



# LUND UNIVERSITY

## Genetic Characterisation of Human ABO Blood Group Variants with a Focus on Subgroups and Hybrid Alleles

Hosseini Maaf, Bahram

Published: 2007-01-01

[Link to publication](#)

*Citation for published version (APA):*

Hosseini Maaf, B. (2007). Genetic Characterisation of Human ABO Blood Group Variants with a Focus on Subgroups and Hybrid Alleles Division of Hematology and Transfusion Medicine, Department of Laboratory Medicine, Lund University

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

LUND UNIVERSITY

PO Box 117  
221 00 Lund  
+46 46-222 00 00

# Genetic Characterisation of Human ABO Blood Group Variants with a Focus on Subgroups and Hybrid Alleles

Doctoral thesis  
by

Bahram Hosseini-Maaf

Division of Hematology and Transfusion Medicine  
Department of Laboratory Medicine  
Lund University, Sweden



LUND  
UNIVERSITY

With the approval of the Faculty of Medicine at Lund University,  
this thesis will be defended on March 16, 2007, at 9:00 in Segerfalksalen,  
Wallenberg Neurocentrum, BMC, Sölvegatan 17, Lund

**Faculty opponent:**  
Professor Peter Pähsson  
Department of Biomedicine and Surgery  
Division of Cell Biology  
Linköping University Sweden

GEN DYNAMIC PROPERTY

© Bahram Hosseini-Maaf

ISBN 978-91-85559-16-9  
Printed by Media-Tryck, Lund, Sweden 2007

GEN DYNAMIC PROPERTY

تقدیم به خانواده ام

Till min familj

GEN DYNAMIC PROPERTY

"the chief part of the organisation of every being is simply due to inheritance; and consequently, though each being assuredly is well fitted for its place in nature, many structures now have no direct relation to the habits of life of each species"

*The Origin of Species*  
Charles Darwin, 1859

## CONTENTS

ORIGINAL PAPERS.....	6
ABSTRACT.....	7
ABBREVIATIONS.....	8
PROLOGUE.....	9
BACKGROUND.....	9
History of blood transfusion and basic principles of the ABO blood group system .....	9
Definition of blood group antigens and antibodies.....	11
Blood group terminology.....	12
Biochemistry of the ABO blood groups .....	15
The molecular structure of ABO glycosyltransferase .....	17
Genetics of the ABO blood group system .....	19
The <i>ABO</i> gene .....	19
Serology versus molecular genetics .....	21
Relationship between phenotype and genotype for the major alleles.....	21
Genetic background of hybrid alleles, rare <i>O</i> alleles and weak subgroups .....	27
AIMS OF THE THESIS.....	33
MATERIAL AND METHODS.....	33
Blood and DNA samples.....	33
<i>ABO</i> genotyping.....	33
Polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) .....	33
Allele-specific primer PCR (PCR-ASP).....	34
Direct DNA sequencing.....	34
Reverse transcriptase PCR (RT-PCR).....	34
Site-directed mutagenesis and cloning .....	35
Protein purification and enzyme kinetics .....	35
Computer modelling and crystallization .....	35
RESULTS AND DISCUSSION .....	36
Single point mutations associated with A or B subgroups (Papers I and V) .....	36
Hybrid <i>ABO</i> alleles explain some other weak ABO subgroups (Papers I, II and V) .....	41
The genetic basis of the <i>A<sub>bantu</sub></i> phenotype is also a hybrid allele (Paper IV) .....	45
Novel <i>O</i> alleles involved in unexpected blood group phenotypes (Paper III) .....	47
Development of an improved <i>ABO</i> genotyping assay based on PCR-ASP amplification across intron 6 (Paper VI) .....	51
CONCLUSION.....	54
FUTURE PERSPECTIVES .....	54
SUMMARY IN SWEDISH (sammanfattning på svenska).....	56
ACKNOWLEDGEMENTS .....	61
REFERENCES .....	63
APPENDIX: Papers I-VI .....	70

## ORIGINAL PAPERS

This thesis is based on the following papers, which will be referred to in the text by their Roman numerals (**I-VI**).

- I. Olsson ML, Irshaid NM, **Hosseini-Maaf B**, Hellberg Å, Moulds MK, Sareneva H, Chester MA. Genomic analysis of clinical samples with serologic ABO blood grouping discrepancies: Identification of 15 novel A and B subgroup alleles. *Blood*, 2001;98(5):1585-1593.
- II. **Hosseini-Maaf B**, Hellberg Å, Rodrigues MJ, Chester MA, Olsson ML. ABO exon and intron analysis in individuals with the  $A_{weak}B$  phenotype reveals a novel  $O^{lv}\text{-}A^2$  hybrid allele that causes four missense mutations in the A transferase. *BMC Genetics*, 2003; 4:17 (11 pages, doi: 10.1186/1471-2156-4-17).
- III. **Hosseini-Maaf B**, Irshaid NM, Hellberg Å, Wagner T, Levene C, Hustinx H, Steffensen R, Chester MA, Olsson ML. New and unusual *O* alleles at the *ABO* locus are implicated in unexpected blood group phenotypes. *Transfusion*, 2005;45(1):70-81.
- IV. **Hosseini-Maaf B**, Smart E, Chester MA, Olsson ML. The  $A_{bantu}$  phenotype in the ABO blood group system is due to a splice-site mutation in a hybrid between a new  $O^l$ -like allelic lineage and the  $A^2$  allele. *Vox Sanguinis*, 2005;88(4):256-264.
- V. **Hosseini-Maaf B**, Letts JA, Persson M, Smart E, Le Pennec P-Y, Hustinx H, Zhao Z, Palcic MM, Evans SV, Chester MA, Olsson ML. Structural basis for red cell phenotypic changes in newly-identified, naturally-occurring subgroup mutants of the human blood group B glycosyltransferase. *Transfusion*, 2007; in press.
- VI. **Hosseini-Maaf B**, Hellberg Å, Chester MA, Olsson ML. An extensive PCR-ASP strategy for clinical ABO blood group genotyping that avoids potential errors caused by null, subgroup and hybrid alleles. *Manuscript*.

## ABSTRACT

ABO is the most important blood group system in transfusion medicine and transplantation immunology. The ABO blood groups differ by the presence or absence of antigens on RBCs and antibodies in plasma. Accurate determination of ABO status is critical. Genomic typing can increase the precision of blood group determination in complicated cases, *e.g.* when variant expression of A or B antigen is encountered.

The overall aim of this study was to compare the molecular diversity of *ABO* alleles with various phenotypes, and to contribute to our knowledge of the *ABO* gene and encoded glycosyltransferases.

Novel alleles (six  $A^{weak}$ , eleven  $B^{weak}$ , seven  $O$ ) were identified containing single-point mutations. Structure/function studies explained the weakening of some B subgroup glycosyltransferases. Two new hybrid  $A^x$  alleles were characterised. Analysis of introns 2-5 revealed 44 previously unknown, allele-related polymorphisms that proved valuable allelic markers. These findings enabled localisation of cross-over regions in two other new hybrids: 1) an  $O^{Iv}$  allele fused with an  $A^2$  allele, 2) the novel  $O^{Ibantu}$ - $A^2$  combination that explained the  $A_{bantu}$  phenotype. Phylogenetic and population analyses indicated that  $O^{Ibantu}$  is a unique and distinct evolutionary lineage so far only found among individuals of African descent.

Of clinical importance, a new approach to *ABO* genotyping was developed that identifies all common alleles, most null and weak *A/B* subgroups as well as hybrid alleles resulting from recombinational crossing-over events.

In summary, 30 novel alleles were identified and characterized, representing 30% of all alleles reported since the start of this study in 2001.

## **ABBREVIATIONS**

a.a.	Amino acid
bp	Base pair
kb	kilobase pairs
DNA	Deoxyribonucleic acid
EC	Enzyme commission
GTA or A-enzyme	3- $\alpha$ -N-acetylgalactosaminyltransferase (EC 2.4.1.40) encoded by the blood group A gene
GTB or B-enzyme	3- $\alpha$ -galactosyltransferase (EC 2.4.1.37) encoded by the blood group B gene
nt.	Nucleotide
ISBT	International Society of Blood Transfusion
MAb	Monoclonal antibody
PCR	Polymerase chain reaction
RBC	Red blood cell
RNA	Ribonucleic acid
RFLP	Restriction fragment length polymorphism
ASP	Allele-specific primer
SNP	Single nucleotide polymorphism
VNTR	Variable number of tandem repeats

## **PROLOGUE**

Elucidation of the human genome has now been completed through “the Human Genome Project” and the genetic makeup of man is known. The biologically interpretation of the information contained in the genetic sequence is an extremely complex process, resulting in the production of proteins that act as structural elements and metabolic mediators, *e.g.* enzymes in each cell and organ of an individual. With few exceptions, genetic variation makes every individual exceptional.

Many aspects of this universal uniqueness between individuals are central to transfusion medicine but blood group variation is probably the most important and a prerequisite for modern health care.

The need for donor-recipient matching when performing blood transfusion became apparent at early stages in the history of transfusion medicine, due to the interaction of substances on transfused red blood cells (RBCs), antigens, with other substances, antibodies, in the recipient’s plasma.

The advances in scientific knowledge have allowed characterisation of many of these antigens and antibodies. The structures of the antigens have been elucidated in many cases, some of the biochemical pathways for their biosynthesis have been determined and currently the primary sequences of the genes encoding these antigens are being revealed. The main subject of this thesis will be limited to the ABO blood groups and the following background describes chronologically the development of our knowledge about this important system from the empirical observations of transfusion reactions to the structural variations in the *ABO* gene itself and its products.

## **BACKGROUND**

### **History of blood transfusion and basic principles of the ABO blood group system**

Blood as a medical term is often related to words including *hemo-* or *hemato-* from the Greek stem "*haima*" for blood. Blood is a circulating tissue composed of fluid plasma and cells (RBCs, white blood cells and platelets).

Blood has been used symbolically in both spoken language as “*blood money, stir up bad blood, run in the blood, blood brother, bad blood, blood feud, bloodthirsty*” and in the literature in different contexts, e.g. “*We few, we happy few, we band of brothers; For he today that sheds his blood with me, Shall be my brother*” (*Shakespeare: Henry V, act 4*).

“*The old concept of the movement of blood in the body was elaborated by Galen (129-200 AD). He incorrectly believed that food from the stomach was digested and taken to the liver, where it was transformed into blood, passed to heart, through arteries and veins to the tissues where it was burned up, as wood is consumed by fire*” (Farr, 1980).

Different cultures put varying significance on blood. Some believed that it could be the treatment for many maladies and others were convinced that blood contained the soul of man. Ancient Egyptians used it for baths to resuscitate the dead. The first human-to-human blood transfusion recorded took place in 1492. Pope Innocent VIII sank in a coma, and the blood of three ten-year old boys was infused into the dying Pope’s veins at the suggestion of a physician. Unfortunately, both the recipient and all the donors died (Barsoum & Kleeman, 2002). There is no documentation of the fate of the physician, however.

In 1628, British physician W Harvey described the circulation of the blood in humans after having pumped water through the blood vessels of a corpse (Harvey, 1628). This observation opened a new era and some decades later, in 1667, J-B Denis (Denis, 1667) in France and R Lower (Lower, 1666) in England separately reported transfusions from animal to animal and animal to humans. However, Denis was accused of murder when a young boy died after he was transfused the blood from a lamb, and transfusion of animal blood to man was prohibited by law after the death of further patients (Farr, 1979). Nothing of importance was reported in this field until December 22 in 1818, nearly 150 years later, when J Blundell performed his first successful blood transfusion of man to man (Blundell, 1829). The tendency of blood to clot was the principal barrier to transfusions, and to storage of blood. Furthermore, serious reactions to transfusion (e.g. fever, hypotension, allergic shock, death) continued to occur until 1900 when Landsteiner published his preliminary results and the year after described the first system of blood groups, the so-called ABO system (Landsteiner, 1900, 1901).

Incompatibility between donor and recipient in these blood groups could explain most of these serious transfusion reactions. Landsteiner based his observation on an unexpected interaction between serum and RBCs from himself and a few of his colleagues at the Institute

of Pathological Anatomy in the University of Vienna. He later received the Nobel Prize for his discovery (Landsteiner, 1930). This elucidation laid the basis of a new field in biology and the foundation of transfusion medicine.

This principal finding led to what has become Landsteiner's rule, *i.e.* serological behaviour of the A, B and O (O was initially named type "C") blood groups (**Table 1**). Subsequently, another Austrian group, Decastello and Sturli, identified the fourth group, *i.e.* AB, in 1902 (von Decastello & Sturli, 1902). Several other investigators independently duplicated Landsteiner's findings. Jansky in Poland named the four blood groups I, II, III and IV based on their frequency in European populations in 1907, and Moss described in 1910 the four blood groups in the reverse order of Jansky (reviewed in Garratty *et al.*, 2000; Farr, 1979).

**Table 1.** Basic principles of the ABO system. Relation between antigens on RBCs and antibodies in serum.

ABO blood group	Red cell antigen	Antibodies in serum
O	None	Anti-A, anti-B, anti-A,B
A	A	Anti-B
B	B	Anti-A
AB	A and B	None

In clinical practice, the ABO blood group system is one of the most important since the A and B epitopes may provoke a strong immune reaction. With the introduction of blood typing and crossmatching techniques, blood transfusion became not only a simple but also a much safer procedure. However, blood transfusion has continued to be associated with risks, notably the transmission of blood-borne infectious diseases such as syphilis, viral hepatitis and human immunodeficiency virus. Furthermore, although ABO typing reduced the occurrence of transfusion reactions, they still occurred, indicating the presence of other genetic differences in blood groups of importance in transfusion medicine, as well as in the later emerging field of organ transplantation.

### Definition of blood group antigens and antibodies

The term blood group is generally based on the presence or absence of certain antigens on the RBC membrane. These are identified by characteristic agglutination reactions with specific antibodies and this field is referred to as blood group serology (Daniels, 2002).

An antigen is in fact any substance that is able to stimulate the production of an antibody under certain conditions. Some blood group antigens, *e.g.* in blood group systems ABO, P1, H, LE and I, are due to differences in carbohydrate structures and they are all found on glycolipids or glycoproteins on the RBC membrane. In the case of the ABO blood groups, the antigens are found on the surface of RBCs and other cells and is recoverable from body fluids and secretions. Most other blood group systems result from polymorphisms on different RBC membrane-associated proteins (Daniels, 2002).

Blood group antibodies can be divided into induced and naturally-occurring antibodies. Induced antibodies are raised by the immune system as a response to specific antigens, *e.g.* following transfusion, pregnancy or transplantation. Natural antibodies are also of immune origin, with the difference that the antigen eliciting their production is unknown. In the ABO system, these antibodies are sometimes referred to as isoagglutinins because they can cause blood cells from certain individuals to clump together. The two isoagglutinins, anti-A and anti-B, occur naturally in humans, contrary to most other blood groups antibodies (Erskine & Socha, 1978; Mollison *et al.*, 1993).

Naturally-occurring antibodies can be detected in serum of persons who have never been transfused or women who never have been pregnant (Mollison *et al.*, 1993). Natural antibodies may occur because of unidentified infections during early life or, in the case of blood group antibodies, introduction of substances that are antigenically similar to blood group substances. Interestingly, Springer *et al.*, as early as 1959, demonstrated by an experiment with chickens bred under sterile conditions, that some of these antibodies appear following stimulation by certain bacteria, such as *Escherichia coli*, while the appearance of others was related to exposure to various kinds of food. The origin of natural antibodies is still a matter of debate, though.

Naturally-occurring antibodies are often IgM and sometimes also consist of an IgG component. With exception for antibodies against ABO, P<sup>k</sup> and P, the majority of naturally-occurring antibodies are only cold reactive and of limited clinical importance.

### **Blood group terminology**

A rather incoherent terminology has evolved as more and more blood group antigens were discovered during several decades. Providing a common acceptable and reliable terminology,

which would cover both the serology and genetics, has been difficult. The International Society of Blood Transfusion (ISBT) Working Committee on Terminology for Red Cell Surface Antigens (<http://www.blood.co.uk/ibgrl/>) was set up in 1980 to establish and define a meaningful nomenclature for different blood groups as shown in **Table 2** (Daniels *et al.*, 2003; 2004; Storry & Olsson, 2004). Every valid blood group antigen is given a six digit identification number, according to the following categorization:

- Blood group *systems*; one or more antigens governed either by a single gene locus or by a complex of two or more very closely linked homologous genes with virtually no recombination occurring between them. There are 29 different systems to date, and the first three digits represent the systems (001-029).
- Blood group *collections*; genetically-, biochemically-, or serologically-related sets of antigens, which do not, at this time, merit system status because of insufficient data (205-211).
- A series of low frequency-antigens (the 700 *series*)
- A series of high frequency-antigens (the 901 *series*)

The symbol for a gene or cluster of genes controlling a blood group system is often the italicized symbol for the system. The *ABO* genotypes should consequently be written in italicised capital letters (with superscripts indicating the subgroup alleles).

However, *ABO* allele nomenclature poses significant problems that are still under consideration by the ISBT. In the absence of an officially agreed terminology, in this thesis, alleles are referred to by their serological activity and an alternative allele name is given in square brackets.

ISBT has recently set up a subgroup from the Terminology Working Committee mentioned above. Its task is to create a unified allele nomenclature for blood groups, (ML Olsson personal communication).

