivating mutations in *FUT1* are linked to mutations in *FUT2*. For instance, there is a mutation T725G which was identified in *FUT1* gene together with a deletion mutation at *FUT2* locus. These individuals with homozygous inactivating mutations at both *FUT1* and *FUT2* loci do not express A, B or H antigens. This typical phenotype was termed Bombay, after the place from where it was initially reported (Bhende *et al.*, 1952; Fernandez-Mateos *et al.*, 1998; Koda *et al.*, 1997). The Bombay phenotype is extremely rare with the estimates of one in a million in Europe to one in ten thousand in the genetically relevant Indian population.

In the early postnatal period, IgM antibodies are generated by the immune system against the non-self ABH antigens which probably is the consequence of presentation of similar glycan epitopes by bacteria and fungi (Springer, 1971). For instance, blood group A individuals produce high IgM antibody titers against B antigens but not against A antigens and vice versa for blood group B and O individuals. Therefore, to escape the severe implications of acute transfusion reactions (Mollison, 1952) the blood transfusion procedures must match the blood groups between the donors and acceptors according to "the Law of Landsteiner". The blood transfusion is therefore performed in the direction from O to O, A, B or AB and not from A, B or AB to O. This makes blood group O the universal donor and AB the universal receiver (Stanley and Cummings, 2009).

The Lewis antigens

The Lewis blood group system is named after a family who suffered from blood group incompatibility that aided in identifying these antigens. As noted earlier, the biosynthesis of Lewis antigens is restricted to type 1 and type 2 chain glycans (Fig. 1.5). The Lewis epitope is characterized by the presence of an α1,3 or α1,4-linked fucose onto the GlcNAcβ of type 1 or type 2 chain structures (Lowe, 1993; Stanley and Cummings, 2009). Initially the term Lewis antigens were restricted to only

the type 1 chain structures, but later the term has been used for both type 1 and -2 chain structures.

The addition of the α1,4-linked fucose to the GlcNAcβ in type 1 chain precursor or type 1 H determinant, respectively, gives rise to the Lewis a (Le^a) or Lewis b (Le^b) antigens. The addition of the Lewis epitope to the type 1 blood group determinants is catalyzed by a α1,4-fucosyltransferase encoded by *FUT3* gene. The expression of Le^a and Le^b antigens is more or less restricted to the same epithelial cells that express the α1,2-linked fucose encoded by *FUT2* gene (Oriol, 1990; Oriol *et al.*, 1986; Watkins, 1980). Therefore soluble forms of Lewis antigens are also secreted in mucins and are easily detectable in saliva. The Lewis antigens are also found on red blood cells, not because of the glycosylation in red cell precursors but through the passive adsorption of Lewis positive glycolipids from plasma lipoproteins (Hanfland and Graham, 1981; Hanfland *et al.*, 1986; Marcus and Cass, 1969).

The type 2 chain glycans that belong to Lewis blood group systems include Lewis x (Le^x) and Lewis y (Le^y), which are biosynthesized by the action of α1,3-fucosyltransferases on the type 2 precursor or on the type 2 H determinant, respectively (Fig. 1.5). While there are only two α1,4-fucosyltransferases (coded by the *FUT3* and *FUT5* genes) there are several α1,3-fucosyltransferases (coded by the *FUT3, FUT4, FUT5, FUT6, FUT7* and *FUT9* genes). The gene product of *FUT3* thus preferentially fucosylates type 1 chain glycans but can also be active on type 2 chain structures (Johnson *et al.*, 1992; Johnson and Watkins, 1992; Kukowska-Latallo *et al.*, 1990; Prieels *et al.*, 1981). Likewise, the *FUT5* encoded fucosyltransferase has the preference for type 2 chain acceptor substrates (Mollicone *et al.*, 1994) but may also fucosylate the type 1 chain structures. Fucosyltransferases encoded by *FUT4, FUT6, FUT7* and *FUT9* genes are exclusive for type 2 chain activity (Kumar *et al.*, 1991; Lowe, 1991; Weston *et al.*, 1992a; Weston *et al.*, 1992b) but all have their own unique acceptor specificities. For instance, the *FUT7* gene encoded α1,3-fucosyltransferase accepts only the α2,3-sialylated type 2 precursor [catalyzed by ST3GalIII/IV/VI gene coded enzymes (Okajima *et al.*, 1999)] making Sialyl Lewis x while the *FUT4* gene coded α1,3-fucosyltransferase has a preference for fucosylating an internal GlcNAc in a polylactosamine chain (Natsuka *et al.*, 1994).

It should be noted, however, that the erythrocyte Lewis phenotype is strictly denoted by the action of the *FUT3* gene product and hence Le^a and Le^b are the *true* Lewis antigens, which are synthesized exclusively by this enzyme (Johnson *et al.*, 1992; Johnson and Watkins, 1992; Kukowska-Latallo *et al.*, 1990; Prieels *et al.*, 1981). The individuals who lack the active fucosyltransferase encoded by *FUT3* gene are thus termed 'Lewis negative' and express the phenotype of Le (a-b-) whereas, those with one functional allele at *FUT3* locus are denoted as 'Lewis positive' with the phenotype Le (a+b-) or Le (a-b+) corresponding to non-secretors or secretors, respectively.

Biological roles of ABO and Lewis blood group system

The ABO and Lewis histo-blood group systems have undergone more than 50 million years of primate evolution. The question of why they are still there is challenged by the fact that human individuals with the Bombay phenotype who are devoid of all A, B and H antigens appear completely healthy which demonstrates that the biological roles

of these antigens, if any, are not intrinsic or directly relevant for the single individual. However, a complete block of fucosylation, as found in some genetically deficient rare cases, leads to a severe immune defect called Leukocyte adhesion deficiency type II (i.e. LAD-2, CDG-IIc or SLC35C1-CDG) which proves the importance of fucosylation for the immune response. The role of α1,3-fucosyltransferases in the biosynthesis of Sialyl Lewis x structures and their involvement for the correct homing of immune cells or spread of cancer cells through the interaction with endogenous selectins (a family of mammalian lectins) will not be discussed in this thesis. Instead focus will be put on the role of ABO and Lewis HBGAs in relation to their interactions with pathogens (Marionneau *et al.*, 2001; Stanley and Cummings, 2009).

Polymorphisms at *ABO, FUT1, FUT2* and *FUT3* genetic loci have been largely reported for humans. The manifestations of these mutations produce distinct and diverse antigenic populations which express a selection of different ABH glycans on the surface of red cells (Clausen and Hakomori, 1989; Mollicone *et al.*, 1986; Oriol *et al.*, 1986) and more importantly on the epithelial lining of tissues facing the external environment including those in respiratory tract, urogenital tract and the gut (Marionneau *et al.*, 2001; Ravn and Dabelsteen, 2000; Rouger *et al.*, 1986). Since many pathogens are known to exploit this diversity of glycans on the surface of the cell, it is beneficent for a particular species to express different glycan profiles so that the infection from one pathogen does not wipe out the whole population (Marionneau *et al.*, 2001; Stanley and Cummings, 2009). Over the past few decades, several observations in the support of this theory have been made where it is shown that certain pathogens actually recognize HBGA glycans. Consequently the individuals who do not present these glycan epitopes are somewhat protected from the infection. For instance, the R45 strain of *Escherichia coli* recognizes glycans from type 4 precursor in non-secretor individuals. These glycans are modified by the addition of an α1,2-linked fucose in secretor individuals thereby hiding the binding epitope and consequently protecting them from infection (Stapleton *et al.*, 1992). Similarly, the precursor structures have been shown to be recognized by *Streptococcus pneumonia* and *Salmonella typhimurium* (Barthelson *et al.*, 1998; Giannasca *et al.*, 1996). More recently, *FUT1* and *FUT2* polymorphisms have also been shown to be important in viral infections caused by caliciviruses (Abrantes *et al.*, 2012; Rydell *et al.*, 2011) and rotaviruses (Ramani *et al.*, 2013; Yazgan *et al.*, 2013). Similar examples exist for *FUT3* polymorphisms as well, which are responsible for Lewis positive and Lewis negative phenotypes.

1.2.5. Norovirus and histo-blood group antigens

Over the last few decades, the fact that noroviruses recognize histo-blood group antigens both *in vitro* and *in vivo* has been quite well established. Earlier clues of a genetic cause being involved in resistance or susceptibility of noroviral infection came from the observation that certain individuals were resistant to the infection despite of the absence of norovirus antibodies (Johnson *et al.*, 1990). This resistance is now attributed to the polymorphisms at the genetic loci responsible for HBGA diversity in humans. A few years later, the successful production of the virus-like particles (Jiang *et al.*, 1992) provided the major breakthrough in revealing several binding

characteristics of noroviruses *in vitro*. While challenge and outbreak studies associated HBGA diversity to resistance or susceptibility to noroviruses, the norovirus VLP binding and crystallographic studies revealed the potentially responsible molecular interactions.

Outbreak and challenge studies

Several outbreak and challenge studies have investigated the role of HBGA polymorphisms in norovirus infection and disease. The *secretor status* or *FUT2* polymorphism is the most extensively studied genetic factor in susceptibility or resistance to norovirus infections. The *ABO* and Lewis (*FUT3*) genetic loci have also been associated with the susceptibility to different strains of noroviruses.

FUT2 polymorphism: The secretor status

In 2002, Marionneau and co-workers demonstrated that norovirus VLPs from GI.1 Norwalk strain recognized the H antigen in epithelial cells from the gastroduodenal junction of secretor positive individuals but not from non-secretors. Moreover, only type 1 and type 3 chain H antigens but not type 2 chain structures were recognized, suggesting that the binding was due to *FUT2* gene dependent HBGAs (Marionneau *et al.*, 2002). Later on, challenge studies on the same GI.1 Norwalk strain demonstrated that non-secretors are resistant to norovirus infection (Graham *et al.*, 1994; Hutson *et al.*, 2005; Lindesmith *et al.*, 2003). In one of these studies (Hutson *et al.*, 2005), the authors genotyped archived sera of a challenge study performed earlier (Graham *et al.*, 1994). Overall, Lindesmith *et al.*, reported 34 out of 55 secretors to be infected with Norwalk virus whereas, Hutson *et al.*, found evidence of infection in 42 out of 43 secretor positive individuals. On the other hand, none of 30 non-secretors were found to be infected, collectively in the two studies. Together these studies were able to demonstrate clearly that the secretor status, derived by *FUT2* polymorphism, is a susceptibility factor in Norwalk virus infection and that non-secretors are protected from the disease.

The first challenge study to investigate the effect of secretor status on susceptibility to clinically dominant GII.4 norovirus infection was recently reported (Frenck *et al.*, 2012). Among 40 prospectively randomized subjects (23 secretors, 17 non-secretors) the norovirus infection was defined by either the presence of the virus in stools or seroconversion to GII.4 antibody or both. The subjects were required to have low levels of norovirus antibody to avoid the pre-existing serological effects on infection. The virus shedding and seroconversion to GII.4 antibody was observed in 10/23 secretors and in 0/17 non-secretors. However, one non-secretor individual shed the virus in one stool sample on day 2 post infection but did not develop illness or serological response. Another non-secretor developed illness but it could not be associated to norovirus due to the absence of GII.4 antibody response or virus in stools (Frenck *et al.*, 2012).

Several outbreak studies have also addressed the subject of *FUT2* polymorphism as susceptibility factor for norovirus infections. One such study was reported from Sweden in 2005 (Thorven *et al.*, 2005). The study included 53 symptomatic and 62 asymptomatic individuals from three nosocomial and three community outbreaks of

noroviruses from GI.6, GII.6 and GII.4 strains. The secretor status was determined by genotyping for polymorphisms of *FUT2* genetic locus on nucleotide 385, 428 and 571 of which mutations were observed only at position 428, characteristic of 20% non-secretors in European and American populations. The authors found that all symptomatic individuals were secretor positive while none of the secretor negative individuals (n=18) showed any gastroenteritis symptoms (Thorven *et al.*, 2005). Following the same trend, studies analyzing GII.4 norovirus outbreaks in China and Denmark found all symptomatic individuals to be secretor positive and none to be secretor negative (Kindberg *et al.*, 2007; Tan *et al.*, 2008a).

Additional studies have demonstrated that the evidence of secretor status to be the *absolute* susceptibility factor for noroviruses in general is diluted by the fact that non-secretors were actually found to be infected in several reports in a strain specific manner. For instance, a very recent study on acute gastroenteritis outbreaks in China during 2009 caused by GII.4 and GII.3 noroviruses reported 6 non-secretor individuals to be infected, based on virus identification in stools through RT-PCR (Jin *et al.*, 2013). Similarly, a retrospective study reported 4/7 non-secretors to be infected with GI.3 norovirus in a waterborne outbreak during 2002 (Rockx *et al.*, 2005c). Another study reported an outbreak in Jönköping, Sweden during October 2007 also caused by GI.3 norovirus strain. Interestingly, 7/15 non-secretor individuals were found to be infected in contrast to 26/68 secretors who showed norovirus gastroenteritis symptoms (Nordgren *et al.*, 2010). The authors could not statistically establish secretor status as a susceptibility factor for this particular noroviral strain. Although the dataset was limited, another study failed to correlate secretor status with susceptibility to GII.2 snow mountain norovirus strain (Lindesmith *et al.*, 2005). Likewise, a recent cross-sectional study conducted among children hospitalized with gastroenteritis in northern Vietnam found 5/28 symptomatically infected children to be non-secretors (Trang *et al.*, 2014). A recent study from Burkina Faso in Africa, reported two cases of GII.4 and GII.7 noroviral infection in non-secretor children suffering from diarrhea (Nordgren *et al.*, 2013). A GII.4 norovirus infection was also identified in a non-secretor individual during an outbreak in Valencia, Spain (Carlsson *et al.*, 2009). Although in most of the outbreak studies summarized above the frequency of noroviral infection in secretor positive individuals is considerably high, the resistance to noroviral infection in non-secretors is not absolute but rather relative. This fact was further established by an IgG antibody prevalence study that was conducted on plasma samples of 105 Swedish blood donors which were previously genotyped for secretor status based on *FUT2* polymorphisms. The study revealed that non-secretors do have antibody titers against GII.4 noroviruses, albeit at significantly lower levels when compared to secretor positive individuals suggesting that non-secretors are relatively but not absolutely protected from GII.4 norovirus infections (Larsson *et al.*, 2006).

Polymorphisms at ABO and FUT3 genetic loci: diversity in ABH and Lewis antigens

Several studies have addressed the influence of polymorphisms at *ABO* and *FUT3* genetic loci on the susceptibility or resistance to different norovirus strains. For instance studies have demonstrated that blood group O individuals are significantly more likely to be infected with GI noroviruses while blood group B individuals are

underrepresented (Bucardo *et al.*, 2009; Hennessy *et al.*, 2003; Hutson *et al.*, 2002; Lindesmith *et al.*, 2003). The inverse has been reported for GII noroviruses where blood group A or B individuals are reported to be at higher risk than blood group O individuals (Jin *et al.*, 2013; Tan *et al.*, 2008a). Similarly, the Lewis status has also been investigated for susceptibility to noroviruses but any associations are not clearly established (Bucardo *et al.*, 2009; Carlsson *et al.*, 2009; Tan *et al.*, 2008a).

The associations described above between HBGA polymorphisms and susceptibility to noroviruses are dependent on the genogroup, and in some cases on the genotype. If a certain group of individuals is susceptible to a particular norovirus strain, that group might very well be resistant to the infections from other norovirus strains. The ignorance of this fact in interpreting the data has resulted in conflicting reports and disputed conclusions (Chan *et al.*, 2008; Halperin *et al.*, 2008; Hutson *et al.*, 2004b; Meyer *et al.*, 2004; St Clair and Patel, 2008; Tan and Jiang, 2008). The basis of this diversity and the way norovirus exploits this variation to evade the host immune response and to infect potentially all HBGA polymorphic populations is further discussed in the following sections.

The binding studies

A number of binding methodologies including both solution based and membrane based assays have been employed to study norovirus HBGA interactions *in vitro*. In the following text some of the important findings are highlighted which, in some cases, explain the associations of polymorphic HBGAs to noroviral infection at a molecular level. The binding trends of different human noroviral strains to HBGAs are summarized in Table 1.1 and Table 1.2.

After the principle observation of Norwalk virus binding to gastroduodenal epithelial cells containing H antigen carried on type 1 but not on type 2 chains (Marionneau *et al.*, 2002), several studies were performed to characterize the binding of GI.1 Norwalk and other norovirus strains to HBGAs using saliva and neoglycoproteins (Harrington *et al.*, 2002; Huang *et al.*, 2003; Huang *et al.*, 2005; Marionneau *et al.*, 2005). These studies used ELISAs with immobilized saliva or neoglycorproteins to study VLP HBGA interactions. In one of the studies VLPs were also immobilized and biotinylated polyacrylamide conjugated HBGAs were used in solution (Harrington *et al.*, 2002). Together these reports were able to establish at least 7 strain specific binding patterns of 14 different norovirus strains to HBGAs. The norovirus strains showing similar binding properties are clustered in the same region of the phylogenetic tree. The secretor gene (*FUT2*) dependent HBGAs or saliva from the secretor positive individuals was recognized exclusively for most of the norovirus strains which was in line with the associations derived from outbreak and challenge studies (see previous subsection). The GI.1 noroviruses strongly bound to saliva from blood group O and A individuals while blood group B saliva was poorly recognized in agreement with the earlier study where red blood cells bearing A and H antigen but not B antigens were haemagglutinated using Norwalk VLPs (Hutson *et al.*, 2003). Similarly, some of the GII norovirus strains (GII.5 MOH, GII.3 Paris Island) recognized saliva from blood group A and B individuals but not from blood group O individuals.

Table 1.1 HBGA and Saliva binding patterns of norovirus VLPs.

Norovirus Strains	Saliva				Synthetic HBGAs / Glycosphingolipids											Binding Assay[Reference]
	Se			se	H			A		B		seLe			Le		
	A	B	O		1	2	3	1	2	1	2	Le^a	Le^x	SLe^x	Le^b	Le^y	
Norwalk (GI.1)	+	-	+	-	+	+	+	+	+	-	/	-	-	-	+	+	ELISA[1,2,14,17,18,20], Crystallography[5,9], CBA[19], QCMD[21]
Aichi124(GI.1)	+	+	+	-	+	+	+	+	+	-	-	-	/	/	+	/	ELISA[24], SPR[24]
West Chester (GI.1)	+	-	+	/	+	/	-	+	/	-	/	-	/	/	-	/	ELISA[25]
C59 (GI.2)	+	-	-	-	/	/	/	/	/	/	/	/	/	/	/	/	ELISA[2]
FUV258 (GI.2)	+	+	+	+	+	-	+	+	+	-	-	+	/	/	+	/	Crystallography[15], ELISA[24], SPR[24]
Honolulu (GI.3)	+	-	+	/	-	/	-	+	/	-	/	+	/	/	-	/	ELISA[25]
Kashiwa645(GI.3)	+	+	+	+	-	+	-	+	+	-	-	+			-		ELISA[24], SPR[24]
DSV (GI.3)	-	-	-	-	/	/	/	-	/	-	/	/	/	/	-	/	ELISA[2]
VA115 (GI.3)	-	-	-	-	/	/	/	-	/	-	/	/	/	/	-	/	ELISA[2]
Chiba407(GI.4)	+	+	+	+	-	-	-	+	+	-	-	+	/	/	+	/	ELISA[24], SPR[24]
TCH-060 (GI.7)	/	/	/	/	+	+	/	+	+	/	/	+	+	/	/	+	ELISA[22], Crystallography[22]
Boxer (GI.8)	+	-	+	+	-	-	-	-		-	/	-	-	/	+	+	ELISA[2]
WUG1(GI.8)	+	+	+	+	-	-	-	+	+	+	+	+	/	/	+	/	ELISA[24], SPR[24]
Hawai (GII.1)	-	-	-	-	-	-	-	+	/	+	/	-	+	/	+	-	ELISA[2,4,14]
Noda485 (GII.1)	-	-	-	-	-	-	-	-	/	-	/	-	/	/	-	/	ELISA[24]
SMV (GII.2)	-	+	-	-	/	/	/	/	/	/	/	/	/	/	/	/	ELISA[14]
BUDS (GII.2)	+	-	-	-	-	-	-	+	/	+	/	-	-	/	-	-	ELISA[2]
Matsudo18 (GII.3)	+	+	-	-	-	-	+	-	/	-	/	-	/	/	-	/	ELISA[24]
Kashiwa336 (GII.3)	+	+	-	-	-	-	+	-	/	-	/	-	/	/	-	/	ELISA[24]
Mexico (GII.3)	+	-	-	-	-	-	-	+	/	+	/	-	-	/	+	-	ELISA[2]
Paris Island (GII.3)	+	+	-	-	-	-	-	+	/	+	/	-	-	/	+	-	ELISA[2]
Toronto (GII.3)	/	/	/	/	-	-	+	+	/	-	/	-	-	/	-	-	ELISA[6]
Chron1 (GII.3)	+	+	+	-	+	-	/	-	/	-	/	-	-	+	+	-	ELISA[20]
Ichikawa754 (GII.5)	+	+	-	-	-	-	-	+	/	+	/	-	/	/	-	/	ELISA[24]
MOH (GII.5)	+	+	-	-	-	-	-	+	/	+	/	-	-	/	-	-	ELISA[1,2]
Ueno7k (GII.6)	+	+	+	+	-	+	+	-	/	+	/	+	/	/	+	/	ELISA[24]
Sanbu445 (GII.6)	+	+	+	-	-	-	+	-	/	-	/	-	/	/	-	/	ELISA[24]
Osaka10-25 (GII.7)	+	+	+	+	-	-	+	-	/	+	/	+	/	/	+	/	ELISA[24]
VA207 (GII.9)	+	-	+	+	-	-	+	-	/	-	/	-	+	+	-	+	ELISA[1,2,8], Crystallography[8]
Vietnam026 (GII.10)	/	/	/	/	/	+	/	+	/	+	/	-	-	/	+	+	Crystallography[12], STD NMR[13]
Wortley (GII.12)	-	-	-	/	+	/	+	+	/	+	/	-	/	/	+	/	ELISA[25]
Chitta (GII.12)	-	+	-	-	-	-	-	-	/	-	/	-	/	/	-	/	ELISA[24]
Hiro (GII.12)	/	/	/	/	/	/	/	/	/	+	/	/	/	/	/	/	Crystallography[12]
OIF (GII.13)	-	-	+	+	-	-	-	-	/	-	/	+	-	/	-	-	ELISA[2]
Kashiwa47 (GII.14)	-	-	-	-	-	-	-	-	/	-	/	-	/	/	-	/	ELISA[24]

P particles/dimers are included only in case of crystallographic studies. The HBGAs from type 1, 2 or 3 pre cursor sugar chains are indicated by numbers 1,2 and 3 in column headers. The grey boxes indicate GSL binding. In case of contradiction among different studies, positive results are selected. Signs '+' and '–' indicate binding and no binding, respectively and '/' means "not determined". Se = Secretor, se = non-secretor, seLe = Lewis positive non-secretor, Le = Lewis positive. The references: 1 - (Huang et al., 2003), 2 - (Huang et al., 2005), 3 - (Bally et al., 2011), 4 - (Bok et al., 2009), 5 - (Bu et al., 2008), 6 - (Cannon et al., 2009), 7 - (Cao et al., 2007), 8 - (Chen et al., 2011), 9 - (Choi et al., 2008), 10 - (de Rougemont et al., 2011), 11 - (Fiege et al., 2012), 12 - (Hansman et al., 2011), 13 - (Hansman et al., 2012), 14 - (Harrington et al., 2002), 15 - (Kubota et al., 2012), 16 - (Lindesmith et al., 2008), 17 - (Lindesmith et al., 2003), 18 - (Marionneau et al., 2005), 19 - (Nilsson et al., 2009), 20 - (Rydell et al., 2009b), 21 - (Rydell et al., 2009a), 22 - (Shanker et al., 2014), 23 - (Shanker et al., 2011), 24 - (Shirato et al., 2008), 25 - (Uusi-Kerttula et al., 2014)

However, the most "generic" binding pattern was observed for clinically dominant GII.4 VLPs (VA387 and Grimsby) which demonstrated binding to all secretor gene dependent HBGAs. Lewis binding pattern was also observed for GII.9 (VA207) noroviruses which showed binding to all Lewis positive HBGAs regardless of secretor status. Interestingly, similar trends of secretor gene independent binding were later observed for GII.4 noroviruses as well.

Lindesmith *et al.*, analyzed the evolutionary relationships between GII.4 norovirus strains over the period of 1987 to 2005 (Lindesmith *et al.*, 2008). The authors were able to express the capsid region of GII.4 strains from 5 representative evolutionary clusters and assay them with saliva from secretors and non-secretors. GII.4 norovirus strain 2002a was found to recognize non-secretor saliva and different binding properties were established for the same GII.4 genocluster over the period of 20 years. On a similar note, de Rougemont and co-workers also studied HBGA interactions of 6 different GII.4 norovirus strains from 1987-2007 (de Rougemont *et al.*, 2011) and found different binding patterns for the GII.4 norovirus strains separated over time. These studies suggested that GII.4 noroviruses have undergone so called "epochal evolution" which has allowed them to alter amino acid composition of the binding site in order to escape host immunity under the selection pressure from different polymorphic HBGAs. This subject is further discussed later in the current section.

Binding patterns other than ABH antigens have also been reported. For instance, GII.3 and GII.4 noroviruses recognizing Sialyl Lewis x (SLe^x) was reported by Rydell and colleagues (Rydell *et al.*, 2009b). The VLP ELISA was used for GII.3 Chron1, GII.4 Dijon and GI.1 Norwalk norovirus strains with immobilized HSA-conjugated SLe^x neoglycoproteins. The binding was observed for GII.3 and GII.4 noroviruses but not for GI.1 Norwalk norovirus strain. No interactions could be identified with the structurally related Lewis x or Sialyl Lewis a antigens suggesting the binding to SLe^x to be specific for this antigen. Similarly, three GII strains have been reported to bind heparan sulfate glycosaminoglycan. The binding was reported to be specific to heparan sulfate as pre-treatment of the cells using heparanase resulted in reduced binding (Tamura *et al.*, 2004), while the same was not observed for chondroitinase. Likewise, the principle finding of GII.4 noroviruses binding to galactosylceramide GSL in microdomains is reported in paper V.

Apart from ELISAs to characterize binding of different strains of noroviruses to HBGAs, several other binding techniques have been used to study the VLP HBGA interactions. Shirato *et al.*, used surface plasmon resonance to characterize binding of immobilized glycans to noroviruses (Shirato *et al.*, 2008). The authors used 5 GI and 11 GII VLPs to investigate the difference in binding of these VLPs to type 1 and type 2 chain glycans. Type 2 chain structures demonstrated faster detachment than type 1 chain HBGAs suggesting that type 1 chain glycans bind noroviruses more tightly than type 2 chain structures. This aspect of *tightness*, or binding strength, is discussed in detail in paper IV. The study from de Rougemont *et al.*, discussed above also used SPR to study HBGA binding to different GII.4 norovirus strains (de Rougemont *et al.*, 2011). The authors found that post 2002 GII.4 norovirus strains (Hunter, Den Haag, Osaka, Yerseke) bind A and B antigen much stronger than older GII.4 strains. More

recently, STD NMR have also been used to study GII.4 norovirus interactions with synthetic oligosaccharides in solution (Fiege *et al.*, 2012).

Table 1.2 HBGA and Saliva binding patterns of norovirus GII.4 VLPs.

Norovirus Strains	Saliva				Synthetic HBGAs / Glycosphingolipids											Binding Assay[Reference]
	Se			se	H			A		B		seLe			Le		
	A	B	O		1	2	3	1	2	1	2	Le^a	Le^x	SLe^x	Le^b	Le^y	
Sydney 2012 (GII.4)	+	+	+	/	+	/	+	+	/	+	/	+	/	/	+	/	ELISA[25]
New Orleans (GII.4)	+	+	+	/	+	/	+	+	/	+	/	-	/	/	+	/	ELISA[25]
Minerva (GII.4)	/	/	/	/	-	-	+	+	/	+	/	-	+	/	+	+	ELISA[6]
TCH05 (GII.4)	/	/	/	/	+	/	/	+	/	/	/	/	/	/	+	/	Crystallography[23]
Narita104 (GII.4)	+	+	+	+	+	+	+	+	+	+	+	-	/	/	+	/	ELISA[24], SPR[24]
VA387 (GII.4)	+	+	+	-	+	-	+	+	/	+	/	-	-	/	+	+	ELISA[1,2], Crystallography[7]
004-95M (GII.4)	+	+	+	/	+	/	+	+	/	+	/	+	/	/	+	/	ELISA[25]
Ast6139 (GII.4)	/	/	/	/	+	+	/	+	/	+	/	+	+	+	+	+	**STD NMR[11]**
Bristol (GII.4)	+	+	+	-	+	-	+	+	/	+	/	-	-	+	+	+	ELISA[10], SPR[10]
Hunter (GII.4)	+	+	+	-	+	-	+	+	/	+	/	-	-	+	+	+	ELISA[10], SPR[10]
Yerseke (GII.4)	+	+	+	-	+	-	+	+	/	+	/	-	-	-	+	+	ELISA[10], SPR[10]
Den Haag (GII.4)	+	+	+	+	+	-	+	+	/	+	/	-	+	+	+	+	ELISA[10], SPR[10]
Osaka (GII.4)	+	+	+	+	+	-	+	+	/	+	/	-	+	+	+	+	ELISA[10], SPR[10]
FH2002 (GII.4)	+	+	+	-	-	-	-	-	/	-	/	/	-	-	-	-	ELISA[16]
FH2002a (GII.4)	+	+	+	+	-	-	-	+	/	-	/	/	+	+	-	-	ELISA[16]
Dijon (GII.4)	+	+	+	-	+	-	+	+	/	+	/	-	+	+	+	+	**TIRF[3], ELISA[10,20], SPR[10], QCMD[21]**
Grimsby (GII.4)	+	+	+	-	-	-	+	+	/	+	/	/	-	-	+	+	ELISA[2,16]
Camberwell (GII.4)	+	+	+	-	-	-	+	-	/	-	/	/	-	-	-	+	ELISA[16]
CHDC5191 (GII.4)	/	/	/	/	+	-	+	-	/	+	/	-	+	/	+	+	ELISA[4]
CHDC4871 (GII.4)	/	/	/	/	-	-	+	-	/	+	/	-	-	/	-	+	ELISA[4]

The symbols and referencing are same as in Table 1.1

Studies on norovirus VLP HBGA interactions have also been carried out using glycosphingolipids (Bally *et al.*, 2012; Bally *et al.*, 2011; Nilsson *et al.*, 2009; Rydell *et al.*, 2009a). Nilsson *et al.*, used GSLs for the first time to study radiolabelled Norwalk VLP binding to type 1 and type 2 chain glycans using a chromatogram binding assay. The authors characterized specific binding to A and H antigens and their extended structures but no binding was observed for B antigens and their extensions (Nilsson *et al.*, 2009) which was in agreement with the previous findings (Hutson *et al.*, 2003). Rydell *et al.*, used Quartz Crystal Microbalance with dissipation monitoring (QCM-D) to study binding of Norwalk GI.1 and Dijon GII.4 norovirus strains to H type 1 and Lewis a GSLs embedded in a planar supported lipid bilayer (Rydell *et al.*, 2009a).

The VLPs of both strains bound H type 1 but no binding to Lewis a was observed. The authors found that a threshold of 2 % (weight) GSL was needed to detect binding of H type 1 to Dijon strain while in case of Norwalk strain 0.25 % (weight) GSL was enough to measure the binding suggesting that Norwalk virus bound H type 1 GSL with higher affinity as compared to the Dijon strain. A more sensitive binding assay to characterize interactions of multivalent proteins with membrane bound ligands was

developed by Bally and colleagues (Bally *et al.*, 2011). The assay was based on total internal reflection fluorescent microscopy (TIRFM), which was used to selectively illuminate the interactions in the vicinity of the model membrane. First, the unlabeled VLPs were immobilized on a planar supported lipid bilayer using GSLs already embedded in the bilayer. The bound fraction was then detected by fluorescently labelled vesicles containing GSLs. Using TIRFM, the authors could actually record the attachment and detachment events of single vesicles to VLPs in real time. Both transient and steady state quantitative binding data could be obtained using conventional data analysis. GII.4 norovirus VLPs were used as model protein and the binding to H type 1 and Lewis b GSLs was studied. The details of underlying binding kinetics of a series of type 1 and type 2 GSLs using this TIRFM technique are discussed in paper IV.

Although the binding studies summarized above are based on VLPs and not the infectious virus particles, the credibility of these studies is established (i) through the near perfect correlation between the infection patterns observed in outbreaks or challenge studies and the VLP binding studies (Lindesmith *et al.*, 2003; Nilsson *et al.*, 2009), and (ii) the reasonable agreement with the binding patterns observed through a handful of binding studies conducted on native virions obtained through purification of stool samples (Harrington *et al.*, 2004; Thorven *et al.*, 2005).