**Table 2.** Human blood group systems recognised by the ISBT (<http://www.blood.co.uk/ibgrl/>).

System name	System no. ISBT	Symbol ISBT	Antigens no.	Gene	Locus	Epitope / carrier, or functions
ABO	001	ABO	4*	<i>ABO</i>	9q34.1-q34.2	Carbohydrate
MNS	002	MNS	43	<i>GYPA</i> <i>GYPB</i> <i>GYPE</i>	4q28.2-q31.1	Glycophorin A&B (CD235a-b) <sup>‡</sup>
P	003	P1	1	<i>P1</i>	22q11.2-qter	Carbohydrate
Rh	004	RH	56	<i>RHD</i> <i>RHCE</i>	1q34-36.2	Protein, CD240D CD240CE
Lutheran	005	LU	19	<i>LU</i>	19q13.2	Glycoproteinprotein (IgSF <sup>†</sup> ), adhesion molecule, CD239
Kell	006	KEL	23	<i>KEL</i>	7q33	Glycoprotein, CD238
Lewis	007	LE	6	<i>FUT3</i>	19p13.3	Carbohydrate
Duffy	008	FY	6	<i>FY</i>	1q22-q23	Glycoprotein, receptor, CD234
Kidd	009	JK	3	<i>JK</i>	18q11-q12	Glycoprotein, urea transporter
Diego	010	DI	19	<i>AE1</i>	17q12-q21	Glycoprotein (band 3, AE1), CD233
Yt (Cartwright)	011	YT	2	<i>ACHE</i>	7q22	Enzyme, acetylcholinesterase
Xg	012	XG	1	<i>XG</i>	Xp22.32	Glycoprotein, CD99
Scianna	013	SC	3	<i>ERMAP</i>	1q36.2-p22.1	Glycoprotein
Dombrock	014	DO	5	<i>DO</i>	12q13.2-p12.1	Glycoprotein, CD297
Colton	015	CO	3	<i>AQPI</i>	7p14	Aquaporin 1 (channel)
Landsteiner/ Weiner	016	LW	3	<i>LW</i>	19p13.3	LW glycoprotein (IgSF <sup>†</sup> ) CD242
Chido-Rogers	017	CH/RG	9	<i>C4A</i> , <i>C4B</i>	6p21.3	Complement protein 4
Hh	018	H	1	<i>FUT1</i>	19q13	Carbohydrate, CD173
Kx	019	XK	1	<i>XK</i>	Xp21.1	Glycoprotein
Gerbich	020	GE	7	<i>GPC</i>	2q14-q21	Glycophorins C and D, CD236
Cromer	021	CROM	10	<i>DAF</i>	1q32	Glycoprotein, CD55
Knops	022	KN	5	<i>CRI</i> <sup>‡</sup>	1q32	Glycoprotein, CR1 or CD35
Indian	023	IN	2	<i>CD44</i>	11p13	Glycoprotein, CD44
Ok	024	OK	1	<i>BSG</i>	19p13.3	Glycoprotein, CD147
Raph	025	RAPH	1	<i>MER2</i>	11p15.5	Transmembrane glycoprotein, CD151
John Milton -Hagen	026	JMH		<i>SEMA7A</i>	15q23-q24	Glycoprotein, CD108
I	027	I		<i>GCNT2</i>	6p24	Carbohydrate
Globoside	028	GLOB		<i>B3GALNT1</i>	3q25	Glycolipide
GIL	029	GIL		<i>AQP3</i>	9p13	Aquaporin 3 (channel)

\* Four ABO antigens are recognised by the ISBT and noted as 001 (A), 002 (B), 003 (AB) and 004 (A<sub>1</sub>)

<sup>‡</sup> Cluster of Differentiation (antigens)

<sup>†</sup> The immunoglobulin superfamily, IgSF, <http://imgt.cines.fr/textes/IMGTindex/superfamily.html>

<sup>‡</sup> Complement component (3b/4b) receptor 1 (Knops blood group)

In 1999 a database, The Blood Group Antigen Gene Mutation Database (dbRBC), was also set up under support of Human Genom Variation Society (HGVS) lead by O Blumenfeld (Blumenfeld & Patnaik, 2004). This database contains information about variations in genes that affect blood group expression (<http://www.ncbi.nlm.nih.gov/projects/mhc/xslcgi.fcgi?cmd=bgbmut/home>).

### Biochemistry of the ABO blood groups

The biochemical basis of the ABO and H antigens is well understood due to intensive studies during the 1950s and 1960s by the pioneering work on ovarian cyst fluids (which contain large amounts of water-soluble blood-group-active glycoproteins) by Morgan & Watkins and Kabat. Subsequently, this led to the explanation of biosynthesis pathways for ABH antigens, **Figure 1** (Kabat, 1956; references and review in Watkins 1980; Lowe, 1995).

The ABO antigens are not limited to erythroid tissues, but are also found in different tissues and on some epithelial cells (Oriol *et al.*, 1992). Therefore, they can sometimes be noted as histo-blood group antigens (Clausen & Hakomori, 1989). On the other hand, they have not been observed in connective tissues, muscles, and the nervous system (Schenkel-Brunner, 2000). However, in this thesis, as in the field of transfusion medicine, these antigens will be referred as blood groups and not histo-blood groups.

The H antigen is the immediate precursor for creation of blood-group-A-and B-active structures. Two genes, *FUT1* (responsible for synthesis of H antigen on type 2 precursor, *e.g.* on RBC) or *FUT2* (responsible for synthesis of H antigen on type 1 precursor, *e.g.* in secretions) on chromosome 19 encode fucosyltransferases (Oriol, 1995; Lowe, 1995). These 2- $\alpha$ -L-fucosyltransferases catalyse the transfer of L-fucose from the guanosinediphosphofucose (GDP-L-fucose) donor to the terminal galactose of at least one of the precursor types. Six precursor types have been identified but only four types (**Table 3**) are known to carry ABO activity (Clausen & Hakomori, 1989; Clausen *et al.*, 1994). These substances subsequently act as acceptor substrate for the enzyme product of the *ABO* gene. The *ABO* gene product, *i.e.* glycosyltransferases A and B (GTA and GTB) use UDP-GalNAc and UDP-Gal, respectively, as substrates.

**Table 3.** H antigens precursors.

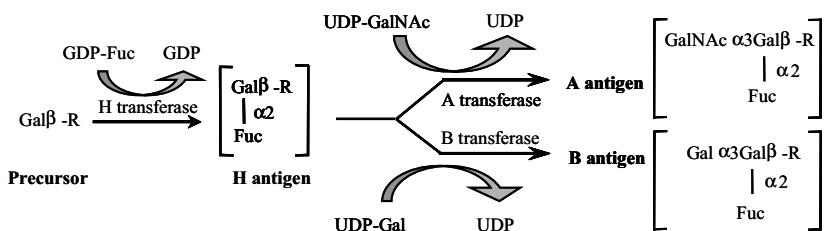
Terminal disaccharide structure of precursor	
Type 1	Gal $\beta$ 1→3GlcNAc $\beta$ 1→R (e.g. endodermal cells)
Type 2	Gal $\beta$ 1→4GlcNAc $\beta$ 1→R (e.g. mesodermal cells, e.g. erythrocytes)
Type 3	Gal $\beta$ 1→3GalNAc $\alpha$ 1→R (e.g. O-linked, linkage to Ser/Thr, repeating A)
Type 4	Gal $\beta$ 1→3GalNAc $\beta$ 1→R (e.g. Glycolipids)

R, indicates the carrier which can be carbohydrate, glycolipid or glycoprotein

ABO(H) antigens are found most commonly on type 1 and type 2 precursors. Type 1 chains are the main carriers of the ABH antigens in plasma and other body fluids. Type 2 precursor chains are the primary acceptors for ABO glycosyltransferases associated with the RBC membrane (Oriol, 1995). Type 3, repetitive A-associated chains are present on RBC glycolipids or secreted in blood group A individuals (Clausen *et al.*, 1985, 1986a-b). Type 4 chain are found in small amounts on RBCs and in larger amounts in kidney cells of persons with blood group A (Breimer & Jovall, 1985). Type 4 ABH structures are only located on glycolipids (Kannagi *et al.*, 1984)

Blood group A individuals possess a 3- $\alpha$ -N-acetylgalactosaminyltransferase (GTA), which catalyzes the transfer of N-acetyl-D-galactosamine from the donor UDP-GalNAc to the H precursor structure Fuca(1-2)Gal $\beta$ -R to provide an A determinant GalNAca(1-3)[Fuca(1-2)]Gal $\beta$ -R. Blood group B individuals possess a 3- $\alpha$ -galactosyltransferase (GTB), which uses the same H structure but catalyzes the transfer of galactose from UDP-Gal to make the B determinant Gal $\alpha$ (1-3)[Fuca(1-2)]Gal $\beta$ -R. Blood group O individuals do not have any functional enzyme and AB individuals have both GTA and GTB (reviewed in Oriol, 1995; Lowe, 1995). The relationship between the A, B and H determinant structures of the ABO blood group is illustrated in **Figure 1**.

**Figure 1.** Schematic representation of the biosynthetic pathway for ABO(H) blood group structures shows the link between H, A and B antigens.



The existence of subgroups of A, *i.e.* A<sub>1</sub> and A<sub>2</sub>, was first recognized by von Dungern and Hirszfeld in 1911 (von Dungern & Hirszfeld , 1911). A<sub>1</sub> RBCs have precursor types 1, 2, 3 and 4, while A<sub>2</sub> erythrocytes appear to have mainly type 1 and type 2, and even small amounts of type 3 (Clausen *et al.*, 1985, 1986b).

The A<sub>1</sub> and A<sub>2</sub> transferases are qualitatively different with pH optima at 5.6 for A<sub>1</sub> and 7-8 for A<sub>2</sub>. The activity of A<sub>1</sub> is higher, it has a lower K<sub>m</sub> for acceptors, and A<sub>1</sub> has a higher isoelectric point than the A<sub>2</sub> enzyme (Watkins, 1980). There is also a quantitative difference between A<sub>1</sub> and A<sub>2</sub> cells that has been illustrated by various techniques, among others, radioimmunoassay with labelled antibodies and lectins, electron microscopy. These studies showed that the number of antigen sites per RBC for A<sub>1</sub> is 8-12x10<sup>5</sup> and for A<sub>2</sub> is 1-4x10<sup>5</sup> (Economidou *et al.*, 1967; Cartron *et al.*, 1974).

Expression of ABO(H) antigens may be affected by different factors, *e.g.* the amount and type of precursor, place of expression and disease. The development of ABO(H) antigens is a complex process and many factors are involved and may affect the final antigen. No attempts will be made to go deeply into the biosynthesis of ABO blood groups in this thesis because of the complexity of the topic and taking into account that the scope of the study is mainly genetic.

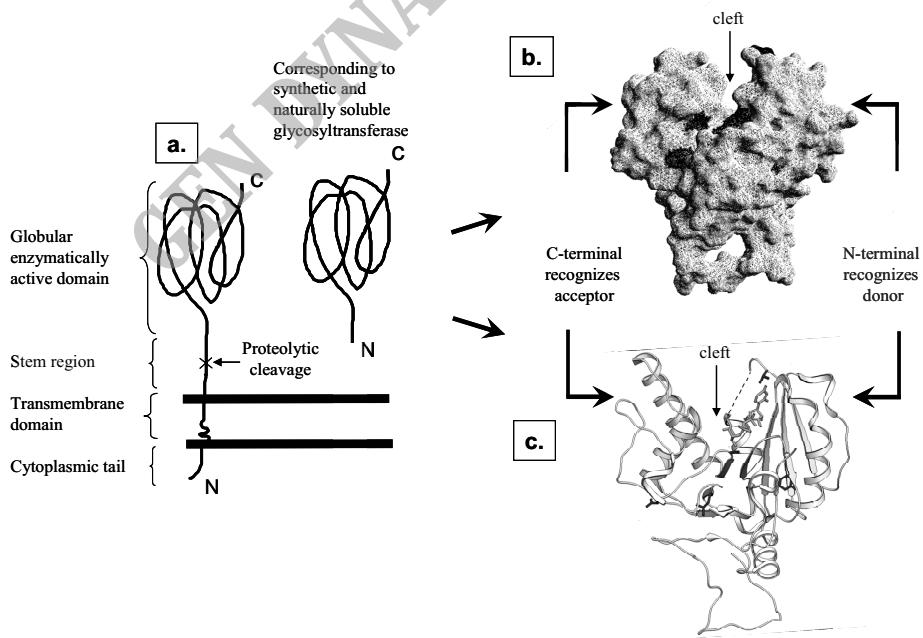
### **The molecular structure of ABO glycosyltransferase**

Glycosyltransferases (GTs), [Enzyme Commission, (EC) 2.4], make up a large family of enzymes that are involved in the biosynthesis of oligosaccharides, polysaccharides, and glycoconjugates. These enzymes show an enormous diversity and are present both in prokaryotes and eukaryotes (Breton *et al.*, 2006). Glycosyltransferases are classified into 87 different families based on substrate/product stereochemistry according to the CAZy database (<http://afmb.cnrs-mrs.fr/~pedro/CAZY/db.html>) (Henrissat, 1998). Knowledge of the sequences of the *ABO* genes (Yamamoto *et al.*, 1990a-b) have established that mammalian GTA (EC 2.4.1.40) and GTB (EC 2.4.1.37) are type II integral membrane proteins containing 354 amino acids (a.a.) and are localised in the lumen of the Golgi apparatus. Type II integral membrane enzymes typically have a short amino-terminal cytoplasmic tail, a hydrophobic membrane domain, a short protease-sensitive stem region and a large catalytic domain that includes the carboxy terminus (Paulson & Colley, 1989), as shown in **Figure 2a**. GTA and

GTB are very similar in the coding regions and the soluble enzyme can be found in serum (Schenkel-Brunner *et al.*, 1972), urine (Chester, 1974) and milk (Ginsburg, 1972). The high-resolution X-ray-crystal structures of the catalytic domains (a.a. residues 63-354) in the GTA and GTB enzymes were resolved by Patenaude and co-workers (Patenaude *et al.*, 2002). **Figure 2b** shows a three-dimensional surface model that predicts the appearance of the active domain in humans. The glycosyltransferase structure has a central cleft where the enzymatic action takes place (**Figure 2b** and **2c**). The enzyme contains an acceptor-recognition domain that binds H antigen and a donor recognition domain that binds UDP-GalNAc or UDP-Gal. GTA and GTB are retaining, metal-dependent family glycosyltransferases (Boix *et al.*, 2001; Pak *et al.*, 2006).

GTA and GTB require the metal ion  $Mn^{2+}$  (manganese) for activity and have a characteristic  $^{211}DVD^{213}$  motif that coordinates attachment of the phosphate in the donor via the metal ion (Breton *et al.* 2006). *In vitro* conversion of O cells to A by GTA as mentioned above, requires the presence of  $Mn^{2+}$  ions but if  $Mn^{2+}$  is changed to  $Mg^{2+}$ ,  $A_1$  glycosyltransferase remains active, but  $A_2$  transferase does not (Schachter *et al.*, 1973).

**Figure 2.** ABO glycosyltransferase. The topology of the transferase is shown in (a) as Golgi-localised membrane-bound enzyme with a cytoplasmic N-terminus and catalytic C-terminal domain (modified from Paulsen & Colley, 1989). The 3D-surface model (b) was created with the Deep View Swiss Pdb Viewer version 3.7 and the 3D-ribbon structure (c) was generated using SETOR and SetoRibbon (unpublished).



## Genetics of the ABO blood group system

The genetic background for a protein polymorphism is straightforward. The occurrence of most such variations is due to a single nucleotide polymorphism (SNP) that results in a single a.a. change, which can directly affect the protein's antigenicity if it is a blood group carrier. The genetics governing carbohydrate structures (*e.g.* ABO antigens) are more complex because the sequential action of a number of gene products, *i.e.* enzymes, is needed to generate antigens of carbohydrate nature. Any polymorphism within these genes may create extra sources of complexity.

In addition, many different genetic backgrounds may underlie one phenotype and different phenotypes appears to be possible result from the same genetic variant. This makes phenotype prediction from an *ABO* genotype difficult (but not impossible).

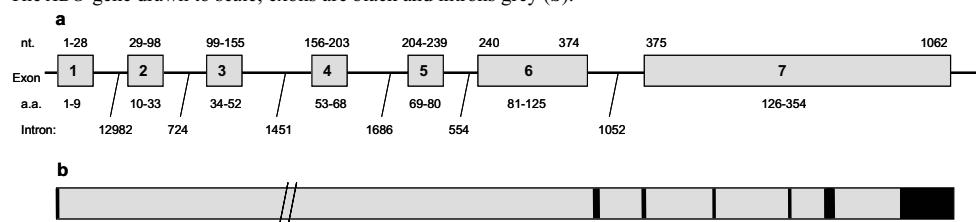
### The *ABO* gene

The ABO blood group locus is located on the terminal portion of the long arm of chromosome 9 and has been determined through quantitative red cell adenylate kinase (AK-1) assay (Ferguson-Smith *et al.*, 1976). This was later confirmed by linkage analysis (Allardice *et al.*, 1986) to position 9q34.1-q34.2.

The first successful isolation and purification of GTA, followed by cloning of cDNA representing the blood group A transferase mRNA, were completed in 1990 from human lung tissue (Clausen *et al.*, 1990; Yamamoto *et al.*, 1990a-b).

The coding region of the *ABO* gene consists of seven exons (exon 1 exists in two variants), spanning about 19.5 kb and ranging in size from 28 to 688 bp (Figure 3). The 1062 bp sequence encodes 354 a.a. corresponding to a 41 kDa protein. The *ABO* gene contains six introns with sizes ranging from 554 to 12982 bp.

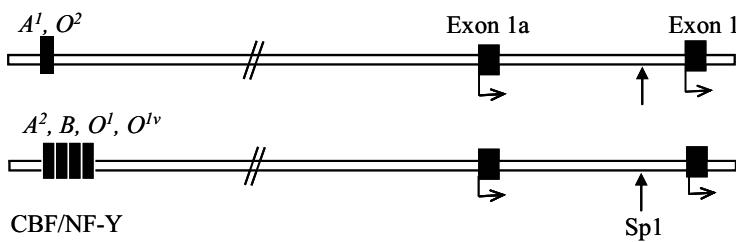
**Figure 3.** Organisation of the *ABO* gene. The seven exons and six introns are not drawn to scale. The numerals above boxes represent the first and last nucleotides of the coding region in each exon and those below boxes show the corresponding a.a. numbers (a). The size of each intron is indicated with a thin oblique bar (a). The *ABO* gene drawn to scale; exons are black and introns grey (b).



The two longest exons (6 and 7) contain 823 of 1062 bp in the transcribed  $A^I$ -mRNA. This covers 77% of the whole encoded protein or 91% of the catalytic domain of the ABO glycosyltransferase. Exons 1-5 encode the amino-terminal cytosolic domain, a transmembrane domain (a.a.17-37), the stem region, and the remaining 9% of the catalytic domain (**Figure 2a**) (Yamamoto *et al.*, 1990a-b; Bennett *et al.*, 1995; Paulson & Colley, 1989). An alternative exon 1a has been identified, located 682 bp upstream of the original exon 1. This exon was found in AC133-CD34+ cultured cells obtained from peripheral blood and represents approximately 2% of all transcripts (Kominato *et al.*, 2002).

Our knowledge of transcriptional regulation of the *ABO* gene is still limited. The presumed promotor of *ABO* was identified by sequencing and functional testing of the 5'-upstream region and it is located between -117 and +31, relative to the upstream translation codon (Kominato *et al.*, 2002; Hata *et al.*, 2002). The promoter region contains CpG islands that can start transcription but neither CAAT or TATA (TATAAAA) boxes were found close to these sites (Yamamoto *et al.*, 1995; Kominato *et al.*, 1997). An enhancer-active minisatellite motif (CBF/NF-Y), located approximately 3.8 kbp upstream of exon 1 in the *ABO* gene, contains either one (alleles  $A^I$  [A101] and  $O^2$  [O03]) or four ( $A^2$  [A201],  $O^I$  [O01],  $O^{IV}$  [O02] and  $B$  [B101]) 43-bp repetitive units (also recognized as a variable number of tandem repeats, VNTR) and appears to play a role in expression (Kominato *et al.*, 1997; Irshaid *et al.*, 1999; Yu *et al.*, 2000), **Figure 4**. A nucleotide substitution G>A at nt. 41 of 43 was identified in all alleles with only one repeat (Irshaid *et al.*, 1999). This region contains a CBF/NF-Y binding site and mutations in this site appeared to decrease expression in an experimental gastric cell line system (Kominato *et al.*, 1997) but a recent report found transcript levels to be independent of VNTR status (Thuresson *et al.*, 2006). A zinc finger transcription factor, Sp1 or Sp1-like protein, may also play a role since there is a binding site in the proximal promoter of *ABO* (Hata *et al.*, 2002, 2003), **Figure 4**.