The crystallographic studies

Since the structural determination of the Norwalk capsid with crystallography in 1999 (Prasad *et al.*, 1999), several norovirus P domain structures in complex with the bound HBGAs have been solved, which include three GI (Bu *et al.*, 2008; Choi *et al.*, 2008; Kubota *et al.*, 2012; Shanker *et al.*, 2014) and four GII norovirus strains (Cao *et al.*, 2007; Hansman *et al.*, 2011; Shanker *et al.*, 2011). The general observation from these studies is that the overall structural fold of P domain dimers is similar in both genogroups and that the binding sites for GI and GII noroviruses are different in a genogroup dependent manner. Additionally, subtle differences in the binding site residues have resulted in altered specificities towards the binding of HBGAs within the same genotype.

The norovirus GII binding site

Earlier computational studies together with mutagenesis data indicated the presence of a potential binding pocket on the surface of the highly variable P2 domain (Chakravarty *et al.*, 2005; Tan *et al.*, 2003). The exact location of the binding site was however, not revealed until the first norovirus P domain structure of GII.4 VA387 strain was solved in complex with HBGA A and B saccharides (Cao *et al.*, 2007). The crystal structure at 2.2 Å resolution demonstrated that the binding site is located at the dimer interface and α1,2-linked fucose, representative of secretor phenotype, was observed to be involved in direct hydrogen bonding and hydrophobic interactions. This also explained on the level of atomic resolution the basis for the preference of GII.4 noroviruses to bind secretor gene dependent HBGAs. However, it was not clear how the extensions of ABH antigens would be recognized by the P domain and which residues would be important for such binding modes.

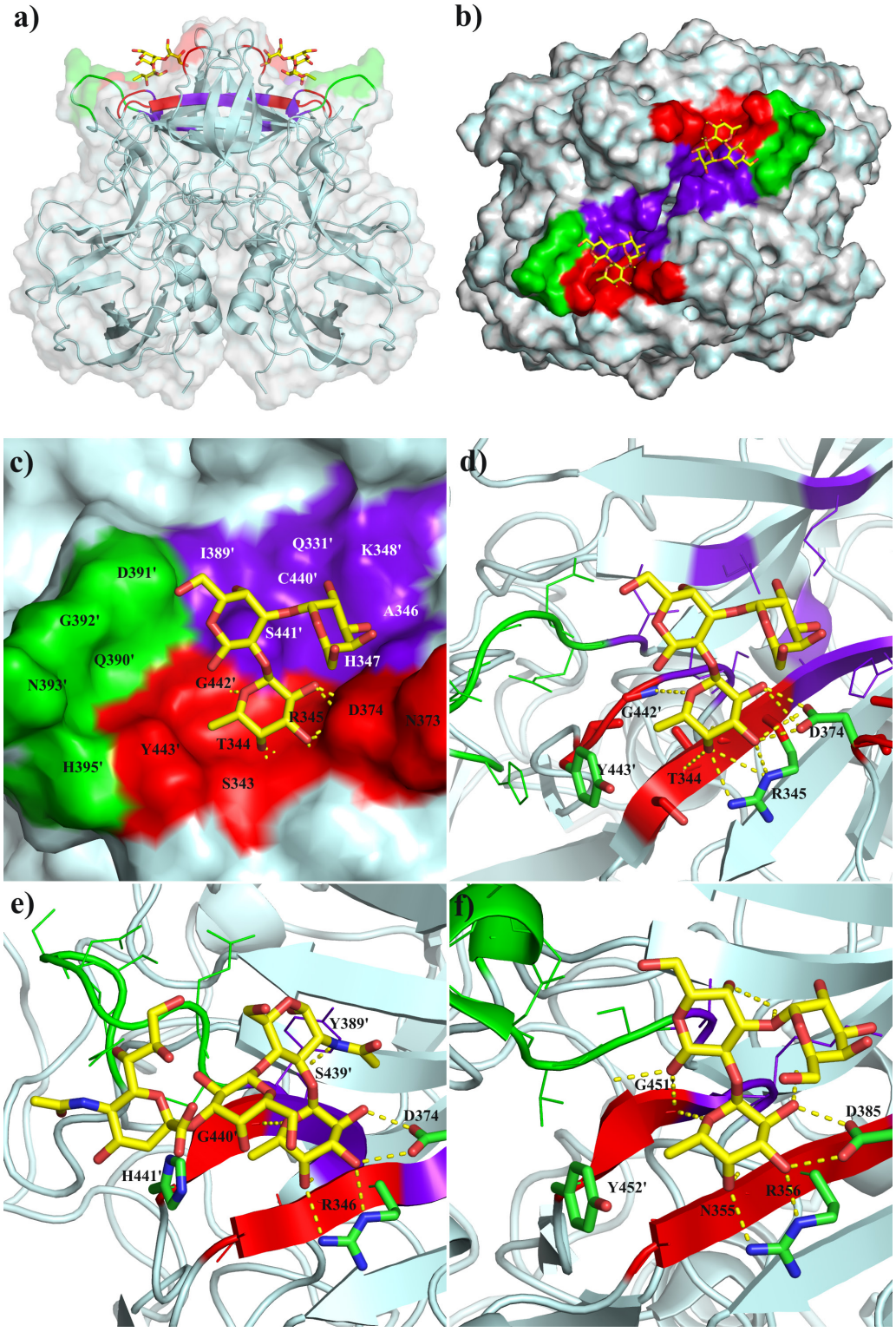

Figure 1.6 The GII norovirus binding site. *The P dimer of VA387 GII.4 norovirus strain is shown as (a) ribbons (with partially transparent surface) and (b) solid surface in complex with B-trisaccharide (Cao et al., 2007). The zoomed in views of binding sites of norovirus VA387 in complex with B-trisaccharide (c, d), GII.9 VA207 (Chen et al., 2011) in complex with Sialyl Lewis x (e) and GII.10 Vietnam (Hansman et al., 2011) in complex with B-trisaccharide (f) are illustrated as studied through crystallography. Panel (c) shows the surface view of the VA387 binding site with the labeled residues identified to be influential for HBGA binding (Tan et al., 2008b). For all figures, the fucose binding pocket is colored red, the flexible loop region is colored green, whereas the adjacent binding pocket for the ABH extensions is colored purple. The interacting fucose binding site residues are shown as sticks and hydrogen bonds are indicated by yellow dashed lines. Note the conservation of R345, D374 (VA387 numbering) and an aromatic residue (Y or H) in the fucose binding pocket which together define the key fucose norovirus interactions. Also, clearly visible is the flexibility in the loop region (green) which defines the specificity of fucosylated HBGAs interacting with GII noroviruses (see the main text).*

The following year, the same group published a mutagenesis study in which some residues in or around the binding site were shown to be essential for binding (Tan *et al.*, 2008b). More importantly, mutations in some residues around the principal fucose binding site resulted in loss of binding to A antigens but not to B or H antigens. This provided clear insights into the regions that were important for the specificity of this particular strain to bind HBGAs. The potential binding modes for the extensions of ABH antigens in complex with VA387 P dimers, for some of which there is still no available crystallographic data, are discussed in Paper I.

Other GII norovirus strains which were studied through crystallography showed exactly the same binding pose for secretor gene dependent HBGAs with the α1,2-linked fucose exhibiting most of the direct interactions with binding site residues. The analysis of the crystal structure of another GII.4 variant from 2004 [TCH05] (Shanker *et al.*, 2011), suggested that, as compared to the 1996 variant, the amino acids in contact with the Lewis fucose might improve the binding to the Lewis b HBGAs. Furthermore, the electrostatic landscape of this 2004 variant was also found to be substantially different from that of the 1996 variant suggesting that the variation in GII.4 noroviruses over time not only affects the binding affinity and specificity but also the manner in which the whole capsid protein is recognized at the cell surface. Some of the residues responsible for the altered electrostatics of P domain dimers were previously identified as evolutionary hotspots for GII.4 noroviruses (Donaldson *et al.*, 2010). This further suggested that these amino acid changes could also have implications in antigenic variation of GII.4 norovirus strains. Following the same trend as GII.4 noroviruses, GII.10 and GII.12 noroviruses have also been demonstrated to recognize α1,2-linked fucose in the principal fucose binding site through recent crystallographic studies (Hansman *et al.*, 2011).

The norovirus GI.1 binding site

In contrast to GII noroviruses, the binding site in GI norovirus strains is not highly specific for any single sugar residue (Bu *et al.*, 2008; Choi *et al.*, 2008; Kubota *et al.*, 2012; Shanker *et al.*, 2014). Furthermore, the location of this binding site is typically found on one of the P-monomers, distinctly different from the GII noroviruses, where the binding site is observed at the dimer interface. The difference in the location and structural characteristics of GI and GII norovirus binding sites explains their different binding specificities. For GII noroviruses, which bind to all ABH antigens, the central sugar residue defining this specificity at a molecular level is α-L-fucose. All the HBGAs, which have been shown to bind GII norovirus strains, consistently fit into a binding pose specific to this sugar moiety. L-fucose has been observed in all structural studies to be involved in strong hydrogen bonding and hydrophobic interactions. However, for GI noroviruses this is not the case. GI noroviruses do not bind to blood group B glycans but only to blood group A and O HBGAs. It has been demonstrated that only two hydroxyls from the Galβ residue of the precursor sugar chain are involved in hydrogen bonding with the binding site residues. The crystallographic data further elucidated that for H antigens, this binding of Galβ residue positions the hydrophobic moieties of α1,2-linked fucose in favorable orientation to exhibit hydrophobic interactions with tryptophan in the binding site. Similarly, when binding A HBGA, these hydrophobic interactions are instead observed by the methyl group of

GalNAcα (Fig. 1.7). Since for B HBGAs the GalNAcα of the A antigen is replaced by Galα, which lacks the methyl group, no binding is observed for this particular antigen.

The structural basis of recognition of secretor gene-independent Lewis antigens by GI and GII noroviruses has also been demonstrated (Chen *et al.*, 2011; Kubota *et al.*, 2012; Shanker *et al.*, 2014). The general conclusion from these studies was that the strains demonstrating binding to Lewis antigens show variability in the loop regions around the binding site, which results in the different binding specificity, while the main binding site residues remain strikingly similar.

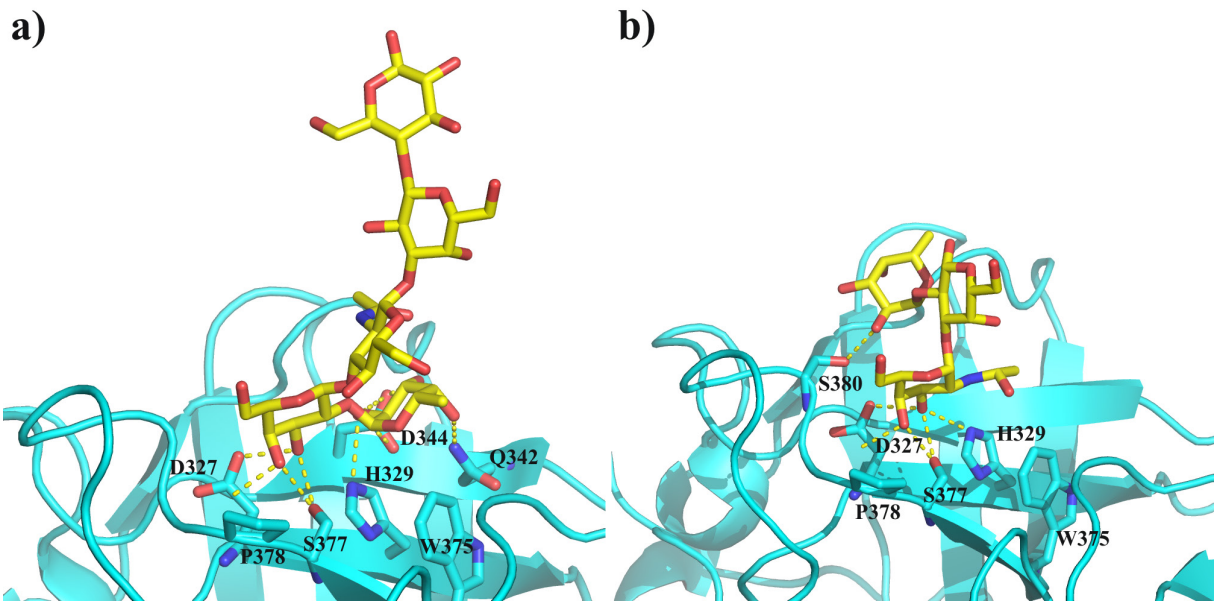

Figure 1.7 The Norwalk (GI.1) binding site. *The binding poses of Norwalk GI.1 norovirus strain in complex with (a) H type 1 and (b) A trisaccharide (Choi et al., 2008) are illustrated with the interacting residues in stick representation. Hydrogen bonds are indicated by yellow dashed lines. Note the hydrogen bonds between 3,4-OH of Galβ and OH of S377 in H type 1 which are replaced by the 3,4-OH of GalNAcα and S377 in case of A trisaccharide. These hydrogen bonds position the subsequent hydrophobic moieties (CH at positions C1 and C2 in fucose of H type 1 and methyl group in the GalNAcα of A trisaccharide) in the vicinity of W375 (~4 Å) for hydrophobic interactions (Choi et al., 2008).*

For instance, the binding mode observed for the secretor gene-independent Lewis antigens in GII noroviruses was identical to the one observed in case of secretor gene dependent HBGAs with the fucose being involved in most of the hydrogen bonding and hydrophobic interactions in the conserved fucose binding site. However, this time the α1,3-linked Lewis fucose exhibited the key interactions instead of the α1,2-linked secretor fucose (Chen *et al.*, 2011). The possibility of recognition of both the secretor and the Lewis fucose residues in HBGAs by the clinically dominant GII.4 noroviruses and its implication on binding strength are further discussed in Paper II.

The norovirus *epochal* evolution within the pandemic GII.4 lineage

The GII.4 noroviruses recognize a wide variety of ABH antigens. As described earlier, the binding mode for these interactions has been consistently based on a fucose binding pocket that has remained reasonably conserved for several decades (Bok *et al.*, 2009). However, sequence variations around this binding pocket and elsewhere in the surface exposed P2 domain are frequently reported and are presumably the manifestations of the observed epochal evolution of GII.4 noroviruses. The phenomenon of epochal evolution refers to mechanism in which the virus undergoes

periods of "epochs" or stasis, where it accumulates several mutations probably in immunocompromised hosts (Debbink *et al.*, 2014), followed by sudden emergence of a new epidemic strain dominating the outbreaks worldwide. This phenomenon has been observed clearly for GII.4 noroviruses where a new pandemic strain has emerged every 2-3 years. The US 1995/96 strain caused the first reported global epidemic outbreak in 1996 followed by Farmington Hills in 2002, Hunter variant in 2004, Minerva or 2006b in 2007-2008, New Orleans in 2009 and the most recent Sydney 2012 at present.

It has been hypothesized that the epochal evolution of GII.4 noroviruses is driven by (i) the antigenic drift in an attempt to escape the host immunity (Debbink *et al.*, 2012a; Debbink *et al.*, 2013; Lindesmith *et al.*, 2013; Lindesmith *et al.*, 2008) and (ii) the relatively lower selection pressure on the residues around the fucose binding pocket in response to highly polymorphic HBGA human populations (Donaldson *et al.*, 2010; Shanker *et al.*, 2011; Tan *et al.*, 2009). Therefore, on one hand the virus manages to escape the herd immunity through changing the composition of antigenic sites and on the other hand the mutations around the binding site give rise to novel binding specificities of GII.4 noroviruses thereby additionally infecting previously unaffected subpopulations and dominating the outbreaks worldwide.

The separate evolutionary analyses of amino acids in S, P1 and P2 domains of GII.4 noroviruses over a period of 20 years has demonstrated that the S domain does not undergo major persistent amino acid changes, whereas the P1 and particularly P2 subdomain show accumulated mutations which establish a linear relationship between the emerging variants meaning that every new variant within the distance of 1-2 years has emerged as a result of mutations in the previous epidemic strain (Lindesmith *et al.*, 2008). The protein residues involved in the direct interactions with fucose have remained largely conserved during this evolution. This is in agreement with the observation that a vast majority of the human population, including secretors and non-secretors, produces fucosylated glycans as HBGAs, which would explain the selection for interacting amino acid residues in the fucose binding site. However, the residues *around* the binding site are relatively less conserved and the accumulated mutations in these regions result in strains with altered binding specificities possibly affecting host susceptibilities (de Rougemont *et al.*, 2011; Lindesmith *et al.*, 2008; Shanker *et al.*, 2011). For instance, the insertion of threonine at position 395 is an epidemiological feature of GII.4 norovirus strains that emerged after the year 2002 (Dingle and Norovirus Infection Control in Oxfordshire Communities, 2004). In one study (de Rougemont *et al.*, 2011), the saliva and synthetic oligosaccharide binding characteristics of different pre-2002 GII.4 norovirus strains were compared with those of post-2002 variants. It was observed that all six GII.4 noroviruses temporally distributed over the period of 20 years recognized saliva from secretor positive individuals. However, two post 2002 variants, Den Haag and Osaka, additionally recognized non-secretor saliva. Moreover, the deletion of T395 resulted in decreased binding to A and B antigens and gained strong binding to Lewis x and Sialyl Lewis x HBGAs (de Rougemont *et al.*, 2011). Finally, it was demonstrated through SPR that post 2002 variants showed higher binding affinity towards A and B antigens as compared to the pre 2002 GII.4 strains (de Rougemont *et al.*, 2011).

The GII.4 noroviruses are reported to be consistently causing pandemic outbreaks worldwide for almost two decades. The most recent GII.4 norovirus epidemic strain, Sydney 2012, has been shown to recognize all ABH antigens along with Lewis a HBGA *in vitro* (Uusi-Kerttula *et al.*, 2014), suggestive of broader target population for this particular norovirus strain. The studies on epochal evolution of GII.4 noroviruses have addressed the aspects of antigenic drift and changes of binding behavior and susceptibility as a consequence of accumulated mutations over time. However, in order to predict the future epidemic norovirus strains, there is globally a need for more effective and complete outbreak surveillance particularly for GII.4 noroviruses.

```
                    331 343       373  389          440/441
CHDC5191/1974       Q...STRAHK...ND...IQDGD-HH...CSGY
CHDC4871/1977       Q...STRAHK...ND...IQDGD-HH...CSGY
Bristol/1987        Q...STRAHK...ND...IQDGD-HH...CSGY
Camberwell/1987     Q...STRAHK...ND...IQDGD-HH...CSGY
Dijon/1996          Q...STRAHK...ND...IQDGN-NH...CSGY
Narita104/1997      Q...STRAHK...ND...IQDGN-NH...CSGY
VA387/1998          Q...STRAHK...ND...IQDGN-NH...CSGY
Grimsby/1998        Q...STRAHK...ND...IQDGN-NH...CSGY
004-95M/1999        Q...STRAHK...ND...IQDGN-NH...CSGY
Ast6139/2001        Q...STRGHK...ND...IQDGN-NH...CSGY
FH2002/2002         Q...STRGHK...ND...IQDGNGTH...CSGY
FH2002a/2004        Q...STRGHK...ND...IQDGNGAH...CSGY
Hunter/2004         Q...STRGHK...ND...IQDGSTTH...CSGY
TCH05/2005          Q...STRGHK...ND...IQDGSTTH...CSGY
Minerva/2006        Q...STRGHK...ND...IQDGSTTH...CSGY
DenHaag/2006        Q...STRGHK...ND...IQDGSTTH...CSGY
Yerseke/2006        Q...STRGHK...ND...IQDGSTTH...CSGY
Osaka/2007          Q...STRGHK...ND...IQDGSTTH...CSGY
New-Orleans/2009    Q...STRGHK...ND...IQDGSTTP...CSGY
Sydney/2012         Q...STRGHK...RD...IQDGGTTH...CSGY
                    *   ***.**   .*   .***.      ****
```

Figure 1.8 Multiple alignments of GII.4 norovirus binding site residues. *The coloring is the same as in Fig. 1.6. The major capsid proteins (VP1) of strains with known binding characteristics (Table 1.2) are considered for the alignments performed by ClustalW2. The binding sites of both Ast6139 and Dijon norovirus strains (bold and underlined; used for in-vitro binding studies in Papers III, IV and V) are similar to both the pre- and post-2002 GII.4 strains including VA387 which has been used in the modeling studies of Paper I and II. The numbering following forward slash applies to post-2002 strains with T/G insertion at position 394 (bold).*

The Ast6139/01/Sp GII.4 norovirus strain

In 2004, a study was conducted to investigate the etiology of acute gastroenteritis among children in Asturias, Spain (Boga *et al.*, 2004). A total of 363 stool samples were analyzed and norovirus was classified as the causative agent for 8% (16 cases) of 179 cases where a known pathogen could be found. The identified Ast6139/01/Sp norovirus strain was classified as GII.4 norovirus based on whole genome phylogenetic analysis where it appeared in the same cluster as the Bristol 1993/UK GII.4 strain.

The multiple alignment of binding site residues of GII.4 norovirus strains with known binding characteristics indicate that Ast6139/01/Sp binding site resembles both pre- and post-2002 norovirus strains (Fig. 1.8). Interestingly, the Ast6139/01/Sp norovirus associated disease symptoms in studied pediatric cases demonstrated the peak during summer time and no case was observed in colder months in contrast to the generally conceived notion of norovirus seasonality (section 1.1.4). The binding characteristics of Ast6139/01/Sp GII.4 norovirus VLPs were more recently reported through STD NMR study (Fiege *et al.*, 2012) demonstrating that this particular GII.4 norovirus strain could afford the binding of all synthetic fucosylated HBGAs. The saturation transfer was observed for mainly the L-fucose moiety of glycans in solution with the VLP. This was in agreement with the observed crystallographic and evolutionary data summarized above, which indicated that fucose is the major binding epitope and that the residues in the fucose binding site are remarkably conserved among GII.4 noroviruses. In case of difucosylated Lewis antigens the saturation transfer was observed for both the secretor and the Lewis fucose residues shedding light on the possibility of alternate binding poses specific to one of these two fucose moieties in the principal fucose binding site. For Lewis b the saturation transfer signal for both fucose residues was comparatively similar. However, for Lewis y the signal for α1,2-linked fucose was stronger than for the α1,3-linked fucose. This possibility of alternate binding poses specific to either the secretor or the Lewis fucose is further explored in detail in Paper II.

1.2.6. Binding strength of protein-carbohydrate interactions

The protein-carbohydrate interactions are characterized by low affinities that generally yield dissociation constants (K_d) in the low mM range (Collins and Paulson, 2004), although exceptions have been reported (Perret *et al.*, 2005). The hallmark of low affinity binding in protein-carbohydrate interactions is apparently the basis of their diverse functional roles in cellular adhesion and signaling processes (see page 13). However, this low affinity, sometimes complemented with non-specificity, may allow for e.g. the viral capsid, to *hover* on the surface of the cell until it reaches the high affinity receptor specific for the viral uptake.

The monovalent protein-carbohydrate interactions

At the heart of *monovalent* (one binding site: one ligand) protein-carbohydrate interactions is the interplay between favorable enthalpy (ΔH) gains (the heat release) and the unfavorable entropic ($T\Delta S$) costs due to conformational restrictions at the atomic level (Dam and Brewer, 2002; Garcia-Hernandez and Hernandez-Arana, 1999). The Gibbs free energy of binding (ΔG) can be represented by the following principle equation:

$$\Delta G = \Delta H - T\Delta S \quad (1)$$

The enthalpy primarily stems from hydrogen bonds, van der Waals and electrostatic interactions between the protein and the carbohydrate residues. Although the intermolecular interactions give rise to the favorable enthalpy gain they also restrict the conformation of the ligand increasing the unfavorable entropic costs. In order to reduce the entropic costs and increase the binding affinity, conformationally

constrained oligosaccharides in complex with protein have been studied. The results demonstrated that the same reduction in the degree of glycan flexibility which decreased the entropic costs also decreased the enthalpy gains for the observed protein-carbohydrate interactions (Navarre *et al.*, 1999). This is called the entropy-enthalpy compensation effect. In severe cases, this effect seems to limit the affinity for protein-carbohydrate interactions.

In comparison to many protein-protein interfaces the protein-carbohydrate interactions are characterized by smaller interaction area which can explain their overall relatively lower affinity. However, this is not the case if the interactions per unit area of the binding interface are analyzed in which case the protein-carbohydrate interactions are actually stronger than the corresponding protein-protein interactions (Garcia-Hernandez *et al.*, 2000). One such analysis of 32 lectin-glycan complexes in comparison to 32 protein homodimers from the Protein Data Bank (PDB) revealed that in case of protein-carbohydrate interactions the intermolecular polar-polar contacts are more frequently observed per unit area of the binding interface. Moreover, the protein-carbohydrate complexes are relatively closely packed due to the shorter hydrogen bond distances in the binding site in comparison to the protein-protein interfaces (Garcia-Hernandez *et al.*, 2000).

The reorganization of solvent or water molecules in the bulk solvent or at the protein-carbohydrate binding interfaces has consequences on the entropy and enthalpy of the system (Lemieux, 1996; Mammen *et al.*, 1998). Water molecules have high tendency to form hydrogen bonds with other water molecules or the polar surfaces in proteins or carbohydrates, whereas the non-polar or hydrophobic regions have lower water densities. The release of water molecules from the carbohydrate binding site of a protein is often a favorable reaction due to increases in the entropy of the system. On the other hand, partial dehydration of protein-ligand interface due to the lack of complementarity in the polar surface of the protein and the ligand may give rise to the overall lower binding affinities (Bryce *et al.*, 2001).

The *multivalency* effect

Several proteins, for e.g. virus capsids, have multiple glycan recognition sites which enable them to bind to many ligand molecules simultaneously. Multivalency for proteins can also be achieved through clustering of the membrane associated proteins on cell surface (Varki *et al.*, 2009b). Similarly, glycan epitopes can also be presented multivalently to the receptor molecules in several ways. For instance, glycans can be branched and the terminal ends could contain multiple copies of the binding epitope or highly glycosylated proteins, for instance mucins in saliva, could allow for multivalent glycan presentation. Glycans in glycosphingolipids could also be multivalently presented in microdomains in the cellular membrane. The ΔG of multivalent binding is basically the sum of the ΔGs of individual monovalent affinities (Mammen *et al.*, 1998). This leaves an exponential effect on the K_d which substantially decreases the dissociation rate of interaction. In this scenario the formation of the first monovalent ligand contact with the receptor could position the other ligands in the vicinity of the binding sites on the multivalent receptor resulting in the formation of the second contact and so on. In a typical case the loss of entropy for binding of the ligand would

be significantly reduced after establishing the first monovalent interaction. The result is a very strong binding of multivalent protein-ligand complex where the participating monovalent interactions stabilize each other. The overall affinity of the multivalent interaction is termed as *avidity*. The avidity of such protein-ligand complex can be 10-1000 folds stronger than the corresponding monovalent affinities (Mammen *et al.*, 1998). One such example of multivalent effect has been reported in the context of norovirus inhibition (Rademacher *et al.*, 2011) and is further discussed in Paper IV.

Epitope crypticity and glycan presentation

The availability of the glycan epitope to the receptor molecule is essential for protein-carbohydrate interactions to take place. Usually two to four sugar residues participate in the protein binding where not only the interacting residues but the internal sugar saccharides can also be detrimental in defining the ligand conformation. In one study, the binding affinity of protein-ligand complex is shown to be influenced by the internal saccharide, not directly involved in binding (Maaheimo *et al.*, 1995). This observation has been supported by an extensive modeling study in which the presence or absence of non-terminal GlcNAc in oligosaccharides was shown to influence the conformation of the whole sugar (Imberty *et al.*, 1995). Similarly, it was also shown that monoclonal anti-blood group antibodies recognize the conformationally dependent micro-epitopes of histo-blood group antigens (Imberty *et al.*, 1996; Mollicone *et al.*, 1996). In paper II, the changes in the conformation of Sialyl Lewis x are shown to influence the estimated binding energy through molecular dynamics simulations. Similarly, in paper IV the preference of GII.4 norovirus VLPs towards type 1 chain sugars is shown while type 2 chain structures are mostly not recognized. This is explained by the conformational change in type 2 chain glycans where the GlcNAc of the type 2 chain precursor is ~180^{o} flipped as compared to type 1 chain structures (section 4.1.1).

The multivalent protein-ligand interactions have very delicate structural requirements and the presentation effects can be highly pronounced in these systems. If the multivalent ligand geometries are not well matched with the receptor binding sites, the enthalpic and entropic effects may diminish binding. For instance, the backbone structures used to synthesize the multivalent carbohydrate ligands might not always present the epitope in the most favorable conformation which might result in lower enthalpic gains and higher entropic costs thereby tipping the entropy-enthalpy balance towards lower avidity. The understanding of these aspects of multivalent interactions is important for the design of the backbone or the carbohydrate linker. A rigid backbone may result in the enthalpy loss. However, too flexible linkers might have consequences on the entropy of the association.

In the lipid bilayer systems the presentation of the saccharide of GSLs is limited by the membrane surface (Nyholm *et al.*, 1989) and by the conformational preferences of saccharide-ceramide linkage (Nyholm and Pascher, 1993). This should lead to a preference of only a subset of oligosaccharide orientations. In these systems the glycan epitope is presented in a substantially different way than in solution which might result in distinct receptor binding patterns. The presentation effects of membrane associated glycolipids observed in Paper IV are further discussed in section 4.3.2.

1.2.7. Glycan interactions in pathogenesis

The carbohydrate structures presented on the surface of the cell in the form of glycoconjugates provide a dense structural code which is deciphered by glycan binding proteins, i.e. lectins in a multitude of intracellular and extracellular physiological and pathophysiological events. Several viral, bacterial and fungal pathogens also exploit this glycan code for host cell attachment through recognizing different carbohydrate structures present on the cell surface (Imberty and Varrot, 2008; Olofsson and Bergstrom, 2005).

Most of the identified viral glycan receptors or attachment factors are negatively charged including sialic acids and glycosaminoglycans. Sialic acids are also recognized by a number of bacterial lectins. For viruses, a well-studied example includes influenza virus which has a preference for α2,6 and α2,3-linked sialic acids for human and avian strains, respectively. This glycan specificity actually explains that the site of infection for most human influenza viruses is in human upper respiratory tract where sialic acid in α2,6-linkage but not the one in α2,3-linkage is abundant. Interestingly and unfortunately, more than 300 human deaths from avian influenza virus have been reported by World Health Organization during the years 2003 to 2013 (WHO, 2013). The studies have shown that the avian influenza virus infections in humans occur in the cells belonging to lower respiratory tract where sialic acid in α2,3 linkage is relatively more abundant (Shinya *et al.*, 2006; van Riel *et al.*, 2006). Other examples of viruses recognizing sialylated glycans include, but are not limited to, coronaviruses (Schwegmann-Wessels and Herrler, 2006) causative agents of respiratory and gastrointestinal syndromes, Simian virus SV40 (Neu *et al.*, 2008) and adenoviruses (Cashman *et al.*, 2004). Heparin sulfate is also recognized by several viral agents including Herpes simplex virus type 1 (Copeland *et al.*, 2008) and papillomavirus L1 (de Witte *et al.*, 2007).

A number of viruses, mostly non-enveloped, have also been shown to recognize neutral glycan epitopes or histo-blood group antigens. Rabbit hemorrhagic disease virus (RHDV) produces fatal hemorrhage in rabbits and was shown to bind A antigen and H type 2 in epithelial cells of respiratory and gastrointestinal tracts (Ruvoen-Clouet *et al.*, 2000). Also it was shown that rabbits with impairment in expression of HBGAs were relatively resistant to the lethal infection (Nystrom *et al.*, 2011). Likewise rotavirus, the most dominating cause of gastroenteritis in children and infants worldwide, was more recently shown to recognize histo-blood group antigens (Huang *et al.*, 2012). This was in addition to the previously reported binding of animal rotaviruses to gangliosides (Delorme *et al.*, 2001). The structural basis of A type 2 binding to a human rotavirus strain was also recently reported (Hu *et al.*, 2012). Similarly, Parvovirus B19, which is the cause of persistent anemia in immunocompromised individuals and very often results in fetal death following infection during pregnancy, has been shown to recognize globotetraosylceramide on thin layer plates (Brown *et al.*, 1993). The globotetra- and globotriaosylceramides are expressed on erythrocytes of blood group P_1 and P_2 individuals. It was also shown, based on serological evidence in the later study, that individuals from blood group p, devoid of these GSLs, are completely resistant to the infection (Brown *et al.*, 1994).

Similar examples of glycan binding bacterial lectins are reviewed elsewhere (Ernst and Magnani, 2009; Imberty and Varrot, 2008).