**Figure 4.** Schematic depiction of the upstream region in different alleles of the *ABO* gene. The figure indicates the relative locations of exon 1, 1a and the repetitive 43 bp units with binding motifs for CBF/NF-Y in the enhancer region and for Sp1 in the proximal promoter. The figure is modified from an original kindly provided by B Thuresson.



The ABO polymorphism is a relatively recent evolutionary event. It probably developed 13 million years ago from a old ancestral gene (Martinko *et al.*, 1993). It has been proposed that A and B antigens were established long before humans and primates diverged (Saitou *et al.*, 1997). A high homology in nt. and a.a. sequences was detected in the *ABO* genes of primates and furthermore human *ABO* shares significant sequence homology with primates (Kominato *et al.*, 1992). Other investigations showed, *e.g.* the murine gene to consist of at least six exons and encoding a *cisAB*-like gene with both GTA and GTB activities but unlike the separate human GTA and GTB (Yamamoto *et al.*, 2001). The genetics of other species with special relevance for human research have been described, *e.g.* the porcine *ABO* gene (Yamamoto & Yamamoto, 2001). The ABO equivalents in other animals like cats (Griot-Wenk *et al.*, 1993) and dogs (Yamamoto *et al.*, 2001) have also been studied but will not be discussed further here. Although the genetic basis of many subgroups has been mapped at the molecular level, samples with unresolved structure are still encountered in a reference laboratory. Numerous point mutations resulting in a number of a.a. changes abolish or weaken the enzymatic activity of ABO glycosyltransferases.

## Serology versus molecular genetics

### **Relationship between phenotype and genotype for the major alleles**

Prior to DNA analysis of the *ABO* gene, serological analyses have been the only practical way to determine blood groups. In general the ABO typing process has two steps which determine either the presence or absence of A and B antigens on the RBCs and also the presence or absence of IgM anti-A and anti-B in the serum or plasma. These two parts of the ABO grouping are sometimes referred to as forward and reverse typing. Human or monoclonal murine anti-A and anti-B reagents are mixed with the RBCs in the first step. The second step involves mixing the serum with suspensions of RBCs known to be group A or B. For the blood group to be interpretable as one of the four major ABO groups A, B, O or AB, the forward and reverse typings must be concordant. Discrepancies can be due to inherited or acquired (*e.g.* transfusion, chimerism, disease etc) phenotype variants. Other ABO subgroups with different A and B characteristics have reduced amounts of A and B antigens on the RBC membrane.

Before this study, there were approximately 70 *ABO* alleles reported to the Blood Group Antigen Gene Mutation Database, “dbRBC” (Chester & Olsson, 2001). The majority of these

alleles are very uncommon; some of them have arisen by simple SNPs, whilst others were formed by recombination between two *ABO* alleles. Certain alleles are more common depending on which populations are involved. The frequency of ABO phenotypes varies notably between different geographic areas and ethnic groups. E.g populations with a high frequency of A phenotype are found mainly in Northern and central Europe. The A<sub>2</sub> phenotype is most common among the Sami people in Northern Scandinavia and the general frequency of A<sub>2</sub> is relatively high in Europe, Southwestern Asia and Africa. On the contrary, A<sub>2</sub> is very rare in native Americans and Australian Aborigines but is also rare in the Eastern Asia (Mourant & Kopec, 1976). On the other hand, the B phenotype is more common in the Far East but almost absent in Amerindians (Mourant & Kopec, 1976). Blood group O is very frequent among native American Indians. Parts of Australia and Africa also show high frequencies of blood group O (Mourant & Kopec, 1976).

This thesis will henceforth consider the A<sup>1</sup>, A<sup>2</sup>, B, O<sup>1</sup>, O<sup>1v</sup> and O<sup>2</sup> alleles as the six major *ABO* alleles. A<sub>1</sub> and A<sub>2</sub> subgroups can be distinguished serologically as shown in **Table 4**. RBCs from A<sub>1</sub> and A<sub>2</sub> persons both react strongly with monoclonal or human anti-A reagents in direct agglutination tests.

Distinction between RBCs of these two subgroups can be made with preparations of the lectins from *Dolichos biflorus* seeds or from albumin glands of vineyard snails, *Helix pomatia*. These lectins react both with A<sub>1</sub> and A<sub>2</sub> RBCs, but at the right dilution, the lectin reagents will readily differentiate A<sub>1</sub> and A<sub>1</sub>B from A<sub>2</sub> and A<sub>2</sub>B. In addition, the plasma from A<sub>1</sub> individuals may contain Anti-H while A<sub>2</sub> person can make anti-A<sub>1</sub>.

**Table 4.** Typical serological reactions in blood samples from persons with A<sub>1</sub> and A<sub>2</sub> phenotype.

Blood Group	Forward typing			Reverse typing			
	Anti-A	Anti-B	Anti-H	Anti-A1	A1	A2	
A <sub>1</sub>	4+	-	-	4+	-	-*	4+
A <sub>2</sub>	4+	-	4+	-	-**	-	4+

\* or 1+ to 4+ representing anti-H

\*\* or 1+ to 4+ representing anti-A<sub>1</sub>

#### A<sup>1</sup> alleles

The A<sup>1</sup> allele represents the "wild-type" A<sub>1</sub> phenotype. The sequence of the A<sup>1</sup> [A101] allele is often denoted as the "consensus sequence" in an *ABO* genotype context. There are variants of A<sup>1</sup> alleles, one of which is very common in Asian populations with a 467C>T [A102]

polymorphism resulting in the substitution Pro156Leu (Yamamoto *et al.*, 1990a). Two minor  $A^1$  alleles [A103 and A104] have been described and differ as follows. The first one also has the 467C>T point mutation but an additional silent mutation 567C>T, the second allele contains a silent polymorphism (also found in  $B$  and  $O^2$  alleles) in nucleotide 297A>G (Ogasawara *et al.*, 1996b).

A105 is like A102 with the same mutation in exon 7, 467C>T but analysis of intron 6 showed additional SNPs compared to A102 (Ogasawara *et al.*, 2001). Another  $A^1$  allele [A106] contains both 297A>G and 467C>T (Ogasawara *et al.*, 2001).

#### $A^2$ alleles

The  $A_2$  subgroup is the most common A phenotype after the  $A_1$  subgroup and the serology is discussed above. The main genetic difference between  $A^1$  [A101] and  $A^2$  alleles [A201] is one point mutation in exon 7, 467C>T (Pro156Leu), and a deletion of one of the three cytosines at nt. 1059-1061 (CCC to CC). The latter mutation results in an extension of the reading frame by 64 nucleotides. This deletion occurs in the codon before the translation stop codon (TGA), resulting in a gene product with an extra 21a.a. at its C-terminus (Yamamoto *et al.*, 1992). The glycosyltransferases encoded by the  $A^2$  allele have lower efficiency, leading to a weaker A phenotype. The enzyme activity has been suggested to be decreased by 30-50 times compared to  $A_1$ , at least according to immunostaining of HeLa cells transfected with A201 constructs (Yamamoto *et al.*, 1992). Some other variants of  $A^2$  alleles have subsequently been elucidated (A202-206). Three  $A^2$ -like alleles were shown to have three different single mutations near the 3' end of exon 7 by Ogasawara,  $A^2$ -2 [A202] contains 1054C>T (Arg352Trp),  $A^2$ -3 [A203] with 1054C>G (Arg352Gly) (Ogasawara *et al.*, 1996a) and  $A^2$ -4 [A205] with both 467C>T and an additional new mutation 1009A>G (Arg337Gly) (Ogasawara *et al.*, 1998). A204 with four common  $B$ -related base substitutions (297A>G, 526C>G, 657C>T and 703G>A) seems to be a hybrid with two extra substitutions 771C>T (silent mutation) and 829G>A (Val277Met) (Ogasawara *et al.*, 1996a). One other rare  $A^2$  subgroup,  $A^2$ -5, [A206] carried only the single deletion (1061delC) without the usual SNP 467C>T (Olsson & Chester, 1996b; Yip, 2000).

#### $B$ alleles

Serologically, blood group B shows strong reaction with anti-B in forward typing. Accordingly, in reverse typing the appearance of agglutination confirms the presence of A

antibodies in the B plasma. High frequencies of group B are found in Central and East Asia and differ significantly compared to Europe, *i.e.* the frequency decreases from east to west (Mourant & Kopec, 1976). When cloning cDNA sequences from cell lines, *B*-1 [B101] was found in the clone FY-59-5 and showed seven single nucleotide substitutions, 297 (silent mutation), 526, 657 (silent mutation), 703, 796, 803 and 930 (silent mutation) throughout exon 6 and 7 and later one extra mutation outside the coding region at the 3' end, at nt. 1096 (Olsson & Chester, 1995). The four a.a. substitutions governed by nt. 526C>G (Arg176Gly), 703G>A (Gly235Ser), 796C>A (Leu266Met) and 803G>C (Gly268Ala) discriminate GTA from GTB (Yamamoto & Hakomori, 1990; Yamamoto *et al.*, 1990a). ABO glycosyltransferases can accordingly be described by using the letters A and B to illustrate the derivation of the a.a. at these four residues. GTA would be represented by AAAA indicating the presence of Arg/Gly/Leu/Gly and similarly BBBB would describe the GTB with Gly/Ser/Met/Ala at residues 176/235/266 and 268. The substitutions at positions 266 and 268 were shown to be responsible for the nucleotide/donor specificity of the transferases (Yamamoto & Hakomori, 1990) and the other residues may have a role in acceptor binding and turnover (Patenaude *et al.*, 2002). Some other variants have been reported afterwards that differ from *B*-1 [B101] by lacking the 930 substitution for *B*-2 [B102], the 657 substitution for *B*-3 [B103] (Ogasawara *et al.*, 1996b), the 526 substitution for *B*-4 [B107] (Ogasawara *et al.*, 1998) and 297 for B108 (Ogasawara *et al.*, 2001).

#### *O* alleles

The serologically-determined pattern of blood group O demonstrates the absence of A and B antigens on the RBC surface in forward blood typing. Reverse blood typing indicates the presence of both anti-A and -B in the plasma.

The first *O* allele [*O*'-1, O01] cloned was shown to be identical to the consensus *A* allele [A101] except for a nucleotide deletion, 261delG, in exon 6. This results in a shift in the reading frame, giving rise to a truncated protein that alters the protein sequence after a.a. 88. A stop codon halts translation after a.a. 117 and the resulting protein is enzymically inactive (Yamamoto *et al.*, 1990a). Some other *O*'-1-like alleles that are characterized by the presence of the 261delG and at least one additional point mutation are presented in **Table 5**.

A second kind of *O* allele has the same inactivating deletion (261delG) as the original *O* allele (*O*'-1 [O01]), but in addition has nine point mutations spread throughout exons 3 to 7

(Yamamoto *et al.*, 1990a; Olsson & Chester, 1996a) and a further 13 mutations have been found amongst the intron 6 (Suzuki *et al.*, 1997; Olsson & Chester, 1998a). This allele is referred to as  $O^{Ivariant}$  ( $O^{Iv}$ -1) [O02] since the inactivating mutation is the same as in the original  $O^I$  allele. **Table 5** shows some more variants of this allele.

Other  $O$  alleles not due to 261delG also exist that are caused by other inactivating mutations along the reading frame. The first allele described of this type,  $O^2$ -1 [O03], has a critical mutation (802G>A) which causes an a.a. change (Gly268Arg) that prevents the enzyme from utilising the nucleotide sugar donor (Yamamoto *et al.*, 1993, 1996). Additionally, two polymorphic sites in exon 2 and 5, 53G>T and 220 C>T, respectively (Amado *et al.*, 2000) and 297A>G in exon 6, 526 C>G in exon 7 (Yamamoto *et al.*, 1993d) and 1096G>A (Olsson & Chester, 1995) have been found. This  $O$  allele comprises approximately 2-5% of  $O$  alleles in Caucasians but seems to be absent, or at least very rare, in other populations (Grunnet *et al.*, 1994; Chester & Olsson, 2001).

It has been debated whether  $O$  alleles are translated into protein at all. Even if definitive proof is lacking, the presence of  $O_2$  protein has been shown by MAb staining of tissue sections (Amado *et al.*, 2000) and a hint towards alloreactivity against polymorphic peptides derived from the N-terminal parts of  $O^I$ - and  $O^{Iv}$ -translated sequences suggested in mixed lymphocyte cultures (Eiz-Vesper *et al.*, 2005).

**Table 5.** Schematic comparison of known *O* allele mutations and deduced a.a. sequence, [ $O^{'-1}, 001$ ],  $O^{variant}(O^{J_Y}-1)[002]$ , and  $O^{2-1}[003]$ .

\* Additionally, two more polymorphic sites were shown in exon 2 (53G>T) and exon 5 (220C>T).

## Genetic background of hybrid alleles, rare *O* alleles and weak subgroups

### Hybrid alleles

From 1996 and onwards, it became increasingly clear that *O* alleles with the 261delG mutation could also include *A<sup>2</sup>*-or *B*- related SNPs. Mixtures of *O<sup>I</sup>* and *O<sup>Iv</sup>* sequences were also observed. The implications and consequences for *ABO* genotyping were serious (Olsson & Chester 2001). Some important *O* hybrids have been summarized in **Table 6**.

**Table 6.** Hybrid alleles reported before the start of this study. Nucleotide changes are compared with the consensus allele.

Exon	2	4	5	6	7	Reference
<b>nt. position</b>						1
1	1	1	2	2	4 5 6 6 6 7 7 7 8 8 9 0	
0	8	8	2	6	6 2 4 5 8 0 7 9 0 2 3 6	
6	8	9	0	1	7 6 6 7 1 3 1 6 2 9 0 1	
<b>Consensus</b>	G	G	C	C	G A C C T C G G C C G G G C	
<b>Hybrid alleles</b>						
<i>O<sup>Iv</sup>-B</i> [O24]	T	A	T	T	delG G G T A A C A	Olsson et al., 1997 [1]
<i>O<sup>R</sup>-2</i> [ <i>O<sup>I</sup>-O<sup>Iv</sup></i> ] (1)					delG G A A T A	[1]
<i>O<sup>R</sup>-3</i> [ <i>O<sup>I</sup>-O<sup>Iv</sup></i> ] (2)			T	delG G A A A T A	[1]	
<i>O<sup>R</sup>-4</i> [ <i>O<sup>I</sup>-O<sup>Iv</sup></i> ]				delG A A T A	Suzuki et al., 1997	
<i>O<sup>R</sup>-5</i> [ <i>O<sup>I</sup>-O<sup>Iv</sup>-O<sup>I</sup></i> ]	A	T		delG T		
<i>O<sup>R</sup>-6</i> [ <i>O<sup>I</sup>-A<sup>2</sup></i> ], [O22]				delG T	delC [1] & Gassner et al., 1996	
O202 [O18]				delG A A T A	Ogasawara et al., 2001 [2]	
<i>O<sup>R</sup>-7</i> [O17], [ <i>O<sup>Iv</sup>-O<sup>I</sup></i> ]				delG G A T A	[2] & Ogasawara et al., 1996	
<b>Amino acid</b>					1 1 2 2 2 2 2 2 2 2 3 3	
3	6	6	7	8	5 7 1 1 2 3 5 6 6 7 1 5	
6	3	3	4	7	6 6 6 9 7 5 7 6 8 7 0 4	
<b>Consensus</b>	V	R	R	P	V T P R F H P G P L G V L P	
<b>Change</b>	F	H		S	L G I S M A M	

### Rare *O* alleles

#### *O<sup>3</sup>* allele

A few other rare *O* alleles have subsequently been found in single samples, such as *O<sup>3</sup>* [O08] that does not have 261delG but instead contains both the common *A<sup>2</sup>* allele polymorphisms 467C>T and C-deletion at nt. 1059-1061 and an insertion of an extra guanosine in the 7-guanosine sequence at nt. 798-804 (Olsson & Chester, 1996c). This insertion is also responsible for *A<sup>el-1</sup>* [Ael01, A109] (see later).

### Other rare *O* alleles

Two other rare *O* alleles lacking 261delG were reported in the Japanese population by Ogasawara. O301 [O14] has the missense mutation 893C>T (Ala268Val) on an A102 background whereas O302 [O15] has the nonsense mutation 927C>A (Tyr309Stop) on an A101 background (Ogasawara *et al.*, 2001).

### Weak subgroups

In addition to the major phenotypes characterised by either strong or absent haemagglutination with anti-A/-A<sub>1</sub> and -B reagents, the ABO blood group system also includes phenotypes in which erythrocytes react weakly with the anti-A and -B reagents, for example A<sub>3</sub>, A<sub>x</sub>, A<sub>finn</sub>, A<sub>el</sub>, B<sub>3</sub>, B<sub>x</sub>, B<sub>v</sub>, B<sub>el</sub>, *cis*-AB and B(A) (Daniels, 2002).

In some cases of weak phenotypes, it is difficult to determine clearly their specific ABO subgroup by conventional serological methods. Indeed this is sometimes a question of interpretation between laboratories and even between individuals in the same laboratory. The weak subgroups are important in more than one way. First, they risk complicating patient and donor blood group determination. At worst, the wrong group can be assigned if *e.g.* a weak antigen is missed. Second, they allow us to characterize and understand glycosyltransferase mechanisms by studying the results of mutations in the underlying alleles. Any defect in the *ABO* gene complicates ABO biosynthesis and the transcription and translation of the gene products (*e.g.* GTA and GTB) can fail. In addition, any factor including substrate or acceptor availability or localisation of the GTA/GTB could be the limiting factor, at least in theory. In general, SNPs in exon 7 have turned out to be significant in generating weaker phenotypes.

### *A*<sup>3</sup> alleles

A typical serological pattern illustrates that RBCs from individuals with the A<sub>3</sub> phenotype agglutinate strongly with anti-A and anti-A,B but show a large number of free cells.

Investigation of enzyme activity in sera of A<sub>3</sub> individuals have shown two different results. In one study, normal A<sub>1</sub> enzyme activity was reported and in another much lower than A<sub>1</sub> (Cartron *et al.*, 1978; Nakamura *et al.*, 1989). One A<sub>3</sub>B individual had a novel point mutation, 871G>A (Asp291Asn), on the *A'* [A101] background (Yamamoto *et al.*, 1993a) and this allele was named A301. Two research groups (Olsson & Chester, 1996b; Barjas-Castro *et al.*, 1997)

reported separately the absence of this missense mutation in seven Swedish and 11 Brazilian serologically defined A<sub>3</sub> samples, thus finding only consensus  $A^I$  sequences in exons 6 and 7. The latter research group instead presented A302 which has two single point mutations, the substitution 829G>A (Val277Met) and 1061delC. In fact, A<sub>3</sub> appears to be more complex compared to other subgroups and the true allelic background is not yet solved.