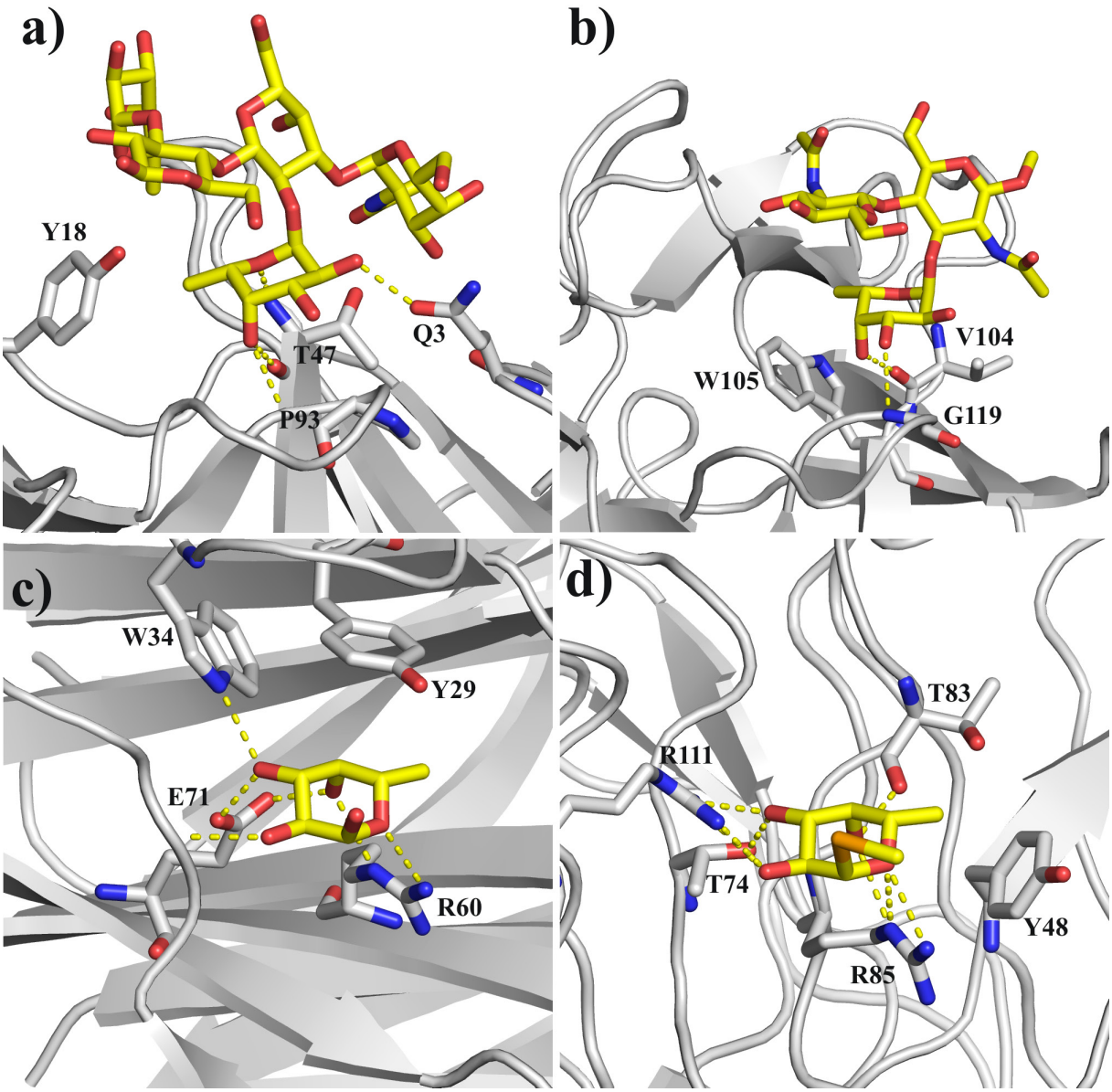

Figure 1.9 Examples of fucose binding sites. *The hydrogen bonds are indicated by yellow dashed lines and interacting residues are labeled and shown in stick representation. Note the common hydrophobic interactions in all 4 cases involving the methyl group of fucose situated in the proximity (~4 Å) of an aromatic residue (W or Y) in the binding pocket.(a) heat labile toxin of enterotoxigenic E. coli strain in complex with blood group A analog (Holmner et al., 2007), (b) fruiting body lectin, CCL2, in complex with fucosylated trisaccharide (Schubert et al., 2012), (c) BambL lectin from the fungal pathogen Burkholderia ambifaria in complex with fucose (Audfray et al., 2012), (d) BC2LC lectin from bacterium Burkholderia cenocepaica in complex with methylselenofucoside (Sulak et al., 2010).*

Fucose specific recognition sites of importance for pathogenesis

The fucose monosaccharide is characteristically different from other carbohydrate residues due to the presence of a methylated C6 instead of a hydroxyl group making it a deoxy hexose sugar with a distinct hydrophobic anchor. This unique feature of fucose has been exploited rigorously by fucose binding bacterial lectins and viral proteins. Most of the fucose specific binding modes studied through crystallography have demonstrated the presence of an aromatic protein residue in the vicinity of the

fucose methyl group which is involved in hydrophobic interactions while the hydrogen bonds are usually formed between the polar binding site residues and the sugar hydroxyl groups (Fig. 1.9). The fucose binding site features are further discussed and reviewed in chapter 4 and in Paper II of this thesis.

The remarkable feature of GII.4 noroviruses, which has been under positive selection for more than 3 decades, is the conservation of fucose binding site residues and consequently the recognition of fucosylated HBGAs (section 1.2.5). RHDV is another example of a virus recognizing fucose as the minimal structural requirement for binding HBGAs (Rademacher *et al.*, 2008). Rotaviruses are shown recently to recognize HBGAs but the binding mode specific to a fucose residue has not been reported so far. Among bacterial lectins, PA-IIL which is a virulence factor for an opportunistic pathogen *Psudomonas aeruginosa* has unusually high affinity for fucose ($k_d \approx 200$ nM for Le^a) (Perret *et al.*, 2005). The structural characterization of PA-IIL in complex with fucose (Mitchell *et al.*, 2002) has demonstrated a unique binding mode where 2 calcium ions mediate the fucose interactions with the lectin. Heat labile toxin (hLT) of enterotoxigenic E. coli strain which has been co-crystallized with blood group A analog also demonstrates a fucose specific binding mode (Holmner *et al.*, 2007). Similarly BambL, a fucose binding lectin from an opportunistic fungal pathogen *Burkholderia ambifaria,* was more recently crystallized with H type 1 and other fucosylated oligosaccharides (Audfray *et al.*, 2012). The structural data demonstrated fucose to be the main interacting residue. The crystallographic binding poses of some of these fucose binding proteins in complex with fucosylated structures are depicted in (Fig. 1.9).

1.2.8. Microdomains or lipid rafts

Cellular membranes are lipid bilayers mainly composed of (glyco) sphingolipids, phospholipids, cholesterol and a selection of several transmembrane proteins. The distribution of these structural units in the fluid bilayer is asymmetrical and results in lateral patches with high density of sphingolipids and cholesterol on the apical surface, which are called *rafts*. The concept of these lipid rafts or microdomains was first put forward by Simons and Ikonen more than a decade ago (Simons and Ikonen, 1997) but the current form of this concept is still debated among some cell biologists (Leslie, 2011). Lipid rafts are believed to be involved in a number of physiologically and immunologically relevant cellular processes mainly including the regulation of signal transduction pathways and protein trafficking (Lingwood and Simons, 2010; Yoshizaki *et al.*, 2008). Furthermore, rafts are also involved in protein-carbohydrate interactions between pathogens and their host cells.

The role of microdomains, particularly in viral entry and egress, has been implicated in a number of enveloped and non-enveloped viruses (Manes *et al.*, 2003). For enveloped viruses, the very early steps of virus infection include viral attachment and membrane fusion followed by viral entry. Most of the studies have investigated cholesterol dependence for these processes in either the cellular or the viral membrane or both. The requirement of cholesterol in either membrane for viral entry is attributed to the dependence on microdomains or lipid rafts. For instance, in case of Human Herpes

Virus 6 (HHV-6) it has been demonstrated that cholesterol in both viral (Huang *et al.*, 2006) and cellular membranes (Tang *et al.*, 2008) are required for successful viral entry. In contrast, for influenza virus it has been shown that cholesterol is required only in viral membrane but not in the cellular membrane (Sun and Whittaker, 2003), while the inverse has been reported for murine leukemia virus (Lu *et al.*, 2002). Similarly, for HIV, the membrane fusion requires the presence of cholesterol in both viral and cellular membrane. Also HIV glycoprotein gp120 recognizes CD4 localized in lipid rafts. The subsequent binding of gp120 to CD4 induces conformational changes in the glycoprotein which exposes the co-receptor binding site. At the same time, the CD4-gp120 binding also promotes clustering of microdomains containing co-receptors. The insertion of gp41 fusion peptide into the host cell is then induced by gp120-CD4-co receptor complex (Manes *et al.*, 2003). Simian Virus 40 (SV40), a non-enveloped virus, is internalized through clathrin independent caveolin-mediated endocytosis and binds to MHC class 1 molecules which are not present in lipid rafts. The virus-MHC-1 complex is then either trafficked to caveolae enriched microdomains or caveolae are recruited around the virus MHC-1 complex. The virus is then compartmentalized to ER. This pathway has been shown to be cholesterol dependent since depletion of cholesterol inhibits viral entry (Manes *et al.*, 2003). Apart from viruses, bacteria and protozoa are also reported to exploit lipid rafts for cellular entry, survival and signaling. These are also comprehensively reviewed in Manes et al., 2003. The principle finding of noroviruses binding to galactosylceramide in microdomains of planar supported lipid bilayers is discussed in Paper V.

1.3. Anti-noroviral strategies

A number of antiviral approaches have been used for different viruses which include vaccination, design of anti-adhesion molecules to block the virus attachment to the host cell and inhibitor design specific for enzymes essential in viral replication. Each virus presents new challenges for the development of antiviral therapeutics and therefore no single common approach has been identified to be successful for several different viruses unlike antibiotics in the treatment of bacterial infection.

In the absence of cell culture systems, but with access to detailed structural information on noroviral HBGA binding modes, the design of norovirus entry inhibitors seems a very interesting and practical approach. More recently, there have been some reports on the use of sugar-like compounds to inhibit Norovirus-HBGA interactions using both VLPs and P particles (Dou *et al.*, 2012; Dou *et al.*, 2011; Feng and Jiang, 2007; Hansman *et al.*, 2012; Pokhrel *et al.*, 2012; Tamura *et al.*, 2003; Zhang *et al.*, 2012; Zhang *et al.*, 2013). In one of the studies, VLP-Saliva ELISA was blocked using a library of 5000 glycomimetic compounds out of which 14 compounds were identified with IC_{50} value of < 15 uM (Feng and Jiang, 2007). An STD NMR study was of particular interest where authors designed multivalent entry inhibitors for VLP-HBGA interactions using Ast6139/01/Sp GII.4 norovirus VLPs. A library of 430 compounds was screened for candidates with inhibition activity against α-methyl fucose monosaccharide, H type 1 and blood group B antigens binding to the VLPs through competitive NMR experiments. The identified highly potent glycomimetic compounds were then conjugated with several fucose residues using polyacrylamide

backbone linker (Rademacher *et al.*, 2011). When measured through SPR, the authors could show that the IC_{50} value for the designed multivalent fucose inhibitors was around 1000 folds less than that observed with monovalent α-methyl fucose. For one of these inhibitors the IC_{50} value was observed to be in pM to nM range if the molecular weight of the whole compound including the polyacrylamide spacer was considered. These compounds are the most promising candidates identified so far for norovirus anti-adhesion therapy.

Vaccine design for noroviruses faces several challenges. The mechanisms of persistence of noroviruses in human populations together with complex long term and short term immunity, antigenic drift in the face of herd immunity, genetic diversity and changing host specificity are to name a few (Atmar and Estes, 2012; Donaldson *et al.*, 2010). The lack of effective and robust animal models and inability to grow the virus in cell culture systems have been the main barriers for successful noroviral vaccine design. Nevertheless, efforts are still underway to neutralize norovirus infection through effective vaccine development and vaccine candidates are currently under evaluation. The phase 1 clinical trials demonstrate that the use of GI.1 VLPs at least when administered orally is safe and elicits humoral, cellular and mucosal immune responses (Ball *et al.*, 1999; Tacket *et al.*, 2003). Several animal models have also been used to study the immunogenicity of vaccine candidates including VLPs and P particles. The norovirus VLPs engineered from the consensus of three naturally occurring GII.4 strains were shown to be immunogenic in rabbits when given intramuscularly (Parra *et al.*, 2012). The authors could show that the infected animals developed high serum antibody titers against the GII.4 and showed weaker cross-reactivity towards GI.1 VLPs. However, the antibody cross-reactivity towards GI.1 noroviruses increased when multivalent vaccine (both GI.1 and GII.4 VLPs) was given instead. A similar GII.4/GI.1 bivalent vaccine candidate has recently been submitted for evaluation in clinical trials. In another study, chimpanzees were intramuscularly vaccinated with GI.1 and GII.4 VLPs separately (Bok *et al.*, 2011). The authors noted that long term homologous protection (6-18 months post vaccination) was achieved in case of GI.1 but not GII.4 VLPs. More studies are needed to address the questions of heterologous protection within and between genogroups for effective vaccine design that could cover maximal human populations.

Moreover, the epochal evolution observed for GII.4 noroviruses has probably enabled them to target previously resistant populations. Single amino acid change in the P2 domain of the capsid region has resulted in altered binding specificities and diverse antigenic strains. This implies that the predominantly circulating GII.4 norovirus strains have to be considered individually for vaccine development. A multivalent vaccine with several GII.4 VLPs therefore seems essential for effective neutralization (Donaldson *et al.*, 2010).

2. Aims

The general aim of the thesis is to aid the development of anti-adhesion strategies against the clinically dominant GII.4 norovirus strains. This can be achieved only from a deeper understanding of multivalent norovirus-ligand interactions, taking place during viral binding to saliva glycoproteins and their subsequent transport to the gastrointestinal tract and the final viral uptake through the target cell plasma membrane. The specific aims of the thesis include;

- To study *in-silico* the binding site of GII.4 norovirus strains for possible interactions with the ABO and Lewis histo-blood group antigens.
- To specifically investigate the potential role of L-fucose, regardless of its positional linkage, in binding to GII.4 noroviruses through *in-silico* studies using difucosylated histo-blood group antigens.
- To describe the specificity of protein-glycan interactions of Ast6139/01/SP GII.4 norovirus VLPs binding to glycosphingolipids carrying ABH and Lewis histo-blood group antigens in a novel lipid bilayer assay system using the TIRFM technique and to study the interaction dynamics by characterizing the underlying kinetics of the system.
- To describe the binding of Dijon 171/96 GII.4 norovirus VLP to galactosylceramide in fluid and non-fluid membranes.

3. Methodological considerations

The detailed description of reagents and methods are found in the respective published papers and manuscripts included in the thesis. Here, a brief summary and some important aspects of the techniques and reagents are discussed without repetition of detailed methodological procedures.

3.1. Virus-like particles (VLPs) and P dimers

The VLPs (section 1.2.5) used in Papers III and IV were prepared from Ast6139/01/Sp GII.4 strain (GenBank accession number: AJ583672.2) isolated in Asturias, Spain (Boga *et al.*, 2004), while those used in Paper V were from the Dijon171/96 GII.4 strain (GenBank accession number: AAL79839) which were isolated from an outbreak in France during 1995-1996 winter season (Nicollier-Jamot *et al.*, 2003). The VLPs were recombinantly expressed in insect cells using a well established baculovirus expression system (Jiang *et al.*, 1992). The coding region of major capsid protein VP1 (ORF2, 1620 bp) was used for directional cloning followed by co-transfection in Sf9 insect cells with the baculovirus transfer vector and subsequent VLP production through the infection of High five cells as described elsewhere (Fiege *et al.*, 2012; Nicollier-Jamot *et al.*, 2003). Since the VLPs are produced in two different laboratories minor differences exist in their protocol, activity and prescribed handling. The Ast6139/01/Sp VLPs were stored at 4 ^{o}C before use while Dijon VLPs were kept at -80 ^{o}C after snap freezing under liquid nitrogen. Moreover, it has been observed that at a given concentration Ast6139 VLPs show at least 5 times higher coverage of bound vesicles than the Dijon VLPs. It should also be pointed out that the coordinates of P dimers used in the modeling studies of Papers I and II were from the data of crystal structure GII.4 norovirus strain, VA387 (Cao *et al.*, 2007) since we did not have access to crystals of the other GII.4 strains.

The choice of GII.4 norovirus VLPs and P dimers is justified by their dominating role in pandemic outbreaks caused by noroviruses among human populations for more than 15 years. Moreover, these strains are unique in their broadness of binding specificities which makes them interesting and useful targets to study VLP-glycan interactions.

3.2. Glycans

Carbohydrate structures used in the experimental binding assays in Papers III, IV and V were membrane associated glycosphingolipids while the ones studied *in-silico* in first two papers were monovalent energy minimized structures. All the glycans considered in the thesis were fucosylated with the exception of galactosylceramide (Paper V) and neo-lacto tetraosylceramide (Paper IV).

3.2.1. Glycosphingolipids (GSLs)

The GSLs used in Papers III, IV and V were purified as described previously (Karlsson, 1987). The preparations were primarily purified as to their glycan structures.

Table 3.1 Glycan structures used in the thesis to study the binding properties of GII.4 noroviruses.

Glycan	Structural Formula	Thesis Papers	Binding Assay or Computational Technique
Membrane bound glycosphingolipids			
Lc4-Cer	Galβ3GlcNAcβ3Galβ4Glcβ-Cer	-	-
H type 1-Cer	Fucα2Galβ3GlcNAcβ3Galβ4Glcβ-Cer	III, IV, V	TIRF, CBA
B type 1-Cer	Galα3(Fucα2)Galβ3GlcNAcβ3Galβ4Glcβ-Cer	III, IV	TIRF, CBA
A type 1-Cer	GalNAcα3(Fucα2)Galβ3GlcNAcβ3Galβ4Glcβ-Cer	III, IV	TIRF, CBA
Lewis a-Cer	Galβ3(Fucα4)GlcNAcβ3Galβ4Glcβ-Cer	III, IV	TIRF, CBA
Lewis b-Cer	Fucα2Galβ3(Fucα4)GlcNAcβ3Galβ4Glcβ-Cer	III, IV	TIRF, CBA
B Lewis b-Cer	Galα3 (Fucα2)Galβ3(Fucα4)GlcNAcβ3Galβ4Glcβ-Cer	-	-
A Lewis b-Cer	GalNAcα3 (Fucα2)Galβ3(Fucα4)GlcNAcβ3Galβ4Glcβ-Cer	III, IV	TIRF, CBA
nLc4-Cer	Galβ4GlcNAcβ3Galβ4Glcβ-Cer	IV	CBA
Lewis x-Cer	Galβ4(Fucα3)GlcNAcβ3Galβ4Glcβ-Cer	IV	TIRF, CBA
Lewis y-Cer	Fucα2Galβ4(Fucα3)GlcNAcβ3Galβ4Glcβ-Cer	III, IV	TIRF, CBA
A Lewis y-Cer	GalNAcα3(Fucα2)Galβ4(Fucα3)GlcNAcβ3Galβ4Glcβ-Cer	IV	TIRF, CBA
Gal-Cer	Galβ-Cer	V	TIRF, CBA
Monovalent Glycans			
H type 1	Fucα2Galβ3GlcNAcβ	I	Molecular Dynamics
B type 1 B-Tri B-Tri-O-CH3	Galα3(Fucα2)Galβ3GlcNAcβ Galα3(Fucα2)Galβ Galα3(Fucα2)Galβ-O-CH3	I I II	Molecular Dynamics
A type 1 A-Tri	GalNAcα3(Fucα2)Galβ3GlcNAcβ GalNAcα3(Fucα2)Galβ	I	Molecular Dynamics
Lewis a (Le^a)	Galβ3(Fucα4)GlcNAcβ	I	Molecular Dynamics
Lewis b (Le^b) Le^b–O-CH3	Fucα2Galβ3(Fucα4)GlcNAcβ Fucα2Galβ3(Fucα4)GlcNAcβ-O-CH3	I II	Molecular Dynamics
ALe^b ALe^b-O-CH3	GalNAcα3(Fucα2)Galβ3(Fucα4)GlcNAcβ GalNAcα3(Fucα2)Galβ3(Fucα4)GlcNAcβ-O-CH3	I II	Molecular Dynamics
BLe^b	Galα3(Fucα2)Galβ3(Fucα4)GlcNAcβ	I	Molecular Dynamics
Lewis x (Le^x) Le^x-O-CH3	Galβ4(Fucα3)GlcNAcβ Galβ4(Fucα3)GlcNAcβ-O-CH3	I II	Molecular Dynamics
Lewis y -O-CH3	Fucα2Galβ4(Fucα3)GlcNAcβ-O-CH3	II	Molecular Dynamics
Sialyl Le^x (SLe^x) SLe^x-O-CH3	Neu5Acα3Galβ4(Fucα3)GlcNAcβ Neu5Acα3Galβ4(Fucα3)GlcNAcβ-O-CH3	I II	Molecular Dynamics
ALe^y-O-CH3	GalNAcα3(Fucα2)Galβ4(Fucα3)GlcNAcβ-O-CH3	II	Molecular Dynamics
H type 3	Fucα2Galβ3GalNAcα	I	Molecular Dynamics
A type 3	GalNAcα3(Fucα2)Galβ3GalNAcα	I	Molecular Dynamics

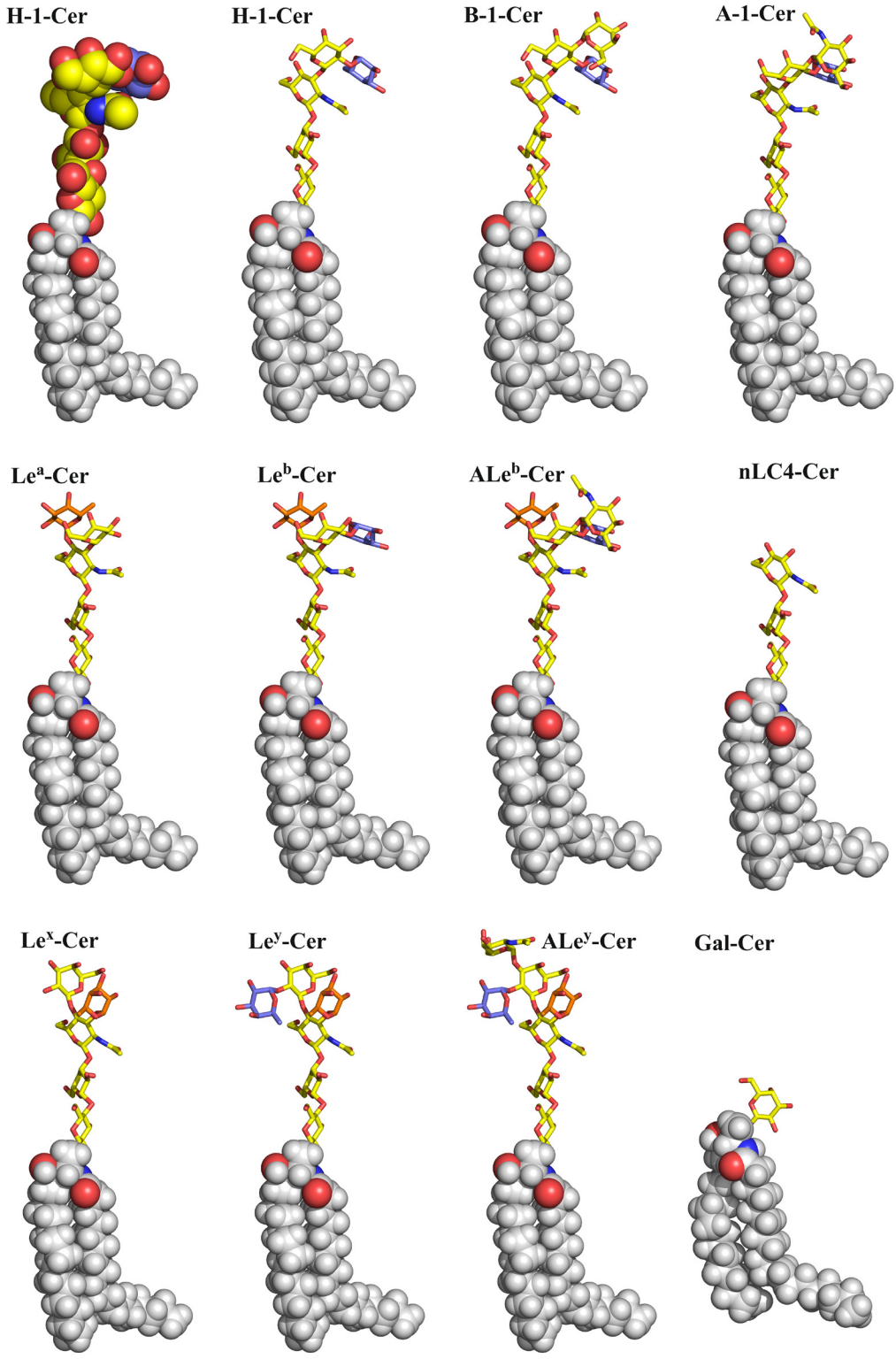

Figure 3.1 On the presentation of sugar epitopes in glycosphingolipids. *The orientation of glycans were prepared based on the crystal structure of B-1 in complex with the lectin domain of F18 fimbrial adhesion FedF (PDB: 4B4R) (Moonens et al., 2012) and GlcCer in complex with GLTP (Samygina et al., 2011). Glycans are shown as sticks (Carbons for α1,2-linked fucose: blue, α1,3-/α1,4-linked fucose: orange; other saccharides: yellow) or spheres (in case of H-1) and ceramide (d18:1,C18:1) as spheres (carbons and hydrogens: white; oxygen: red and nitrogen: blue). It should be noted that other conformations of the Glc-ceramide linkage will lead to different presentations of the fucose epitopes (Nyholm and Pascher, 1993).*

Structural differences in the ceramide composition among these GSLs are expected not only due to the differences in the sources from which the GSLs were prepared but also from the complexity of each tissue and cell type. All the type 1 chain GSLs were purified from human meconium samples of either a single individual or individuals pooled according to their ABO blood group phenotype. The rest of the GSLs were obtained from pooled samples of human erythrocytes or from canine small intestine (Nilsson *et al.*, 2009). The GalCer glycolipid was obtained in a monospecies form from Avanti and also kindly provided from prof Jan-Erik Månsson, University of Gothenburg. The purity and identity of these GSLs was confirmed through ^{1}H-NMR spectroscopy and mass spectrometry analyses (Angstrom *et al.*, 1982; Karlsson and Larson, 1981; McKibbin *et al.*, 1982). The glycosphingolipids used in, or related to, the thesis are listed in Table 3.1 and depicted in Fig. 3.1.

3.2.2. Oligosaccharides for *in-silico* studies

The glycan structures used in Paper I and II were obtained either from the SWEET2 server (Bohne *et al.*, 1999) or through the LEAP utility provided by Amber Tools version 1.4. The SWEET2 server generates glycosidic dihedrals based on the information from the available crystallographic data (Lutteke *et al.*, 2005), whereas the LEAP utility uses the set of parameters provided by GLYCAM, the AMBER forcefield for carbohydrates (Kirschner *et al.*, 2008). In both the cases, the final structural optimization was performed using GLYGAL program (Nahmany *et al.*, 2005) or molecular mechanics MM3 or MM4 forcefield in TINKER (Ponder, 2014), respectively.

3.3. Inhibitors

A multivalent entry inhibitor P1 (Rademacher *et al.*, 2011) was provided by prof. Thomas Peters to be tested in Paper IV to block the interactions of GSL-containing vesicles with the VLPs. The inhibitor was constructed using a polyacrylamide backbone linker attached to an estimated 20 L-fucose residues. Based on geometric calculations the P1 inhibitor could potentially reach the two binding sites of the dimeric units in the VLPs. Also, it was shown in the same paper, through SPR, that the inhibitor blocked VLP-glycan interactions with 1000 fold better IC_{50} value as compared to monovalent α-methyl-fucose. Other free monovalent oligosaccharides were also used as inhibitors in Paper IV. Specifically, H type 1 pentaose, A Lewis b hexaose, Sialyl Lewis x hexaose and α-methyl fucose monosaccharide were tested for inhibition activity. The synthetic glycans were purchased from Elicityl Oligotech, France.

3.4. *In silico* modeling of protein-carbohydrate interactions

Molecular modeling is a collection of computational methods and techniques which can be applied to biological systems in order to study the behavior of atoms through mathematical estimation of the forces of interactions leading atomic motions. Subsequently, the behavior of the whole system can be studied based on such

estimations. Leaving aside the *ab initio* molecular modeling, studies on protein-carbohydrate interactions performed using molecular docking and molecular dynamics techniques have provided valuable results aiding the drug design and discovery process (Durrant and McCammon, 2011; Ge *et al.*, 2013; Meng *et al.*, 2011; Stark and Powers, 2012).

At the heart of both these approaches are mathematical algorithms to calculate forces and potential energies on each and every atom in the system. *Forcefield,* is then termed as the set of parameters and expressions defining these potential energies, which are based on a combination of bonded and non-bonded interactions including van der Waals and electrostatic interactions among others (Durrant and McCammon, 2011; Fadda and Woods, 2010). Some docking programs use empirically modified forms of the forcefield as their scoring function which are used to calculate the interaction energies of receptor ligand complex. The results from molecular modeling studies depend largely on the design of these forcefield parameters which in turn depends upon the atomic interaction model used to derive these parameters. Clearly, there is a tradeoff between the accuracy of the interaction model and the CPU time. The more detailed and accurate the model is, the more CPU time it will take to process the calculations.

3.4.1. Molecular docking

The aim of docking is to the accurately predict the binding pose of a ligand in the receptor binding site in order to estimate the binding strength based on the scoring function (Halperin *et al.*, 2002; Taylor *et al.*, 2002). A typical docking procedure starts by searching for "suitable" protein sites which can potentially bind ligand in an energetically favorable pose. The conformational space of the ligand is then explored in the proposed binding regions of the receptor protein and each protein-ligand conformation is then ranked according to the interactions energies calculated using the scoring function employed by that program. Usually, the receptor protein conformation is taken from crystallographic or NMR studies. The process of docking the ligand to receptor is not trivial. The key step in the whole workflow is the estimation of interaction energy. The factors that influence its mathematical description include enthalpic and entropic estimations, solvation effects and charge calculations. Some of the more commonly used docking programs include Glide (Friesner *et al.*, 2004; Halgren *et al.*, 2004), Autodock (Morris *et al.*, 1998), FlexX (Rarey *et al.*, 1996) and GOLD (Jones *et al.*, 1997). The dockings using Glide XP, after energy minimization with the OPLS2001 forcefield, has been used in Paper I and Paper II to score the proposed binding poses of HBGA extensions in complex with P dimers of GII.4 norovirus VA387 strain (Cao *et al.*, 2007).

3.4.2. Molecular dynamics

In the summer of 1977, there was an article published in Nature from the group of Martin Karplus in the department of chemistry at the Harvard University, reporting a 8.8 picosecond molecular dynamics simulation of bovine pancreatic trypsin inhibitor (450 atoms) in implicit solvent model (McCammon *et al.*, 1977). This was the first molecular dynamic study which was conducted on a bio-molecular system. Today, this

simulation would not take more than a couple of minutes on an Intel Core i7 personal laptop which indicates the enormous progress both in terms of computational power and modeling technique over the last four decades. In the recognition of pioneering work in the field, Martin Karplus, Michael Levitt and Arieh Warshel were honored with a nobel prize in chemistry in 2013.

Molecular dynamics approach has been extensively and successfully applied to model protein-carbohydrate interactions over the years. The simulation presents a dynamic view of, for example, protein atoms through estimating atomic motions in a biological system for a given period of time. This is done in a deterministic fashion by solving the Newtonian equations of motion simultaneously for all the atoms in the system in order to calculate the position and velocities for the next following time point. The process is repeated many times in order to generate a trajectory of 3D conformations for certain duration of time. Usually, 1 femtosecond time steps are used. Since, the calculations performed at each time point are computationally demanding, even with today's computer power, the commonly sampled simulation time is in the nanosecond-microsecond range. Recently millions of dollars were spent and a dedicated computer architecture had to be employed to break the millisecond barrier (Borrell, 2008; Shaw *et al.*, 2009). Forcefields, as described above incorporate mathematical expressions and parameters for potential energy function which could be used to deduce the force on each atom. Some of these forcefields are compatible with proteins, carbohydrates and nucleic acids while others are specific for only one class of biomolecule such as sugars (Fadda and Woods, 2010). For molecular dynamics study to produce good results, the choice of an appropriate forcefield is clearly an important aspect to be considered.

The fundamental working principle of molecular dynamics simulation is based on the *Ergodic hypothesis*, according to which, if simulated for sufficiently long time, the resulting average state of the system should be comparable to the time independent statistical average, indicating convergence. This means that converged molecular dynamics simulation would provide an ensemble of structures comparable to the experimental data. This also means that the ideal molecular dynamic simulation should not depend upon the initial state of the system. However, with limitations on the length of the simulation time, the use of crystallographic data to represent the initial state of the system can decrease the number of time steps before convergence is achieved.

Molecular dynamics simulations have aided our understanding of the dynamic nature of proteins, sugars and nucleic acids at timescales pertaining to small atomic vibrations of the system to larger conformational changes, including protein folding (Shaw *et al.*, 2010). The time-averaged and thermodynamic properties of the system such as entropy and interaction energy estimation are also possible to deduce from molecular dynamics simulations. Apart from their usage as a rational tool to understand the biological process at molecular level, molecular dynamics has been successfully used to refine the results from X-ray or NMR techniques. Commonly used molecular dynamics packages include AMBER (Cornell *et al.*, 1995), CHARMM (Brooks *et al.*, 1983), GROMOS (Scott *et al.*, 1999) and NAMD (Nelson *et al.*, 1996).

It is important to consider that before the computational techniques discussed above could be applied, the feasibility of the system under study must be analyzed. As these approaches are *theoretical* therefore the limitations of the techniques should be considered. First, the gaps in our understanding of molecular interactions and inaccuracies of mathematical models demand continuous refinement and improvement of the employed algorithms, second, the demands on computational power needed to perform the calculations behind these techniques increases largely with the number of atoms involved in the system. Considering these and other factors for a good study design, based on solid experimental observations, will produce better results.

Molecular dynamics with both implicit and explicit solvation models has been used in Papers I and II to model the interactions of HBGAs in the fucose binding site of VA387 GII.4 norovirus P dimers. The AMBER suite of programs was used to carry out simulations with the forcefields ff99/ff99SB (Hornak *et al.*, 2006; Wang *et al.*, 2000) for proteins and GLYCAM06 (Kirschner *et al.*, 2008) for carbohydrates. In Paper II the binding energy estimates from Glide and MM-GBSA approach are also compared.

3.5. Chromatogram Binding Assay (CBA)

CBA is nowadays a classical binding assay in which the GSLs are chromatographed and assayed for binding to protein of interest over a thin layer plate. The binding is then detected using either immune staining in case of non-labelled protein or autoradiography in case of iodine-125 labelled proteins (Karlsson and Stromberg, 1987). The assay was first used to study binding of cholera toxin to GM1 ganglioside (Magnani *et al.*, 1980) in which the polar compound polyvinylpyrrolidone (PVP) was used to coat the plates before assaying for the proteins. Although the technique worked successfully for characterizing the interactions of monoclonal antibodies (Magnani *et al.*, 1981), an optimization step where PVP was replaced by poly-isobutyl methacrylate (P28) was introduced later (Brockhaus *et al.*, 1981; Hansson *et al.*, 1983). The suggested role of plastic coating was to avoid dissociation of the silica gel from the solid support and to introduce a hydrophobic film to promote better presentation of hydrophilic oligosaccharides in the amphipathic GSLs (Hansson *et al.*, 1984). This treatment presumably makes CBA fairly analogous but not identical to membrane based binding systems. CBA has been successfully utilized to study the GSL binding characteristics of several proteins, e.g. mono- and polyclonal antibodies, lectins, bacterial adhesins and toxins and also a few viruses including norovirus (Nilsson *et al.*, 2009; Rydell *et al.*, 2009a), parvovirus B19 (Brown *et al.*, 1993; Nasir *et al.*, 2014) and Sendai virus (Hansson *et al.*, 1984) to name a few. It is here interesting to note that in case of norovirus (both GII.4 and GI.1 strain) and parvovirus B19 no contradiction has been observed between the results obtained from membrane based assays and CBA. Parvovirus B19 VLPs were initially shown to recognize globoseries GSLs, which were identified through CBA as potential viral receptors (Brown *et al.*, 1993). Later a study reported that the membrane associated form of globoseries GSLs is not recognized by the virus (Kaufmann *et al.*, 2005) thereby questioning their role as viral receptors. However, this controversy has recently been resolved and the parvovirus B19 VLPs are indeed shown to recognize both

globotriaosyl and globotetraosyleceramides (globoside) on planar supported lipid bilayers (Nasir *et al.*, 2014).

CBA was used in Papers IV and V to study the binding of Ast6139/01/Sp and Dijon GII.4 norovirus VLPs to GSLs. The binding assay with purified GSLs and total GSL fractions from small intestine was performed essentially as described (Karlsson and Stromberg, 1987). The antibody detection principle in CBA used in Papers IV and V was recently introduced (Nilsson *et al.*, 2009).

3.6. Total Internal Reflection fluorescent microscopy (TIRFM)

Although the glycan profile of mucins in individual saliva is similar to the carbohydrates expressed on intestinal epithelial cells (Ravn and Dabelsteen, 2000), their complex glycan profiles are not identical, and the studies performed on saliva assays suffer from differences in the presentation of the glycan epitope when compared to membrane bound ligands. The fluidity of the membrane together with the potential lateral diffusion of membrane proteins and GSLs, in the context of microdomains or lipid rafts (section 1.2.7), adds layers of complexities to protein-carbohydrate interactions, which are beyond the working scope of conventional binding assays. The glycan binding profile of a virus obtained from conventional binding assays is relevant in terms of studying specificity, for example in case of noroviruses, how the virus attaches to saliva glycoproteins and how the resulting saliva "capsule" is transmitted to the site of infection. However, in order to stipulate the role of certain glycoproteins or specific GSLs in viral entry one should consider the membrane based assays which, not only unravel the very early steps of viral infection on a cellular membrane but, are also appropriate tools to screen for *true* receptors or attachment factors. This was one of the motives for the development of TIRFM methodology used in the thesis (Bally *et al.*, 2011).

3.6.1. The working principle

TIRFM uses evanescent waves to selectively illuminate the region just above the glass surface resulting in zero background from solution. When an incident beam of light is projected at the glass/water interface of the sample at an angle greater than the critical angle, the resulting beam of light is totally internally reflected. This leaves the electromagnetic evanescent waves on the surface of the glass, which exponentially decay as a function of perpendicular distance from the interface. Therefore the penetration depth of these waves is only ~ 150 nm which sets the axial resolution of the TIRF microscope. This unique feature of selective illumination of TIRFM has been exploited rigorously since its first use in 1950s (Ambrose, 1956) to study the properties of events at the cellular membranes (Axelrod, 1981) and of transmembrane biomolecules (Sako *et al.*, 2000; Zenisek *et al.*, 2000).

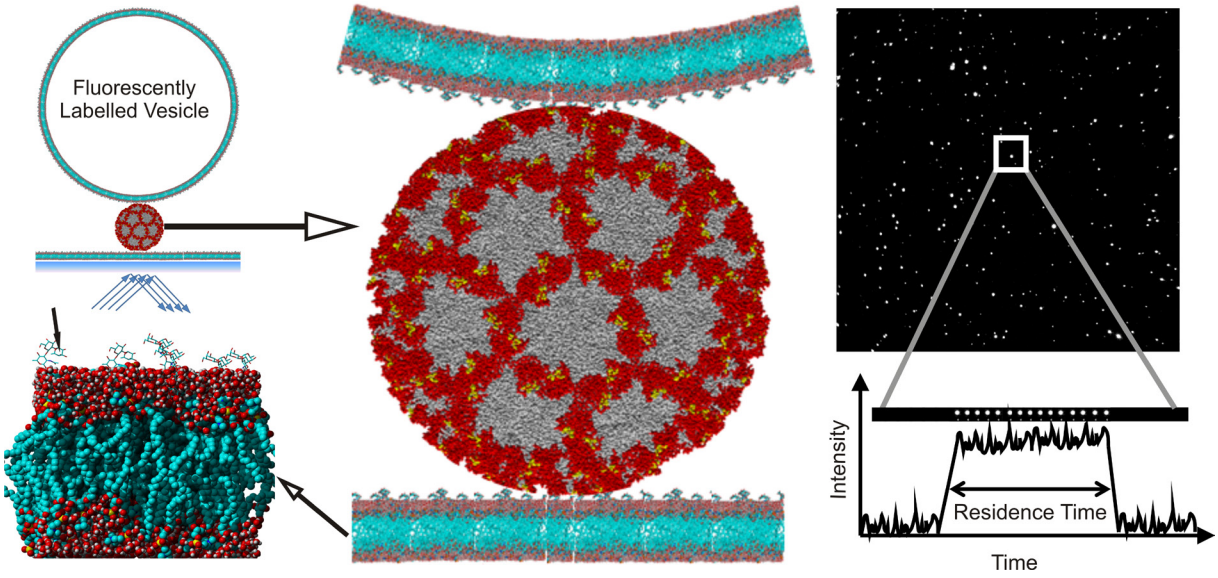

Figure 3.2 The TIRFM technique. *The working principle of TIRFM based binding assay used in the thesis is shown (top left). The VLPs are immobilized on lipid bilayer containing H-1 GSL. The fluorescently labeled vesicles, which also contain the GSLs of interest, are used as probes to interact with the bound VLPs. This setup is used to detect the vesicles bound to VLPs as spots in TIRF images (right) where each spot represents a vesicle-VLP complex. The association and dissociation kinetics are recorded based on the residence time analysis. The membrane structure is obtained from (Tieleman et al., 1999) and is the result of 2.6 ns of molecular dynamics simulation done in the presence of water molecules which are shown as red balls in the bilayer image with lipids molecules colored cyan. The 5% H-1 glycans were set manually into the POPC lipid head groups. The blowup view of crystal structure of GI.1 Norwalk virus capsid (Prasad et al., 1999) (P domain: red, shell domain: grey, binding sites: yellow) is also shown (middle) together with the parts of vesicular and solid supported bilayers. The glycolipids can be seen as small protrusion from the surface of the membrane. The relative molecular scales in this illustration are real and all reconstructions are done based on the molecular structures.*

3.6.2. The experimental setup

The setup for the TIRFM technique is discussed at length in Papers III and IV. In summary, the wells in a glass bottom 96 well microtiter plate were used for the binding/inhibition assays. For binding experiments, the bilayer was formed in wells through spontaneous rupture of vesicles containing 5 % (weight percentage) H type 1 GSLs in the buffer (PBS: 100 mM NaCl, 1 mM $CaCl_2$; pH 7.2) suspension. This was followed by the VLP incubation in the wells. The unbound fraction was washed away and binding to the irreversibly bound VLPs were detected through fluorescently labelled vesicles containing GSLs of interest (Fig. 3.2). For inhibition experiments, additional steps of inhibitor incubation were performed after establishing the VLP binding setup. The measurements at the microscope were taken at least 30 minutes after the injection of vesicles and/or inhibitors.

The average diameter of vesicles used both for detection and binding was ~ 145-170 nm regardless of the chemical composition and size of GSLs used (unpublished data). The vesicles were prepared from GSL fractions which were mixed with the lipid solution (Avanti Polar Lipids Inc.), lyophilized under light N_2 stream and vacuum treated for at least 2 hours. Thus obtained lipid films were then hydrated in the buffer suspension, vortexed for at least 3 minutes and extruded 21 times through a 100 nm polycarbonate membrane (Whatman, UK). The vesicle solution was stored at 4^{o}C until use within a week. (For further details see methods in Papers III and IV)

3.6.3. Measurements and analysis

The measurements for TIRFM were performed with Nikon Eclipse Ti-E inverted microscope with 60X magnification oil immersion objective (Numerical Aperture: 1.49). The final readout of TIRFM binding assay is a collection of images with spatially resolved spots. Each spot represents a VLP-liposome complex. The population of spots thus obtained is assumed to be predominantly composed of *single* VLP-vesicle complexes; however two small vesicles could also fuse to form a larger aggregate to bind the VLPs. The coverage of the irreversibly bound VLPs at the surface of the bilayer is controlled primarily by the concentration of the GSLs in the vesicles used to form the bilayer together with the VLP concentration and incubation time for binding. Since 5 % of GSLs were used in the vesicles for bilayer formation, assuming one to one binding ratio between VLPs and GSLs, at most 5 % of the total lipid molecules in the planar bilayer could theoretically bind the VLPs. However, this is likely to be an overestimate due to the fact that multivalency of the VLPs and lateral diffusion of GSLs in the fluid bilayer increase the number of acquired GSL-VLP contacts. This number was estimated to be between 6 and 12 based on simple geometric considerations (Bally *et al.*, 2011).

The data on vesicle association was obtained simply through counting the number of arrived vesicles as a function of time. The slope of resulting cumulative plot represents the rate of association which can be used to calculate the association rate constant. For dissociation analysis, the *residence time* is defined as the duration for which a vesicle remains bound to the surface fixed VLP. There were two sets of measurements performed in order to cover both *shorter* and *longer* residence times pertaining to lower and higher dissociation energies of bound vesicles, respectively. To control for the unspecific binding of vesicles to the bilayer surface, the minimum residence time was fixed to 3 seconds for lower energy and 30 seconds for higher energy time windows. Vesicles staying less than the minimum residence time were removed from the analysis. Through counting the number of vesicles still bound as a function of residence time one can arrive at the dissociation plot. Ideally, this curve should be exponential with the exponent representing the rate of dissociation. However, it had been noted previously that this was not the case. Instead, the resulting curve shows decay with multiple rates of dissociation and cannot fit with the single exponential curve. This heterogeneity is attributed primarily, but not exclusively, to the broad size distribution of the detecting vesicles (Bally *et al.*, 2011). Therefore, the multiple rates of dissociation are the consequence of several distinct populations of vesicles with different propensities and hence different tendencies to make GSL mediated contacts. In the original paper, describing the TIRF methodology (Bally *et al.*, 2011), this distribution of dissociation energies was approximated by the simplest rectangular distribution. The complete mathematical treatment of vesicle dissociation dealing with this heterogeneity and the dissociation energy distribution function extending the previous work is done in Paper III and discussed further in section 4.3.3. The image analysis is performed in Matlab as described in detail elsewhere (Gunnarsson *et al.*, 2008).

3.6.4. Pros and cons of TIRFM binding assay

Among several advantages of the TIRFM technique is the low material consumption and high sensitivity. As noted earlier, the technique has the potential to detect association and dissociation events of vesicles binding to single VLPs over a lipid membrane. In fact, it has been recently shown that TIRFM technique can be used as virus biosensor with the limit of detection in low fM regime sufficient to detect noroviruses, for example, from stool samples (Bally *et al.*, 2013). Furthermore, the high sensitivity of TIRFM technique was recently put to work in order to detect weaker carbohydrate-carbohydrate interactions (Kunze *et al.*, 2013). The TIRFM setup is simple and the readout is rich in quantitative kinetics data, including apparent rate constants and estimates of activation energies. On one hand these data provide means to discriminate different VLP-GSL interactions based on their binding strength and on the other hand, the ability to monitor the real-time kinetics behavior of the GSL population over different durations of residence times. This separates the TIRFM technique from the rest of the assays which largely rely on endpoint measurements. Finally, TIRFM technique may be greatly useful for lead optimization in drug design where the effect of inhibitors on the VLP-vesicle kinetics data might be important in filtering out the most interesting leads.

The mathematical analyses performed to model the interactions of VLPs with GSL-containing vesicles might be challenging for experimentalists, especially in the context of dissociation energy calculations which are directly related to binding energies of VLP-vesicle complex. Although TIRFM technique can distinguish between different GSL-containing vesicles based on relative differences between dissociation energies of activation, the calculation of the absolute value of binding energies (ΔG) is not possible under current settings. Isothermal titration calorimetry still remains the most relevant technique to study the binding energy of protein-ligand complexes. Also, the TIRFM setup, although simple, is currently not fast enough to be suitable for a high-throughput inhibitor screening protocol.

3.7. Quartz crystal microbalance with dissipation (QCM-D) monitoring

QCM-D is a particularly suitable and well established technique to study the protein-ligand interactions on the surface of lipid bilayer as it provides specific signatures for bilayer formation and the binding of biomolecules. The working principle of QCM-D relies on the piezo-electric properties of a quartz crystal (Janshoff *et al.*, 2000). More specifically, if an alternating electric field is applied across the electrodes of a quartz crystal, properly cut at a specific angle, mechanical oscillations or vibrations can be induced at fundamental or overtone resonant frequency. The resonant frequency, f, is mathematically related to the thickness and subsequently the mass of the crystal. The increase in adsorbed mass increases the thickness resulting in decrease of the resonant frequency. The changes in the mass, which include the mass of the solvent attached to the film, can be detected by observing the frequency shift Δf of the oscillating crystal (Rodahl *et al.*, 1995). This is particularly useful, for instance in case of the bilayer formation process where the vesicles first accumulate on the surface of the crystal

(SiO_2) resulting in a negative frequency shift followed by their rupturing and release of the solvent which subsequently increases the resonant frequency. Similar frequency shifts are also observed when, for instance, a VLP is bound to the GSL-containing bilayer (Rydell *et al.*, 2009a).

The nature of the bound mass, i.e. its viscoelastic properties, can also be studied by monitoring the energy dissipation or the extent of damping, D, of the oscillating crystal (Höök *et al.*, 1998). This gives information about the rigidity of the adsorbed mass which can be useful for instance in distinguishing the intact bound vesicles which induce more damping from the fluid lipid bilayer (Keller and Kasemo, 1998). The combined analysis of f and D makes QCM-D suitable for studying real-time attachment and detachment kinetics of unlabeled biomolecules.

QCM-D has been used to study the interactions of noroviruses with glycosphingolipids (Rydell *et al.*, 2009a). A similar setup has been used in Paper V to study the interactions of membrane bound galactosylceramide with the Dijon norovirus VLPs. The bilayers were formed at 37^{o}C and the binding to Dijon VLPs was measured both at 22^{o}C and 37^{o}C.

4. Results and discussion

When this thesis was started in 2010, the only available structural information on the GII.4 norovirus-glycan interactions was a crystal structure of VA387 P dimer in complex with an HBGA B-trisaccharide (Cao *et al.*, 2007). The binding studies had shown previously that the ABH antigens are indeed recognized by VA387 and other GII.4 noroviruses (section 1.2.5) but the structural aspects of these interactions were unknown. It was realized quite early in our modelling studies with these HBGAs (Paper I) that L-fucose could potentially play a central role in these interactions not only in α1,2 linkage but also in α1,3 and α1,4-linkage (Paper II). The following year a crystal structure of VA209 (GII.9) norovirus P dimer was solved in complex with Lewis y and Sialyl Lewis x (Chen *et al.*, 2011). The reported crystallographic pose was in complete agreement with the observations made through the modeling studies for GII.4 VA387 P dimer with α1,3/α1,4-linked fucose residues in the fucose binding site.

The next step was to characterize these interactions of GII.4 noroviruses with extended ABH structures over a lipid membrane. During the initial phase of the thesis, the research group was involved in the development of a technique to study the nature of interactions between individual VLPs and glycosphingolipids in lipid bilayers (planar membranes and vesicles) using total internal reflection fluorescent microscopy (TIRFM) technique (Bally *et al.*, 2011). This seemed a well-suited system to study the interactions of VLPs with membrane associated glycosphingolipids. The detailed kinetics data was obtained through TIRFM for Ast6139/01/Sp GII.4 VLPs interacting with membrane GSLs (Paper IV) and in the thesis work, the mathematical treatment of TIRFM raw data, particularly the dissociation analysis, was also extended (Paper III).

Intrigued by the role of galactosylceramide (GalCer) as a means of transcytosis and thus a facilitator for HIV infection (Lingwood and Branch, 2011), the nature of potential interactions with GII.4 noroviruses with GalCer was also investigated. Particularly, GalCer in membrane domains but not in fluid membranes was shown to bind Dijon 171/96 GII.4 norovirus VLPs indicating a need for multivalency in binding (Paper V).

4.1. Paper I – Potential interactions of GII.4 norovirus VA387 P dimers with ABH extensions

The susceptibility to GII.4 norovirus infections is mainly dependent on the expression of a secretor (*FUT2*) gene dependent minimal glycan structure (Fucα2Galβ3GlcNAc) and crystallographic data has demonstrated that the "secretor" fucose is the major interacting residue in VA387/B-trisaccharide (Galα3(Fucα2)Gal) interactions. Therefore, α1,2-L-fucose is expected to bind with high affinity and to contribute with the majority of the interaction energy. Glide XP dockings of α-L-fucose monosaccharide in complex with P dimers of VA387 reproduced the crystallographic binding pose with a Glide XP score of -6.0 kcal/mol while the score for the crystallographic pose of B-trisaccharide in complex with VA387 was -7.3 kcal/mol.

This demonstrated that theoretically α-L-fucose indeed contributed the major part of the binding strength of interactions. The Glide XP dockings for A- and B-trisaccharides in complex with VA387 were however not successful. Therefore, molecular dynamics simulations of ABH extensions in complex with GII.4 norovirus VA387 P dimers both in implicit (250 ps) and explicit solvation models (2.3 ns) were carried out to explore the computational fit of these ligands in the fucose binding site. The initial protein-ligand complexes were obtained through superimposition of α1,2-linked fucose of the ABH antigens on the secretor fucose of the VA387/B-trisaccharide crystal structure complex (Cao *et al.*, 2007).

4.1.1. Clashing in type 2 structures

All ABH antigens considered in this study shared an α1,2-linked fucose residue. The binding poses produced through computational fitting assumed the same interactions for α1,2-linked fucose as reported by the crystallographic data. The initial ABH structures were immediately accommodated in the VA387 binding site except H type 2 which showed steric clashes with residues Q390 and D391 around the fucose binding pocket. Preliminary computational fittings with other type 2 structures also indicated clashing in the same region of the binding site. The GlcNAc in type 2 structures is conformationally flipped by almost 180^{o} as compared to the type 1 structures which is the apparent reason for this clashing. These clashes were however relieved after brief minimization in line with a somewhat reduced affinity for ABH type 2 antigens as compared to the ABH type 1 antigens in binding to GII.4 noroviruses (Paper IV).

4.1.2. Interaction dynamics of ABH extensions in the VA387 binding site

The analysis of simulated protein-ligand complexes demonstrated that the amino acid residues around the fucose binding pocket are indeed involved in several hydrogen bonds and hydrophobic interactions which were not directly apparent from crystallographic data. However, a mutagenesis study had reported these interactions to be important and in some cases essential for binding ABH extensions (Tan *et al.*, 2008b). For instance, residues I389, Q331 and K348 were shown to be essential for binding to HBGA of A but not of B synthetic oligosaccharides using VA387 P proteins. The analysis of simulated complexes showed that these residues were indeed involved in hydrogen bonds and hydrophobic interactions with the GalNAcα residue in A antigens. These interactions were not reported in the crystal structure in which the A-trisaccharide structure was erroneously reported with strained geometry. In the simulations, the residue I389 was exclusively involved in hydrophobic interactions with the methyl group of terminal GalNAcα in A structures. For residue Q331, less frequent hydrogen bonds were observed with the terminal Galβ of the B structures along with the GalNAcα of A antigens. The interactions of K348 were observed in both HBGA A and B structures and therefore the mutagenesis data for this specific interaction could not be explained by the simulations. However, *in-silico* mutagenesis study with the perspective of systemic analysis of intramolecular interactions might hold some clues.

The residues surrounding the fucose binding site were also predicted to be important in binding the extended ABH structures (Tan *et al.*, 2008b). Specifically the mutations of residues Q390, D391 and G392 to alanine reduced the optical density in ELISA observed for binding A and B antigens to VA387 P proteins. The simulations indicated that these residues are indeed consistently involved in some hydrogen bonding interactions with the α1,4-linked fucose residue in difucosylated ABH antigens bound with the secretor fucose in the fucose binding site. However, for monofucosylated structures some of these interactions are instead observed with either the internal Galβ residue or with the GlcNAcβ residue at the reducing end.

4.1.3. Glide XP scoring of simulated protein-ligand complexes

Protein-ligand complexes from the later part of implicit solvent molecular dynamics simulations were scored using Glide XP "score and refine" function. B Lewis b, B type 1, A Lewis b and A type 1 were suggested to be the best binders with the Glide XP scores in the range of -8 to -12 kcal/mol. B antigens were scored slightly higher than the A structures indicating that the acetamido group of the terminal GalNAcα residue adds negatively to the binding energy. This might be due to the inherent flexibility of the acetamido group as observed by the detailed analysis of simulated complexes and indicated by the weaker electron density in the crystal structure (Cao *et al.*, 2007). The Glide XP scores of H structures suggested them to be the weakest binders with H type 2 scoring almost equally as the single docked fucose monosaccharide (-6.0 kcal/mol). The difference between H and A/B structures was explained by the interactions of terminal Gal/GalNAc residues. Lewis b tetrasaccharide with the Glide score of -8 kcal/mol was suggested to be an intermediate binder.

Together these results offered an explanation of the binding and mutagenesis data at the atomic level. It is striking that the α1,2-fucose binding site in GII.4 noroviruses can accommodate a range of structures from small trisaccharides to pentasaccharides all binding with the same binding mode specific to α-L-fucose. Furthermore, terminal and internal sugar structures bind to the amino acid residues around the fucose binding pocket. The mutations in this region have therefore resulted in altered binding specificities. This feature of the fucose binding site has been instrumental in shaping the dynamics of epochal evolution in GII.4 noroviruses (For detailed discussion and references see section 1.2.5).

4.2. Paper II – The Lewis fucose mediating GII.4 norovirus binding

It was realized already during the simulations of Paper I that also the α1,3/α1,4-linked fucose residues could potentially interact with the GII.4 norovirus binding site with exactly the same interaction profile as reported for secretor fucose in the crystal structure (Cao *et al.*, 2007). To computationally analyze this possibility, molecular dynamics simulations with a series of type 1 and type 2 difucosylated carbohydrate structures in complex with VA387 P dimers were carried out with explicit water molecules for at least 5 ns. The protein-ligand complexes were prepared as in Paper I.

The accepted binding poses so-obtained were specific for either α1,3/α1,4-linked or α1,2-linked fucose and were termed as Lewis or secretor poses, respectively.

4.2.1. The secretor and the Lewis poses[1] found in lectins and antibodies

The observation of the secretor and the Lewis poses in binding of difucosylated structures is not limited to noroviruses (Chen *et al.*, 2011; Shanker *et al.*, 2011). Similar trends have been observed in a number of lectins and antibodies with both the secretor and the Lewis poses for difucosylated structures such as Lewis b and Lewis y. To substantiate this observation, an analysis of Protein Data Bank (PDB) coordinate files was carried out in which the protein-ligand complexes with Lewis b or Lewis y were studied. The results revealed 6 out of 8 difucosylated structures were recognized in the Lewis pose. Since the number of available protein-ligand complexes in PDB is limited no conclusion could be drawn on the preference of the Lewis or the secretor pose. The observed Lewis and secretor poses might have implication in cross-reactivity of anti HBGA and Lewis antibodies. For instance, the cases of Lewis b/y antibodies cross reacting with H type 1/2 might be explained by the binding of these antigens in the Lewis and the secretor poses, respectively. Finally, the characteristics of fucose binding sites observed in these lectins and antibodies were strikingly similar to GII.4 norovirus binding sites which further strengthen the possibility that these poses might be observed in the same noroviral strain depending upon the presentation of the epitope.

4.2.2. Stable 3D trajectories for the secretor and the Lewis poses

The analysis of 5 ns molecular dynamics simulations revealed no major conformational changes or steric hindrances in either of the two binding poses. The interactions observed in Paper I for the secretor poses of Lewis b and A Lewis b were retained in the current simulations as well. However, the simulations of the Lewis poses for Lewis b and A Lewis b structures demonstrated that the region with amino acid residues Q390, D391 and G392 was observed to be involved in hydrogen bonds with the secretor fucose instead of the Lewis fucose. For type 2 chain structures Lewis x, Lewis y and A Lewis y, similar interactions were observed. The GlcNAcβ residue in type 2 chain structures is flipped by ~180^{o} therefore minor differences in the interactions of this residue with VA387 P dimers were observed. The chain continuation points faced the bulk solvent, in both binding poses. In fact for the Lewis poses of both type 1 and type 2 chain structures, the O1 atoms of GlcNAcβ residues at the reducing end of the two ligands in same dimer pointed towards each other. This indicates a possibility of a branched structure, for e.g. in saliva mucins, binding to the same dimeric unit of GII.4 noroviruses.

[1] The terms "secretor" and the "Lewis" poses are used unconservatively in this context since, for instance, biosynthesis of Lewis y can be derived from *FUT1* and *FUT4,5,6* or *FUT9* gene products which are distinct from *FUT2* (secretor genotype) and *FUT3* (the Lewis genotype) coded enzymes that make the true secretor and the Lewis antigens, respectively (section 1.2.4).

4.2.3. The proposed interactions of Sialyl Lewis x with VA387 P dimer

The observation of Sialyl Lewis x binding to GII.4 noroviruses has been reported multiple times (de Rougemont *et al.*, 2011; Fiege *et al.*, 2012; Rydell *et al.*, 2009b). The binding mode and the interaction details of this binding specificity were up to some extent characterized through an STD NMR study (Fiege *et al.*, 2012) during the writing of this paper. Also during this study, the structure of VA207 (GII.9) norovirus P dimers in complex with Sialyl Lewis x and Lewis y were solved (Chen *et al.*, 2011). The crystal structure demonstrated that both the antigens were recognized by VA207 in the Lewis pose. The interaction profile of the α1,3-linked fucose in the binding site was almost indistinguishable from that of secretor fucose in VA387 binding site. The α2,3 Neu5Ac linkage of Sialyl Lewis x antigen was found to be in *gauche* conformation [φ (C1-C2-O3-C3) ≈ 190^{o}] but the Neu5Acα residue demonstrated no interactions with the binding site. Shortly afterwards, the STD NMR study reported the conformational flexibility in this α2,3 linkage which was observed to switch between *gauche* and *trans* [φ (C1-C2-O3-C3) ≈ -70^{o}] conformations in solution.

The simulations of Sialyl Lewis x in complex with VA387 P dimer also demonstrated this switch between the two conformations of the α2,3 linkage. Moreover, it was observed that in *gauche* conformation Neu5Acα residue in Sialyl Lewis x exhibited several water mediated interactions with amino acids N393, G392 and D391 around the fucose binding pocket. Since this loop region is both sequentially and structurally distinct from VA207 GII.9 norovirus, no such interaction was observed between the Neu5Acα residue and the VA207 binding site.

4.2.4. Glide XP and MM-GBSA scoring of simulated protein-ligand complexes

The simulated protein-ligand complexes were scored using Glide XP scoring and MM-GBSA methods. A total of 10 frames from the last parts of simulated trajectory were used for scoring and binding energy estimation. Both the secretor and Lewis poses for Lewis b and Lewis y showed similar scores and binding energies, which were comparable to corresponding values for Lewis x. This was in agreement with the reported binding of Lewis x to GII.4 noroviruses (de Rougemont *et al.*, 2011; Fiege *et al.*, 2012). Also, it suggested that the α-L-fucose, other than the one in the binding pocket, theoretically does not add substantial binding strength. The secretor poses of A Lewis b and A Lewis y had slightly better Glide scores and MM-GBSA binding energies than the Lewis poses. This was explained by the fact that terminal GalNAcα residue in the secretor pose interacts mainly with the amino acids I389, Q331 and K348 adjacent to the fucose binding pocket. When switched to the Lewis pose, these interactions are lost. However, the Lewis poses of HBGA A structures were still better in Glide XP scores and MM-GBSA binding energies than Lewis b and Lewis y. Sialyl Lewis x in gauche conformation of α2,3 linkage was estimated to be the best binder among all the ligands in the Lewis pose by both Glide scores and MM-GBSA interaction energies. The difference between the predicted scores and energies of Lewis x and Sialyl Lewis x were explained by the protein interactions of sialic acid in

gauche conformation. For trans conformation of the α2,3 linkage of sialic acid, when no interactions of Neu5Acα residue with protein are observed, the estimates of Glide scores and MM-GBSA interaction energies are almost similar for both Lewis x and Sialyl Lewis x.

The results reported in this study have implications in understanding the nature of norovirus-glycan interactions both in the design of anti-adhesion molecules and the dynamics of norovirus evolution. This is especially interesting in the context of recent reports demonstrating that new strains of GII.4 noroviruses have evolved to recognize secretor gene independent Lewis antigens and non-secretor saliva (de Rougemont *et al.*, 2011; Fiege *et al.*, 2012; Uusi-Kerttula *et al.*, 2014).

4.3. Paper III and IV – On the study of the interaction kinetics of membrane associated HBGAs binding to GII.4 norovirus VLPs using TIRFM technique

Dynamic studies on model membrane systems targeting norovirus-glycan interaction are virtually lacking. The power of the TIRFM technique comes from its ability to use planar supported lipid bilayer and lipid vesicles to probe the multivalent interactions of unlabeled VLPs with GSLs at single particle resolution. This allows for the calculation of apparent kinetics data both on association and dissociation of the vesicle from the VLPs bound on a supported lipid bilayer. The interactions of Ast6139/01/Sp (AST) GII.4 norovirus VLPs with GSLs were characterized using TIRFM methodology and CBA was used as a complementary method for such characterization. Furthermore, the TIRFM technique in the present study was also utilized to estimate the inhibition activity (IC_{50}) of multivalent and monovalent inhibitors targeted against norovirus-glycan interactions.