#### $A^x$ alleles

These alleles are responsible for the rare subgroup A<sub>x</sub> (Fischer & Hahn, 1935) and the RBCs typically show a positive reaction with anti-A,B and anti-H but notably weaker (if any) reactions with anti-A. There are four base substitutions involved in these alleles; 646T>A (Phe216Ile), 681G>A (silent), 771C>T (silent), 829G>A (Val277Met).  $A^x$ -1 [Ax01, A108] has the missense mutation 646T>A on an  $A^I$  background (Yamamoto 1993b, Olsson & Chester 1998a). Combination of 646T>A and 681G>A on the  $A^I$  background was demonstrated and designated as Ax04 [A113] (Ogasawara *et al.*, 2001). Olsson & Chester presented additional alleles associated with the A<sub>x</sub> phenotype, namely  $A^x$ -2 [Ax03] and  $A^x$ -03 [Ax02] which contained all four defining substitutions but  $A^x$ -03 had a 297A>G mutation (silent) in addition. These authors concluded that  $A^x$ -2 and -3 alleles are in fact hybrid alleles between common *ABO* alleles ( $A$  and  $O^{IV}$  or  $B/O^2$  and  $O^{IV}$ ) differing by their crossing-over regions in intron 6 (Olsson & Chester, 1998a).

#### $A^{el}$ alleles

Forward typing in this subgroup does not show any agglutination by anti-A or anti-A,B but RBCs react strongly with anti-H. The reverse typing shows that plasma lacks antibodies against A<sub>1</sub> and A<sub>2</sub> RBCs. The key analysis is a positive eluate after adsorption with human polyclonal anti-A or anti-A,B.  $A^{el}$ -1 [Ael01] has a single G insertion in the seven-guanosine sequence at nt. 798 to 804. This novel insertion results in a frame-shift that alters the a.a. sequences after the glycine at position 268 which is located immediately at the nucleotide-sugar binding site of the enzyme (Olsson *et al.*, 1995a). The encoded transferase is expected to be 37 a.a. longer than the normal consensus allele, and 16 a.a. longer than  $A^2$ -encoded transferase. It is surprising but very consistent in the literature that this anomalous transferase has the ability to synthesize A antigens at all (Ogasawara *et al.*, 1996a). Experiments with gold-labelled monoclonal anti-A showed that approximately 1% of all A<sub>el</sub> RBC express A antigen strongly whereas the rest of the cells remain almost unlabelled (Hansen *et al.*, 1998).

Heier's group had hypothesized that similar findings in other subgroups could be based on recombination tendency and that eventually somatic cell clones in or outside the haematopoietic cells may result in small amounts of cells with another phenotype in the same individual, thus causing a pseudo-chimeric pattern (Heier *et al.* 1994).

One other allelic member in this subgroup is  $A^{el-2}$  [Ael02] that contains two SNPs, 646T>A and 681G>A on the A102 background with two a.a. substitutions, Pro156Leu and Phe216Ile compared to consensus (Ogasawara *et al.*, 1996a). This is actually very similar to an  $A^x$  allele. It is probably the a.a. substitution Phe216Ile in combination with the otherwise neutral Pro156Leu that dramatically decreases the enzymatic activity of the GTA.

#### Other $A^{weak}$ alleles

The serological characteristics of some other minor phenotypes included among A subgroups, e.g.  $A_{end}$ ,  $A_m$ ,  $A_y$ ,  $A_{finn}$ ,  $A_{bantu}$  and the collective description  $A_w$  are more unclear. Guiding principles exist (**Table 7**), but any pattern of agglutination is due to several factors, e.g. the condition of the RBCs, the antibodies used to test them and the experience of the analyst. Thus, it can be discussed if all samples categorized and reported as  $A_{end}$  e.g. can be differentiated from all other A subgroups.

**Table 7.** Serological reaction patterns adopted from current text books. A negative reaction is noted by 0 and positive reactions are denoted from + (very weak agglutination) to 4+ (maximal agglutination).

Subgroup of A	RBC reactions with				ABO substances in saliva	Anti-A <sub>1</sub> in serum
	Anti-A	Anti-A,B	Anti-A1	Anti-H		
$A_1$	4+	4+	4+	0	A, H	No
$A_2$	4+	4+	0	4+	A, H	Sometimes
$A_{int}$	4+	4+	2+/3+	2+/3+	A, H	No
$A_3$	2+/+ <sup>mf</sup>	2+/+ <sup>mf</sup>	0	4+	A, H	No
$A_x$	0/+	2+/+	0	4+	H	Often
$A_{el}$	0 <sup>§</sup>	0 <sup>§</sup>	0	4+	H	Sometimes
$A_{end}$	+	+	0	4+	H	Sometimes
$A_{finn}$	+	+	0	4+	H	Yes
$A_{bantu}$	+(+)	+(+)	0	4+	H	Yes
$A_m$	0/+	0/+	0	4+	A, H	No
$A_y$	0 <sup>§</sup>	0 <sup>§</sup>	0	4+	A, H	No

<sup>§</sup> anti-A can adsorbed and eluted from these cells despite absence of agglutination.

<sup>mf</sup> mixed field agglutination

$B^{weak}$  alleles

The above uncertainty is particularly true about  $B_{weak}$  phenotypes which are especially difficult to classify. B variants are much more uncommon than A variants, while this may just reflect the relatively low frequency of the B blood group in many populations. Accordingly, most B subgroups so far have been characterized in Asian populations.

The molecular basis underlying many of these B subgroups has been studied and some  $B^{weak}$  alleles elucidated so far. Guidelines exist and are summarized in **Table 8** (Daniels, 2002). Typically, these phenotypes often appear to result from missense mutations at the *ABO* locus causing single a.a. changes in the GTB.

**Table 8.** Characteristics of some more frequent,  $B_{weak}$  phenotypes

Subgroup of B	RBC reactions with				ABO substance in saliva	Anti-B in serum	GTB in serum
	Anti-A	Anti-B	Anti-A,B	Anti-H			
B	0	++++	++++	++	B, H	None	Yes
$B_3$	0	mf	mf	+++	B, H	None	Yes
$B_x$	0	w	w	+++	( $B_x$ ), H	Yes	None
$B_m$	0	*/w	0/w	+++	B, H	None	Yes
$B_{el}$	0	0	0	+++	H	Sometimes	None

mf, mixed field; w, very weak agglutination

\*Anti-B may be adsorbed onto and eluted from these cells ( $B_x$ ), may require inhibition of agglutination of  $B_x$  cells for detection

The  $B_3$  phenotype was proposed to depend on the missense mutation 1054C>T (Arg352Trp) on a  $B$ -1 [B101] background and named B301 (Yamamoto *et al.*, 1993a). A  $B^x$  allele [Bx01 or B104] responsible for the  $B_x$  phenotype in a Japanese sample had a point mutation at nt. 871G>A (Asp291Asn) (Ogasawara *et al.*, 1996a) which was also found in an  $A_3$  sample (Yamamoto *et al.*, 1993a).

The  $B_{el}$  phenotype was divided into two suballeles,  $B^{el}$ -1 and  $B^{el}$ -2 [Bel01 and Bel02], which had substitutions at 641T>G (Met214Arg) and 669G>T (Glu223Asp), respectively (Ogasawara *et al.*, 1996a). Presumably, these missense mutations reduce the enzymatic activities of the GTB. Ogasawara *et al.* also found another  $B^3$  allele, B302, which differs from  $B$  consensus, by two nucleotide substitutions, 646T>A (Phe216Ile) and 657T>C (same nucleotide as A consensus). Thus, again the same mutation found in  $A_x$  also appears to be responsible for a  $B_3$  phenotype. It can be speculated if this allele was really a hybrid,  $B-O^{lv}-B$

### CisAB alleles

H Seyfried described this rare phenotype first in 1964 (Seyfried *et al.*, 1964). It represents a very interesting phenomenon that proved that it is possible for a child with blood group O to have a parent with blood group AB. Seyfried hypothesized that both A and B determinants of the AB blood group could be located on the same chromosome. Later on, this idea confirmed in that the existence of an exceptional *ABO* allele encoding a glycosyltransferase is indeed able to produce both A and B enzymes at the same time. This rare blood group is characterized by the presence of A at levels comparable to the A<sub>2</sub> phenotype, weakened B and prominent H antigen expression on RBCs. An anti-B reactive only at room temperature and lower is often present on reverse typing of *cis*-AB individuals and most also has a cold-reacting anti-A<sub>1</sub> (Pacuszka *et al.*, 1975; Daniels, 2002). *Cis*-AB01 was sequenced and showed the substitution 803G>C (Gly268Ala) on the *A*<sup>I</sup>-2 [A102] background and thus can be described as AAAB (Yamamoto *et al.*, 1993c). Another allele named *cis*-AB02 was discovered when a Vietnamese man who was to undergo organ transplantation showed irregular blood grouping results. The sequencing showed that his *ABO* genes was nearly identical to the normal *B* allele except for a 796A>C (Met266Leu) substitution. This enzyme can also be designated BBAB (Mifsud *et al.*, 2000).

### *B(A)* alleles

As the name suggests, this phenomenon was found in individuals who express normal B antigens on RBCs with anti-A in serum but whose RBCs also react with some potent monoclonal anti-A reagents (Goldstein *et al.*, 1989). The first reported *B(A)* allele (Yamamoto *et al.*, 1993b) showed that the nucleotide constellation in exon 7 was BABB, [noted in bold face: 526C>**G** (Arg176Gly), 703**G**>A (Gly235Ser), 796C>**A** (Leu266**Met**), 803G>**C** (Gly268**Ala**)]. This allele was designated B(A)01. Another *B(A)* allele named *B(A)*02, had a composition of BBBB and further showed a novel nonsynonymous substitution at 700C>G which produces an a.a. substitution, Pro234Ala. It is important to note that this substitution occurs next to the second (residue 235) of the four A- vs. B-characteristic a.a. residues just mentioned above. Accordingly, this reduced the GTB activity and enhanced the normally small amounts of GTA activity that are present in serum (Yu *et al.*, 1999).

## AIMS OF THE THESIS

The aim of this thesis was to continue earlier work at this department on the elucidation and description of the diversity of *ABO* alleles. Furthermore, the purpose was also to contribute to a deeper knowledge about the relationship between the structure and function of *ABO* subgroups.

More specifically the aims were:

- To elucidate new alleles in the *ABO* blood group gene in samples either with suspected A and B subgroups or other blood grouping discrepancies
- To assess the impact of polymorphisms in the *ABO* gene in some weak subgroup phenotypes by 3D-modelling, enzyme kinetics and crystal structure resolution of the mutant glycosyltransferase
- To develop and evaluate a novel *ABO* genotyping method for safer blood grouping in clinical practice

## MATERIAL AND METHODS

### Blood and DNA samples

Blood samples with common *ABO* phenotypes, undefined or rare phenotypes in **Papers I-VI** were obtained either from apparently healthy blood donors or in some cases patients whose samples had been referred to the Nordic Reference Laboratory for Genetic Blood Group Typing, and were from different geographic/ethnic backgrounds. This included (but was not limited to) Swedish, Swiss, South African, Brazilian, Chinese, Israeli, English, Turkish and Jordanian samples. The blood was obtained in tubes containing ethylenediaminetetraacetic acid (EDTA) or acid-citrate dextrose (ACD) as anticoagulants. DNA was isolated using a salting-out procedure described by Miller *et al.*, 1988.

### *ABO* genotyping

#### Polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP)

A PCR-RFLP method was used for the initial *ABO* genotype screening. The principle for this method is based on using the PCR technique to amplify exons 6 and 7. The amplified fragments cover the two major exons and constitute 91% of the catalytic domain. Subsequent digestion of the fragments simultaneously using the restriction enzymes *Hpa*II and *Kpn*I

results in a specific fragment pattern for the common *ABO* alleles and thus interpretation of at least 15 different *ABO* genotype can be obtained (Olsson & Chester, 1995). The limitation in this method is that it can only detect mutations at these specific points (restriction enzyme sites) in exon 6 and 7. Accordingly, the major *ABO* alleles ( $A^1$ ,  $A^2$ ,  $B$ ,  $O^1/O^{1v}$  and  $O^2$ ) can be identified as well as a few subgroup alleles. As with other *ABO* genotyping methods, other rare or novel mutations would be missed.

#### **Allele-specific primer PCR (PCR-ASP)**

PCR was used throughout (**Papers I-VI**) both to amplify gene-specific fragments (regardless of allele) and allele-specific fragments (PCR-ASP). A method linking mutations in exons 6 and 7 by PCR with allele-specific primers (ASP) across intron 6 was used when exclusion of recombinant hybrid alleles was required (first mentioned in **Paper I** but mainly used in **Paper VI**). Oligonucleotide primers used in this thesis were synthesized by DNA Technology APS. The primers and conditions used for PCR are described in the respective articles and more extensively in **Paper VI**.

#### **Direct DNA sequencing**

In all studies PCR was used to obtain sufficient material for analysis either by gene-specific or allele-specific primers (ASP). The former primers allow amplification of all known variants of the particular gene being examined. The latter approach only allows amplification of segments having a polymorphism specific for the primer. After electrophoretic separation of the reaction products on agarose gels these were visualised by staining with ethidium bromide, excised from the gel and purified using the Qiaquick gel extraction kit (for details see **Papers I-VI**). The Big Dye Terminator Cycle Sequencing kit and an ABI PRISM 310 or 3130 Genetic Analyser were used for direct DNA sequencing by capillary electrophoresis and automated fluorescence-based detection according to the manufacturer's instructions.

Sequence analysis was performed with SeqEd software 1.03.

#### **Reverse transcriptase PCR (RT-PCR)**

In **Paper IV**, RNA was prepared from peripheral blood samples using both TRIzol<sup>®</sup> LS Reagent and RNeasy kit. Complementary DNA (cDNA) was synthesized using random hexamers, random octamers or oligo-d(T) primers from the GeneAmp RNA PCR kit according to the manufacturer's instructions. Two rounds of PCR amplification, *i.e.* nested

PCR, were performed improving both the specificity and the amount of required fragments for further analysis.

### **Site-directed mutagenesis and cloning**

A Quick Change II XL Site-directed Mutagenesis Kit (Stratagene) was used to generate the required mutant enzymes. The plasmid pCWΔlac harbouring the -10GTB gene was used as a template. This sequence for -10GTB lacks the N-terminal transmembrane domain-encoding region and encodes a GTB of residues 63-354 containing the catalytic domain. The PCR products generated (**Paper V**) were treated with restriction enzyme *Dpn*I to digest template methylated, non-mutated plasmid DNA. The mutated *Dpn*I-treated fragments were cloned and transformed into *E. coli* XL10-Gold cells. Conformation of mutagenesis was done by sequencing as described above. Positive mutants were transformed into *E. coli* BL21 Gold cells for expression. These experiments were performed at the Carlsberg Laboratory, Copenhagen, Denmark.

### **Protein purification and enzyme kinetics**

The recombinant enzyme was purified using a two-step protocol comprising ion-exchange chromatography followed by UDP-hexanolamine affinity column chromatography as has been described elsewhere (Seto *et al.*, 1999). Enzyme kinetic studies were performed using a radiochemical assay (Palcic *et al.*, 1988).

In this assay a Sep-Pak reverse-phase cartridge is used to isolate radiolabelled reaction products that are produced when label is transferred from a radioactive donor to a hydrophobic acceptor (Palcic *et al.*, 1988). The kinetic parameters  $k_{cat}$  and  $K_m$  were obtained by non-linear regression analysis of the Michaelis-Menten equation with the Graph Pad PRISM 3.0 program (GraphPad Software). A unit of enzyme is the amount that converts one micromole of substrate to product in one minute. Protein concentrations were estimated using the Bradford method using bovine gamma globulin as a protein standard.

### **Computer modelling and crystallization**

The GTB mutant structures were crystallized by collaborators in Canada using conditions similar to the native GTB enzymes as described previously (Patenaude *et al.*, 2002; Breton *et al.*, 2006). Data were collected on a detector (Pflugrath, 1999) and X-rays were produced and subsequently the crystals were frozen. All structures were solved by using molecular

replacement techniques with wildtype GTB (Protein Data Bank accession code 1LZ7) as a starting model and were refined using the CCP4 program suite (The CCP4 program suite, 1994).

## RESULTS AND DISCUSSION

The molecular genetic background of the common and a few of the variant phenotypes at the ABO blood group locus was described by Yamamoto and his colleagues in the 1990s. Until 2001, this studies was started, several research groups had reported new findings, and the number of *ABO* alleles registered in dbRBC\* had increased to more than 70 (Chester & Olsson, 2001). However, the molecular basis for a number of variant phenotypes in the ABO blood group system had not been yet defined.

### Single point mutations associated with A or B subgroups (Papers I and V)

To date, we have investigated the *ABO* gene in numerous samples of clinical or research purposes. For the studies presented in **Papers I** and **V** approximately 400 samples were included. These collected data constitute one of the most extensive sets of clinical samples of its sort and have been referred to our laboratory during the last decade. In most cases, serological discrepancies were the reasons for the referral to the Nordic Reference Laboratory for Genetic Blood Group Typing. On the basis of medical history, phenotype condition and even initial genotyping results (for some more recent samples from the referring centres), all 324 samples investigated in **Paper I** were categorized in different subdivisions, e.g. ABO discrepancies due to weak A or B expression in pregnancy or haemological disorders, acquired B syndrome or acquired A-like antigen, suspected chimeras or weak A/B expression without known reason. DNA was prepared and genotype determined according to published methods (Olsson & Chester, 1995; Olsson & Chester, 1996b; Hosseini-Maaf et al, 1998).

Even if **Paper I** was a rather clinically oriented article, the main focus in this thesis is on alleles and mutations associated with ABO subgroups. The 1<sup>st</sup> and 2<sup>nd</sup> authors are both MDs who were responsible for the clinical interpretations and consequences of our findings.

All samples were found in the initial *ABO* genotype screening by PCR-RFLP and PCR-ASP to be heterozygous for at least one *A*- or *B*-like allele. Similarly, all samples were examined at

---

\* <http://www.ncbi.nlm.nih.gov/projects/mhc/xslcgi.fcgi?cmd=bmut/home> (Blumenfeld & Patnaik, 2004)

the CBF-NF/Y enhancer regions approximately 4 kbp upstream from the translation start codon, and the PCR results were as expected without any deviation (Irshaid *et al.*, 1999). In total, the phenotype and genotype mis-correlation led us to sequence 42 samples (20 with suspected A subgroups and 22 with suspected B subgroups) for exons 1-7 of the *ABO* gene and two proposed regulatory motifs (Kominato *et al.*, 1997). All samples were homozygous for the consensus sequence in exons 1-5, except those containing an *O<sup>Iv</sup>* allele or *A<sup>w</sup>-3* (see below). As expected, *O<sup>Iv</sup>* alleles were polymorphic at nt. 106 (exon 3), 188/189 (exon 4) and 220 (exon 5) (Olsson & Chester, 1996a) and the *O<sup>2</sup>* alleles contained polymorphic sites at nt. 53 (exon 2) and 220 (exon 5) (Amado *et al.*, 2000).

The summarized sequencing results of seventeen novel A and B subgroup alleles with single point mutations are shown in **Table 9**.

**Table 9.** The seventeen novel A and B subgroup alleles reported in **Papers I** and **V**. Nucleotide and a.a. changes are given in comparison to the common *A* or *B* alleles.

Allele description	Nucleotide change(s)	Amino acid change(s)	Pos/ Neg*	Invariant‡	Conserved‡	Geographic origin	Accession no.
<i>A<sup>w</sup>-1</i>	407C>T	Thr136Met			X	United Kingdom	AF321010
<i>A<sup>w</sup>-2</i>	350G>C	Gly117Ala		X		Various**	AF324009
<i>A<sup>w</sup>-3</i>	203G>C	Arg68Thr	Pos			Scandinavia	AF324008
<i>A<sup>w</sup>-4</i>	721C>T	Arg241Trp	Pos			Belgium/ Germany	AF324011
<i>A<sup>w</sup>-5</i>	965A>G	Glu322Gly	Neg	X		Finland	AF324012
<i>A<sup>w</sup>-6</i>	996G>A	Trp332stop			X	New Zealand	AF324013
<i>B<sup>w</sup>-2</i>	873C>G	Asp291Glu	Neg	X		France	AF324018
<i>B<sup>w</sup>-3</i>	721C>T	Arg241Trp	Pos			Sweden	AF324014
<i>B<sup>w</sup>-4</i>	548A>G	Asp183Gly				Sweden	AF324016
<i>B<sup>w</sup>-5</i>	539G>A	Arg180His	Pos		X	USA	AF324015
<i>B<sup>w</sup>-6</i>	1036A>G	Lys346Glu	Pos	X		Finland	AF324019
<i>B<sup>w</sup>-7</i>	1055G>A	Arg352Gln	Pos	X		USA	AF324020
<i>B<sup>w</sup>-8</i>	863T>G	Met288Arg	Pos	X		Turkey	AF324017
<i>B<sup>w</sup>-15</i>	565A>G	Met189Val		X		Turkey	DQ646577
<i>B<sup>w</sup>-16</i>	575T>C	Ile192Thr			X	Pakistan	DQ646578
<i>B<sup>w</sup>-17</i>	784G>A	Asp262Asn	Neg	X		South Africa	DQ646580
<i>B<sup>w</sup>-18</i>	802G>A	Ala268Thr			X	France	DQ646581

\* Charged residues involved in a.a. changes, Pos = positive and Neg = negative

\*\* New Zealand, England, Australia and USA

‡ According to published data on nine homologous genes (Gastinel *et al.*, 2001).

All these novel alleles carried single missense mutation except the  $A^x$ -6 allele which carried a nonsense mutation. All mutations except one were located in exon 7 and affect the enzymatically active domain.

The mutation 203G>C provides the first example of a mutation outside exon 6-7 that is associated with phenotype change. It has been later shown by transfections to HeLa cells that this mutation appears to weaken enzyme activity (Seltsam & Blasczyk, 2005). Screening for subgroup mutations found in this study (**Paper I**) by PCR-ASP in samples of normal ABO phenotypes was negative except for one sample with the additional presence of the 203G>C (Arg68Thr) mutation in an  $A^2$  allele. The screening undertaken in **Paper V** was negative.

#### How can the effects of these mutations be explored?

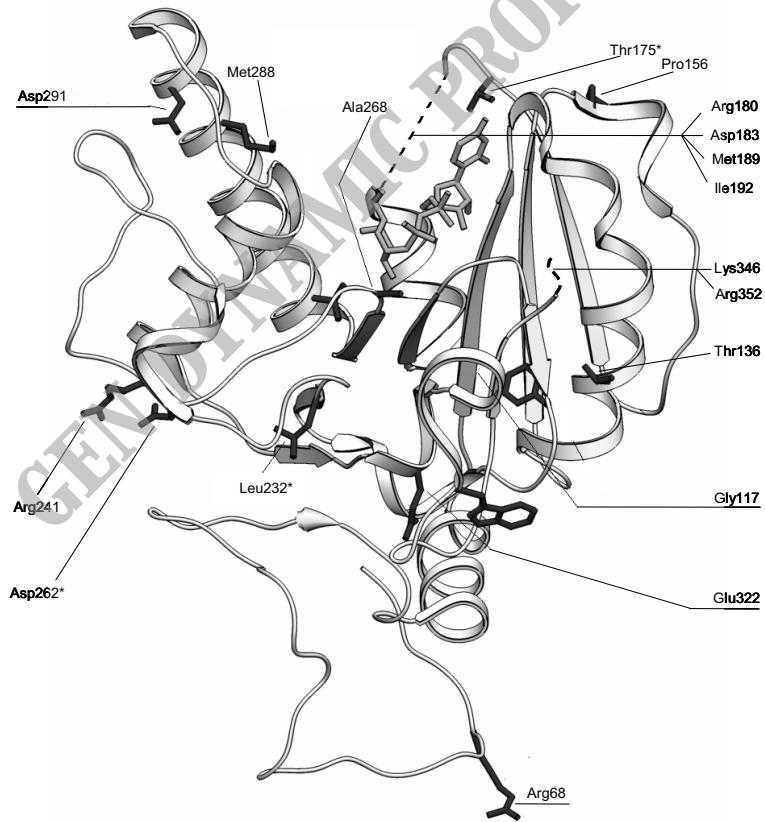
The obvious question raised was how these mutations affected the enzyme structure and/or activity and furthermore the quality/quantity of antigens on RBC. There are different methods to examine this, *e.g.* transfection to cell lines, measuring the enzyme activity and kinetics, and using flow cytometry. An early attempt to use molecular modelling was done but not shown by Yamamoto but this was presumably limited by the available software (Yamamoto *et al.*, 1993a). In collaboration with colleagues in Canada, a computed molecular model was used to show the presumed effect of the mutations in the 3D structure of the B glycosyltransferase. Together with Monica Palcic's group in Denmark, we purified the enzyme and measured the enzyme kinetics for some of the  $B_{\text{weak}}$  subgroups (**Paper V**). Computed modelling to show the location of all these SNPs is illustrated in **Figure 5**. We investigated two new missense mutations found in  $B^w$ -17 (Asp262Asn) and  $B^w$ -19 (Phe216Ile) in **Paper V**. The calculations of computed modelling and kinetic constants supported each other in these cases and the predictions were good enough to understand the potential effects of SNPs in these weak B blood groups.

We also showed computed modelling of one recently published  $B^w$ -11 (Leu232Pro) (Deng *et al.*, 2005) in **Paper V**.

Of special interest was the missense mutation 721C>T (Arg241Trp), a mutation that can lead to both a weak A and B phenotype. This is the first example of the same point mutation in two different alleles affecting the expression of either antigen: another example is the 502C>G mutation which results in both weak A and B antigens (Seltsam *et al.*, 2002; Lin *et al.* 2003). Later, the 539G>A mutation found here was shown also to cause both weak B and A phenotypes (Svensson *et al.*, 2005).

In one case, we succeeded in determining the crystal structure of a subgroup glycosyltransferase, the Arg268Thr mutant enzyme in *B<sup>w</sup>-18* with the 802G>A mutation. Unfortunately, the structure resolution was not clear enough to determine precisely the effect of this mutation but enzyme kinetic data indicated a dramatically decreasing  $k_{cat}$  in comparison with the native GTB. This mutation is interesting, however, because the effect on GTA (Gly268Arg) is an *O* allele (*O<sup>2</sup>*) whilst the same mutation in *B* (Gly268Thr in GTB) caused only moderate weakening of the phenotype.

**Figure 5.** Ribbon diagram of the general B enzyme structure showing the sites of the naturally-occurring mutations in this study. Asterisks indicate those alleles which have been described recently and shown here for the second time. Main chain atoms are shown as a thin white bar with  $\beta$ -pleated sheets as grey ribbons. The connectivity of the main chain across the 21- and 10-residue-disordered internal or C-terminal loop, respectively, are represented as a black dashed line. UDP-Gal modelled in the active site is shown in the enzyme cleft. The figure was modified from an original kindly provided by JA Letts.



The genetic background for the  $B_w$  phenotype appears to be very heterogeneous and usually arises through seemingly random missense mutations throughout the last two exons of the *ABO* gene. However, the targeted a.a. residues are well conserved during evolution and based on modelling of the resulting structural changes in the glycosyltransferase. The mutations are likely to disrupt molecular bonds of importance for enzymatic function.

Six of the novel mutations, *i.e.* 548A>G, 539G>A, 565A>G, 575T>C, 1036A>G and 1055G>A were situated on disordered loops. These six mutations are either within the internal disordered loop (a.a. residues 175-195) or the C-terminus (residues 344-355) which may cross the active site and are thought to be involved in catalysis, although it is unknown which a.a. residues in the loops may contact the donor substrate or be involved in proper loop functioning. However, the effect of a.a. mutations in the disordered loops of other species has been studied and it is known that corresponding disordered loops are important for catalysis in other glycosyltransferases of other species (Boix *et al.*, 2001; Unligil *et al.*, 2000; Morera *et al.*, 1999 and Persson *et al.*, 2001). Furthermore, a mutation in this region can directly impact catalytic activity. Other glycosyltransferases from both related and unrelated species also demonstrate the presence of such a disordered loop of amino acids (Yazer & Palcic, 2005).

An  $A_x$  sample from New Zealand was the only sample with a nonsense mutation, 996G>A, predicting loss of 23 a.a. due to a premature stop codon ( $A^{x-6}$ ). Whether this sample is really a classical  $A_x$  phenotype or differs compared to regular  $A_x$  samples due to 646T>A (Phe216Ile) is unclear, but it was noted by the referring centre as a weak  $A_x$ .

This study also showed ABO discrepancies associated with weakened A antigen expression in leukaemia and pregnancy, or acquired B antigens (Table 4 in **Paper I**). The relationship between leukaemia and weak expression of ABO antigens was first described in 1957 (van Loghem JJ *et al.*, 1957), while the acquired B phenomenon has been described in association with bacterial infections of the gastrointestinal tract (Cameron *et al.*, 19590).

*ABO* genotyping is an important complement to serology for correct determination of donor and patient ABO blood group status. As an independent tool in the clinical laboratory it can be used to confirm a weakly expressed A or B antigen in inherited subgroups (Olsson & Chester, 1996b) but also *e.g.* to exclude B allele markers in the acquired B antigen phenotype (Yip *et al.*, 1996). The findings of these studies may help in improving *ABO* genotyping

methods to detect A and B subgroups and in solving discrepancies and detecting acquired and weak phenotypes.

### **Hybrid *ABO* alleles explain some other weak *ABO* subgroups (Papers I, II and V)**

Samples from Poland and USA in **Paper I** were sent to Lund with phenotypic backgrounds that were suspected by the referring centres to be  $A_{\text{weak}}B$ ,  $B(A)$ , respectively. Both samples were shown to have mutations previously characterised as associated with the  $A_x$  phenotype. The results of sequencing these samples (**Table 10**) in **Paper I** showed two novel variants of the  $A^x$  hybrid allele with the same mutations as in  $A^x$ -2 [ $A-O^{IV}$ ] and  $A^x$ -3 [ $B-O^{IV}$ ], *i.e.* nt. 646T>A (Phe216Ile), 681G>A, 771C>T, 829G>A (Val277Met) in exon 7. In addition,  $A^x$ -3 contained a 297A>G mutation in exon 6 (Olsson & Chester, 1998a). Sequencing of intron 6 showed two novel crossing-over regions in these hybrids, designated as  $A^x$ -4 and  $A^x$ -5.

**Table 10.** Summary of the novel  $A^x$  and  $B^w$  hybrid alleles (**Paper I** and **V**).

Allele description	Nucleotide change(s)	Amino acid change(s)	Geographic origin	Accession no.
$A^x$ -4 §	646T>A	Phe216Ile	Poland	AF324006
	681G>A			
	771C>T			
	829G>A	Val277Met		
$A^x$ -5 §	646T>A	Phe216Ile	USA	AF324007
	681G>A			
	771C>T			
	829G>A	Val277Met		
$B^w$ -19	646T>A	Phe216Ile	China	DQ646582
	657T>C			
	681G>A			

§ These alleles differ by their intron 6 polymorphisms showing different cross-over regions between *A* and  $O^{IV}$ , compared to the previously reported hybrids, namely, nt. 189-225 in intron 6 for  $A^x$ -2, nt. 236-445 in intron 6 for  $A^x$ -4, and between nt. 298 in exon 6 and nt. 41 in intron 6 for  $A^x$ -5.

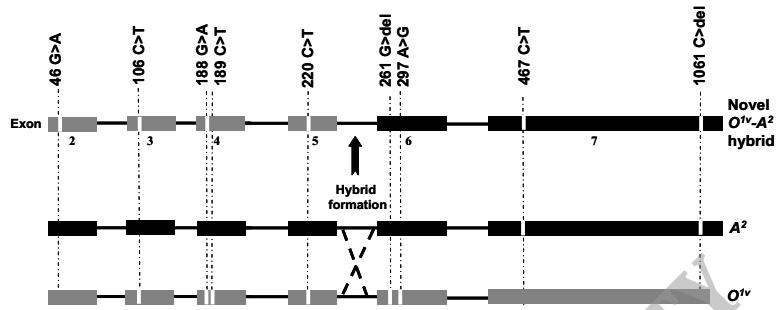
In **Paper II**, blood samples from five individuals with different geographic/ethnic background were analysed. These individuals were all confirmed to be of Black ethnic origin. Each donor had the unusual  $A_wB$  phenotype but appeared to have the  $A^2B$  genotype without any new mutations in exons 6 or 7 associated with weak A or B expression.

A *Bst*UI-based PCR-RFLP test (Olsson & Chester, 1996a) results in cleavage between 188C and 189G in exon 4 of all known alleles except  $O^{lv}$  (and some minor variants of  $O^I$  and  $O^{lv}$ ), which has two mutations at these positions (188G>A and 189C>T). The fact that these five individuals already genotyped as  $A^2B$  and were also heterozygous for an  $O^{lv}$ -allele-specific mutation led us to sequence all exons to find out the reason for this discrepancy.

All samples showed normal sequences in exon 1-5 as expected, apart from for the  $O^{lv}$ -related 106G>T (exon 3), 188G>A (exon 4) and 220C>T (exon 5). The additional presence of a novel missense mutation in nt. 46G>A (Ala16Thr) in exon 2 was noted, whereas 189C>T was absent, as were all  $O^{lv}$ -related mutations in exon 6 and 7 (261delG, 297A>G, 646T>A, 681G>A, 771C>T and 829G>A). Based on these results, we suspected a possible hybrid allele formed somewhere between the end of exon 5 and the beginning of exon 6. The lack of useful polymorphic markers in the *ABO* gene in this region made us sequence introns 2, 3, 4 and 5 in all major alleles  $A^I$  [A101],  $A^2$  [A201],  $B$  [B101],  $O^I$  [ $O^I$ -1, O01],  $O^{lv}$  [O02] and  $O^2$  [O03]) and the results revealed 7, 12, 17 and 8 polymorphic positions, respectively. Intron 6 had been analysed previously (Olsson & Chester, 1998a). Intron 1 was not examined due to its huge size (approximately 13,000 bp). At the same time, Seltsam and colleagues published findings confirming our intron data (Seltsam *et al.*, 2003). Polymorphic sites have been identified in five of the noncoding regions (introns 2-6). Most of these were located in introns 4 and 6. The polymorphic nt. positions are very valuable genetic markers in the *ABO* gene and allowed us to identify a cross-over event in intron 5 where the  $O^{lv}$ -related sequence was fused with an  $A^2$  allele to form a hybrid allele (**Figure 6**). This novel  $O^{lv}$ - $A^2$  hybrid was identified according to *ABO* exon/intron analysis in all the five alleles studied with the A<sub>w</sub>B blood group phenotype.

Screening of more than 100 samples (**Paper II**) for the polymorphic positions 46G>A or 188G>A (without 189C>T) resulted in the finding of two new *O* alleles, one  $O^{lv}$  allele having 188G>A but lacking 189C>T and one otherwise normal  $O^I$  allele having 188G>A only. Later, an  $O^{lv}$ -like allele with the addition of 46G>A in exon 2 was identified by our laboratory in a Dutch donor of African descent during the ISBT 2006 workshop on blood group genotyping (Daniels *et al.*, workshop report in preparation).

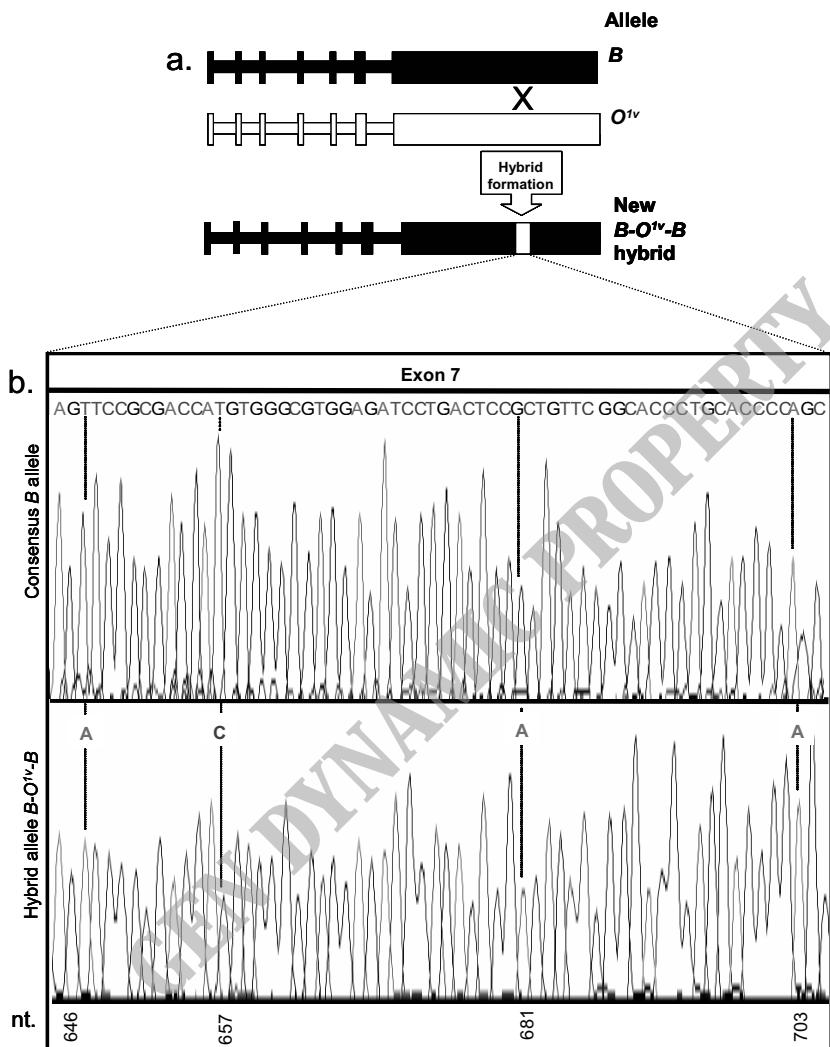
**Figure 6.** Representation of exons 2-7 and intervening introns in  $A^2$  and  $O^{lv}$  alleles and the new  $O^{lv}\text{-}A^2$  hybrid allele in **Paper II**. Exons are represented as boxes and introns as black lines. Black boxes indicate  $A^2$  and grey boxes  $O^{lv}$  sequences. The dashed X indicates the cross-over region. White lines represent exonic SNPs used for allele identification.