4.3.1. AST VLPs recognizes GSLs on thin layer plates (Paper IV)

The results from chromatogram binding assay (CBA) demonstrated that AST VLPs recognize all tested type 1 chain GSLs. For the GSLs belonging to the type 2 chain, Lewis y but neither Lewis x nor A Lewis y were recognized by the VLPs. This is particularly interesting in the context of biosynthesis of these HBGAs. The type 1 chain HBGAs are biosynthesized predominantly dependent on *FUT2* (secretor) and *FUT3* (Lewis) gene products whereas those from type 2 chain are predominantly catalyzed by enzymes coded for by *FUT1* (α1,2-linked fucose) and *FUT4/5/6/7/9* (α1,3-linked fucose) (see section 1.2.4). Since the tissue specificity of these fucosyltransferases defines the expression of resulting HBGAs in saliva or other body secretions, or on mucosal cells, these results could have implications on the susceptibility to GII.4 noroviruses. The only *true* Lewis non-secretor antigen which was recognized weakly by AST VLPs on thin layer plates was Lewis a. The discussion on this structure will follow in the subsequent sections.

4.3.2. Characterization of AST VLP-GSL interactions using TIRFM (Paper IV)

The qualitative binding pattern observed from CBA demonstrated the preference of VLPs towards type 1 chain GSLs. The same GSLs were also tested in TIRF for binding activity which resulted in identical binding behavior for these GSLs towards AST VLPs. The VLPs were bound to H type 1 in the bilayer followed by the detection with different GSL-containing vesicles. The association and dissociation events of these vesicles binding to AST VLPs were then recorded in two time windows to capture both stronger (longer) and weaker (shorter) interactions. The analysis of association data revealed that vesicles with Lewis b and H type 1 GSLs had the highest rate of association along with those with Lewis a while the vesicles with A Lewis b demonstrated the slowest rate of attachment. The rest of the fucosylated GSL-containing vesicles had intermediate rates of association. The vesicles with GSLs Lewis x and A Lewis y did not associate as concluded from the fact that their rate of association was no better than the negative control (without VLPs). The corresponding K_{on} values for different GSL-containing vesicles were also calculated.

The association analysis provides one side of the picture in protein-carbohydrate interactions which is perhaps not relevant in drawing any conclusions on the binding strength of these interactions. In the current context, the different association rates demonstrate the differences in the activation energy of binding i.e. the energy barrier which must be broken by the liposomes in order to attach to the VLPs. The process of attachment is governed by random diffusion of the liposomes in the vicinity of the VLPs and starts with the formation of one or a few weak GSL mediated contacts. In case of the vesicles containing GSLs which are easily recognized by the VLPs the contact formation and subsequent attachment will occur depending upon the structural (binding site and ligand conformation) and thermodynamic (entropy and enthalpy) feasibility of the binding event which defines the activation energy barrier. However, for the vesicles which do not have the appropriate GSLs, that are recognized by the VLPs, the random diffusion of vesicles will not result in any contact formation.

Once the ligand (here GSL-containing vesicles) is bound to the receptor (VLPs), the estimates on the binding energy can be made. The time the vesicle stays bound to the VLP is then directly proportional to the binding strength with which it is recognized. This gives us the parameter of *residence time* as a tool for analysis of the binding strength of the interactions. Fortunately, in TIRFM technique the calculation of this parameter is quite trivial since one can record the time the vesicle docks onto the VLP and the time when the vesicle leaves. The vesicles that were bound before the measurement and remained bound until the end of the experiment are then termed as *irreversibly bound* vesicles which are recognized by the VLPs with the highest binding affinity. The residence time not only defines the strength of interactions, it can also be used to filter out the unspecific binding events by applying a threshold value. The vesicles staying less than a *minimum* residence time can be excluded from the analysis to filter for these events.

The dissociation analysis of the GSL-containing vesicles binding to the VLPs revealed that vesicles with A Lewis b, which were demonstrated to have the slowest rate of association, had the highest number of irreversibly bound vesicles and slowest rate of dissociation. On the other hand vesicles with Lewis a, which had rather high rate of association, dissociated faster than the rest of the GSL-containing vesicles. Similarly, the vesicles with Lewis b showed slower dissociation when compared to the ones with Lewis y, the corresponding type 2 GSL. It is interesting that the only type 2 structure which is recognized by AST VLPs, i.e. Lewis y, binds weaker than all the corresponding type 1 chain structures. Unfortunately, we did not have access to Sialyl Lewis x glycolipid.

These results also demonstrated that AST VLPs have the preference for type 1 chain HBGAs which is in line with our modeling data where we observed steric hindrances in accommodating the type 2 structures in VA387 binding site (see section 4.1.1). The observation of no binding of Lewis x and A Lewis y to GII.4 norovirus in a membrane based system might appear as a contradiction to the reported literature (Fiege *et al.*, 2012) including the modeling data in Paper II. However, it must be realized that in both cases (Fiege *et al.*, and Paper II) the binding was either observed in solution or modelled as a monovalent interaction where the binding epitope is suitably presented without any restrains to the protein. However, in the membrane based settings the glycans are naturally presented with some conformational restraints at the ceramide linkage which affect the subsequent projection of the binding epitope (Nyholm *et al.*, 1989). This means that the conformational space of a lipid linked glycan epitope is limited (Nyholm and Pascher, 1993) and might not contain *the* conformation required for optimal binding. This might explain the nonbinding behavior of the type 2 chain GSLs membranes (Fig. 3.1).

The characterization of AST VLP binding to membrane associated GSLs also demonstrated that Lewis a, the secretor gene independent GSL, binds weakly to the VLPs. It is known that the resistance to norovirus infection is largely driven by the polymorphisms in the *FUT2* gene. The individuals expressing inactive form of *FUT2* gene coded fucosyltransferase are inherently relatively protected from infection. However, as noted previously this resistance is not absolute and several cases of non-secretor infections have been reported (see section 1.2.5). Furthermore, studies have shown that secretor positive individuals are more susceptible to GI than to GII noroviruses (section 1.2.5). Since GII noroviruses have a conserved fucose binding site which might result in binding to other fucosylated structures, including those expressed by non-secretors (e.g. Lewis a), it might explain somewhat weaker relationship between GII norovirus susceptibility and secretor status. However, the resistance of non-secretor individuals to GII.4 norovirus infection can also be analyzed by considering the following facts. 1) Only 20% of European and American populations are comprised of non-secretors (population bias). 2) The variety of HBGA glycan epitopes in secretors, available for virus attachment, is considerably larger than that of non-secretors where only Lewis a epitope is present (individual bias). 3) Lewis a on the model membrane is shown to be the weakest binder for GII.4 noroviruses. Therefore, as a consequence of the natural population bias in 1, it might be speculated that norovirus has *adapted* to the secretor population where HBGA expression,

although diverse, is dependent on the fucosylation, a niche that has been exploited rigorously by GII.4 noroviruses after decades of evolution. Although the attachment and virus uptake in secretor negative individuals might be possible through Lewis a epitope, its role as receptor mediating cellular entry remains questionable due to its weak binding avidity in the lipid bilayer.

4.3.3. The desorption energy of activation (Paper III)

A cellular membrane is characterized by the heterogeneous mix of glycolipids and glycoproteins of different chemical compositions and sizes (Alberts *et al.*, 2013). Another level of heterogeneity is introduced by the microdomains or lipid rafts which are characterized by the clustering of GSLs in specific regions of the membrane. Finally, the ability of different GSLs to laterally diffuse in the membranes further adds to the complexity of the system. When using model membranes to probe multivalent interactions in a fluid membrane environment, one faces the challenges of dealing with some aspects of this heterogeneity. TIRFM assay was no exception (see section 3.6.3). The heterogeneity in dissociation energy of activation observed in TIRFM was attributed primarily to the broad size distribution and propensity of the vesicles which presumably results in different multivalent options to make GSL mediated contacts with VLPs. This may lead to subpopulations of identical types of vesicles with different dissociation energies which implies that multiple rates of dissociation would be observed based on these subpopulations of vesicles. The distribution of corresponding dissociation energies pertaining to different subpopulations of vesicles was previously estimated (Bally *et al.*, 2011) using the rectangular distribution function. The principle equation used for dissociation analysis was;

$$\ln\left(\frac{-dn/dt}{n(0)}\right) = \ln\left(\frac{k_B T}{\Delta E}\right) - \ln t \qquad (1)$$

where $-dn/dt$ is the dissociation rate of the vesicles, $n(0)$ is the initial number of bound vesicles, k_B is the Boltzmann's constant, T is the temperature, ΔE represents the range of dissociation energies and t denotes time. Eq. (1) implies that the term $\ln\left(\frac{-\mathrm{dn/dt}}{\mathrm{n(0)}}\right)$ depends linearly on the logarithm of time with the slope of -1. During the current analysis when binding of a series of different GSL-containing vesicles with AST VLPs was studied the resulting dissociation kinetics for some of the vesicles did not accurately fit the mathematical model. The approximation of rectangular distribution works reasonably well for "bell shaped" distributions (e.g. normal) but in case of underlying skewed distributions the approximation is no longer accurate. Therefore, the rectangular box distribution function was modified with a 2nd degree polynomial, eq. (2), which could approximate the skewed distribution of the dissociation energies in the system.

$$f(E) = a - b(E - E_*)^2 \qquad (2)$$

where a, b and E_* are the fitting parameters. The final equation for dissociation analysis was then derived as;

$$\frac{n(t)}{n(0)} \cong \frac{b}{3}E(t)^3 - bE_*E(t)^2 - \left(a - bE_*{}^2\right)E(t) - \frac{b}{3}E_{\max}{}^3 + bE_*E_{\max}{}^2 + (a - bE_*{}^2)E_{\max} \qquad (3)$$

where $E(t)$ is the dissociation energy which is related to the bound fraction of liposomes, $\frac{n(t)}{n(0)}$, through degree 3 polynomial. The energy distribution plots indeed demonstrated that different GSL-containing vesicles have quite different distribution of dissociation energies and could not have been accurately estimated using the simple rectangular distribution. If the vesicle sizes are assumed to be similar for different GSL-containing vesicles along with the distribution of GSLs in the membrane then among different vesicles, the differences in activation energies of detachment predominantly stem from the atomic interactions, such as hydrogen bonds and hydrophobic contacts, between GSL and VLP which are studied through crystallography (see section 1.2.5), modeled in Paper I and Paper II and reflected in the energetic landscape of this multivalent protein-ligand system.

It should be noted that the analysis performed here gives piecewise information on the *actual* heterogeneity in the system. Since the experimental time is limited the binding reaction in reality could be shorter than the minimum residence time probed by the experimental settings or longer than the actual experimental time, the effect of such binding events cannot be tracked and hence is not estimated in the present calculations. However, the current mathematical treatment is suitable for comparative analyses where, for instance, the effects of induced microdomains or lipid rafts on VLP-GSL interactions could be studied.

4.3.4. Inhibition of AST VLP-GSL interactions on the membrane (Paper IV)

The TIRF setup was also used to test monovalent and multivalent sugar compounds for inhibitory activity against AST VLPs. The difference in rates of association before and after inhibition was used to calculate the IC_{50} values for inhibitors. The results demonstrated that monovalent α-methyl-fucose inhibited GSL containing vesicles at IC_{50} ~ 10 mM. In contrast, the multivalent JG34 (P1) inhibitor, which contains on average 20 fucose residues on a polyacrylamide backbone linker, showed at least 100 times lower IC_{50} value. Among monovalent HBGAs, A Lewis b showed at least 10 times lower IC_{50} value than the rest of the glycans which demonstrated IC_{50} value of ~10 mM, including H type 1, Sialyl Lewis x and α-methyl-fucose. This was in line with the binding experiments suggesting A Lewis b as the strongest binder. The almost similar IC_{50} values for α-methyl-fucose and monofucosylated antigens further supports the role of α-L-fucose monosaccharide in the binding of HBGAs to human GII.4 noroviruses.

4.4. Paper V – Human noroviruses recognize galactosyl-ceramide (GalCer) in GSL microdomains

The studies on histopathology of noroviruses have suggested that intestinal epithelial cells in upper small intestine are the likely sites of infection (see section 1.1.8). The true target cell however, has not yet been identified. The apical plasma membrane domains of intestinal epithelial cells are rich in glucosyl- and galactosyl ceramide (GalCer) (Simons and van Meer, 1988) along with HBGA fucosylated GSLs (Bjork *et al.*, 1987; Breimer *et al.*, 2012), which have already been identified as potential receptors or attachment factors (see section 1.2.5). In this study the binding of Dijon 171/96 GII.4 VLPs to GalCer was investigated. The CBA was performed with Dijon VLPs and the total non-acid fractions from human intestinal epithelial cells of small and large intestines of secretor and non-secretor blood group O individuals. The results demonstrated that monoglycosylceramides, a mixture of GlcCer and GalCer, were recognized by Dijon VLPs together with H type 1 GSL from secretor individual. A separate Dijon VLP CBA with pure GalCer and H type 1 further confirmed the binding to these GSLs.

It is known that GalCer is present in the phase-separated domains as it has been isolated from detergent resistant membranes (Brown and Rose, 1992; Thompson and Tillack, 1985). The GalCer rich microdomains have also been observed in model membrane systems. To visualize such microdomains, atomic force microscopy (AFM) was utilized allowing high resolution imaging of GalCer domains. GalCer bilayers were produced through lipid vesicle fusion on silicon dioxide substrate with 10% (w/w) GSL content. At a higher temperature (37^{o}C), i.e. greater than the liquid-gel phase transition temperatures, homogenous bilayers were obtained and the coverage of the microdomains was only 0.4%, However, this coverage increased to 7 % in the gel phase when the miscibility was diminished by lowering the temperature (22^{o}C). The GalCer rich domains thus visualized had an average area of 21 um^{2}.

The effect of microdomains on GalCer binding to the Dijon VLPs was investigated through quartz crystal microbalance with dissipation monitoring (QCM-D), which is an effective technique in tracking the real time bilayer formation and interactions of membrane associated GSLs with multivalent proteins like VLPs. The bilayer formation was performed at 37^{o}C which was confirmed by the signatures of frequency and dissipation shifts reflecting the accumulation and subsequent rupturing of the vesicles to form a high quality bilayer. VLP binding was then monitored at 22^{o}C, presumably in the presence of domains, after heating the system to 50^{o}C and gradually cooling it down. The binding was confirmed with the decrease and increase in frequency (Δf = -8.7Hz) and dissipation ($\Delta D = 0.85*10^{-6}$), respectively and reached saturation within 15 minutes without rinsing. The binding at 37^{o}C (liquid phase) was also monitored which resulted in a substantially decreased binding signal (Δf = -1.2Hz). These results clearly demonstrated that the microdomains are required for binding of Dijon VLPs to GalCer in lipid bilayer. For the control experiment, H type 1 bilayers (5% GSL; w/w) were formed at room temperature and the binding to Dijon VLPs was also monitored through QCM-D. The resulting frequency shifts were in the excess of -20 Hz both at room temperature and at 37^{o}C. Also, saturation was not

reached before rinsing. The fact that GalCer binding was saturated within 15 minutes without rinsing suggests that all the conformationally favorable GalCer molecules were bound to the VLPs during this time and that only a fraction of the total GalCer population was recognized by the VLPs, since the total amount of GalCer was more than H type 1.

Finally, fluorescence microscopy was used to further characterize the binding of Dijon VLPs to membrane associated GalCer. The domains were clearly visualized in the bilayers with 20 % or 35 % GalCer at room temperature (gel-phase) while above 37^{o}C (liquid-phase) the bilayers were observed to be homogenous in agreement with the AFM experiments. Moreover, the fluorescence recovery after photobleaching (FRAP) experiments demonstrated the GalCer domains to be in the gel phase while the rest of the bilayer was mobile and post bleaching lateral diffusion of fluorescent lipids was observed in the bilayer, however restricted by the immobile GalCer microdomains. The binding of Dijon VLPs to GalCer rich domains was detected through fluorescently labeled vesicles. Strikingly, the VLPs demonstrated binding to the edges but not to the center of the domains. This can be explained by presentation effects of the glycolipids in the microdomains. If the GSLs are too densely packed then it is likely that steric hindrance by the surrounding GSLs could inhibit the binding altogether whereas on the edges of the domain the favorable conformational pose could be found with relative ease.

The microdomains or lipid rafts are not only involved in cell signaling or trafficking but could in some cases serve as the sites suitable for pathogen invasion (see section 1.2.7). The finding of GII.4 norovirus binding to GalCer microdomains is particularly interesting in the context of plasma membrane GSL composition of intestinal epithelia which includes GalCer and is dominated by fucosylated type 1 chain glycans, known to recognize human GII.4 noroviruses. Furthermore, it has been shown that GalCer acts as a substitute receptor for HIV (Harouse *et al.*, 1991) and binds to membrane glycoprotein gp120 in the cells lacking CD4 (Cook *et al.*, 1994; Fantini *et al.*, 1993). The virus binds to neural, colonic or vaginal CD4 negative cells through GalCer or to sperm cell through a GalCer glycerolipid analogue, i.e. galactosylalkylacylglycerol (Gadella *et al.*, 1998). The so-infected cells can then transport the virus to CD4 positive lymphocytes or to macrophages where further virus propagation leads to Autoimmune deficiency syndrome (AIDS). The results of reported herein are both novel and thought provoking in the context of norovirus infection in which not only GalCer can act as a receptor or attachment factor but also the GSL clustering might be a key player.

4.5. Concluding remarks

This thesis gives a detailed account on the host carbohydrate - Norovirus VLP interactions both at atomic and molecular levels in monovalent *in silico* binding system and in multivalent experimental (*in vitro*) VLP binding assays based on model bilayers as cell membrane mimics. The computational studies (Papers I and II) provide insights into the dynamics of atomic interactions based on the crystallographic studies and provide reasonable explanations for the binding data at an atomic level. These

studies also suggest that the Lewis fucose monosaccharide in secretor gene dependent and independent HBGAs can interact with the fucose binding site of the VLP even in the presence of α1,2-linked fucose. The role of different sugar moieties in relation to their potential impact on binding energies is also analyzed. These data could aid the process of anti-adhesion molecular design for GII.4 noroviruses.

The TIRFM technique (Papers III and IV) has revealed novel kinetic features of multivalent GSL-VLP interactions. These kinetics data provide valuable insights in the *real time* binding behavior of specific GSL populations to GII.4 VLPs. This enabled us to distinguish between GSLs with higher and lower desorption energy of activation which is not possible through conventional end-point measurements. Furthermore, the incorporation of the native bilayer-like environment makes it possible to study specialized features of the lipid membrane, e.g. microdomains. Some aspects of heterogeneity of cellular membrane are also modeled in the current TIRFM setup and the subsequent mathematical treatment, although challenging (Paper III), has enabled us to distinguish between different GSL-containing vesicles based on the desorption energy of activation. The mathematical treatment also gives an appreciation of the underlying complexity of the native cellular membrane and the subsequent interactions which take place in extremely heterogeneous environments. The conformational difference in glycan epitope presentation in solution and over the lipid bilayer gives rise to different GSL binding patterns for the VLPs. The results may be useful in understanding the evolutionary dynamics of the pandemic GII.4 lineage and in the development of antiviral strategies.

For the first time the interactions of galactosylceramide with GII.4 norovirus VLPs have been identified. The observation that these interactions take place in microdomains, when GalCer is in gel-phase while the rest of the bilayer is mobile, may have profound implications on the study of norovirus infection and treatment since both GalCer and microdomains have been shown to play crucial role in the pathogenesis of other viral pathogens, such as HIV.

5. Outlook and future perspective

5.1. TIRFM binding assay

The unique features of the TIRFM technique including the ability to probe membrane dynamics, single vesicle sensitivity, quantitative analysis of binding kinetics and low material consumption, make this assay suitable for studying virus host interactions. The natural extension of TIRFM technique in this application would be to use different live viruses with different glycosphingolipid and glycoprotein receptor structures imbedded in model vesicles or vesicles from target cells. Currently, studies on infectious parvovirus B19 virions binding to globoseries glycosphingolipids using the TIRFM setup are ongoing. Similar studies with caliciviruses are intended in the future.

Another application area of TIRFM is to probe for weak carbohydrate - carbohydrate interactions which are essential for cellular adhesion. The first study with Lewis x-Lewis x interactions was reported last year using the TIRFM technique (Kunze *et al.*, 2013). However, an immediate extension would be to study the roles of different glycan and ceramide structures in defining these interactions. Comparisons could be made based on both the degree of hydroxylation and saturation of sphingosine base or fatty acid lipid chains of sphingolipids as it could be associated with the hydrogen bonding capability of the lipids which may have implications on the functional role of these lipids along with the way these lipids are localized in the plasma membrane (Karlsson, 1982). Also, the cholesterol and sphingomyelin effects can be studied through the current TIRFM setup. These studies are currently in planning phase.

5.2. All atom molecular dynamics membrane simulations

To complement the TIRFM methodology and to investigate the atomic details of the interactions at the surface of the bilayer, molecular dynamics membrane simulations are being planned. The presentation effects of glycan epitopes in the membrane and their effect on the binding of the GII.4 P dimers will be investigated at the level of atomic interactions. In this context the recent development of lipid forcefields from AMBER and Schrödinger seems very promising (Dickson *et al.*, 2014; Maciejewski *et al.*, 2014).

6. Acknowledgements

I cannot resist the temptation of being extremely grateful to *Allah*, the Exalted for enabling me, by His grace, to complete this thesis. Paying rightful gratitude is beyond my ability.

I thank all my colleagues and friends who have provided me with a great time during the last few years and made my research an enjoyable and successful experience. I am especially grateful to;

Göran Larson, my supervisor. Thank you for teaching me how to do research, how to think scientifically and perform experiments. Thank you for taking care of me during all these years. I still remember your diagnosis of my hyperthyroidism during summer of 2011 when you skipped lunch for me to help me with tests and later with the medicine as well. Your continuous encouragement and willingness to help in scientific and non-scientific matters is of great value to me. Thank you for excellent and unforgettable times. I am lucky to have you as my supervisor.

Per-Georg Nyholm, my co-supervisor. It all started with you in 2008. Thank you for teaching me structural biology, for supporting me during all these years. Your words of encouragement mean a lot to me. Thank you for your continuous guidance. It is not unrealistic to state that I would not be writing this thesis without your guidance.

Fredrik Höök, my co-supervisor. Thank you for showing me the physical perspective of biomolecules and for stimulating scientific discussions and supervision on TIRF microscopy and binding kinetics. It was a pleasure.

Vladimir Zhdanov, thank you for teaching me how to handle mathematics, for your prompt and comprehensive feedbacks and for great scientific discussions.

The Glycobiology group, both former and current members. Thank you Gustaf for training me on CBA and for great scientific ideas and collaboration. Thanks Alejandro for being my friend in science and in life. Thanks for all the great times we had in the lab, during the courses and in Norway. Thank you Jonas for helping me with chemistry, for nice discussions and for good times. Thanks Fredrik for good discussions both about science and tennis. Thank you Mette, Ammi, Inger and Camilla.

Biognos AB. Martin Frank for showing me the strength of molecular modeling and for helping me with CAT and other analyses. Thank you for making simulations look so easy. Chaitanya and Fransesco for great times and challenging pool games. Thank you for helping me with the understanding of molecular dynamics and for good scientific discussions.

Marta Bally and Angelika Kunze, my colleagues from Chalmers and Sahlgrenska. Thank you for training me on TIRFM technique and teaching me how to handle the microscope. Thanks Marta for great discussions on TIRFM manuscripts and for helping me out with the analyses. Thank you Angelika for showing me the inhibition TIRFM technique.

Lennart Svenson and Johan Nordgren in Linköping. Thank you for very good discussion on norovirus and rotavirus and for stimulating collaboration.

Fransesco Parra (Spain) for providing the VLPs and Thomas Peters (Germany) for a great conference (noro2012) and for providing the P1 inhibitor.

My halftime seminar committee: Susanna Cardell, Gunnar C. Hanson and Tore Samuelsson. Thanks for valuable and constructive feedback.

Lena Kjellén and Barbro Lowisin for very well organized Medigly meetings with great scientific atmosphere. Thank you Sigvard Olofsson for refreshing Smögen meetings.

Ulla and Ulrica for all the administrative help.

The Sahlgrenska Academy at the Gothenburg University for providing me this opportunity to study as a Ph.D student in a great and vibrant scientific environment.

Swedish research council (8266 to Göran Larson), Vinnova (Fredrik Höök and Göran Larson) and the Sahlgrenska Academy at the University of Gothenburg for financial support. Travel grants from Sahlgrenska Academy were also kindly awarded for international conferences (noro2012 and SFG annual meeting 2014).

My dear father, mother, Mamoon, Mani, aunti and my 7 siblings who could not be here in Sweden with me but their love and prayers were always by my side.

Thank you Salmana and Munifa!

7. References

Abrantes, J., van der Loo, W., Le Pendu, J., & Esteves, P. J. (2012). Rabbit haemorrhagic disease (RHD) and rabbit haemorrhagic disease virus (RHDV): a review. *Vet Res, 43*, 12.

Adler, J. L., & Zickl, R. (1969). Winter vomiting disease. *J Infect Dis, 119*(6), 668-673.

Agus, S. G., Dolin, R., Wyatt, R. G., Tousimis, A. J., & Northrup, R. S. (1973). Acute infectious nonbacterial gastroenteritis: intestinal histopathology. Histologic and enzymatic alterations during illness produced by the Norwalk agent in man. *Ann Intern Med, 79*(1), 18-25.

Ahmed, S. M., Lopman, B. A., & Levy, K. (2013). A systematic review and meta-analysis of the global seasonality of norovirus. *PLoS One, 8*(10), e75922.

Alberts, B., Bray, D., Hopkin, K., Johnson, A., Lewis, J., Raff, M., et al. (2013). *Essential cell biology*: Garland Science.

Ambrose, E. J. (1956). A surface contact microscope for the study of cell movements. *Nature, 178*(4543), 1194.

Ando, T., Noel, J. S., & Fankhauser, R. L. (2000). Genetic classification of "Norwalk-like viruses. *J Infect Dis, 181 Suppl 2*, S336-348.

Angstrom, J., Falk, K. E., Karlsson, K. A., & Larson, G. (1982). Chemical fingerprinting of glycosphingolipids in meconium of a human blood group O Le(a-b+) secretor. *Biochim Biophys Acta, 710*(3), 428-436.

Appleton, H. (1987). Small round viruses: classification and role in food-borne infections. *Ciba Found Symp, 128*, 108-125.

Apweiler, R., Hermjakob, H., & Sharon, N. (1999). On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim Biophys Acta, 1473*(1), 4-8.

Atmar, R. L., Bernstein, D. I., Harro, C. D., Al-Ibrahim, M. S., Chen, W. H., Ferreira, J., et al. (2011). Norovirus vaccine against experimental human Norwalk Virus illness. *N Engl J Med, 365*(23), 2178-2187.

Atmar, R. L., & Estes, M. K. (2001). Diagnosis of noncultivatable gastroenteritis viruses, the human caliciviruses. *Clin Microbiol Rev, 14*(1), 15-37.

Atmar, R. L., & Estes, M. K. (2006). The epidemiologic and clinical importance of norovirus infection. *Gastroenterol Clin North Am, 35*(2), 275-290, viii.

Atmar, R. L., & Estes, M. K. (2012). Norovirus vaccine development: next steps. *Expert Rev Vaccines, 11*(9), 1023-1025.

Atmar, R. L., Opekun, A. R., Gilger, M. A., Estes, M. K., Crawford, S. E., Neill, F. H., et al. (2008). Norwalk virus shedding after experimental human infection. *Emerg Infect Dis, 14*(10), 1553-1557.

Atmar, R. L., Opekun, A. R., Gilger, M. A., Estes, M. K., Crawford, S. E., Neill, F. H., et al. (2013). Determination of the 50% Human Infectious Dose for Norwalk Virus. *J Infect Dis*.

Audfray, A., Claudinon, J., Abounit, S., Ruvoen-Clouet, N., Larson, G., Smith, D. F., et al. (2012). Fucose-binding lectin from opportunistic pathogen Burkholderia ambifaria binds to both plant and human oligosaccharidic epitopes. *J Biol Chem, 287*(6), 4335-4347.

Axelrod, D. (1981). Cell-substrate contacts illuminated by total internal reflection fluorescence. *J Cell Biol, 89*(1), 141-145.

Ayukekbong, J. A., Andersson, M. E., Vansarla, G., Tah, F., Nkuo-Akenji, T., Lindh, M., et al. (2013). Monitoring of seasonality of norovirus and other enteric viruses in Cameroon by real-time PCR: an exploratory study. *Epidemiol Infect*, 1-10.

Bailey, D., Karakasiliotis, I., Vashist, S., Chung, L. M., Rees, J., McFadden, N., et al. (2010). Functional analysis of RNA structures present at the 3' extremity of the murine norovirus genome: the variable polypyrimidine tract plays a role in viral virulence. *J Virol, 84*(6), 2859-2870.

Baker, K., Morris, J., McCarthy, N., Saldana, L., Lowther, J., Collinson, A., et al. (2011). An outbreak of norovirus infection linked to oyster consumption at a UK restaurant, February 2010. *J Public Health (Oxf), 33*(2), 205-211.

Ball, J. M., Graham, D. Y., Opekun, A. R., Gilger, M. A., Guerrero, R. A., & Estes, M. K. (1999). Recombinant Norwalk virus-like particles given orally to volunteers: phase I study. *Gastroenterology, 117*(1), 40-48.

Bally, M., Dimitrievski, K., Larson, G., Zhdanov, V. P., & Hook, F. (2012). Interaction of virions with membrane glycolipids. *Phys Biol, 9*(2), 026011.

Bally, M., Graule, M., Parra, F., Larson, G., & Hook, F. (2013). A virus biosensor with single virus-particle sensitivity based on fluorescent vesicle labels and equilibrium fluctuation analysis. *Biointerphases, 8*(1), 5.

Bally, M., Gunnarsson, A., Svensson, L., Larson, G., Zhdanov, V. P., & Hook, F. (2011). Interaction of single viruslike particles with vesicles containing glycosphingolipids. *Phys Rev Lett, 107*(18), 188103.

Barcena, J., Verdaguer, N., Roca, R., Morales, M., Angulo, I., Risco, C., et al. (2004). The coat protein of Rabbit hemorrhagic disease virus contains a molecular switch at the N-terminal region facing the inner surface of the capsid. *Virology, 322*(1), 118-134.

Baron, R. C., Greenberg, H. B., Cukor, G., & Blacklow, N. R. (1984). Serological responses among teenagers after natural exposure to Norwalk virus. *J Infect Dis, 150*(4), 531-534.

Barthelson, R., Mobasseri, A., Zopf, D., & Simon, P. (1998). Adherence of Streptococcus pneumoniae to respiratory epithelial cells is inhibited by sialylated oligosaccharides. *Infect Immun, 66*(4), 1439-1444.

Bertolotti-Ciarlet, A., White, L. J., Chen, R., Prasad, B. V., & Estes, M. K. (2002). Structural requirements for the assembly of Norwalk virus-like particles. *J Virol, 76*(8), 4044-4055.

Bertozzi, C. R., & Rabuka, D. (2009). Structural Basis of Glycan Diversity. In A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart & M. E. Etzler (Eds.), *Essentials of Glycobiology* (2nd ed.). Cold Spring Harbor (NY).

Bhella, D., Gatherer, D., Chaudhry, Y., Pink, R., & Goodfellow, I. G. (2008). Structural insights into calicivirus attachment and uncoating. *J Virol, 82*(16), 8051-8058.

Bhende, Y. M., Deshpande, C. K., Bhatia, H. M., Sanger, R., Race, R. R., Morgan, W. T., et al. (1952). A "new" blood group character related to the ABO system. *Lancet, 1*(6714), 903-904.

Bishop, W. P., & Ulshen, M. H. (1988). Bacterial gastroenteritis. *Pediatr Clin North Am, 35*(1), 69-87.

Bjork, S., Breimer, M. E., Hansson, G. C., Karlsson, K. A., & Leffler, H. (1987). Structures of blood group glycosphingolipids of human small intestine. A relation between the expression of fucolipids of epithelial cells and the ABO, Le and Se phenotype of the donor. *J Biol Chem, 262*(14), 6758-6765.

Blumenfeld, O. O., & Patnaik, S. K. (2004). Allelic genes of blood group antigens: a source of human mutations and cSNPs documented in the Blood Group Antigen Gene Mutation Database. *Hum Mutat, 23*(1), 8-16.