In **Paper V**, samples from a Chinese family submitted to us were subjected to an extensive serological and genetic investigation. In this family, the propositus was phenotyped as  $A_1B_3$  and his mother was initially typed as group O but was later found to have very weak expression of B antigen. We found in the initial genotyping that the samples were heterozygous for a B-like allele and for either an  $A^1$  allele [A101] in the son or an O allele [O01] in the mother. Sequencing of these samples (shown in **Figure 7a**) revealed that the B-like allele contained a segment of  $O^{lv}$  [O02]-related polymorphisms in the B-specific DNA fragments, nts. 646, 657 and 681, only the first of which would lead to an a.a. change, Phe216Ile. The new hybrid allele was named  $B^{w\text{-}19}$  and is effectively a new hybrid allele with B and  $O^{lv}$  characteristics (Table 3 in **Paper V**). It is produced by an apparent double cross-over in exon 7 (**Figure 7a**).

The structural consequences of the Phe216Ile change due to the 646T>A mutation in this hybrid allele has been demonstrated by molecular modelling. The result showed an empty space in the core of the glycosyltransferase that resulting from the substitution of an a.a. with a smaller side chain (Figure 4a in **Paper V**). This mutation occurs in a part of the enzyme that is responsible for donor binding, and a mutation that generates space in the hydrophobic core of the protein would likely destabilize the protein fold and result in lower activity. The kinetic data for the Phe216Ile mutant support the molecular modelling and phenotyping results. The crystallization was unsuccessful for this subgroup GTB, which may be due to the destabilizing effect of the mutation.

**Figure 7.** A schematic representation showing how the  $B-O^{1v}$ - $B$  hybrid may have been formed (a). DNA direct sequencing chromatogram showing exon 7 in the *ABO* gene of a normal  $B$  allele compared to the new hybrid allele (b).



### **The genetic basis of the A<sub>bantu</sub> phenotype is also a hybrid allele (Paper IV)**

About 50 years ago, Shapiro observed the existence of a large proportion of blood group A individuals whose cells did not react with anti-A<sub>1</sub> and only weakly with anti-A (Shapiro, 1951). This A subgroup, designated A<sub>bantu</sub>, has its highest frequency in a restricted population but globally seems to be a rare allele. In about 1000 B.C., a bi-directional migration of Bantu people began that created two main language families, Eastern and Western Bantu. The former migrated to Zimbabwe, Mozambique and down into South Africa and the latter migrated into Angola and Namibia. The A<sub>bantu</sub> blood group has been identified in almost 4% of Black Africans of blood group A and in up to 8% of San and Khoikhoi, the ethnic groups from which the A<sub>bantu</sub> phenotype may have originated (Jenkins, 1974). In **Paper IV**, we investigated blood samples from a total of seven individuals serologically characterized as A<sub>bantu</sub>: six from KwaZulu-Natal in South Africa and one from a person of African descent living in England. Genomic *ABO* typing indicated that all samples were heterozygous for what appeared to be the common A<sup>2</sup>-I [A201] allele and for an O allele with the 261delG mutation.

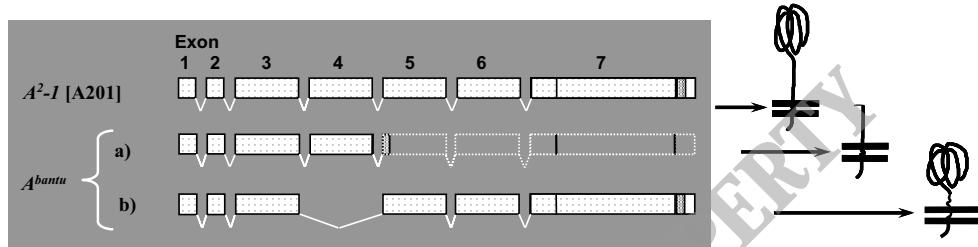
Direct sequencing of exons 1-7 was performed. All samples lacked any of the previously reported mutations associated with weak A and B phenotypes except for the silent 297A>G SNP in exon 6. A deviation found compared to the usual A<sup>2</sup> sequence was a single nt. deletion (AGGT>AGT) at the exon 4 and intron 4 junction, *i.e.* a presumed splice-site mutation. However, since it is not clear what the consequences of this mutation would be, different hypotheses concerning the potential effects on the GTA were formed. These are illustrated in **Figure 8**. As concluded from identical splice-site mutations in other genes, skipping of the preceding exon in the mRNA is predicted. RT-PCR analysis based on synthesis of *ABO*-cDNA from RNA extracts failed in spite of repeated attempts. We could detect only O-mRNA but no A<sup>bantu</sup>-mRNA.

**Figure 8.** Potential consequences of the deletion at the intron 4 / exon 4 junction.

Schematic representation of  $A^2$ -I [A201] cDNA (top row) shows the characteristic sequence associated with the blood group  $A^2$  allele, i.e. the 467C>T and 1061delC mutations (vertical black lines in exon 7). The latter causes an extension of the consensus ( $A^1$ ) sequence by 21 a.a. due to frame shift.

a) Disorder in reading frame (resulting in a premature stop): The deletion results in the exclusion of one of two consecutive guanosines. If the first of the two Gs is considered to be deleted, the mutation is 203delG in exon 4. This would lead to a frame shift and result in a stop codon. This would not allow expression of an enzyme capable of creating even weak A activity but is in effect an  $O$  allele.

b) Exclusion of exon 4 from the mRNA: The deletion would act as a splice-site mutation that affect intron 4 rather than exon 4, i.e. deletion of the second guanosine. The consequence of the resulting splicing would be skipping of exon 4. Removal of the 48 nucleotides representing exon 4 does not change the reading frame but simply shorten the stem region of the GTA by 16 a.a.



An unexpected finding in this study was a novel  $O$  allele which contains the same nt. deletion (AGGT>AGT) at the exon 4/intron 4 junction and an extra mutation in 297A>G as in the  $A^bantu$  allele. The sequence of the  $A^bantu$  alleles was identical to this novel  $O$  allele sequence in introns 2-4. This sequence differs from the common  $A^2$ -I allele [A201] by 3, 8 and 11 polymorphic sites in introns 2, 3 and 4, respectively. On the other hand, the sequences of introns 5 and 6 of the  $A^bantu$  allele were identical to the common  $A^2$  allele but differed from this novel  $O$  allele (named  $O^{1bantu}$ , O54) sequence at 1 and 12 position(s), respectively. This clearly indicated that the  $A^bantu$  allele is a hybrid allele with a crossing-over point after nt. 1623 in intron 4 but before nt. 498 in intron 5. Interestingly, the Bantu mutation was found in 9% of  $O^1$  alleles in Black South Africans but not in Swedish or Jordanian blood donors, or non- $O^1$  alleles.

In conclusion, a novel allele was found in all  $A_{bantu}$  samples tested, possibly causing decreased mRNA stability and/or changing the coding region. The same mutation in African  $O^1$  alleles indicates that  $A_{bantu}$  has arisen by gene conversion or crossing over between  $O^{1bantu}$  and  $A^2$  alleles. A phylogenetic analysis based on approximately 6500 bp of sequence (exons 2-7, including intervening introns) from six common alleles ( $A^1$ ,  $A^2$ ,  $B$ ,  $O^1$ ,  $O^{1v}$  and  $O^2$ ) indicated that  $O^{1bantu}$  is a distinct evolutionary lineage separated from all other lineages but most closely related to  $O^{1v}$ .

### **Novel *O* alleles involved in unexpected blood group phenotypes (Paper III)**

The sequence of the most common allele [ $O^1$ -1, O01] is identical to the  $A^1$  sequence, apart from one nonsense mutation in exon 6, *i.e.* 261delG which leads to a truncated protein lacking enzymatic activity (Yamamoto & Hakomori, 1990; Yamamoto *et al.*, 1990a). Various *O* alleles including hybrid  $O^{1v}$ -*B* and  $O^1$ - $A^2$  alleles coding for non-functional glycosyltransferases have been described by several groups but the vast majority show the 261delG mutation (for a review, see Chester & Olsson, 2001) only very few non-deletional *O* alleles had been described by the start of this study:  $O^3$  (Olsson & Chester, 1996c) and four Japanese alleles (Ogasawara *et al.*, 2001), two of which are identical to the  $A^x$  hybrids discussed above. All these alleles are based on the consensus A sequence. Fundamentally, however any change which leads to the inactivation of the ABO glycosyltransferase can be considered as a lethal null mutation. In this study, we included the investigation of samples referred to our laboratory in Lund because of suspected A subgroups following discrepant serologic blood grouping results but where ABO genotyping or sequencing surprisingly found only *O* alleles. A correlation between these weak A phenotypes and genotypes by ABO genotyping analysis (PCR-RFLP and PCR-ASP) showed an overrepresentation of the otherwise infrequent  $O^2$  allele (Yamamoto *et al.*, 1993d; Grunnet *et al.*, 1994). Thus, we examined 57 existing  $O^2$  alleles in detail both from regular group O donors and from the above-mentioned samples with suspected A subgroup phenotypes. A flow chart summarizes the samples that were included in this study (Figure 1, Paper III). All 57  $O^2$  alleles differed from the consensus allele [A101] in exons 2, 5, 6 and 7 at nt. 53, 220, 297, 526, 802 and nt. 1096 located in the 3'-untranslated region as expected (Yamamoto *et al.*, 1993d; Olsson & Chester, 1995; Amado *et al.*, 2000). We confirmed in this study that all  $O^2$  alleles contain the mutations 53G>T and 220C>T. This had only been described in few samples in a single paper before (Amado *et al.*, 2000).

Two new  $O^2$  alleles were identified, having the 689G>A mutations resulting in Gly229Asp with or without 649C>T (Arg217Cys). These  $O^2$  variants were called  $O^2$ -2 [O48] and  $O^2$ -3 [O49], respectively. There was no difference in the coding sequences of  $O^2$  alleles found in normal group O donors or in samples referred to us for ABO discrepancies and suspected A subgroups. Furthermore, the same  $O^2$  allele could produce different phenotypes (O or weak A) in different members of the same family.

Yet three novel *O* alleles surfaced in six samples with suspected A subgroups. Further analysis including sequencing of coding regions showed that these were *A<sup>I</sup>*-like alleles having nonsense mutations, 88dupG [*O<sup>4</sup>*], 322C>T [*O<sup>5</sup>*] or 542G>A [*O<sup>6</sup>*] causing premature truncation at codons 56, 107 or 181, respectively. All of these alleles would be expected to be totally incapable of producing an active enzyme. Consequently, when present with another inactive allele with 261delG they should produce a normal O phenotype. This is clearly not the case (Table 2 in **Paper III**). Kominato *et al.* suggested that perhaps A antigen expression could be the result of the use of an alternative translational start codon (Kominato *et al.*, 2002) or that alternative splicing might explain the effect of early mutations.

In another unusual case in **Paper III**, we investigated an Australian blood donor who by DNA sequencing, showed homozygosity for a silent mutation (768C>A) and denoted as O26. The same mutation was described in the Basque population and proposed to be a specific Basque marker (Roubinet *et al.*, 2001). Surprisingly our screening study showed that this mutation was quite frequent among Swedish donors and has a frequency of about 9% of all *O* alleles. We also examined blood samples from two other populations (Jordanian and Swiss) and the results indicate a widespread distribution of this mutation, but at lower frequencies than in Sweden. Thus, this appears to be the third most common *O* allele among Swedes, following *O<sup>I</sup>* and *O<sup>Iv</sup>* whilst clearly more frequent than *O<sup>2</sup>*.

A recently-described *O<sup>2</sup>* variant, 488C>T (Thr163Met), denoted Aw08 (Seltsam *et al.*, 2003) was also found in our study and classified as *O<sup>2-4</sup>*. Surprisingly, the original and variant *O<sup>2</sup>* alleles could all be associated with either O or A<sub>weak</sub> phenotypes. Our data (**Paper III**) disagree with the idea that 488C>T would result in an a.a. substitution with the ability to restore some enzymatic function as suggested by Seltsam and co-workers (Seltsam *et al.*, 2003). Recent work from Yazer *et al.* indicated that they could not find any reason (based on sensitive adsorption-elution studies and flow cytometry) not to use blood from group O donors with an *O<sup>2</sup>* allele to recipients with anti-A (Yazer *et al.*, 2006). Structural studies have also supported that the crucial mutation 802G>A (Gly268Arg), the most significant mutation in the *O<sup>2</sup>* allele, completely blocks the donor sugar recognition site (Lee *et al.*, 2005), although extremely low GTA-like enzyme activities could be measured.

The presence of the *A<sup>el</sup>*-characteristic insertion at nt. 804 and the *A<sup>2</sup>*-characteristic polymorphisms at nt. 467 and 1061 indicated the presence of an *O<sup>3</sup>* allele in a Swedish blood

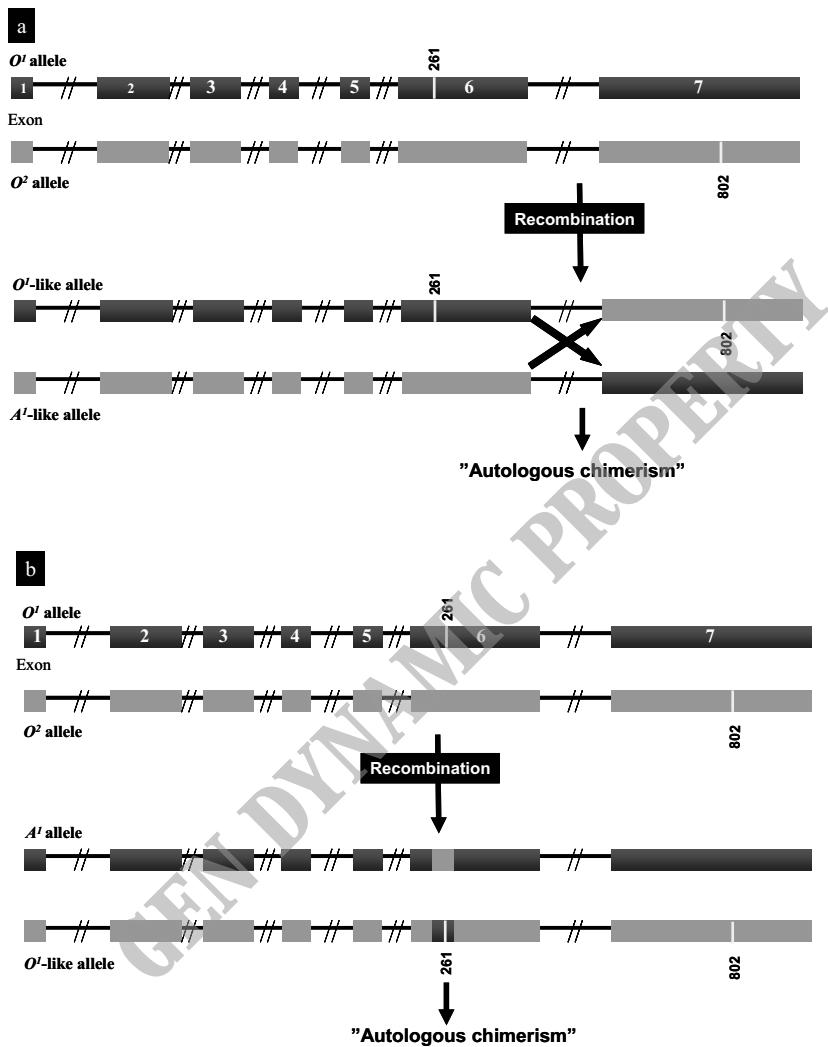
donor (Olsson & Chester, 1996c). In this study we observed yet another  $O^3$  allele but resulting in a different phenotype (suspected A<sub>weak</sub>) than the original one (clear-cut group O). No other polymorphisms were detected after sequence analysis of exon 1-7.

How can non-deletional  $O$  alleles result in small amounts of A antigen?

Thus, we found that apparent  $O$  alleles lacking 261delG (non-deletional  $O$  alleles) may cause weak A expression on erythrocytes and/or inhibit anti-A production. The *Chi* sequence (5'-GCTGGCGG-3') motif has been described in bacteria and been shown to influence the recombination frequency in *Escherichia coli* and has also been implicated as a hot spot for recombination in human genes (Smith, 1994; Stahl & Meyers, 1995). Several investigators have published numerous examples of inheritable, recombinant *ABO* alleles generated during meiosis (Suzuki *et al.*, 1997; Olsson & Chester, 1998b; Ogasawara *et al.*, 1996a, 2001).

According to the data in this study,  $O$  alleles without 261delG may cause weak A expression on RBCs in the presence of an another  $O$  allele with 261delG, which in turn may result in weak anti-A in plasma. The challenge is that it does not always happen when different individuals with the same genotype are compared, even within the same family. To explain this interesting phenomenon, we hypothesized that exchange of genetic material between alleles during mitosis may result in so-called "autologous chimerism" that restores glycosyltransferase activity in some cells. The duplicated *Chi* sequences in the 3' end of intron 6 in the *ABO* gene may also increase the probability of recombination in somatic cells (during mitosis). Gene conversion/recombination events in somatic cell clones in or outside the bone marrow can then result in "autologous chimerism" which produces low levels of A antigen and inhibits the natural production of anti-A (**Figure 9**). However, we have been unable to provide experimental evidence to support this hypothesis and alternative explanations may well exist.

**Figure 9.** Recombination or gene conversion events that may explain how *O* alleles can result in weak A phenotypes. Probable creation of enzymically active gene products from inactive alleles by recombination (a), gene conversion (b).



## **Development of an improved *ABO* genotyping assay based on PCR-ASP amplification across intron 6 (Paper VI)**

The majority of individuals possess the standard blood group A, B, and O phenotypes. In addition, a number of A and B subgroups, B(A) and *cis*-AB exist. Generally, these subgroups are characterised phenotypically by the occurrence and strength of the corresponding A and B antigens on the RBCs and in secretions such as saliva.

The invention of the PCR technique and the elucidation of the sequence (Yamamoto & Kominato, 1990; Yamamoto *et al.*, 1990a-b) and organisation (Yamamoto *et al.*, 1995; Bennett *et al.*, 1995) and the increasing number of alleles (now almost 170 alleles reported to dbRBC) of the *ABO* gene has opened another dimension that complements serological analysis.