Boga, J. A., Melon, S., Nicieza, I., De Diego, I., Villar, M., Parra, F., et al. (2004). Etiology of sporadic cases of pediatric acute gastroenteritis in asturias, Spain, and genotyping and characterization of norovirus strains involved. *J Clin Microbiol, 42*(6), 2668-2674.

Bohne, A., Lang, E., & von der Lieth, C. W. (1999). SWEET - WWW-based rapid 3D construction of oligo- and polysaccharides. *Bioinformatics, 15*(9), 767-768.

Bok, K., Abente, E. J., Realpe-Quintero, M., Mitra, T., Sosnovtsev, S. V., Kapikian, A. Z., et al. (2009). Evolutionary dynamics of GII.4 noroviruses over a 34-year period. *J Virol, 83*(22), 11890-11901.

Bok, K., Parra, G. I., Mitra, T., Abente, E., Shaver, C. K., Boon, D., et al. (2011). Chimpanzees as an animal model for human norovirus infection and vaccine development. *Proc Natl Acad Sci U S A, 108*(1), 325-330.

Borrell, B. (2008). Chemistry: power play. *Nature, 451*(7176), 240-243.

Breimer, M. E., Hansson, G. C., Karlsson, K. A., Larson, G., & Leffler, H. (2012). Glycosphingolipid composition of epithelial cells isolated along the villus axis of small intestine of a single human individual. *Glycobiology, 22*(12), 1721-1730.

Brockhaus, M., Magnani, J. L., Blaszczyk, M., Steplewski, Z., Koprowski, H., Karlsson, K. A., et al. (1981). Monoclonal antibodies directed against the human Leb blood group antigen. *J Biol Chem, 256*(24), 13223-13225.

Brockhausen, I., Schachter, H., & Stanley, P. (2009). O-GalNAc Glycans. In A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart & M. E. Etzler (Eds.), *Essentials of Glycobiology* (2nd ed.). Cold Spring Harbor (NY).

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983). CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of computational chemistry, 4*(2), 187-217.

Brown, D. A., & Rose, J. K. (1992). Sorting of GPI-anchored proteins to glycolipid-enriched membrane subdomains during transport to the apical cell surface. *Cell, 68*(3), 533-544.

Brown, K. E., Anderson, S. M., & Young, N. S. (1993). Erythrocyte P antigen: cellular receptor for B19 parvovirus. *Science, 262*(5130), 114-117.

Brown, K. E., Hibbs, J. R., Gallinella, G., Anderson, S. M., Lehman, E. D., McCarthy, P., et al. (1994). Resistance to parvovirus B19 infection due to lack of virus receptor (erythrocyte P antigen). *N Engl J Med, 330*(17), 1192-1196.

Bryce, R. A., Hillier, I. H., & Naismith, J. H. (2001). Carbohydrate-protein recognition: molecular dynamics simulations and free energy analysis of oligosaccharide binding to concanavalin A. *Biophys J, 81*(3), 1373-1388.

Bu, W., Mamedova, A., Tan, M., Xia, M., Jiang, X., & Hegde, R. S. (2008). Structural basis for the receptor binding specificity of Norwalk virus. *J Virol, 82*(11), 5340-5347.

Bucardo, F., Kindberg, E., Paniagua, M., Grahn, A., Larson, G., Vildevall, M., et al. (2009). Genetic susceptibility to symptomatic norovirus infection in Nicaragua. *J Med Virol, 81*(4), 728-735.

Bucardo, F., Nordgren, J., Carlsson, B., Kindberg, E., Paniagua, M., Mollby, R., et al. (2010). Asymptomatic norovirus infections in Nicaraguan children and its association with viral properties and histo-blood group antigens. *Pediatr Infect Dis J, 29*(10), 934-939.

Bull, R. A., Tanaka, M. M., & White, P. A. (2007). Norovirus recombination. *J Gen Virol, 88*(Pt 12), 3347-3359.

Cannon, J. L., Lindesmith, L. C., Donaldson, E. F., Saxe, L., Baric, R. S., & Vinje, J. (2009). Herd immunity to GII.4 noroviruses is supported by outbreak patient sera. *J Virol, 83*(11), 5363-5374.

Cao, S., Lou, Z., Tan, M., Chen, Y., Liu, Y., Zhang, Z., et al. (2007). Structural basis for the recognition of blood group trisaccharides by norovirus. *J Virol, 81*(11), 5949-5957.

Carlsson, B., Kindberg, E., Buesa, J., Rydell, G. E., Lidon, M. F., Montava, R., et al. (2009). The G428A nonsense mutation in FUT2 provides strong but not absolute protection against symptomatic GII.4 Norovirus infection. *PLoS One, 4*(5), e5593.

Cashman, S. M., Morris, D. J., & Kumar-Singh, R. (2004). Adenovirus type 5 pseudotyped with adenovirus type 37 fiber uses sialic acid as a cellular receptor. *Virology, 324*(1), 129-139.

Caul, E. O. (1995). Hyperemesis hiemis--a sick hazard. *J Hosp Infect, 30 Suppl*, 498-502.

Centers for Disease, C., & Prevention. (2002). Outbreak of acute gastroenteritis associated with Norwalk-like viruses among British military personnel--Afghanistan, May 2002. *MMWR Morb Mortal Wkly Rep, 51*(22), 477-479.

Centers for Disease, C., & Prevention. (2006). Multisite outbreak of norovirus associated with a franchise restaurant--Kent County, Michigan, May 2005. *MMWR Morb Mortal Wkly Rep, 55*(14), 395-397.

Centers for Disease, C., & Prevention. (2008). Norovirus outbreak in an elementary school--District of Columbia, February 2007. *MMWR Morb Mortal Wkly Rep, 56*(51-52), 1340-1343.

Chakravarty, S., Hutson, A. M., Estes, M. K., & Prasad, B. V. (2005). Evolutionary trace residues in noroviruses: importance in receptor binding, antigenicity, virion assembly, and strain diversity. *J Virol, 79*(1), 554-568.

Chan, M. C., Ho, W. S., & Sung, J. J. (2011). In vitro whole-virus binding of a norovirus genogroup II genotype 4 strain to cells of the lamina propria and Brunner's glands in the human duodenum. *J Virol, 85*(16), 8427-8430.

Chan, M. C., Wong, Y. P., Sung, J. J., & Leung, W. K. (2008). Histo-blood group antigens and susceptibility to infection with norovirus genogroup II genotype 4. *J Infect Dis, 198*(6), 940; author reply 942-943.

Cheetham, S., Souza, M., McGregor, R., Meulia, T., Wang, Q., & Saif, L. J. (2007). Binding patterns of human norovirus-like particles to buccal and intestinal tissues of gnotobiotic pigs in relation to A/H histo-blood group antigen expression. *J Virol, 81*(7), 3535-3544.

Cheetham, S., Souza, M., Meulia, T., Grimes, S., Han, M. G., & Saif, L. J. (2006). Pathogenesis of a genogroup II human norovirus in gnotobiotic pigs. *J Virol, 80*(21), 10372-10381.

Chen, R., Neill, J. D., Noel, J. S., Hutson, A. M., Glass, R. I., Estes, M. K., et al. (2004). Inter- and intragenus structural variations in caliciviruses and their functional implications. *J Virol, 78*(12), 6469-6479.

Chen, Y., Tan, M., Xia, M., Hao, N., Zhang, X. C., Huang, P., et al. (2011). Crystallography of a Lewis-binding norovirus, elucidation of strain-specificity to the polymorphic human histo-blood group antigens. *PLoS Pathog, 7*(7), e1002152.

Choi, J. M., Hutson, A. M., Estes, M. K., & Prasad, B. V. (2008). Atomic resolution structural characterization of recognition of histo-blood group antigens by Norwalk virus. *Proc Natl Acad Sci U S A, 105*(27), 9175-9180.

Clausen, H., & Hakomori, S. (1989). ABH and related histo-blood group antigens; immunochemical differences in carrier isotypes and their distribution. *Vox Sang, 56*(1), 1-20.

Clay, S., Maherchandani, S., Malik, Y. S., & Goyal, S. M. (2006). Survival on uncommon fomites of feline calicivirus, a surrogate of noroviruses. *Am J Infect Control, 34*(1), 41-43.

Collins, B. E., & Paulson, J. C. (2004). Cell surface biology mediated by low affinity multivalent protein-glycan interactions. *Curr Opin Chem Biol, 8*(6), 617-625.

Cook, D. G., Fantini, J., Spitalnik, S. L., & Gonzalez-Scarano, F. (1994). Binding of human immunodeficiency virus type I (HIV-1) gp120 to galactosylceramide (GalCer): relationship to the V3 loop. *Virology, 201*(2), 206-214.

Copeland, R., Balasubramaniam, A., Tiwari, V., Zhang, F., Bridges, A., Linhardt, R. J., et al. (2008). Using a 3-O-sulfated heparin octasaccharide to inhibit the entry of herpes simplex virus type 1. *Biochemistry, 47*(21), 5774-5783.

Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., et al. (1995). A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society, 117*(19), 5179-5197.

Cylwik, B., Lipartowska, K., Chrostek, L., & Gruszewska, E. (2013a). Congenital disorders of glycosylation. Part II. Defects of protein O-glycosylation. *Acta Biochim Pol, 60*(3), 361-368.

Cylwik, B., Naklicki, M., Chrostek, L., & Gruszewska, E. (2013b). Congenital disorders of glycosylation. Part I. Defects of protein N-glycosylation. *Acta Biochim Pol, 60*(2), 151-161.

Dam, T. K., & Brewer, C. F. (2002). Thermodynamic studies of lectin-carbohydrate interactions by isothermal titration calorimetry. *Chem Rev, 102*(2), 387-429.

Daughenbaugh, K. F., Fraser, C. S., Hershey, J. W., & Hardy, M. E. (2003). The genome-linked protein VPg of the Norwalk virus binds eIF3, suggesting its role in translation initiation complex recruitment. *EMBO J, 22*(11), 2852-2859.

de Jong, J. C., Claas, E. C., Osterhaus, A. D., Webster, R. G., & Lim, W. L. (1997). A pandemic warning? *Nature, 389*(6651), 554.

de Rougemont, A., Ruvoen-Clouet, N., Simon, B., Estienney, M., Elie-Caille, C., Aho, S., et al. (2011). Qualitative and quantitative analysis of the binding of GII.4 norovirus variants onto human blood group antigens. *J Virol, 85*(9), 4057-4070.

de Wit, M. A., Widdowson, M. A., Vennema, H., de Bruin, E., Fernandes, T., & Koopmans, M. (2007). Large outbreak of norovirus: the baker who should have known better. *J Infect, 55*(2), 188-193.

de Witte, L., Zoughlami, Y., Aengeneyndt, B., David, G., van Kooyk, Y., Gissmann, L., et al. (2007). Binding of human papilloma virus L1 virus-like particles to dendritic cells is mediated through heparan sulfates and induces immune activation. *Immunobiology, 212*(9-10), 679-691.

Debbink, K., Donaldson, E. F., Lindesmith, L. C., & Baric, R. S. (2012a). Genetic mapping of a highly variable norovirus GII.4 blockade epitope: potential role in escape from human herd immunity. *J Virol, 86*(2), 1214-1226.

Debbink, K., Lindesmith, L. C., Donaldson, E. F., & Baric, R. S. (2012b). Norovirus immunity and the great escape. *PLoS Pathog, 8*(10), e1002921.

Debbink, K., Lindesmith, L. C., Donaldson, E. F., Costantini, V., Beltramello, M., Corti, D., et al. (2013). Emergence of new pandemic GII.4 Sydney norovirus strain correlates with escape from herd immunity. *J Infect Dis, 208*(11), 1877-1887.

Debbink, K., Lindesmith, L. C., Ferris, M. T., Swanstom, J., Beltramello, M., Corti, D., et al. (2014). Within Host Evolution Results in Antigenically Distinct GII.4 Noroviruses. *J Virol*.

Delorme, C., Brussow, H., Sidoti, J., Roche, N., Karlsson, K. A., Neeser, J. R., et al. (2001). Glycosphingolipid binding specificities of rotavirus: identification of a sialic acid-binding epitope. *J Virol, 75*(5), 2276-2287.

Desai, R., Hembree, C. D., Handel, A., Matthews, J. E., Dickey, B. W., McDonald, S., et al. (2012). Severe outcomes are associated with genogroup 2 genotype 4 norovirus outbreaks: a systematic literature review. *Clin Infect Dis, 55*(2), 189-193.

Dickson, C. J., Madej, B. D., Skjevik, Å. A., Betz, R. M., Teigen, K., Gould, I. R., et al. (2014). Lipid14: The Amber Lipid Force Field. *Journal of Chemical Theory and Computation*.

Dingle, K. E., & Norovirus Infection Control in Oxfordshire Communities, H. (2004). Mutation in a Lordsdale norovirus epidemic strain as a potential indicator of transmission routes. *J Clin Microbiol, 42*(9), 3950-3957.

Dolin, R., Blacklow, N. R., DuPont, H., Buscho, R. F., Wyatt, R. G., Kasel, J. A., et al. (1972). Biological properties of Norwalk agent of acute infectious nonbacterial gastroenteritis. *Proc Soc Exp Biol Med, 140*(2), 578-583.

Donaldson, E. F., Lindesmith, L. C., Lobue, A. D., & Baric, R. S. (2008). Norovirus pathogenesis: mechanisms of persistence and immune evasion in human populations. *Immunol Rev, 225*, 190-211.

Donaldson, E. F., Lindesmith, L. C., Lobue, A. D., & Baric, R. S. (2010). Viral shape-shifting: norovirus evasion of the human immune system. *Nat Rev Microbiol, 8*(3), 231-241.

Dou, D., He, G., Mandadapu, S. R., Aravapalli, S., Kim, Y., Chang, K. O., et al. (2012). Inhibition of noroviruses by piperazine derivatives. *Bioorg Med Chem Lett, 22*(1), 377-379.

Dou, D., Tiew, K. C., He, G., Mandadapu, S. R., Aravapalli, S., Alliston, K. R., et al. (2011). Potent inhibition of Norwalk virus by cyclic sulfamide derivatives. *Bioorg Med Chem, 19*(20), 5975-5983.

Doyle, T. J., Stark, L., Hammond, R., & Hopkins, R. S. (2009). Outbreaks of noroviral gastroenteritis in Florida, 2006-2007. *Epidemiol Infect, 137*(5), 617-625.

Duizer, E., Schwab, K. J., Neill, F. H., Atmar, R. L., Koopmans, M. P., & Estes, M. K. (2004). Laboratory efforts to cultivate noroviruses. *J Gen Virol, 85*(Pt 1), 79-87.

Durrant, J. D., & McCammon, J. A. (2011). Molecular dynamics simulations and drug discovery. *BMC Biol, 9*, 71.

Eden, J. S., Hewitt, J., Lim, K. L., Boni, M. F., Merif, J., Greening, G., et al. (2014). The emergence and evolution of the novel epidemic norovirus GII.4 variant Sydney 2012. *Virology, 450-451*, 106-113.

Eden, J. S., Tanaka, M. M., Boni, M. F., Rawlinson, W. D., & White, P. A. (2013). Recombination within the pandemic norovirus GII.4 lineage. *J Virol, 87*(11), 6270-6282.

Ernst, B., & Magnani, J. L. (2009). From carbohydrate leads to glycomimetic drugs. *Nat Rev Drug Discov, 8*(8), 661-677.

Esko, J. D., Kimata, K., & Lindahl, U. (2009). Proteoglycans and Sulfated Glycosaminoglycans. In A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart & M. E. Etzler (Eds.), *Essentials of Glycobiology* (2nd ed.). Cold Spring Harbor (NY).

Evans, M. R., Meldrum, R., Lane, W., Gardner, D., Ribeiro, C. D., Gallimore, C. I., et al. (2002). An outbreak of viral gastroenteritis following environmental contamination at a concert hall. *Epidemiol Infect, 129*(2), 355-360.

Fadda, E., & Woods, R. J. (2010). Molecular simulations of carbohydrates and protein-carbohydrate interactions: motivation, issues and prospects. *Drug Discov Today, 15*(15-16), 596-609.

Fankhauser, R. L., Monroe, S. S., Noel, J. S., Humphrey, C. D., Bresee, J. S., Parashar, U. D., et al. (2002). Epidemiologic and molecular trends of "Norwalk-like viruses" associated with outbreaks of gastroenteritis in the United States. *J Infect Dis, 186*(1), 1-7.

Fantini, J., Cook, D. G., Nathanson, N., Spitalnik, S. L., & Gonzalez-Scarano, F. (1993). Infection of colonic epithelial cell lines by type 1 human immunodeficiency virus is associated with cell

surface expression of galactosylceramide, a potential alternative gp120 receptor. *Proc Natl Acad Sci U S A, 90*(7), 2700-2704.
Farkas, T., Sestak, K., Wei, C., & Jiang, X. (2008). Characterization of a rhesus monkey calicivirus representing a new genus of Caliciviridae. *J Virol, 82*(11), 5408-5416.
Feng, X., & Jiang, X. (2007). Library screen for inhibitors targeting norovirus binding to histo-blood group antigen receptors. *Antimicrob Agents Chemother, 51*(1), 324-331.
Fernandez-Mateos, P., Cailleau, A., Henry, S., Costache, M., Elmgren, A., Svensson, L., et al. (1998). Point mutations and deletion responsible for the Bombay H null and the Reunion H weak blood groups. *Vox Sang, 75*(1), 37-46.
Fiege, B., Rademacher, C., Cartmell, J., Kitov, P. I., Parra, F., & Peters, T. (2012). Molecular details of the recognition of blood group antigens by a human norovirus as determined by STD NMR spectroscopy. *Angew Chem Int Ed Engl, 51*(4), 928-932.
Freeze, H. H., & Elbein, A. D. (2009). Glycosylation Precursors. In A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart & M. E. Etzler (Eds.), *Essentials of Glycobiology* (2nd ed.). Cold Spring Harbor (NY).
Frenck, R., Bernstein, D. I., Xia, M., Huang, P., Zhong, W., Parker, S., et al. (2012). Predicting susceptibility to norovirus GII.4 by use of a challenge model involving humans. *J Infect Dis, 206*(9), 1386-1393.
Fretz, R., Schmid, D., Jelovcan, S., Tschertou, R., Krassnitzer, E., Schirmer, M., et al. (2009). An outbreak of norovirus gastroenteritis in an Austrian hospital, winter 2006-2007. *Wien Klin Wochenschr, 121*(3-4), 137-143.
Friesema, I. H., Vennema, H., Heijne, J. C., de Jager, C. M., Morroy, G., van den Kerkhof, J. H., et al. (2009). Norovirus outbreaks in nursing homes: the evaluation of infection control measures. *Epidemiol Infect, 137*(12), 1722-1733.
Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., et al. (2004). Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem, 47*(7), 1739-1749.
Futerman, A. H., & van Meer, G. (2004). The cell biology of lysosomal storage disorders. *Nat Rev Mol Cell Biol, 5*(7), 554-565.
Gadella, B. M., Hammache, D., Pieroni, G., Colenbrander, B., van Golde, L. M., & Fantini, J. (1998). Glycolipids as potential binding sites for HIV: topology in the sperm plasma membrane in relation to the regulation of membrane fusion. *J Reprod Immunol, 41*(1-2), 233-253.
Gaensslen, R. E., Bell, S. C., & Lee, H. C. (1987). Distributions of genetic markers in United States populations: I. Blood group and secretor systems. *J Forensic Sci, 32*(4), 1016-1058.
Gallimore, C. I., Cubitt, D., du Plessis, N., & Gray, J. J. (2004). Asymptomatic and symptomatic excretion of noroviruses during a hospital outbreak of gastroenteritis. *J Clin Microbiol, 42*(5), 2271-2274.
Gao, Y., Jin, M., Cong, X., Duan, Z., Li, H. Y., Guo, X., et al. (2011). Clinical and molecular epidemiologic analyses of norovirus-associated sporadic gastroenteritis in adults from Beijing, China. *J Med Virol, 83*(6), 1078-1085.
Garcia-Hernandez, E., & Hernandez-Arana, A. (1999). Structural bases of lectin-carbohydrate affinities: comparison with protein-folding energetics. *Protein Sci, 8*(5), 1075-1086.
Garcia-Hernandez, E., Zubillaga, R. A., Rodriguez-Romero, A., & Hernandez-Arana, A. (2000). Stereochemical metrics of lectin-carbohydrate interactions: comparison with protein-protein interfaces. *Glycobiology, 10*(10), 993-1000.
Ge, H., Wang, Y., Li, C., Chen, N., Xie, Y., Xu, M., et al. (2013). Molecular dynamics-based virtual screening: accelerating the drug discovery process by high-performance computing. *J Chem Inf Model, 53*(10), 2757-2764.
Ghazarian, H., Idoni, B., & Oppenheimer, S. B. (2011). A glycobiology review: carbohydrates, lectins and implications in cancer therapeutics. *Acta Histochem, 113*(3), 236-247.
Giannasca, K. T., Giannasca, P. J., & Neutra, M. R. (1996). Adherence of Salmonella typhimurium to Caco-2 cells: identification of a glycoconjugate receptor. *Infect Immun, 64*(1), 135-145.
Glass, R. I., Parashar, U. D., & Estes, M. K. (2009). Norovirus gastroenteritis. *N Engl J Med, 361*(18), 1776-1785.

Gomez, E. B. (2008). Lessons learned from an elementary school norovirus outbreak. *J Sch Nurs, 24*(6), 388-397.

Graham, D. Y., Jiang, X., Tanaka, T., Opekun, A. R., Madore, H. P., & Estes, M. K. (1994). Norwalk virus infection of volunteers: new insights based on improved assays. *J Infect Dis, 170*(1), 34-43.

Gray, J. J., Cunliffe, C., Ball, J., Graham, D. Y., Desselberger, U., & Estes, M. K. (1994). Detection of immunoglobulin M (IgM), IgA, and IgG Norwalk virus-specific antibodies by indirect enzyme-linked immunosorbent assay with baculovirus-expressed Norwalk virus capsid antigen in adult volunteers challenged with Norwalk virus. *J Clin Microbiol, 32*(12), 3059-3063.

Green, K. Y., Ando, T., Balayan, M. S., Berke, T., Clarke, I. N., Estes, M. K., et al. (2000). Taxonomy of the caliciviruses. *J Infect Dis, 181 Suppl 2*, S322-330.

Greig, J. D., & Lee, M. B. (2009). Enteric outbreaks in long-term care facilities and recommendations for prevention: a review. *Epidemiol Infect, 137*(2), 145-155.

Grotto, I., Huerta, M., Balicer, R. D., Halperin, T., Cohen, D., Orr, N., et al. (2004). An outbreak of norovirus gastroenteritis on an Israeli military base. *Infection, 32*(6), 339-343.

Guerrant, R. L., & Bobak, D. A. (1991). Bacterial and protozoal gastroenteritis. *N Engl J Med, 325*(5), 327-340.

Guix, S., Asanaka, M., Katayama, K., Crawford, S. E., Neill, F. H., Atmar, R. L., et al. (2007). Norwalk virus RNA is infectious in mammalian cells. *J Virol, 81*(22), 12238-12248.

Gunnarsson, A., Jonsson, P., Marie, R., Tegenfeldt, J. O., & Hook, F. (2008). Single-molecule detection and mismatch discrimination of unlabeled DNA targets. *Nano Lett, 8*(1), 183-188.

Gutiérrez-Escolano, A. L., Brito, Z. U., del Angel, R. M., & Jiang, X. (2000). Interaction of cellular proteins with the 5′ end of Norwalk virus genomic RNA. *Journal of virology, 74*(18), 8558-8562.

Hakomori, S. (1990). Bifunctional role of glycosphingolipids. Modulators for transmembrane signaling and mediators for cellular interactions. *J Biol Chem, 265*(31), 18713-18716.

Hakomori, S. (2003). Structure, organization, and function of glycosphingolipids in membrane. *Current opinion in hematology, 10*(1), 16-24.

Hakomori, S. (2004). Carbohydrate-to-carbohydrate interaction, through glycosynapse, as a basis of cell recognition and membrane organization. *Glycoconj J, 21*(3-4), 125-137.

Hakomori, S. I. (2000). Cell adhesion/recognition and signal transduction through glycosphingolipid microdomain. *Glycoconj J, 17*(3 -4), 143-151.

Halgren, T. A., Murphy, R. B., Friesner, R. A., Beard, H. S., Frye, L. L., Pollard, W. T., et al. (2004). Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J Med Chem, 47*(7), 1750-1759.

Halim, A., Ruetschi, U., Larson, G., & Nilsson, J. (2013). LC-MS/MS characterization of O-glycosylation sites and glycan structures of human cerebrospinal fluid glycoproteins. *J Proteome Res, 12*(2), 573-584.

Halperin, I., Ma, B., Wolfson, H., & Nussinov, R. (2002). Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins, 47*(4), 409-443.

Halperin, T., Vennema, H., Koopmans, M., Kahila Bar-Gal, G., Kayouf, R., Sela, T., et al. (2008). No association between histo-blood group antigens and susceptibility to clinical infections with genogroup II norovirus. *J Infect Dis, 197*(1), 63-65.

Hamano, M., Kuzuya, M., Fujii, R., Ogura, H., & Yamada, M. (2005). Epidemiology of acute gastroenteritis outbreaks caused by Noroviruses in Okayama, Japan. *J Med Virol, 77*(2), 282-289.

Hanfland, P., & Graham, H. A. (1981). Immunochemistry of the Lewis-blood-group system: partial characterization of Le(a)-, Le(b)-, and H-type 1 (LedH)-blood-group active glycosphingolipids from human plasma. *Arch Biochem Biophys, 210*(1), 383-395.

Hanfland, P., Kordowicz, M., Peter-Katalinic, J., Pfannschmidt, G., Crawford, R. J., Graham, H. A., et al. (1986). Immunochemistry of the Lewis blood-group system: isolation and structures of Lewis-c active and related glycosphingolipids from the plasma of blood-group O Le(a-b-) nonsecretors. *Arch Biochem Biophys, 246*(2), 655-672.

Hansen, S., Stamm-Balderjahn, S., Zuschneid, I., Behnke, M., Ruden, H., Vonberg, R. P., et al. (2007). Closure of medical departments during nosocomial outbreaks: data from a systematic analysis of the literature. *J Hosp Infect, 65*(4), 348-353.

Hansman, G. S., Biertumpfel, C., Georgiev, I., McLellan, J. S., Chen, L., Zhou, T., et al. (2011). Crystal structures of GII.10 and GII.12 norovirus protruding domains in complex with histo-blood group antigens reveal details for a potential site of vulnerability. *J Virol, 85*(13), 6687-6701.

Hansman, G. S., Jiang, X. J., & Green, K. Y. (2010). *Caliciviruses: molecular and cellular virology*: Horizon Scientific Press.

Hansman, G. S., Shahzad-Ul-Hussan, S., McLellan, J. S., Chuang, G. Y., Georgiev, I., Shimoike, T., et al. (2012). Structural basis for norovirus inhibition and fucose mimicry by citrate. *J Virol, 86*(1), 284-292.

Hansson, G. C., Karlsson, K. A., Larson, G., McKibbin, J. M., Blaszczyk, M., Herlyn, M., et al. (1983). Mouse monoclonal antibodies against human cancer cell lines with specificities for blood group and related antigens. Characterization by antibody binding to glycosphingolipids in a chromatogram binding assay. *J Biol Chem, 258*(7), 4091-4097.

Hansson, G. C., Karlsson, K. A., Larson, G., Stromberg, N., Thurin, J., Orvell, C., et al. (1984). A novel approach to the study of glycolipid receptors for viruses. Binding of Sendai virus to thin-layer chromatograms. *FEBS Lett, 170*(1), 15-18.

Hardy, M. E. (2005). Norovirus protein structure and function. *FEMS Microbiol Lett, 253*(1), 1-8.

Hardy, M. E., Kramer, S. F., Treanor, J. J., & Estes, M. K. (1997). Human calicivirus genogroup II capsid sequence diversity revealed by analyses of the prototype Snow Mountain agent. *Arch Virol, 142*(7), 1469-1479.

Harouse, J. M., Bhat, S., Spitalnik, S. L., Laughlin, M., Stefano, K., Silberberg, D. H., et al. (1991). Inhibition of entry of HIV-1 in neural cell lines by antibodies against galactosyl ceramide. *Science, 253*(5017), 320-323.

Harrington, P. R., Lindesmith, L., Yount, B., Moe, C. L., & Baric, R. S. (2002). Binding of Norwalk virus-like particles to ABH histo-blood group antigens is blocked by antisera from infected human volunteers or experimentally vaccinated mice. *J Virol, 76*(23), 12335-12343.

Harrington, P. R., Vinje, J., Moe, C. L., & Baric, R. S. (2004). Norovirus capture with histo-blood group antigens reveals novel virus-ligand interactions. *J Virol, 78*(6), 3035-3045.

Harris, J. P., Edmunds, W. J., Pebody, R., Brown, D. W., & Lopman, B. A. (2008). Deaths from norovirus among the elderly, England and Wales. *Emerg Infect Dis, 14*(10), 1546-1552.

Hart, G. W., & Akimoto, Y. (2009). The O-GlcNAc Modification. In A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart & M. E. Etzler (Eds.), *Essentials of Glycobiology* (2nd ed.). Cold Spring Harbor (NY).

Hattrup, C. L., & Gendler, S. J. (2008). Structure and function of the cell surface (tethered) mucins. *Annu. Rev. Physiol., 70*, 431-457.

Helenius, A., & Aebi, M. (2001). Intracellular functions of N-linked glycans. *Science, 291*(5512), 2364-2369.

Hennessy, E. P., Green, A. D., Connor, M. P., Darby, R., & MacDonald, P. (2003). Norwalk virus infection and disease is associated with ABO histo-blood group type. *J Infect Dis, 188*(1), 176-177.

Hjelm, B. E., Kilbourne, J., & Herbst-Kralovetz, M. M. (2013). TLR7 and 9 agonists are highly effective mucosal adjuvants for norovirus virus-like particle vaccines. *Hum Vaccin Immunother, 10*(3).

Hjertqvist, M., Johansson, A., Svensson, N., Abom, P. E., Magnusson, C., Olsson, M., et al. (2006). Four outbreaks of norovirus gastroenteritis after consuming raspberries, Sweden, June-August 2006. *Euro Surveill, 11*(9), E060907 060901.

Hoffmann, D., Mauroy, A., Seebach, J., Simon, V., Wantia, N., & Protzer, U. (2013). New norovirus classified as a recombinant GII.g/GII.1 causes an extended foodborne outbreak at a university hospital in Munich. *J Clin Virol, 58*(1), 24-30.

Holmner, A., Askarieh, G., Okvist, M., & Krengel, U. (2007). Blood group antigen recognition by Escherichia coli heat-labile enterotoxin. *J Mol Biol, 371*(3), 754-764.

Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., & Simmerling, C. (2006). Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins, 65*(3), 712-725.

Hu, L., Crawford, S. E., Czako, R., Cortes-Penfield, N. W., Smith, D. F., Le Pendu, J., et al. (2012). Cell attachment protein VP8* of a human rotavirus specifically interacts with A-type histo-blood group antigen. *Nature, 485*(7397), 256-259.

Huang, H., Li, Y., Sadaoka, T., Tang, H., Yamamoto, T., Yamanishi, K., et al. (2006). Human herpesvirus 6 envelope cholesterol is required for virus entry. *J Gen Virol, 87*(Pt 2), 277-285.

Huang, P., Farkas, T., Marionneau, S., Zhong, W., Ruvoen-Clouet, N., Morrow, A. L., et al. (2003). Noroviruses bind to human ABO, Lewis, and secretor histo-blood group antigens: identification of 4 distinct strain-specific patterns. *J Infect Dis, 188*(1), 19-31.

Huang, P., Farkas, T., Zhong, W., Tan, M., Thornton, S., Morrow, A. L., et al. (2005). Norovirus and histo-blood group antigens: demonstration of a wide spectrum of strain specificities and classification of two major binding groups among multiple binding patterns. *J Virol, 79*(11), 6714-6722.

Huang, P., Xia, M., Tan, M., Zhong, W., Wei, C., Wang, L., et al. (2012). Spike protein VP8* of human rotavirus recognizes histo-blood group antigens in a type specific manner. *J Virol, 86*(9), 4833-4843.

Hutson, A. M., Airaud, F., LePendu, J., Estes, M. K., & Atmar, R. L. (2005). Norwalk virus infection associates with secretor status genotyped from sera. *J Med Virol, 77*(1), 116-120.

Hutson, A. M., Atmar, R. L., & Estes, M. K. (2004a). Norovirus disease: changing epidemiology and host susceptibility factors. *Trends Microbiol, 12*(6), 279-287.

Hutson, A. M., Atmar, R. L., Graham, D. Y., & Estes, M. K. (2002). Norwalk virus infection and disease is associated with ABO histo-blood group type. *J Infect Dis, 185*(9), 1335-1337.