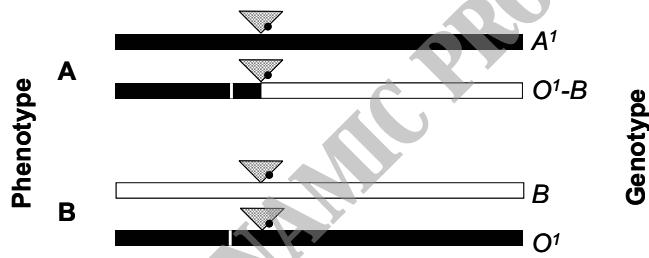
Many alleles responsible for the rarer subgroups have been examined and an increasing number of polymorphisms have been found (Chester & Olsson, 2001; Yamamoto, 2004; Yip, 2002). Knowledge and improvement of *ABO* genotyping methods are very important for practical implications in transfusion medicine. A number of different screening methods have been developed and used with varying degrees of success (reviewed in Olsson & Chester, 2001). Improvement of these *ABO* genotyping methods for clinical use is much needed for several reasons.

Recombination and hybrid formation between common alleles in the *ABO* gene are recent findings that threaten to complicate *ABO* genotype interpretation. These hybrid alleles may lead to unexpected *ABO* phenotypes, including rare subgroups. Hybrids are formed by interaction between two alleles, *e.g.*  $O^I$  or  $O^{IV}$  alleles with  $A^2$  or B alleles. These phenomena are interesting and complex because they can give rise to O alleles if the nt. 261-containing exon 6 is donated by the  $O^I$  or  $O^{IV}$  allele. However, if nt. 261 is consensus (*A*, *B* or  $O^2$ ) and combined with an O allele, then there are various consequences depending on the cross-over point. Hybrid alleles obviously exhibit a whole range of serological activities. Different investigators have published genotyping methods without any attention to addressing the problem that can arise due to hybrid alleles (for a detailed description see review in Olsson & Chester, 2001). All *ABO* genotyping methods to date can fail in the prediction of *ABO* phenotype in the presence of hybrid alleles or the newly found O alleles and are potentially dangerous and difficult to use in the clinical routine laboratory because of interpretation

problems (see example in **Figure 10**). In addition, all *ABO* genotyping methods published to date will fail in predicting the correct phenotype when A or B subgroup alleles including *cisAB* and *B(A)* are present.

We tested two versions of a commercially available *ABO* genotyping kit and confirmed its limitations in detecting hybrid genes and many subgroup alleles. This kit (2<sup>nd</sup> version) which is actually CE-marked predicts the common alleles and a few subgroups, but fails to define hybrid alleles and most non-deletional *O* alleles except *O<sup>2</sup>* and *O<sup>3</sup>*, the latter being one of the improvements from the previous version.

**Figure 10.** Two samples with A or B phenotypes would illustrate an identical genotyping pattern in most, if not all current *ABO* genotyping methods used routinely today. Triangles illustrate the intron 6 with black circles where the *Chi* sequence may act as a hotspot. The white vertical line shows the 261delG mutation. Filled rectangles show consensus sequence shared between *A<sup>1</sup>* and *O<sup>1</sup>*. Open rectangles represent *B* sequences. Thus, heterozygosity for both 261delG in exon 6 and *B*-SNPs in exon 7 are present in both samples, a fact that renders genotyping results potentially identical.



Over the past ten years, we have developed and expanded a PCR-ASP method that is flexible and capable of detecting an increasing number of known alleles. This method gives flexibility to different users and can be adjusted according to their requirements for detection of the vast array of *ABO* alleles (see **Paper VI**). For example, the suggested approach is flexible so that each user can add or remove primers according to the population tested or as new alleles continue to be described. In analogy with many HLA typing systems, it is also possible to choose the relevant level of resolution (standard or high) depending on the purpose of the analysis.

The goal of this part of the study was to develop and establish an uncomplicated method which can detect different alleles including hybrid, non-deletional null and subgroup alleles. The proposed method is able to detect all major and a large number of rare alleles due to its ability to *cis*-link the mutations in exons 6 and 7 and thus predict the *ABO* genotype with high

accuracy, even in the presence of hybrids or subgroup alleles with SNPs in common with the major alleles (e.g.  $A^x$ -1 to  $A^x$ -5 and  $O^{Iv}$  sharing 646T>A). Most combinations of the following alleles can be discriminated:  $A^1$ ,  $A^2$ ,  $A^x$ ,  $A^{el}$ , other  $A^{weak}$ ,  $B$ , a number of  $B^{weak}$  alleles and various null alleles including  $O^1$ ,  $O^{Iv}$ ,  $O^{I-B}$ ,  $O^{I-A2}$ ,  $O^2$ ,  $O^3$ ,  $O^4$ ,  $O^5$ ,  $O^6$  and also several  $cisAB$  and  $B(A)$  variants.

In a first step, all the major alleles are defined and at the same time multiplex PCR-ASP screening for most of the known but unusual or rare hybrid, non-deletional null and A or B subgroup alleles is undertaken in four separate reactions. In most cases, this screening will turn out negative and the genotype can be defined immediately. If one of the four screening mixes are positive, the user has the choice to report the genotype at this level of resolution (e.g.  $A^wO^1$ ,  $A^1B^w$ ,  $B O^{hybrid}$ ,  $A^2O^{non-del}$ ). If, for some reason, a better definition of the genotype is desirable, a second round of high-resolution mixes is performed in which all the multiplexed primers pairs in each screening mix have been separated into a number of reactions. The result will further detail the genotype (e.g. the above-mentioned genotype examples may now instead be expressed as  $A^{bantu}O^1$ ,  $A^1B^w$ -18,  $B O^{Iv-B}$ ,  $A^2O^5$ ). We believe that this technique is the most cost-efficient, flexible and informative *ABO* genotyping method reported to date. Compared to the commercial method which is based on a published PCR-ASP method originally utilizing eight PCR tubes (Gassner *et al.*, 1996) but now upgraded to 16 tubes, the number of genotypes resolved correctly by the method presented here is superior. In principle, no *ABO* genotyping method should be used unless hybrid alleles and known non-deletional *O* alleles can be detected properly, especially if the method is used in a clinical reference laboratory and the results are used to predict the blood group of donors or patients.

## **CONCLUSIONS**

The results of the studies presented as part of this Ph.D. thesis have led to a number of conclusions, some of which can be summarized as follows.

- The molecular genetic basis for ABO subgroups is very heterogeneous but single missense mutations targeting evolutionarily conserved a.a. residues is a common theme.
- Hybrids based on common alleles constitute another principal background for both A and B subgroup phenotypes.
- Some A<sub>weak</sub> phenotypes which appear to be characteristic among people of African descent are due to recombinant hybrid formation involving previously unknown O alleles, so far only found in Black individuals.
- Surprisingly, non-deletional O alleles can result in either common O or suspected A<sub>weak</sub> phenotypes but the reason for this has not yet been elucidated.
- Clarification of the relationship between mutation, structural change in the glycosyltransferase and enzymatic function can be achieved and is helpful when trying to understand ABO subgroups as well as the importance of different domains in the enzyme.
- Hybrids, non-deletional null and A or B subgroup alleles constitute major challenges for ABO genotyping.
- Improvement of the current standard of ABO genotyping was made possible by development of a new approach for allele-specific amplification across intron 6 and incorporation of multiplexed primers sets screening for rare or previously difficult-to-detect alleles.

## **FUTURE PERSPECTIVES**

The number of SNPs known at the *ABO* locus has dramatically increased during the last decade. This has raised a number of new questions about how these allelic changes affect the quantity, structure, function and localisation of GTA and GTB, and subsequently the A and B antigens on RBCs. Future studies will also have to address the mechanisms by which the *ABO* transcription is regulated, especially in the most relevant compartment, namely the bone marrow. Until now, most work has been carried out in cell lines and most of them of gastric or colonic origin.

New imaging technologies like cellular computer tomography may help us close in on what is actually going on, in a cell biological and biochemical sense, in the Golgi apparatus where the

glycosyltransferases reside. In addition to computer-based 3D predictions and x-ray crystallography as used in this study, further improvements in software and resolution may yield further insight into the molecular enzymology of ABO. Another attractive method could be using electron microscopy, which may facilitate the understanding of what is happening with ABO antigens in different genotypes or in single erythroid progenitors and their successors including the mature RBCs.

Current and future studies will focus on applying these techniques for the characterisation of  $A_{weak}$  and  $B_{weak}$  samples in which the mutations are located in other, principally interesting areas of the enzyme, *e.g.* close to the DXD motif (Persson *et al.*, 2007) or in the disordered loops. Also, there are still samples sent to reference laboratories for which the atypical ABO phenotype has not yet been explained at the genetic level. One of the most challenging examples is the most common weak A subgroup,  $A_3$ . Despite the finding of a complete  $A^I$  consensus sequence, this inheritable variant is completely different from the standard  $A_1$  phenotype. One possibility in this case would be to look at mRNA stability and quantities in combination with analysis of regulatory regions outside the coding region.

Another potential aim in future studies could be to address the hypothesis in **Paper III** regarding gene conversion or recombination events in somatic cell clones in or outside the bone marrow. Our suggestion about a phenomenon of so-called "autologous chimerism" as well as other theories about how apparent  $O$  alleles predicted to result in severely truncated proteins without enzymatic activity, can sometimes give rise to weak A expression require further work. The mechanisms responsible for these apparent anomalies in correlation between serological and genetic blood group typing, is very interesting to us. As yet, we have not provided enough experimental evidence to confirm our theory, so it would be interesting to explore this hypothesis. It may be that other principally different methods are required for this goal. Preliminary work using flow cytometry or erythropoietic culture models have not yet been successful. Cloning of *ABO* transcripts from cell populations or even single cells may be used to identify low levels recombinant transcripts that can translate into functional GTA.

In general, the ABO blood group system continues to serve as one of the most useful models to study both human genetics, carbohydrate synthesis and glycosyltransferase functional anatomy.

## SUMMARY IN SWEDISH (sammanfattning på svenska)

De första blodtransfusionerna kan härledas till 1600-talet då man försökte transfundera blod från djur till djur, djur till människa och människa till människa, vilket ofta resulterade i komplikationer och i många fall patientens död.

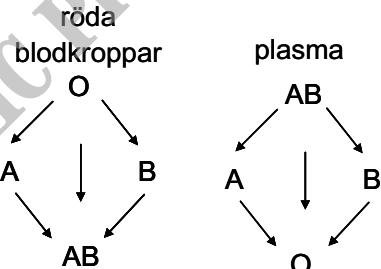
I början av 1900-talet upptäckte Karl Landsteiner (1868-1943) blodgruppssystemet ABO (läses A, B, noll) efter empiriska försök där han blandade blod från sig själv och sina kollegor (utanför kroppen).

Transfusionsmedicin är därmed en ung vetenskap inom den medicinska läran. Däremot har ordet blod länge använts i olika sammanhang, inte bara medicinska, utan även i talspråk såsom t.ex. blodtörstig, blodsbröder, blodsband och blodshämnd etc.

Landsteiner beskrev principen för den antigen-antikroppsreaktionen vilken numera benämns hemagglutination och som används rutinmässigt idag vid de transfusionsmedicinska laboratorierna, blodcentralerna.

Antigenen utgörs av kolhydratstrukturer på cellerna medan antikroppar är immunförsvarets verktyg som kan attackera främmande kroppar som t.ex. bakterier eller blodgivarens blodceller.

Bilden här intill visar de möjliga ABO-blodgruppskombinationerna vid transfusion av röda blodkroppar (erytrocyter) och plasma.



Fenotypning innebär att man fastställer vilka antigen som förekommer hos individen i fråga, och tolkas oftast som "finns" (positivt) eller "finns inte" (negativt). Förekomsten av de vanliga ABO-fenotyperna ( $A_1$ ,  $A_2$ , B,  $A_1B$ ,  $A_2B$  och O) varierar liksom andra blodgruppsfenotyper i olika populationer. Denna variation benämns polymorfism. ABO-systemets antikroppar är oftast av IgM-typ. Dessa bildas normalt (kallas även naturliga/spontana antikroppar eller isoagglutininer) mot de ABO-antigen som saknas hos individen i fråga. ABO-systemet är på grund av dessa naturligt förekommande antikroppar det kliniskt viktigaste blodgruppssystemen. För att utföra en säker blodtransfusion krävs det att ett laboratorium kontrollerar vilken ABO-blodgrupp individen tillhör och om denna har bildat antikroppar mot något annat blodgruppssystem. Därefter kan passande givarblod transfunderas. Även vid transplantation av organ som njure, lever, hjärta och lungor är ABO-systemet bland det viktigaste att ta hänsyn till.

Det finns enligt International Society of Blood Transfusion (ISBT) ca. 270 erkända blodgrupsantigener på erytrocyternas yta. Blodgrupsantigenen delas in i olika blodgruppssystem, hittills finns 29 beskrivna. Gemensamt för alla blodgruppsmarkörer är att de sitter på utsidan av de röda blodkropparnas membran. ABO-blodgrupsantigen finns förutom på dessa celler även på många andra celltyper. Exempel på andra viktiga system är Rh, Duffy, Kell, Kidd och P.

I vissa situationer, bland annat vid bestämning av ärftligt svagare ABO-uttryck, s.k. subgrupper som  $A_x$  och  $A_{el}$ , förvärvade blodgruppsegenskaper (t.ex. förvärvat B-antigen, där en individ med blodgrupp A kan framstå som AB) eller vid olika sjukdomar, kan dock resultaten vara svårtolkade och framstår som svagt positiva ("finns men svagare än normalt"). Detta kan bero på någon ärftlig variant av *ABO*-genen eller förändrat uttryck i olika typer av tumörceller och vid vissa svåra infektioner som sannolikt beror på nedsatt aktivitet i enzymet som tillverkas med *ABO*-genen som mall (se nedan).

Intensiv forskning på 1950-60-talet klargjorde de biokemiska grunderna för ABO-systemet och resulterade i en bättre förståelse av detsamma. *ABO*-genen kartlades till den långa armen av kromosom 9. Yamamoto och medarbetare visade 1990 att *ABO*-genen kodar för ett glykosyltransferas. Detta innebär att *ABO*-genen ger upphov till ett protein av enzymtyp som i sin tur bidrar till att bygga enkla socker i en viss ordning så att en kolhydratkedja bildas. ABO-systemets antigener består av olika specifika kombinationer av dessa kolhydrater. Olika mutationer, d.v.s. ärftliga och bestående förändringar i cellens genetiska material (DNA), har upptäckts i *ABO*-genen vilket i sin tur kan leda till minskning av enzymaktiviteten och följdaktligen till en minskning av antalet antigen på erytrocyter. Detta kan göra blodgruppering vansklig att tolka och risken finns att fel blodrekommendation utfärdas.

Enstaka punktmutationer i *ABO*-genen ansvarar för många kända fenotyper. *A*- och *B*-generna ger upphov till enzymen A- respektive B-transferas. Blodgrupp O beror däremot på dubbel uppsättning av *O*-gener innehållande någon typ av defekt (t.ex. punktdeletion), som orsakar ett inaktivt (ofunktionellt) enzym eller t.o.m. avsaknad därav. Det finns även ett stort antal undergrupper inom A/B-blodgrupperna, t.ex.  $A_3$ ,  $A_x$ ,  $A_{el}$ ,  $A_{weak}$ ,  $B_3$ ,  $B_{el}$  och  $B_{weak}$  som uppvisar svagare reaktioner vid försök till blodgruppering. Studier avseende A- och B-specificiteten har förfinats på senare tid genom de nya DNA-baserade teknikerna, bl.a. PCR-teknik (en "kopieringsapparat för DNA"), kloning etc. Mina studier i denna avhandling har kunnat utföras

tack vare alla dessa framsteg inom molekylärbiologin tillsammans med upptäckten av *ABO*-genen.

I stort handlar avhandlingen om en fördjupad studie av bakomliggande orsaker till och konsekvenser av den överraskande mångfald som ses i *ABO*-genen. Länge trodde man att det fanns fyra genetiska varianter, s.k. *ABO*-alleler ( $A^1$ ,  $A^2$ ,  $B$  och  $O$ ), men redan 1993 stod det klart att situationen var betydligt mer komplex än så. Syftet i de fem första vetenskapliga artiklarna har varit att klarlägga den molekylärgenetiska bakgrundens och blodgruppsolikheter inom den svenska populationen och i andra befolkningsgrupper med fokus på svaga *ABO*-blodgrupper. I den sista artikeln försökte vi att utveckla en säker och kostnadseffektiv metod för att detektera alla dessa vanliga och sällsynta alleler.

En omfattande insamling av kliniska fall, remitterade till det Nordiska Referenslaboratoriet för Genomisk Blodgruppstypning vid Blodcentralen Skåne på Universitetssjukhuset i Lund, undersöktes i de olika delprojekten. I delarbetet I och V identifierades olika nya punktmutationer, både s.k. missense- (som ger upphov till ett aminosyrabyte) och nonsense-mutationer (som förstör genens möjlighet att ge upphov till ett funktionellt enzym). Därmed karakteriseras den genetiska bakgrundens hos 17 nya undergrupper inom *ABO*-systemet.

Efter detta mödosamma arbete, började jag/vi fundera på om hur man skulle kunna bevisa att dessa mutationer verkligen ligger bakom det avvikande uttrycket av antigenerna på erytrocyter. Det fanns ett antal olika analyser som kunde användas i detta syfte men jag valde att genomföra mina undersökningar i Köpenhamn där professor Monica Palcics grupp forskar inom området som handlar om enzymers aktivitet och specificitet samt hur enzymerna kan visualiseras med s.k. datormodellering eller röntgenkristallografi. Jag fick lära mig att uttrycka olika typer av *ABO*-undergrupper genom att införa konstgjorda mutationer (motsvarande våra nyupptäckta *ABO*-subgrupper) i en syntetisk *ABO*-gen i *E.coli*-bakterier. Stora mängder enzym/protein uttrycktes och egenskaperna hos dessa glykosyltransferas-mutanterna kunde fastställas. Slutligen skickades proteinet iväg för röntgenkristallografi (en sorts visualisering av den tredimensionella strukturen) detta skedde hos en kanadensisk grupp, ledd av professor Stephen Evans, som samarbetar med Köpenhamnsgruppen. Kollegorna i Kanada lyckades att kristallisera ett av våra enzymer och analyserade dessutom flera andra nya mutationer via datorsimulering. Analyserna visade att de nya mutationerna med stor sannolikheten är orsaken till de svaga reaktioner som man funnit på blodcentralernas laboratorier.

I delarbete II beskrivs en hybridallel (en blandning mellan två alleler) hos personer av afrikanskt ursprung med en variant av blodgruppen AB varvid svagt uttryck av A-antigenet på blodkropparna gjorde blodgruppering svårtolkad. Resultatet av en lång rad av undersökningar på *ABO*-genen visade att en fusion av en vanlig *O*-allel ( $O^1v$ ) och en vanlig  $A^2$ -allel hade bildats. Denna nya alell verkade orsaka ett mindre effektivt enzym. För att med säkerhet kunna definiera allelen som en hybrid, undersöktes och kartlades intronerna (de icke-kodande regioner som finns mellan genens exoner, d.v.s. de kodande regionerna) och deras genetiska variation (polymorfism) hos personer med olika blodgrupper. Med denna kartläggning kunde vi bevisa var sammanblandningen (hybridformation) av dessa alleler hade skett.