Hutson, A. M., Atmar, R. L., Marcus, D. M., & Estes, M. K. (2003). Norwalk virus-like particle hemagglutination by binding to h histo-blood group antigens. *J Virol, 77*(1), 405-415.

Hutson, A. M., Estes, M. K., & Atmar, R. L. (2004b). Re: Nosocomial outbreak of norovirus gastroenteritis and investigation of ABO histo-blood group type in infected staff and patients. *J Hosp Infect, 58*(2), 163-164.

Huynen, P., Mauroy, A., Martin, C., Savadogo, L. G., Boreux, R., Thiry, E., et al. (2013). Molecular epidemiology of norovirus infections in symptomatic and asymptomatic children from Bobo Dioulasso, Burkina Faso. *J Clin Virol, 58*(3), 515-521.

Hyde, J. L., & Mackenzie, J. M. (2010). Subcellular localization of the MNV-1 ORF1 proteins and their potential roles in the formation of the MNV-1 replication complex. *Virology, 406*(1), 138-148.

Höök, F., Rodahl, M., Brzezinski, P., & Kasemo, B. (1998). Energy dissipation kinetics for protein and antibody-antigen adsorption under shear oscillation on a quartz crystal microbalance. *Langmuir, 14*(4), 729-734.

Imberty, A., Mikros, E., Koca, J., Mollicone, R., Oriol, R., & Perez, S. (1995). Computer simulation of histo-blood group oligosaccharides: energy maps of all constituting disaccharides and potential energy surfaces of 14 ABH and Lewis carbohydrate antigens. *Glycoconj J, 12*(3), 331-349.

Imberty, A., Mollicone, R., Mikros, E., Carrupt, P. A., Perez, S., & Oriol, R. (1996). How do antibodies and lectins recognize histo-blood group antigens? A 3D-QSAR study by comparative molecular field analysis (CoMFA). *Bioorg Med Chem, 4*(11), 1979-1988.

Imberty, A., & Varrot, A. (2008). Microbial recognition of human cell surface glycoconjugates. *Curr Opin Struct Biol, 18*(5), 567-576.

Janshoff, A., Galla, H. J., & Steinem, C. (2000). Piezoelectric Mass-Sensing Devices as Biosensors-An Alternative to Optical Biosensors? *Angew Chem Int Ed Engl, 39*(22), 4004-4032.

Jiang, X., Wang, M., Graham, D. Y., & Estes, M. K. (1992). Expression, self-assembly, and antigenicity of the Norwalk virus capsid protein. *J Virol, 66*(11), 6527-6532.

Jin, M., He, Y., Li, H., Huang, P., Zhong, W., Yang, H., et al. (2013). Two gastroenteritis outbreaks caused by GII Noroviruses: host susceptibility and HBGA phenotypes. *PLoS One, 8*(3), e58605.

Johnson, P. C., Mathewson, J. J., DuPont, H. L., & Greenberg, H. B. (1990). Multiple-challenge study of host susceptibility to Norwalk gastroenteritis in US adults. *J Infect Dis, 161*(1), 18-21.

Johnson, P. H., Donald, A. S., Feeney, J., & Watkins, W. M. (1992). Reassessment of the acceptor specificity and general properties of the Lewis blood-group gene associated alpha-3/4-fucosyltransferase purified from human milk. *Glycoconj J, 9*(5), 251-264.

Johnson, P. H., & Watkins, W. M. (1992). Purification of the Lewis blood-group gene associated alpha-3/4-fucosyltransferase from human milk: an enzyme transferring fucose primarily to type 1 and lactose-based oligosaccharide chains. *Glycoconj J, 9*(5), 241-249.

Jones, G., Willett, P., Glen, R. C., Leach, A. R., & Taylor, R. (1997). Development and validation of a genetic algorithm for flexible docking. *J Mol Biol, 267*(3), 727-748.

Kanagawa, M., Yu, C. C., Ito, C., Fukada, S., Hozoji-Inada, M., Chiyo, T., et al. (2013). Impaired viability of muscle precursor cells in muscular dystrophy with glycosylation defects and amelioration of its severe phenotype by limited gene expression. *Hum Mol Genet, 22*(15), 3003-3015.

Kanerva, M., Maunula, L., Lappalainen, M., Mannonen, L., von Bonsdorff, C. H., & Anttila, V. J. (2009). Prolonged norovirus outbreak in a Finnish tertiary care hospital caused by GII.4-2006b subvariants. *J Hosp Infect, 71*(3), 206-213.

Kapikian, A., Feinstone, S., Purcell, R., Wyatt, R., Thornhill, T., Kalica, A., et al. (1975). Detection and identification by immune electron microscopy of fastidious agents associated with respiratory illness, acute nonbacterial gastroenteritis, and hepatitis A. *Perspect Virol, 9*, 9-47.

Kapikian, A. Z. (2000). The discovery of the 27-nm Norwalk virus: an historic perspective. *J Infect Dis, 181 Suppl 2*, S295-302.

Kapikian, A. Z., Wyatt, R. G., Dolin, R., Thornhill, T. S., Kalica, A. R., & Chanock, R. M. (1972). Visualization by immune electron microscopy of a 27-nm particle associated with acute infectious nonbacterial gastroenteritis. *J Virol, 10*(5), 1075-1081.

Kaplon, J., Guenau, E., Asdrubal, P., Pothier, P., & Ambert-Balay, K. (2011). Possible novel nebovirus genotype in cattle, France. *Emerg Infect Dis, 17*(6), 1120-1123.

Karlsson, K.-A. (1982). Glycosphingolipids and surface membranes. *Biological membranes, 4*, 1-74.

Karlsson, K. A. (1987). Preparation of total nonacid glycolipids for overlay analysis of receptors for bacteria and viruses and for other studies. *Methods Enzymol, 138*, 212-220.

Karlsson, K. A. (1995). Microbial recognition of target-cell glycoconjugates. *Curr Opin Struct Biol, 5*(5), 622-635.

Karlsson, K. A., & Larson, G. (1981). Molecular characterization of cell surface antigens of fetal tissue. Detailed analysis of glycosphingolipids of meconium of a human O Le(a--b+) secretor. *J Biol Chem, 256*(7), 3512-3524.

Karlsson, K. A., & Stromberg, N. (1987). Overlay and solid-phase analysis of glycolipid receptors for bacteria and viruses. *Methods Enzymol, 138*, 220-232.

Karst, S. M., Wobus, C. E., Lay, M., Davidson, J., & Virgin, H. W. t. (2003). STAT1-dependent innate immunity to a Norwalk-like virus. *Science, 299*(5612), 1575-1578.

Katayama, K., Shirato-Horikoshi, H., Kojima, S., Kageyama, T., Oka, T., Hoshino, F., et al. (2002). Phylogenetic analysis of the complete genome of 18 Norwalk-like viruses. *Virology, 299*(2), 225-239.

Katpally, U., Wobus, C. E., Dryden, K., Virgin, H. W. t., & Smith, T. J. (2008). Structure of antibody-neutralized murine norovirus and unexpected differences from viruslike particles. *J Virol, 82*(5), 2079-2088.

Kaufmann, B., Baxa, U., Chipman, P. R., Rossmann, M. G., Modrow, S., & Seckler, R. (2005). Parvovirus B19 does not bind to membrane-associated globoside in vitro. *Virology, 332*(1), 189-198.

Kele, B., Abrok, M. P., & Deak, J. (2009). Sporadic norovirus infections among hospitalized and non-hospitalized 0-3-year-old infants. *Scand J Infect Dis, 41*(1), 67-69.

Keller, C., & Kasemo, B. (1998). Surface specific kinetics of lipid vesicle adsorption measured with a quartz crystal microbalance. *Biophysical journal, 75*(3), 1397-1402.

Kelly, R. J., Rouquier, S., Giorgi, D., Lennon, G. G., & Lowe, J. B. (1995). Sequence and expression of a candidate for the human Secretor blood group alpha(1,2)fucosyltransferase gene (FUT2).

Homozygosity for an enzyme-inactivating nonsense mutation commonly correlates with the non-secretor phenotype. *J Biol Chem, 270*(9), 4640-4649.

Khan, R. R., Lawson, A. D., Minnich, L. L., Martin, K., Nasir, A., Emmett, M. K., et al. (2009). Gastrointestinal norovirus infection associated with exacerbation of inflammatory bowel disease. *J Pediatr Gastroenterol Nutr, 48*(3), 328-333.

Kindberg, E., Akerlind, B., Johnsen, C., Knudsen, J. D., Heltberg, O., Larson, G., et al. (2007). Host genetic resistance to symptomatic norovirus (GGII.4) infections in Denmark. *J Clin Microbiol, 45*(8), 2720-2722.

Kirschner, K. N., Yongye, A. B., Tschampel, S. M., Gonzalez-Outeirino, J., Daniels, C. R., Foley, B. L., et al. (2008). GLYCAM06: a generalizable biomolecular force field. Carbohydrates. *J Comput Chem, 29*(4), 622-655.

Koda, Y., Soejima, M., Johnson, P. H., Smart, E., & Kimura, H. (1997). Missense mutation of FUT1 and deletion of FUT2 are responsible for Indian Bombay phenotype of ABO blood group system. *Biochem Biophys Res Commun, 238*(1), 21-25.

Koopmans, M., & Duizer, E. (2004). Foodborne viruses: an emerging problem. *Int J Food Microbiol, 90*(1), 23-41.

Krenzer, M. E. (2012). Viral gastroenteritis in the adult population: the GI peril. *Crit Care Nurs Clin North Am, 24*(4), 541-553.

Kroneman, A., Harris, J., Vennema, H., Duizer, E., van Duynhoven, Y., Gray, J., et al. (2008a). Data quality of 5 years of central norovirus outbreak reporting in the European Network for food-borne viruses. *J Public Health (Oxf), 30*(1), 82-90.

Kroneman, A., Vega, E., Vennema, H., Vinje, J., White, P. A., Hansman, G., et al. (2013). Proposal for a unified norovirus nomenclature and genotyping. *Arch Virol, 158*(10), 2059-2068.

Kroneman, A., Verhoef, L., Harris, J., Vennema, H., Duizer, E., van Duynhoven, Y., et al. (2008b). Analysis of integrated virological and epidemiological reports of norovirus outbreaks collected within the Foodborne Viruses in Europe network from 1 July 2001 to 30 June 2006. *J Clin Microbiol, 46*(9), 2959-2965.

Kubota, T., Kumagai, A., Ito, H., Furukawa, S., Someya, Y., Takeda, N., et al. (2012). Structural basis for the recognition of Lewis antigens by genogroup I norovirus. *J Virol, 86*(20), 11138-11150.

Kukowska-Latallo, J. F., Larsen, R. D., Nair, R. P., & Lowe, J. B. (1990). A cloned human cDNA determines expression of a mouse stage-specific embryonic antigen and the Lewis blood group alpha(1,3/1,4)fucosyltransferase. *Genes Dev, 4*(8), 1288-1303.

Kumar, R., Potvin, B., Muller, W. A., & Stanley, P. (1991). Cloning of a human alpha(1,3)-fucosyltransferase gene that encodes ELFT but does not confer ELAM-1 recognition on Chinese hamster ovary cell transfectants. *J Biol Chem, 266*(32), 21777-21783.

Kunze, A., Bally, M., Hook, F., & Larson, G. (2013). Equilibrium-fluctuation-analysis of single liposome binding events reveals how cholesterol and Ca2+ modulate glycosphingolipid trans-interactions. *Sci Rep, 3*, 1452.

Kuusi, M., Nuorti, J. P., Maunula, L., Minh Tran, N. N., Ratia, M., Karlsson, J., et al. (2002). A prolonged outbreak of Norwalk-like calicivirus (NLV) gastroenteritis in a rehabilitation centre due to environmental contamination. *Epidemiol Infect, 129*(1), 133-138.

Landsteiner, K. (1900). Zur Kenntnis der antifermentativen, lytischen und agglutinierenden Wirkungen des Blutserums und der Lymphe. *Zbl bakt, 27*(10), 357-362.

Larsson, M. M., Rydell, G. E., Grahn, A., Rodriguez-Diaz, J., Akerlind, B., Hutson, A. M., et al. (2006). Antibody prevalence and titer to norovirus (genogroup II) correlate with secretor (FUT2) but not with ABO phenotype or Lewis (FUT3) genotype. *J Infect Dis, 194*(10), 1422-1427.

Lay, M. K., Atmar, R. L., Guix, S., Bharadwaj, U., He, H., Neill, F. H., et al. (2010). Norwalk virus does not replicate in human macrophages or dendritic cells derived from the peripheral blood of susceptible humans. *Virology, 406*(1), 1-11.

Le Pendu, J., Cartron, J. P., Lemieux, R. U., & Oriol, R. (1985). The presence of at least two different H-blood-group-related beta-D-gal alpha-2-L-fucosyltransferases in human serum and the genetics of blood group H substances. *Am J Hum Genet, 37*(4), 749-760.

Lemieux, R. U. (1996). How water provides the impetus for molecular recognition in aqueous solution. *Accounts of chemical research, 29*(8), 373-380.

Leslie, M. (2011). Mysteries of the cell. Do lipid rafts exist? *Science, 334*(6059), 1046-1047.
Leuenberger, S., Widdowson, M. A., Feilchenfeldt, J., Egger, R., & Streuli, R. A. (2007). Norovirus outbreak in a district general hospital--new strain identified. *Swiss Med Wkly, 137*(3-4), 57-81.
Lichtenstein, R. G., & Rabinovich, G. A. (2013). Glycobiology of cell death: when glycans and lectins govern cell fate. *Cell Death Differ, 20*(8), 976-986.
Lindesmith, L., Moe, C., Lependu, J., Frelinger, J. A., Treanor, J., & Baric, R. S. (2005). Cellular and humoral immunity following Snow Mountain virus challenge. *J Virol, 79*(5), 2900-2909.
Lindesmith, L., Moe, C., Marionneau, S., Ruvoen, N., Jiang, X., Lindblad, L., et al. (2003). Human susceptibility and resistance to Norwalk virus infection. *Nat Med, 9*(5), 548-553.
Lindesmith, L. C., Costantini, V., Swanstrom, J., Debbink, K., Donaldson, E. F., Vinje, J., et al. (2013). Emergence of a norovirus GII.4 strain correlates with changes in evolving blockade epitopes. *J Virol, 87*(5), 2803-2813.
Lindesmith, L. C., Donaldson, E., Leon, J., Moe, C. L., Frelinger, J. A., Johnston, R. E., et al. (2010). Heterotypic humoral and cellular immune responses following Norwalk virus infection. *J Virol, 84*(4), 1800-1815.
Lindesmith, L. C., Donaldson, E. F., & Baric, R. S. (2011). Norovirus GII.4 strain antigenic variation. *J Virol, 85*(1), 231-242.
Lindesmith, L. C., Donaldson, E. F., Lobue, A. D., Cannon, J. L., Zheng, D. P., Vinje, J., et al. (2008). Mechanisms of GII.4 norovirus persistence in human populations. *PLoS Med, 5*(2), e31.
Lingwood, C. A. (2011). Glycosphingolipid functions. *Cold Spring Harbor perspectives in biology, 3*(7), a004788.
Lingwood, C. A., & Branch, D. R. (2011). The role of glycosphingolipids in HIV/AIDS. *Discov Med, 11*(59), 303-313.
Lingwood, D., & Simons, K. (2010). Lipid rafts as a membrane-organizing principle. *Science, 327*(5961), 46-50.
Liu, G., Kahan, S. M., Jia, Y., & Karst, S. M. (2009). Primary high-dose murine norovirus 1 infection fails to protect from secondary challenge with homologous virus. *J Virol, 83*(13), 6963-6968.
Lopman, B. A., Adak, G. K., Reacher, M. H., & Brown, D. W. (2003a). Two epidemiologic patterns of norovirus outbreaks: surveillance in England and wales, 1992-2000. *Emerg Infect Dis, 9*(1), 71-77.
Lopman, B. A., Reacher, M., Gallimore, C., Adak, G. K., Gray, J. J., & Brown, D. W. (2003b). A summertime peak of "winter vomiting disease": surveillance of noroviruses in England and Wales, 1995 to 2002. *BMC Public Health, 3*, 13.
Lopman, B. A., Reacher, M. H., Vipond, I. B., Hill, D., Perry, C., Halladay, T., et al. (2004a). Epidemiology and cost of nosocomial gastroenteritis, Avon, England, 2002-2003. *Emerg Infect Dis, 10*(10), 1827-1834.
Lopman, B. A., Reacher, M. H., Vipond, I. B., Sarangi, J., & Brown, D. W. (2004b). Clinical manifestation of norovirus gastroenteritis in health care settings. *Clin Infect Dis, 39*(3), 318-324.
Lowe, J. B. (1991). Molecular cloning, expression, and uses of mammalian glycosyltransferases. *Semin Cell Biol, 2*(5), 289-307.
Lowe, J. B. (1993). The blood group-specific human glycosyltransferases. *Baillieres Clin Haematol, 6*(2), 465-492.
Lu, X., Xiong, Y., & Silver, J. (2002). Asymmetric requirement for cholesterol in receptor-bearing but not envelope-bearing membranes for fusion mediated by ecotropic murine leukemia virus. *J Virol, 76*(13), 6701-6709.
Lutteke, T., Frank, M., & von der Lieth, C. W. (2005). Carbohydrate Structure Suite (CSS): analysis of carbohydrate 3D structures derived from the PDB. *Nucleic Acids Res, 33*(Database issue), D242-246.
Lyman, W. H., Walsh, J. F., Kotch, J. B., Weber, D. J., Gunn, E., & Vinje, J. (2009). Prospective study of etiologic agents of acute gastroenteritis outbreaks in child care centers. *J Pediatr, 154*(2), 253-257.
Lynch, M., Painter, J., Woodruff, R., Braden, C., Centers for Disease, C., & Prevention. (2006). Surveillance for foodborne-disease outbreaks--United States, 1998-2002. *MMWR Surveill Summ, 55*(10), 1-42.

Maaheimo, H., Renkonen, R., Turunen, J. P., Penttila, L., & Renkonen, O. (1995). Synthesis of a divalent sialyl Lewis x O-glycan, a potent inhibitor of lymphocyte-endothelium adhesion. Evidence that multivalency enhances the saccharide binding to L-selectin. *Eur J Biochem, 234*(2), 616-625.

Maciejewski, A., Pasenkiewicz-Gierula, M., Cramariuc, O. T., Vattulainen, I., & Róg, T. (2014). Refined OPLS All-Atom Force Field for Saturated Phosphatidylcholine Bilayers at Full Hydration. *The Journal of Physical Chemistry B*.

Magnani, J. L., Brockhaus, M., Smith, D. F., Ginsburg, V., Blaszczyk, M., Mitchell, K. F., et al. (1981). A monosialoganglioside is a monoclonal antibody-defined antigen of colon carcinoma. *Science, 212*(4490), 55-56.

Magnani, J. L., Smith, D. F., & Ginsburg, V. (1980). Detection of gangliosides that bind cholera toxin: direct binding of 125I-labeled toxin to thin-layer chromatograms. *Anal Biochem, 109*(2), 399-402.

Makary, P., Maunula, L., Niskanen, T., Kuusi, M., Virtanen, M., Pajunen, S., et al. (2009). Multiple norovirus outbreaks among workplace canteen users in Finland, July 2006. *Epidemiol Infect, 137*(3), 402-407.

Mammen, M., Choi, S.-K., & Whitesides, G. M. (1998). Polyvalent interactions in biological systems: implications for design and use of multivalent ligands and inhibitors. *Angewandte Chemie International Edition, 37*(20), 2754-2794.

Manes, S., del Real, G., & Martinez, A. C. (2003). Pathogens: raft hijackers. *Nat Rev Immunol, 3*(7), 557-568.

Marcus, D. M., & Cass, L. E. (1969). Glycosphingolipids with Lewis blood group activity: uptake by human erythrocytes. *Science, 164*(3879), 553-555.

Marionneau, S., Airaud, F., Bovin, N. V., Le Pendu, J., & Ruvoen-Clouet, N. (2005). Influence of the combined ABO, FUT2, and FUT3 polymorphism on susceptibility to Norwalk virus attachment. *J Infect Dis, 192*(6), 1071-1077.

Marionneau, S., Cailleau-Thomas, A., Rocher, J., Le Moullac-Vaidye, B., Ruvoen, N., Clement, M., et al. (2001). ABH and Lewis histo-blood group antigens, a model for the meaning of oligosaccharide diversity in the face of a changing world. *Biochimie, 83*(7), 565-573.

Marionneau, S., Ruvoen, N., Le Moullac-Vaidye, B., Clement, M., Cailleau-Thomas, A., Ruiz-Palacois, G., et al. (2002). Norwalk virus binds to histo-blood group antigens present on gastroduodenal epithelial cells of secretor individuals. *Gastroenterology, 122*(7), 1967-1977.

Martella, V., Campolo, M., Lorusso, E., Cavicchio, P., Camero, M., Bellacicco, A. L., et al. (2007). Norovirus in captive lion cub (Panthera leo). *Emerg Infect Dis, 13*(7), 1071-1073.

Martella, V., Lorusso, E., Decaro, N., Elia, G., Radogna, A., D'Abramo, M., et al. (2008). Detection and molecular characterization of a canine norovirus. *Emerg Infect Dis, 14*(8), 1306-1308.

Matthews, J. E., Dickey, B. W., Miller, R. D., Felzer, J. R., Dawson, B. P., Lee, A. S., et al. (2012). The epidemiology of published norovirus outbreaks: a review of risk factors associated with attack rate and genogroup. *Epidemiol Infect, 140*(7), 1161-1172.

Mattison, K., Shukla, A., Cook, A., Pollari, F., Friendship, R., Kelton, D., et al. (2007). Human noroviruses in swine and cattle. *Emerg Infect Dis, 13*(8), 1184-1188.

Mattner, F., Sohr, D., Heim, A., Gastmeier, P., Vennema, H., & Koopmans, M. (2006). Risk groups for clinical complications of norovirus infections: an outbreak investigation. *Clin Microbiol Infect, 12*(1), 69-74.

Mayet, A., Andreo, V., Bedubourg, G., Victorion, S., Plantec, J., Soullie, B., et al. (2011). Food-borne outbreak of norovirus infection in a French military parachuting unit, April 2011. *Euro Surveill, 16*(30).

McCammon, J. A., Gelin, B. R., & Karplus, M. (1977). Dynamics of folded proteins. *Nature, 267*(5612), 585-590.

McFadden, N., Bailey, D., Carrara, G., Benson, A., Chaudhry, Y., Shortland, A., et al. (2011). Norovirus regulation of the innate immune response and apoptosis occurs via the product of the alternative open reading frame 4. *PLoS Pathog, 7*(12), e1002413.

McKibbin, J. M., Spencer, W. A., Smith, E. L., Mansson, J. E., Karlsson, K. A., Samuelsson, B. E., et al. (1982). Lewis blood group fucolipids and their isomers from human and canine intestine. *J Biol Chem, 257*(2), 755-760.

Meng, X. Y., Zhang, H. X., Mezei, M., & Cui, M. (2011). Molecular docking: a powerful approach for structure-based drug discovery. *Curr Comput Aided Drug Des, 7*(2), 146-157.
Mesquita, J. R., Barclay, L., Nascimento, M. S., & Vinje, J. (2010). Novel norovirus in dogs with diarrhea. *Emerg Infect Dis, 16*(6), 980-982.
Meyer, E., Ebner, W., Scholz, R., Dettenkofer, M., & Daschner, F. D. (2004). Nosocomial outbreak of norovirus gastroenteritis and investigation of ABO histo-blood group type in infected staff and patients. *J Hosp Infect, 56*(1), 64-66.
Milbrath, M. O., Spicknall, I. H., Zelner, J. L., Moe, C. L., & Eisenberg, J. N. (2013). Heterogeneity in norovirus shedding duration affects community risk. *Epidemiol Infect, 141*(8), 1572-1584.
Miller-Podraza, H., Andersson, C., & Karlsson, K.-A. (1993). New method for the isolation of polyglycosylceramides from human erythrocyte membranes. *Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism, 1168*(3), 330-339.
Mitchell, E., Houles, C., Sudakevitz, D., Wimmerova, M., Gautier, C., Perez, S., et al. (2002). Structural basis for oligosaccharide-mediated adhesion of Pseudomonas aeruginosa in the lungs of cystic fibrosis patients. *Nat Struct Biol, 9*(12), 918-921.
Molinari, M. (2007). N-glycan structure dictates extension of protein folding or onset of disposal. *Nat Chem Biol, 3*(6), 313-320.
Mollicone, R., Cailleau, A., Imberty, A., Gane, P., Perez, S., & Oriol, R. (1996). Recognition of the blood group H type 2 trisaccharide epitope by 28 monoclonal antibodies and three lectins. *Glycoconj J, 13*(2), 263-271.
Mollicone, R., Candelier, J. J., Reguigne, I., Couillin, P., Fletcher, A., & Oriol, R. (1994). Molecular genetics of alpha-L-fucosyltransferase genes (H, Se, Le, FUT4, FUT5 and FUT6). *Transfus Clin Biol, 1*(2), 91-97.
Mollicone, R., Davies, D. R., Evans, B., Dalix, A. M., & Oriol, R. (1986). Cellular expression and genetic control of ABH antigens in primary sensory neurons of marmoset, baboon and man. *J Neuroimmunol, 10*(3), 255-269.
Mollison, P. (1952). *Blood transfusion in clinical medicine*: Wiley Online Library.
Moonens, K., Bouckaert, J., Coddens, A., Tran, T., Panjikar, S., De Kerpel, M., et al. (2012). Structural insight in histo-blood group binding by the F18 fimbrial adhesin FedF. *Mol Microbiol, 86*(1), 82-95.
Morillo, S. G., Luchs, A., Cilli, A., & do Carmo Sampaio Tavares Timenetsky, M. (2012). Rapid detection of norovirus in naturally contaminated food: foodborne gastroenteritis outbreak on a cruise ship in Brazil, 2010. *Food Environ Virol, 4*(3), 124-129.
Morioka, S., Sakata, T., Tamaki, A., Shioji, T., Funaki, A., Yamamoto, Y., et al. (2006). A food-borne norovirus outbreak at a primary school in Wakayama Prefecture. *Jpn J Infect Dis, 59*(3), 205-207.
Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., et al. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of computational chemistry, 19*(14), 1639-1662.
Muhsen, K., Kassem, E., Rubinstein, U., Schachter, Y., Kremer, A., Goren, S., et al. (2013). Incidence and characteristics of sporadic norovirus gastroenteritis associated with hospitalization of children less than 5 years of age in Israel. *Pediatr Infect Dis J, 32*(6), 688-690.
Nahmany, A., Strino, F., Rosen, J., Kemp, G. J., & Nyholm, P. G. (2005). The use of a genetic algorithm search for molecular mechanics (MM3)-based conformational analysis of oligosaccharides. *Carbohydr Res, 340*(5), 1059-1064.
Nasir, W., Nilsson, J., Olofsson, S., Bally, M., & Rydell, G. E. (2014). Parvovirus B19 VLP recognizes globoside in supported lipid bilayers. *Virology, 456*, 364-369.
Natsuka, S., Gersten, K. M., Zenita, K., Kannagi, R., & Lowe, J. B. (1994). Molecular cloning of a cDNA encoding a novel human leukocyte alpha-1,3-fucosyltransferase capable of synthesizing the sialyl Lewis x determinant. *J Biol Chem, 269*(24), 16789-16794.
Navarre, N., Amiot, N., Van Oijen, A., Imberty, A., Poveda, A., Jimenez-Barbero, J., et al. (1999). Synthesis and conformational analysis of a conformationally constrained trisaccharide, and complexation properties with concanavalin A. *CHEMISTRY-WEINHEIM-EUROPEAN JOURNAL-, 5*, 2281-2294.

Nayak, M. K., Balasubramanian, G., Sahoo, G. C., Bhattacharya, R., Vinje, J., Kobayashi, N., et al. (2008). Detection of a novel intergenogroup recombinant Norovirus from Kolkata, India. *Virology, 377*(1), 117-123.

Nelson, M. T., Humphrey, W., Gursoy, A., Dalke, A., Kalé, L. V., Skeel, R. D., et al. (1996). NAMD: a parallel, object-oriented molecular dynamics program. *International Journal of High Performance Computing Applications, 10*(4), 251-268.

Neu, U., Woellner, K., Gauglitz, G., & Stehle, T. (2008). Structural basis of GM1 ganglioside recognition by simian virus 40. *Proc Natl Acad Sci U S A, 105*(13), 5219-5224.

Nguyen-Van-Tam, J. S., & Sellwood, C. (2007). Avian influenza and the threat of the next human pandemic. *J Hosp Infect, 65 Suppl 2*, 10-13.

Nicollier-Jamot, B., Pico, V., Pothier, P., & Kohli, E. (2003). Molecular cloning, expression, self-assembly, antigenicity, and seroepidemiology of a genogroup II norovirus isolated in France. *J Clin Microbiol, 41*(8), 3901-3904.

Nilsson, J., Rydell, G. E., Le Pendu, J., & Larson, G. (2009). Norwalk virus-like particles bind specifically to A, H and difucosylated Lewis but not to B histo-blood group active glycosphingolipids. *Glycoconj J, 26*(9), 1171-1180.

Nilsson, M., Hedlund, K. O., Thorhagen, M., Larson, G., Johansen, K., Ekspong, A., et al. (2003). Evolution of human calicivirus RNA in vivo: accumulation of mutations in the protruding P2 domain of the capsid leads to structural changes and possibly a new phenotype. *J Virol, 77*(24), 13117-13124.

Nordgren, J., Kindberg, E., Lindgren, P. E., Matussek, A., & Svensson, L. (2010). Norovirus gastroenteritis outbreak with a secretor-independent susceptibility pattern, Sweden. *Emerg Infect Dis, 16*(1), 81-87.

Nordgren, J., Nitiema, L. W., Ouermi, D., Simpore, J., & Svensson, L. (2013). Host genetic factors affect susceptibility to norovirus infections in Burkina Faso. *PLoS One, 8*(7), e69557.

Nyholm, P. G., & Pascher, I. (1993). Orientation of the saccharide chains of glycolipids at the membrane surface: conformational analysis of the glucose-ceramide and the glucose-glyceride linkages using molecular mechanics (MM3). *Biochemistry, 32*(5), 1225-1234.

Nyholm, P. G., Samuelsson, B. E., Breimer, M., & Pascher, I. (1989). Conformational analysis of blood group A-active glycosphingolipids using HSEA-calculations. The possible significance of the core oligosaccharide chain for the presentation and recognition of the A-determinant. *J Mol Recognit, 2*(3), 103-113.

Nystrom, K., Le Gall-Recule, G., Grassi, P., Abrantes, J., Ruvoen-Clouet, N., Le Moullac-Vaidye, B., et al. (2011). Histo-blood group antigens act as attachment factors of rabbit hemorrhagic disease virus infection in a virus strain-dependent manner. *PLoS Pathog, 7*(8), e1002188.

Ohwaki, K., Nagashima, H., Aoki, M., Aoki, H., & Yano, E. (2009). A foodborne norovirus outbreak at a hospital and an attached long-term care facility. *Jpn J Infect Dis, 62*(6), 450-454.

Okada, M., Tanaka, T., Oseto, M., Takeda, N., & Shinozaki, K. (2006). Genetic analysis of noroviruses associated with fatalities in healthcare facilities. *Arch Virol, 151*(8), 1635-1641.

Okajima, T., Fukumoto, S., Miyazaki, H., Ishida, H., Kiso, M., Furukawa, K., et al. (1999). Molecular cloning of a novel alpha2,3-sialyltransferase (ST3Gal VI) that sialylates type II lactosamine structures on glycoproteins and glycolipids. *J Biol Chem, 274*(17), 11479-11486.

Okhuysen, P. C., Jiang, X., Ye, L., Johnson, P. C., & Estes, M. K. (1995). Viral shedding and fecal IgA response after Norwalk virus infection. *J Infect Dis, 171*(3), 566-569.

Olofsson, S., & Bergstrom, T. (2005). Glycoconjugate glycans as viral receptors. *Ann Med, 37*(3), 154-172.

Oogane, T., Hirata, A., Funatogawa, K., Kobayashi, K., Sato, T., & Kimura, H. (2008). Food poisoning outbreak caused by norovirus GII/4 in school lunch, Tochigi prefecture, Japan. *Jpn J Infect Dis, 61*(5), 423-424.

Oriol, R. (1990). Genetic control of the fucosylation of ABH precursor chains. Evidence for new epistatic interactions in different cells and tissues. *J Immunogenet, 17*(4-5), 235-245.

Oriol, R., Le Pendu, J., & Mollicone, R. (1986). Genetics of ABO, H, Lewis, X and related antigens. *Vox Sang, 51*(3), 161-171.

Ozawa, K., Oka, T., Takeda, N., & Hansman, G. S. (2007). Norovirus infections in symptomatic and asymptomatic food handlers in Japan. *J Clin Microbiol, 45*(12), 3996-4005.

Papafragkou, E., Hewitt, J., Park, G. W., Greening, G., & Vinje, J. (2013). Challenges of culturing human norovirus in three-dimensional organoid intestinal cell culture models. *PLoS One, 8*(6), e63485.

Park, K., Yeo, S., Jeong, H., Baek, K., Kim, D., Shin, M., et al. (2012). Updates on the genetic variations of norovirus in sporadic gastroenteritis in Chungnam Korea, 2009-2010. *Virol J, 9*, 29.

Parra, G. I., Bok, K., Taylor, R., Haynes, J. R., Sosnovtsev, S. V., Richardson, C., et al. (2012). Immunogenicity and specificity of norovirus Consensus GII.4 virus-like particles in monovalent and bivalent vaccine formulations. *Vaccine, 30*(24), 3580-3586.

Parrino, T. A., Schreiber, D. S., Trier, J. S., Kapikian, A. Z., & Blacklow, N. R. (1977). Clinical immunity in acute gastroenteritis caused by Norwalk agent. *N Engl J Med, 297*(2), 86-89.

Pasquini, J., Guarna, M., Besio-Moreno, M., Iturregui, M., Oteiza, P., & Soto, E. (1989). Inhibition of the synthesis of glycosphingolipids affects the translocation of proteolipid protein to the myelin membrane. *Journal of neuroscience research, 22*(3), 289-296.

Patel, M. M., Hall, A. J., Vinje, J., & Parashar, U. D. (2009). Noroviruses: a comprehensive review. *J Clin Virol, 44*(1), 1-8.

Patel, M. M., Widdowson, M. A., Glass, R. I., Akazawa, K., Vinje, J., & Parashar, U. D. (2008). Systematic literature review of role of noroviruses in sporadic gastroenteritis. *Emerg Infect Dis, 14*(8), 1224-1231.

Perret, S., Sabin, C., Dumon, C., Pokorna, M., Gautier, C., Galanina, O., et al. (2005). Structural basis for the interaction between human milk oligosaccharides and the bacterial lectin PA-IIL of Pseudomonas aeruginosa. *Biochem J, 389*(Pt 2), 325-332.

Pletneva, M. A., Sosnovtsev, S. V., & Green, K. Y. (2001). The genome of hawaii virus and its relationship with other members of the caliciviridae. *Virus Genes, 23*(1), 5-16.

Podraza, H. M., Stenhagen, G., Larsson, T., Andersson, C., & Karlsson, K.-A. (1997). Screening for the presence of polyglycosylceramides in various tissues: partial characterization of blood group-active complex glycosphingolipids of rabbit and dog small intestines. *Glycoconjugate journal, 14*(2), 231-239.

Pokhrel, L., Kim, Y., Nguyen, T. D., Prior, A. M., Lu, J., Chang, K. O., et al. (2012). Synthesis and anti-norovirus activity of pyranobenzopyrone compounds. *Bioorg Med Chem Lett, 22*(10), 3480-3484.

Ponder, J. (2014). TINKER - Software tools for molecular design.

Ponterio, E., Petrizzo, A., Di Bartolo, I., Buonaguro, F. M., Buonaguro, L., & Ruggeri, F. M. (2013). Pattern of activation of human antigen presenting cells by genotype GII.4 norovirus virus-like particles. *J Transl Med, 11*, 127.

Prasad, B. V., Hardy, M. E., Dokland, T., Bella, J., Rossmann, M. G., & Estes, M. K. (1999). X-ray crystallographic structure of the Norwalk virus capsid. *Science, 286*(5438), 287-290.

Prasad, B. V., Hardy, M. E., Jiang, X., & Estes, M. K. (1996). Structure of Norwalk virus. *Arch Virol Suppl, 12*, 237-242.

Prieels, J. P., Monnom, D., Dolmans, M., Beyer, T. A., & Hill, R. L. (1981). Co-purification of the Lewis blood group N-acetylglucosaminide alpha 1 goes to 4 fucosyltransferase and an N-acetylglucosaminide alpha 1 goes to 3 fucosyltransferase from human milk. *J Biol Chem, 256*(20), 10456-10463.

Rademacher, C., Guiard, J., Kitov, P. I., Fiege, B., Dalton, K. P., Parra, F., et al. (2011). Targeting norovirus infection-multivalent entry inhibitor design based on NMR experiments. *Chemistry, 17*(27), 7442-7453.

Rademacher, C., Krishna, N. R., Palcic, M., Parra, F., & Peters, T. (2008). NMR experiments reveal the molecular basis of receptor recognition by a calicivirus. *J Am Chem Soc, 130*(11), 3669-3675.

Rajan, V. P., Larsen, R. D., Ajmera, S., Ernst, L. K., & Lowe, J. B. (1989). A cloned human DNA restriction fragment determines expression of a GDP-L-fucose: beta-D-galactoside 2-alpha-L-fucosyltransferase in transfected cells. Evidence for isolation and transfer of the human H blood group locus. *J Biol Chem, 264*(19), 11158-11167.

Ramachandran, P., Boontheung, P., Pang, E., Yan, W., Wong, D. T., & Loo, J. A. (2008). Comparison of N-linked Glycoproteins in Human Whole Saliva, Parotid, Submandibular, and Sublingual

Glandular Secretions Identified using Hydrazide Chemistry and Mass Spectrometry. *Clin Proteomics, 4*(3-4), 80-104.
Ramani, S., Cortes-Penfield, N. W., Hu, L., Crawford, S. E., Czako, R., Smith, D. F., et al. (2013). The VP8* domain of neonatal rotavirus strain G10P[11] binds to type II precursor glycans. *J Virol, 87*(13), 7255-7264.
Rarey, M., Kramer, B., Lengauer, T., & Klebe, G. (1996). A fast flexible docking method using an incremental construction algorithm. *J Mol Biol, 261*(3), 470-489.
Ravn, V., & Dabelsteen, E. (2000). Tissue distribution of histo-blood group antigens. *APMIS, 108*(1), 1-28.
Rini, J., Esko, J., & Varki, A. (2009). Glycosyltransferases and Glycan-processing Enzymes. In A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart & M. E. Etzler (Eds.), *Essentials of Glycobiology* (2nd ed.). Cold Spring Harbor (NY).
Rockx, B., Baric, R. S., de Grijs, I., Duizer, E., & Koopmans, M. P. (2005a). Characterization of the homo- and heterotypic immune responses after natural norovirus infection. *J Med Virol, 77*(3), 439-446.
Rockx, B., De Wit, M., Vennema, H., Vinje, J., De Bruin, E., Van Duynhoven, Y., et al. (2002). Natural history of human calicivirus infection: a prospective cohort study. *Clin Infect Dis, 35*(3), 246-253.
Rockx, B. H., Bogers, W. M., Heeney, J. L., van Amerongen, G., & Koopmans, M. P. (2005b). Experimental norovirus infections in non-human primates. *J Med Virol, 75*(2), 313-320.
Rockx, B. H., Vennema, H., Hoebe, C. J., Duizer, E., & Koopmans, M. P. (2005c). Association of histo-blood group antigens and susceptibility to norovirus infections. *J Infect Dis, 191*(5), 749-754.
Rodahl, M., Höök, F., Krozer, A., Brzezinski, P., & Kasemo, B. (1995). Quartz crystal microbalance setup for frequency and Q-factor measurements in gaseous and liquid environments. *Review of Scientific Instruments, 66*(7), 3924-3930.
Roddie, C., Paul, J. P., Benjamin, R., Gallimore, C. I., Xerry, J., Gray, J. J., et al. (2009). Allogeneic hematopoietic stem cell transplantation and norovirus gastroenteritis: a previously unrecognized cause of morbidity. *Clin Infect Dis, 49*(7), 1061-1068.
Rohayem, J. (2009). Norovirus seasonality and the potential impact of climate change. *Clin Microbiol Infect, 15*(6), 524-527.
Rouger, P., Poupon, R., Gane, P., Mallissen, B., Darnis, F., & Salmon, C. (1986). Expression of blood group antigens including HLA markers in human adult liver. *Tissue Antigens, 27*(2), 78-86.
Ruvoen-Clouet, N., Ganiere, J. P., Andre-Fontaine, G., Blanchard, D., & Le Pendu, J. (2000). Binding of rabbit hemorrhagic disease virus to antigens of the ABH histo-blood group family. *J Virol, 74*(24), 11950-11954.
Rydell, G. E., Dahlin, A. B., Hook, F., & Larson, G. (2009a). QCM-D studies of human norovirus VLPs binding to glycosphingolipids in supported lipid bilayers reveal strain-specific characteristics. *Glycobiology, 19*(11), 1176-1184.
Rydell, G. E., Kindberg, E., Larson, G., & Svensson, L. (2011). Susceptibility to winter vomiting disease: a sweet matter. *Rev Med Virol, 21*(6), 370-382.
Rydell, G. E., Nilsson, J., Rodriguez-Diaz, J., Ruvoen-Clouet, N., Svensson, L., Le Pendu, J., et al. (2009b). Human noroviruses recognize sialyl Lewis x neoglycoprotein. *Glycobiology, 19*(3), 309-320.
Saito, F., Blank, M., Schroder, J., Manya, H., Shimizu, T., Campbell, K. P., et al. (2005). Aberrant glycosylation of alpha-dystroglycan causes defective binding of laminin in the muscle of chicken muscular dystrophy. *FEBS Lett, 579*(11), 2359-2363.
Sako, Y., Minoghchi, S., & Yanagida, T. (2000). Single-molecule imaging of EGFR signalling on the surface of living cells. *Nat Cell Biol, 2*(3), 168-172.
Sakuma, Y., Kubushiro, K., Tsukazaki, K., & Nozawa, S. (2003). Involvement of H type 1 carbohydrate antigen in cell adhesion to vascular endothelial cells of human endometrial cancer. *Anticancer Res, 23*(2B), 1277-1282.
Samygina, V. R., Popov, A. N., Cabo-Bilbao, A., Ochoa-Lizarralde, B., Goni-de-Cerio, F., Zhai, X., et al. (2011). Enhanced selectivity for sulfatide by engineered human glycolipid transfer protein. *Structure, 19*(11), 1644-1654.

Sarter, K., Mierke, C., Beer, A., Frey, B., Fuhrnrohr, B. G., Schulze, C., et al. (2007). Sweet clearance: involvement of cell surface glycans in the recognition of apoptotic cells. *Autoimmunity, 40*(4), 345-348.

Schedin-Weiss, S., Winblad, B., & Tjernberg, L. O. (2014). The role of protein glycosylation in Alzheimer disease. *FEBS J, 281*(1), 46-62.

Schreiber, D. S., Blacklow, N. R., & Trier, J. S. (1973). The mucosal lesion of the proximal small intestine in acute infectious nonbacterial gastroenteritis. *N Engl J Med, 288*(25), 1318-1323.

Schreiber, D. S., Blacklow, N. R., & Trier, J. S. (1974). The small intestinal lesion induced by Hawaii agent acute infectious nonbacterial gastroenteritis. *J Infect Dis, 129*(6), 705-708.

Schubert, M., Bleuler-Martinez, S., Butschi, A., Walti, M. A., Egloff, P., Stutz, K., et al. (2012). Plasticity of the beta-trefoil protein fold in the recognition and control of invertebrate predators and parasites by a fungal defence system. *PLoS Pathog, 8*(5), e1002706.

Schwegmann-Wessels, C., & Herrler, G. (2006). Sialic acids as receptor determinants for coronaviruses. *Glycoconj J, 23*(1-2), 51-58.

Scipioni, A., Mauroy, A., Vinje, J., & Thiry, E. (2008). Animal noroviruses. *Vet J, 178*(1), 32-45.

Scott, W. R., Hünenberger, P. H., Tironi, I. G., Mark, A. E., Billeter, S. R., Fennen, J., et al. (1999). The GROMOS biomolecular simulation program package. *The Journal of Physical Chemistry A, 103*(19), 3596-3607.

Shanker, S., Choi, J. M., Sankaran, B., Atmar, R. L., Estes, M. K., & Prasad, B. V. (2011). Structural analysis of histo-blood group antigen binding specificity in a norovirus GII.4 epidemic variant: implications for epochal evolution. *J Virol, 85*(17), 8635-8645.

Shanker, S., Czako, R., Sankaran, B., Atmar, R. L., Estes, M. K., & Prasad, B. V. (2014). Structural Analysis of Determinants to HBGA Binding Specificity in GI Noroviruses. *J Virol*.

Sharon, N., & Lis, H. (1989). Lectins as cell recognition molecules. *Science, 246*(4927), 227-234.

Sharon, N., & Lis, H. (1993). Carbohydrates in cell recognition. *Sci Am, 268*(1), 82-89.

Shaw, D. E., Dror, R. O., Salmon, J. K., Grossman, J., Mackenzie, K. M., Bank, J. A., et al. (2009). *Millisecond-scale molecular dynamics simulations on Anton.* Paper presented at the High Performance Computing Networking, Storage and Analysis, Proceedings of the Conference on.

Shaw, D. E., Maragakis, P., Lindorff-Larsen, K., Piana, S., Dror, R. O., Eastwood, M. P., et al. (2010). Atomic-level characterization of the structural dynamics of proteins. *Science, 330*(6002), 341-346.

Shinya, K., Ebina, M., Yamada, S., Ono, M., Kasai, N., & Kawaoka, Y. (2006). Avian flu: influenza virus receptors in the human airway. *Nature, 440*(7083), 435-436.

Shirato, H. (2011). Norovirus and histo-blood group antigens. *Jpn J Infect Dis, 64*(2), 95-103.

Shirato, H., Ogawa, S., Ito, H., Sato, T., Kameyama, A., Narimatsu, H., et al. (2008). Noroviruses distinguish between type 1 and type 2 histo-blood group antigens for binding. *J Virol, 82*(21), 10756-10767.

Shrikhande, G. V., Scali, S. T., da Silva, C. G., Damrauer, S. M., Csizmadia, E., Putheti, P., et al. (2010). O-glycosylation regulates ubiquitination and degradation of the anti-inflammatory protein A20 to accelerate atherosclerosis in diabetic ApoE-null mice. *PLoS One, 5*(12), e14240.

Simmonds, P., Karakasiliotis, I., Bailey, D., Chaudhry, Y., Evans, D. J., & Goodfellow, I. G. (2008). Bioinformatic and functional analysis of RNA secondary structure elements among different genera of human and animal caliciviruses. *Nucleic Acids Res, 36*(8), 2530-2546.

Simmons, G., Garbutt, C., Hewitt, J., & Greening, G. (2007). A New Zealand outbreak of norovirus gastroenteritis linked to the consumption of imported raw Korean oysters. *N Z Med J, 120*(1264), U2773.

Simons, K., & Ikonen, E. (1997). Functional rafts in cell membranes. *Nature, 387*(6633), 569-572.

Simons, K., & van Meer, G. (1988). Lipid sorting in epithelial cells. *Biochemistry, 27*(17), 6197-6202.

Skehel, J. J., & Wiley, D. C. (2000). Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu Rev Biochem, 69*, 531-569.

Smith, A. E., & Helenius, A. (2004). How viruses enter animal cells. *Science, 304*(5668), 237-242.

Smith, A. J., McCarthy, N., Saldana, L., Ihekweazu, C., McPhedran, K., Adak, G. K., et al. (2012). A large foodborne outbreak of norovirus in diners at a restaurant in England between January and February 2009. *Epidemiol Infect, 140*(9), 1695-1701.

Sosnovtsev, S. V., Belliot, G., Chang, K. O., Prikhodko, V. G., Thackray, L. B., Wobus, C. E., et al. (2006). Cleavage map and proteolytic processing of the murine norovirus nonstructural polyprotein in infected cells. *J Virol, 80*(16), 7816-7831.

Souza, M., Azevedo, M. S., Jung, K., Cheetham, S., & Saif, L. J. (2008). Pathogenesis and immune responses in gnotobiotic calves after infection with the genogroup II.4-HS66 strain of human norovirus. *J Virol, 82*(4), 1777-1786.

Springer, G. F. (1971). Blood-group and Forssman antigenic determinants shared between microbes and mammalian cells. *Prog Allergy, 15*, 9-77.

St Clair, K. J., & Patel, S. (2008). Intital descriptive and analytical data on an outbreak of norovirus infection at marine corps recruit depot Parris Island, South Carolina. *J Infect Dis, 198*(6), 941-942; author reply 942-943.

Stanley, P., & Cummings, R. D. (2009). Structures Common to Different Glycans. In A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart & M. E. Etzler (Eds.), *Essentials of Glycobiology* (2nd ed.). Cold Spring Harbor (NY).

Stanley, P., Schachter, H., & Taniguchi, N. (2009). N-Glycans. In A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart & M. E. Etzler (Eds.), *Essentials of Glycobiology* (2nd ed.). Cold Spring Harbor (NY).

Stapleton, A., Nudelman, E., Clausen, H., Hakomori, S., & Stamm, W. E. (1992). Binding of uropathogenic Escherichia coli R45 to glycolipids extracted from vaginal epithelial cells is dependent on histo-blood group secretor status. *J Clin Invest, 90*(3), 965-972.

Stark, J. L., & Powers, R. (2012). Application of NMR and molecular docking in structure-based drug discovery. *Top Curr Chem, 326*, 1-34.

Steinhoff, M. C., & John, T. J. (1980). Non bacterial gastroenteritis. *Indian J Pediatr, 47*(387), 317-320.

Stincarelli, M., Arvia, R., De Marco, M. A., Clausi, V., Corcioli, F., Cotti, C., et al. (2013). Reassortment ability of the 2009 pandemic H1N1 influenza virus with circulating human and avian influenza viruses: public health risk implications. *Virus Res, 175*(2), 151-154.

Straub, T. M., Honer zu Bentrup, K., Orosz-Coghlan, P., Dohnalkova, A., Mayer, B. K., Bartholomew, R. A., et al. (2007). In vitro cell culture infectivity assay for human noroviruses. *Emerg Infect Dis, 13*(3), 396-403.

Subekti, D. S., Tjaniadi, P., Lesmana, M., McArdle, J., Iskandriati, D., Budiarsa, I. N., et al. (2002). Experimental infection of Macaca nemestrina with a Toronto Norwalk-like virus of epidemic viral gastroenteritis. *J Med Virol, 66*(3), 400-406.

Sulak, O., Cioci, G., Delia, M., Lahmann, M., Varrot, A., Imberty, A., et al. (2010). A TNF-like trimeric lectin domain from Burkholderia cenocepacia with specificity for fucosylated human histo-blood group antigens. *Structure, 18*(1), 59-72.

Sun, X., & Whittaker, G. R. (2003). Role for influenza virus envelope cholesterol in virus entry and infection. *J Virol, 77*(23), 12543-12551.

Tacket, C. O., Sztein, M. B., Losonsky, G. A., Wasserman, S. S., & Estes, M. K. (2003). Humoral, mucosal, and cellular immune responses to oral Norwalk virus-like particles in volunteers. *Clin Immunol, 108*(3), 241-247.

Tamura, M., Natori, K., Kobayashi, M., Miyamura, T., & Takeda, N. (2003). Inhibition of attachment of virions of Norwalk virus to mammalian cells by soluble histone molecules. *Arch Virol, 148*(9), 1659-1670.

Tamura, M., Natori, K., Kobayashi, M., Miyamura, T., & Takeda, N. (2004). Genogroup II noroviruses efficiently bind to heparan sulfate proteoglycan associated with the cellular membrane. *J Virol, 78*(8), 3817-3826.

Tan, D. M., Liu, W., & Deng, L. L. (2010). [Norovirus infection in adults with sporadic gastroenteritis during 2007-2008 in Nanning Municipal]. *Zhongguo Yi Miao He Mian Yi, 16*(2), 132-135.

Tan, M., Huang, P., Meller, J., Zhong, W., Farkas, T., & Jiang, X. (2003). Mutations within the P2 domain of norovirus capsid affect binding to human histo-blood group antigens: evidence for a binding pocket. *J Virol, 77*(23), 12562-12571.

Tan, M., & Jiang, X. (2008). Association of histo-blood group antigens with susceptibility to norovirus infection may be strain-specific rather than genogroup dependent. *J Infect Dis, 198*(6), 940-941; author reply 942-943.

Tan, M., Jin, M., Xie, H., Duan, Z., Jiang, X., & Fang, Z. (2008a). Outbreak studies of a GII-3 and a GII-4 norovirus revealed an association between HBGA phenotypes and viral infection. *J Med Virol, 80*(7), 1296-1301.

Tan, M., Xia, M., Cao, S., Huang, P., Farkas, T., Meller, J., et al. (2008b). Elucidation of strain-specific interaction of a GII-4 norovirus with HBGA receptors by site-directed mutagenesis study. *Virology, 379*(2), 324-334.

Tan, M., Xia, M., Chen, Y., Bu, W., Hegde, R. S., Meller, J., et al. (2009). Conservation of carbohydrate binding interfaces: evidence of human HBGA selection in norovirus evolution. *PLoS One, 4*(4), e5058.

Tang, H., Kawabata, A., Takemoto, M., Yamanishi, K., & Mori, Y. (2008). Human herpesvirus-6 infection induces the reorganization of membrane microdomains in target cells, which are required for virus entry. *Virology, 378*(2), 265-271.

Taylor, M. E., & Drickamer, K. (2011). *Introduction to glycobiology*: Oxford university press.

Taylor, R. D., Jewsbury, P. J., & Essex, J. W. (2002). A review of protein-small molecule docking methods. *J Comput Aided Mol Des, 16*(3), 151-166.

Thompson, T., & Tillack, T. W. (1985). Organization of glycosphingolipids in bilayers and plasma membranes of mammalian cells. *Annual review of biophysics and biophysical chemistry, 14*(1), 361-386.

Thorne, L. G., & Goodfellow, I. G. (2014). Norovirus gene expression and replication. *J Gen Virol, 95*(Pt 2), 278-291.

Thorven, M., Grahn, A., Hedlund, K. O., Johansson, H., Wahlfrid, C., Larson, G., et al. (2005). A homozygous nonsense mutation (428G-->A) in the human secretor (FUT2) gene provides resistance to symptomatic norovirus (GGII) infections. *J Virol, 79*(24), 15351-15355.

Tian, P., Jiang, X., Zhong, W., Jensen, H. M., Brandl, M., Bates, A. H., et al. (2007). Binding of recombinant norovirus like particle to histo-blood group antigen on cells in the lumen of pig duodenum. *Res Vet Sci, 83*(3), 410-418.

Tieleman, D. P., Berendsen, H. J., & Sansom, M. S. (1999). Surface binding of alamethicin stabilizes its helical structure: molecular dynamics simulations. *Biophys J, 76*(6), 3186-3191.

Trang, N. V., Vu, H. T., Le, N. T., Huang, P., Jiang, X., & Anh, D. D. (2014). Association between Norovirus and Rotavirus Infection and Histo-Blood Group Antigen Types in Vietnamese Children. *J Clin Microbiol*.

Trivedi, T. K., Desai, R., Hall, A. J., Patel, M., Parashar, U. D., & Lopman, B. A. (2013). Clinical characteristics of norovirus-associated deaths: a systematic literature review. *Am J Infect Control, 41*(7), 654-657.

Troeger, H., Loddenkemper, C., Schneider, T., Schreier, E., Epple, H. J., Zeitz, M., et al. (2009). Structural and functional changes of the duodenum in human norovirus infection. *Gut, 58*(8), 1070-1077.

Turcios-Ruiz, R. M., Axelrod, P., St John, K., Bullitt, E., Donahue, J., Robinson, N., et al. (2008). Outbreak of necrotizing enterocolitis caused by norovirus in a neonatal intensive care unit. *J Pediatr, 153*(3), 339-344.

Uusi-Kerttula, H., Tamminen, K., Malm, M., Vesikari, T., & Blazevic, V. (2014). Comparison of human saliva and synthetic histo-blood group antigens usage as ligands in norovirus-like particle binding and blocking assays. *Microbes Infect*.

Wadl, M., Scherer, K., Nielsen, S., Diedrich, S., Ellerbroek, L., Frank, C., et al. (2010). Food-borne norovirus-outbreak at a military base, Germany, 2009. *BMC Infect Dis, 10*, 30.

van Asten, L., Siebenga, J., van den Wijngaard, C., Verheij, R., van Vliet, H., Kretzschmar, M., et al. (2011). Unspecified gastroenteritis illness and deaths in the elderly associated with norovirus epidemics. *Epidemiology, 22*(3), 336-343.

van Kooyk, Y., & Rabinovich, G. A. (2008). Protein-glycan interactions in the control of innate and adaptive immune responses. *Nat Immunol, 9*(6), 593-601.

van Riel, D., Munster, V. J., de Wit, E., Rimmelzwaan, G. F., Fouchier, R. A., Osterhaus, A. D., et al. (2006). H5N1 Virus Attachment to Lower Respiratory Tract. *Science, 312*(5772), 399.

Wang, J., Cieplak, P., & Kollman, P. A. (2000). How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *Journal of Computational Chemistry, 21*(12), 1049-1074.

Wang, Q. H., Han, M. G., Cheetham, S., Souza, M., Funk, J. A., & Saif, L. J. (2005). Porcine noroviruses related to human noroviruses. *Emerg Infect Dis, 11*(12), 1874-1881.

Varki, A., Esko, J. D., & Colley, K. J. (2009a). Cellular Organization of Glycosylation. In A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart & M. E. Etzler (Eds.), *Essentials of Glycobiology* (2nd ed.). Cold Spring Harbor (NY).

Varki, A., Etzler, M. E., Cummings, R. D., & Esko, J. D. (2009b). Discovery and Classification of Glycan-Binding Proteins. In A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart & M. E. Etzler (Eds.), *Essentials of Glycobiology* (2nd ed.). Cold Spring Harbor (NY).

Varki, A., & Sharon, N. (2009). Historical Background and Overview. In A. Varki, R. D. Cummings, J. D. Esko, H. H. Freeze, P. Stanley, C. R. Bertozzi, G. W. Hart & M. E. Etzler (Eds.), *Essentials of Glycobiology* (2nd ed.). Cold Spring Harbor (NY).

Watkins, W. M. (1980). Biochemistry and Genetics of the ABO, Lewis, and P blood group systems. *Adv Hum Genet, 10*, 1-136, 379-185.

Verbelen, V., Bodeus, M., Garrino, M. G., Scipioni, A., Kabamba, B., Daube, G., et al. (2004). Hospital outbreak of gastroenteritis due to Norovirus in Belgium. *Acta Clin Belg, 59*(1), 30-33.

Westhoff, T. H., Vergoulidou, M., Loddenkemper, C., Schwartz, S., Hofmann, J., Schneider, T., et al. (2009). Chronic norovirus infection in renal transplant recipients. *Nephrol Dial Transplant, 24*(3), 1051-1053.

Weston, B. W., Nair, R. P., Larsen, R. D., & Lowe, J. B. (1992a). Isolation of a novel human alpha (1,3)fucosyltransferase gene and molecular comparison to the human Lewis blood group alpha (1,3/1,4)fucosyltransferase gene. Syntenic, homologous, nonallelic genes encoding enzymes with distinct acceptor substrate specificities. *J Biol Chem, 267*(6), 4152-4160.

Weston, B. W., Smith, P. L., Kelly, R. J., & Lowe, J. B. (1992b). Molecular cloning of a fourth member of a human alpha (1,3)fucosyltransferase gene family. Multiple homologous sequences that determine expression of the Lewis x, sialyl Lewis x, and difucosyl sialyl Lewis x epitopes. *J Biol Chem, 267*(34), 24575-24584.

White, L. J., Ball, J. M., Hardy, M. E., Tanaka, T. N., Kitamoto, N., & Estes, M. K. (1996). Attachment and entry of recombinant Norwalk virus capsids to cultured human and animal cell lines. *J Virol, 70*(10), 6589-6597.

WHO. (2013). Commulative number of H5N1 cases 2003-2013.

Wikswo, M. E., Cortes, J., Hall, A. J., Vaughan, G., Howard, C., Gregoricus, N., et al. (2011). Disease transmission and passenger behaviors during a high morbidity Norovirus outbreak on a cruise ship, January 2009. *Clin Infect Dis, 52*(9), 1116-1122.

Vinje, J., Green, J., Lewis, D. C., Gallimore, C. I., Brown, D. W., & Koopmans, M. P. (2000). Genetic polymorphism across regions of the three open reading frames of "Norwalk-like viruses". *Arch Virol, 145*(2), 223-241.

Vivancos, R., Keenan, A., Sopwith, W., Smith, K., Quigley, C., Mutton, K., et al. (2010). Norovirus outbreak in a cruise ship sailing around the British Isles: investigation and multi-agency management of an international outbreak. *J Infect, 60*(6), 478-485.

Wobus, C. E., Karst, S. M., Thackray, L. B., Chang, K. O., Sosnovtsev, S. V., Belliot, G., et al. (2004). Replication of Norovirus in cell culture reveals a tropism for dendritic cells and macrophages. *PLoS Biol, 2*(12), e432.

Wobus, C. E., Thackray, L. B., & Virgin, H. W. t. (2006). Murine norovirus: a model system to study norovirus biology and pathogenesis. *J Virol, 80*(11), 5104-5112.

Wolf, S., Reetz, J., & Otto, P. (2011). Genetic characterization of a novel calicivirus from a chicken. *Arch Virol, 156*(7), 1143-1150.

Vongpunsawad, S., Venkataram Prasad, B. V., & Estes, M. K. (2013). Norwalk Virus Minor Capsid Protein VP2 Associates within the VP1 Shell Domain. *J Virol, 87*(9), 4818-4825.

Wu, J., Zhu, J., Yin, H., Buckanovich, R. J., & Lubman, D. M. (2014). Analysis of Glycan Variation on Glycoproteins from Serum by the Reverse Lectin-Based ELISA Assay. *J Proteome Res*.

Wyatt, R. G., Dolin, R., Blacklow, N. R., DuPont, H. L., Buscho, R. F., Thornhill, T. S., et al. (1974). Comparison of three agents of acute infectious nonbacterial gastroenteritis by cross-challenge in volunteers. *J Infect Dis, 129*(6), 709-714.

Wyatt, R. G., Greenberg, H. B., Dalgard, D. W., Allen, W. P., Sly, D. L., Thornhill, T. S., et al. (1978). Experimental infection of chimpanzees with the Norwalk agent of epidemic viral gastroenteritis. *J Med Virol, 2*(2), 89-96.

Yamashita, T., Wada, R., Sasaki, T., Deng, C., Bierfreund, U., Sandhoff, K., et al. (1999). A vital role for glycosphingolipid synthesis during development and differentiation. *Proceedings of the National Academy of Sciences, 96*(16), 9142-9147.

Yap, J., Qadir, A., Liu, I., Loh, J., Tan, B. H., & Lee, V. J. (2012). Outbreak of acute norovirus gastroenteritis in a military facility in Singapore: a public health perspective. *Singapore Med J, 53*(4), 249-254.

Yazgan, H., Keles, E., Gebesce, A., Demirdoven, M., & Yazgan, Z. (2013). Blood groups and rotavirus gastroenteritis. *Pediatr Infect Dis J, 32*(6), 705-706.

Yoneda, M., Okayama, A., & Kitahori, Y. (2014). Epidemiological characteristics of norovirus associated with sporadic gastroenteritis among children from the 2006/2007 to 2011/2012 season in Nara Prefecture, Japan. *Intervirology, 57*(1), 31-35.

Yoshizaki, F., Nakayama, H., Iwahara, C., Takamori, K., Ogawa, H., & Iwabuchi, K. (2008). Role of glycosphingolipid-enriched microdomains in innate immunity: microdomain-dependent phagocytic cell functions. *Biochim Biophys Acta, 1780*(3), 383-392.

Zahorsky, J. (1929). Hyperemesis hiemis or the winter vomiting disease. *Arch Pediatr, 46*, 391.

Zanetta, J. P., Kuchler, S., Lehmann, S., Badache, A., Maschke, S., Thomas, D., et al. (1992). Glycoproteins and lectins in cell adhesion and cell recognition processes. *Histochem J, 24*(11), 791-804.

Zenisek, D., Steyer, J. A., & Almers, W. (2000). Transport, capture and exocytosis of single synaptic vesicles at active zones. *Nature, 406*(6798), 849-854.

Zhang, X. F., Dai, Y. C., Zhong, W., Tan, M., Lv, Z. P., Zhou, Y. C., et al. (2012). Tannic acid inhibited norovirus binding to HBGA receptors, a study of 50 Chinese medicinal herbs. *Bioorg Med Chem, 20*(4), 1616-1623.

Zhang, X. F., Tan, M., Chhabra, M., Dai, Y. C., Meller, J., & Jiang, X. (2013). Inhibition of histo-blood group antigen binding as a novel strategy to block norovirus infections. *PLoS One, 8*(7), e69379.

Zheng, D. P., Ando, T., Fankhauser, R. L., Beard, R. S., Glass, R. I., & Monroe, S. S. (2006). Norovirus classification and proposed strain nomenclature. *Virology, 346*(2), 312-323.