I delarbete III undersöktes ett stort antal prover, och 5 nya *O*-alleler i *ABO*-genen detekterades, något som kan få stor betydelse för säkerheten vid genetisk blodgruppering. Vi upptäckte bl.a. vad alla hittills trott var en sällsynt *O*-variant i den baskiska populationen, i själva verket var den 3:e vanligaste *O*-allelen överhuvudtaget, åtminstone bland kaukasier (t.ex. svenskar). I detta arbete lades även en hypotes fram för att förklara varför vissa *O*-alleler (d.v.s. förstörda gener utan förmåga att ge upphov till funktionella enzymer) faktiskt kunde ge upphov till svagt A-uttryck på röda blodkroppar trots allt. Dessvärre har denna hypotes ännu inte gått att bevisa men utgör en möjlighet för ev. framtida projekt.

I delarbete IV, rapporterade vi att den sedan flera decennier välkända *ABO*-undergruppen Abantu överraskande nog, även den, beror på en hybridbildning: denna gång mellan den vanliga  $A^2$ -allelen och en aldrig tidigare beskriven *O*-allel som utgör en helt egen evolutionär (historisk) utvecklingslinje inom *ABO*-allelfamiljen. Allelen förekommer huvudsakligen bland afrikanska folkgrupper inkl. de s.k. Bantu-folken, precis som varit känt länge i fallet  $A_{abantu}$  och därfor kallades alellen  $O^{1bantu}$ . Även här spelade intron-polymorfismerna en avgörande roll för att kunna bestämma var hybridformationen skett. Medan andra svaga *ABO*-undergrupper är sällsynta, är  $A_{abantu}$  den A-undergrupp som har högst frekvens i någon befolkning, upp till 8% av vissa folkgrupper i södra Afrika.

Genetiska metoder finner allt fler applikationer inom det transfusionsmedicinska fältet. Det sista delarbetet (VI), är en sammanfattningsartikel om nästan tio års arbete (från mitt 20-poängsarbete 1997 för M.Sc. fram till idag). Under många år har Martin L Olsson (min handledare), Åsa Hellberg (min käresta) och jag ständigt förbättrat och justerat en metod för att kunna detektera de allt fler vanliga och sällsynta *ABO*-alleler som bör kunna upptäckas (idag

minst 168 st att välja bland). Denna analys har använts parallellt med andra metoder som publicerades från Lund 1995-6, vilket gett en god validering under kliniska betingelser. En jämförelse med en kommersiellt tillgänglig analys visade också att vår nya metod är överlägsen, framför allt vid s.k. hybrid-, ovanliga O och subgruppssalleler.

Att detektera dessa komplexa alleler är viktigt när blodgivare och/eller -mottagares prover ger svårtolkade resultat vid blodgruppering. Då kan DNA-baserad analys användas som ett komplement för att minska risken för falska provsvar, som skulle kunna leda till felaktiga konklusioner och transfusionsrekommendationer.

Sammanfattningsvis har studierna i denna avhandling lett till en bättre förståelse kring de funktionella och genetiska sambanden inom ABO-systemet. Det långsiktiga målet är att med någon typ av automatiserade plattformar kunna fastställa blodgruppen helt genetiskt, så som redan görs för vävnadstypen (HLA) inför transplantationer.

## **ACKNOWLEDGEMENTS**

Although I spent many years of hard work to complete this thesis, it all depended on the fortune of having many generous friends and colleagues around me who supported me during this study. I am especially grateful to the following persons, knowing that there are many others who are not mentioned here.

First and foremost, my ever-enthusiastic supervisor Professor Martin L Olsson, for giving me the opportunity to work with such exciting projects, for guidance through scientific thinking and sharing his knowledge.

Britt Thuresson, PCR expert, for guidance and teaching me scientific methodology, especially in the field of molecular biology, and not least for taking me down to earth at occasions when too many “good ideas” were occupying my mind.

My colleagues at the Department of whom I particularly will mention Alan Chester, Jill Storry and Åsa Hellberg who always helped me with language revision and adopted a critical attitude towards the thesis.

My friend and colleague Elisabet SW, who shared with us her nice house many times during our kick-off meetings (Landet). The most enjoyable times have been those everyday-moments in the corner (dum-forum) of the lab where we could allow us to be dumb and discuss freely about different things.

Nidal and all other people in the research group who inspired me and generously provided both time for discussions (genlunch) and good advice.

I would especially like to thank my roommate and good colleague, Gunilla Hochberg for all support and patience throughout these years, which made it possible to write this thesis. Li and Alicja are thanked, for all help during this long time in the same office.

I would like to express my appreciation towards all the colleagues at the Blood Centre who helped me in all weather when I needed to be cheered up, especially Kerstin x 3 (B, H & T) and of course Elisabeth, Ingela Å, and Eva Marie.

Professor Tor Olofsson, Louise, May-Louise and Bodil in the laboratory at the Department of Hematology and Transfusion Medicine, BMC, for support and for introducing me to RNA preparation.

Professor Monica Palcic, Mattias Persson and James Letts for valuable discussions and help with enzyme purification and computer modelling at the Carlsberg laboratory in Copenhagen.

Charlie, Hannah and Gerd for helping me with sequencing at the Department of Clinical Chemistry.

All my co-authors in different parts of the world for fruitful collaboration.

The friends for reminding me about life outside the lab, Ramtin, Per, Mohammad, Mikael, Stefan (lektorn), Anna, Percy, Sören, Marika, Anica (?-alliansen), Leif, Åsa, Erik, Jytte, Ewa, Magda, Ann-sofie, Orlando, Pasha, Ola, Birger, Örjan, Hamid, Nori, Mahmood and Mattias.

My dear Family, my beloved Parents, Bahareh, Azita and specially Göran and Ingrid, for always believing in me and for being there when Åsa & I needed them.

And most of all, my friend, roommate and my love for her support and good cooperation both at work and at home. Finally the youngest one, my sunshine Arvin (Seed-Nisse) who gave me new strength every day.

This project was supported by the Medical Faculty at Lund University, the Swedish Research Council, governmental ALF research grants to Lund University Hospital, and Region Skåne, Sweden, the Claes Högman SAGMAN-stipendium, the Inga and John Hain Foundation for Medical Research, Georg Danielssons Fund for haematological disease, Tore Nilsons Foundation for Medical Research.

## REFERENCES

- Amado M, Bennett EP, Carneiro F, Clausen H. Characterization of the histo-blood group O(2) gene and its protein product. *Vox Sang* 79:219-226, 2000.
- Allardice PW, Kaita H, Lewis M, Mcalpine PJ, Wong P, Anderson J and Gblett ER. Segregation of marker loci in families with an inherited paracentric insertion of chromosome 9. *Am J Hum Genet* 39:612-617, 1986.
- Barjas-Castro ML, Saad ST. Absence of the G871A mutation in A<sub>3</sub> blood donors from Brazil. *Transfusion* 37:564, 1997.
- Barsoum N, Kleeman C. Now and then, the history of parenteral fluid administration. *Am J of Nephrol* 22: 284-289, 2002.
- Bennett EP, Steffensen R, Clausen H, Weghuis DO, Geurts van Kessel A. Genomic cloning of the human histo-blood group ABO locus. *Biochem Biophys Res Commun* 206:318-325, 1995.
- Blumenfeld OO, Patnaik SK. Allelic genes of blood group antigens: a source of human mutations and cSNPs documented in the Blood Group Antigen Gene Mutation Database. *Hum Mutat* 23:8-16, 2004.
- Blundell J. A successful case of transfusion. *Lancet*. 431, 1829.
- Boix E, Swaminathan GJ, Zhang Y, Natesh R, Brew K, Acharya KR. Structure of UDP complex of UDP-galactose:beta-galactoside-alpha -1,3-galactosyltransferase at 1.53-A resolution reveals a conformational change in the catalytically important C terminus. *J Biol Chem* 276:48608-48614, 2001.
- Breimer ME, Jovall PA. Structural characterization of a blood group A heptagalcosylceramide with globo-series structure. The major glycolipid based blood group A antigen of human kidney. *FEBS Lett* 1;179:165-172, 1985.
- Breton C, Snajdrova L, Jeanneau C, Koca J, Imberty A. Structures and mechanisms of glycosyltransferases *Glycobiology* 16:29R-37R, 2006.
- Cameron C, Graham F, Dunsford I. Acquisition of a B-like antigen by red blood cells. *Br. Med. J.* 32-34, 1959.
- Cartron JP, Gerbal A, Hughes-Jones NC, Salmon C. 'Weak A' phenotypes. Relationship between red cell agglutinability and antigen site density. *Immunology* 27:723-727, 1974.
- Cartron JP, Badet J, Mulet C, Salmon C. Study of the alpha-N-acetylgalactosaminyltransferase in sera and red cell membranes of human A subgroups. *J Immunogenet* 5:107-116, 1978.
- Chester MA. The occurrence of a beta-galactosyltransferase in normal human urine. *FEBS Lett* 46:59-62, 1974.
- Chester MA, Olsson ML. The ABO blood group gene - A locus of considerable genetic diversity. *Transfus Med Rev* 11: 295-313, 2001.
- Clausen H, Watanabe K, Kannagi R, Levery SB, Nudelman E, Arao-Tomono Y, Hakomori S. Blood group A glycolipid (A<sub>x</sub>) with globo-series structure which is specific for blood group A<sub>1</sub> erythrocytes: one of the chemical bases for A<sub>1</sub> and A<sub>2</sub> distinction. *Biochem Biophys Res Commun* 124:523-529, 1984.
- Clausen H, Levery SB, Nudelman E, Tsuchiya S, Hakomori S. Repetitive A epitope (type 3 chain A) defined by blood group A<sub>1</sub>-specific monoclonal antibody TH-1: chemical basis of qualitative A<sub>1</sub> and A<sub>2</sub> distinction. *Proc Natl Acad Sci U S A* 82:1199-1203, 1985.
- Clausen H, Levery SB, Kannagi R, Hakomori S. Novel blood group H glycolipid antigens exclusively expressed in blood group A and AB erythrocytes (type 3 chain H). I. Isolation and chemical characterization. *J Biol Chem* 261:1380-1387, 1986a.
- Clausen H, Holmes E, Hakomori S. Novel blood group H glycolipid antigens exclusively expressed in blood group A and AB erythrocytes (type 3 chain H). II. Differential conversion of different H substrates by A<sub>1</sub> and A<sub>2</sub> enzymes, and type 3 chain H expression in relation to secretor status. *J Biol Chem* 261:1388-1392, 1986b.
- Clausen H, Hakomori S. ABH and related histo-blood group antigens; immunochemical differences in carrier isotypes and their distribution. *Vox Sang* 56:1-20, 1989.

- Clausen H, White T, Takio K, Titani K, Stroud M, Holmes E, Karkov J, Thim L, Hakomori S. Isolation to homogeneity and partial characterization of a histo-blood group A defined Fuc-alpha1-2Gal alpha1-3-N-acetylgalactosaminyltransferase from human lung tissue. *J Biol Chem* 265:1139-1145, 1990.
- Clausen H, Bennett EP, Grunnet N. Molecular genetics of ABO histo-blood groups. *Transfus Clin Biol* 1:79-89, 1994.
- Daniels G. Human blood groups. Blackwell Scientific, Oxford, UK, 2002.
- Daniels GL, Cartron JP, Fletcher A, Garratty G, Henry S, Jorgensen J, Judd WJ, Levene C, Lin M, Lomas-Francis C, Moulds JJ, Moulds JM, Moulds M, Overbeeke M, Reid ME, Rouger P, Scott M, Sistonen P, Smart E, Tani Y, Wendel S, Zelinski T. International Society of Blood Transfusion Committee on terminology for red cell surface antigens. Vancouver Report. *Vox Sang* 84:244-247, 2003.
- Daniels GL, Fletcher A, Garratty G, Henry S, Jorgensen J, Judd WJ, Levene C, Lomas-Francis C, Moulds JJ, Moulds JM, Moulds M, Overbeeke M, Reid ME, Rouger P, Scott M, Sistonen P, Smart E, Tani Y, Wendel S, Zelinski T. International Society of Blood Transfusion. Blood group terminology 2004, from the International Society of Blood Transfusion committee on terminology for red cell surface antigens. *Vox Sang* 87:304-316, 2004.
- Decastello A von & Sturli A. Über die Isoagglutinie im Serum gesunder und kranker menschen. *München Med Wchnschr* 26:1090-1095, 1902. (Translation in Camp & Ellis, 1966)
- Deng ZH, Yu Q, Lian YL, Wu GG, Su YQ, Zhang X. Identification of a novel B variant allele at the ABO locus in Chinese Han individuals with B subgroup. *Ann Clin Lab Sci* 35:265-269, 2005.
- Denis J-B. A letter concerning a new way of curing sundry diseases by transfusion of blood. *Philos Trans. R. Soc* 2:489-504, 1667.
- Dungern E von, Hirschfeld L. Über Vererbung gruppenspezifischer Strukturen des Blutes III. *Z Immunforsch* 8:526-562, 1911. (Translation in Camp & Ellis, 1966)
- Economidou J, Hughes-Jones NC, Gardner B. The reactivity of subunits of IgM anti-B. *Immunology* 13:235-240, 1967.
- Eiz-Vesper B, Seltsam A, Blasczyk R. ABO glycosyltransferases as potential source of minor histocompatibility antigens in allogeneic peripheral blood progenitor cell transplantation. *Transfusion* 45:960-968, 2005.
- Erskine AG, Socha WW. The principles and practice of blood grouping. The C.V. Mosby Company, Missouri, USA, 1978.
- Farr A.D. Blood group serology the first four decades (1900-1939). *Medical History* 23:215-226, 1979.
- Farr A.D. The first human blood transfusion. *Medical History* 24:143-162, 1980.
- Ferguson-Smith MA, Aitken DA, Turleau C, de Grouchy J. Localisation of the human ABO: Np-1: AK-1 linkage group by regional assignment of AK-1 to 9q34. *Hum Genet* 34:35-43, 1976.
- Fischer W, Hahn F. Über auffallende Schwäche der gruppenspezifischen Reaktionsfähigkeit bei einem Erwachsenen. *Z Immunforsch* 84:177-188, 1935.
- Garratty G, Dzik W, Issitt PD, Lublin DM, Reid ME, Zelinski T. Terminology for blood group antigens and genes-historical origins and guidelines in the new millennium. *Transfusion* 40:477-489, 2000.
- Gassner C, Schmarda A, Nussbaumer W, Schonitzer D. ABO glycosyltransferase genotyping by polymerase chain reaction using sequence-specific primers. *Blood* 88: 1852-1856, 1996.
- Gastinel LN, Bignon C, Misra AK, Hindsgaul O, Shaper JH, Joziassse DH. Ine alpha1,3-galactosyltransferase catalytic domain structure and its relationship with ABO histo-blood group and glycosphingolipid glycosyltransferases. *EMBO J.* 15;20:638-649, 2001.
- Ginsburg V. Enzymatic basis for blood groups in man. *Adv Enzymol Relat Areas Mol Biol* 36:131-149, 1972.
- Goldstein J, Lenny L, Davies D, Voak D. Further evidence for the presence of A antigen on group B erythrocytes through the use of specific exoglycosidases. *Vox Sang* 57:142-146, 1989.
- Griot-Wenk M, Pahlsson P, Chisholm-Chait A, Spitalnik PF, Spitalnik SL, Giger U. Biochemical characterization of the feline AB blood group system. *Anim Genet* 24:401-407, 1993.

- Grunnet N, Steffensen R, Bennett EP, Clausen H. Evaluation of histo-blood group ABO genotyping in a Danish population: frequency of a novel O allele defined as O<sup>2</sup>. *Vox Sang* 67:210-215, 1994.
- Hansen T, Namork E, Olsson ML, Chester MA, Heier HE. Different genotypes causing indiscernible patterns of A expression on A(el) red blood cells as visualized by scanning immunogold electron microscopy. *Vox Sang* 75:47-51, 1998.
- Harvey W. *Exercitatio anatomica de moto cordis et sanguinis in animalibus*. Frankfurt, 1628.
- Hata Y, Kominato Y, Yamamoto FI, Takizawa H. Characterization of the human ABO gene promoter in erythroid cell lineage. *Vox Sang* 82: 39-46, 2002.
- Hata Y, Kominato Y, Takizawa H, Tabata S, Michino J, Nishino K, Yasumura S, Yamamoto F. Transcription starting from an alternative promoter leads to the expression of the human ABO histo-blood group antigen. *Transfusion* 43:656-662, 2003.
- Heier HE, Namork E, Calkovska Z, Sandin R, Kornstad L. Expression of A antigens on erythrocytes of weak blood group A subgroups. *Vox Sang* 66:231-236, 1994.
- Henissat B. Glycosidase families. *Biochem Soc Trans* 26:153-156, 1998.
- Irshaid NM, Chester MA, Olsson ML. Allele-related variation in minisatellite repeats involved in the transcription of the blood group ABO gene. *Transfusion Med* 9:219-226, 1999.
- Jenkins T. Blood group A<sub>bantu</sub> population and family studies. *Vox Sang* 26:537-550, 1974.
- Kabat EA. Blood group substances. Their chemistry and immunochemistry. Academic Press, New York, 1956.
- Kannagi R, Levery SB, Hakomori S. Blood group H antigen with globo-series structure. Isolation and characterization from human blood group O erythrocytes. *FEBS Lett* 175:397-401, 1984.
- Kominato Y, McNeill PD, Yamamoto M, Russell M, Hakomori S, Yamamoto F. Animal histo-blood group ABO genes. *Biochem Biophys Res Commun* 189:154-164, 1992.
- Kominato Y, Tsuchiya T, Hata N, Takizawa H, Yamamoto F. Transcription of human ABO histo-blood group genes is dependent upon binding of transcription factor CBF/NF-Y to minisatellite sequence. *J Biol Chem* 272:25890-25898, 1997.
- Kominato Y, Hata Y, Takizawa H, Matsumoto K, Yasui K, Tsukada J. Alternative promoter identified between a hypermethylated upstream region of repetitive elements and a CpG island in human ABO histo-blood group genes. *J Biol Chem* 277: 37936-37948, 2002.
- Landsteiner K. Zur kenntnis der antifermentativen, lytischen und agglutinierenden Wirkungen des Bluterums und der Lymphe. *Zbl Bakter* 27:357-366, 1900. (Translation in Camp & Ellis, 1966)
- Landsteiner K. Über Agglutinationserscheinungen normalen menschlichen Blutes. *Klein. Wschr.* 14,1132-1134, 1901. (Translation in Camp & Ellis, 1966)
- Landsteiner K. On Individual Differences in Human Blood. Physiology or Medicine 1922-1941:234-245, Nobel Lectures 1930.
- Lee HJ, Barry CH, Borisova SN, Seto NO, Zheng RB, Blancher A, Evans SV, Palcic MM. Structural basis for the inactivity of human blood group O<sup>2</sup> glycosyltransferase. *J Biol Chem* 280:525-529, 2005.
- Lin PH, Li L, Lin-Tsai SJ, Lin KT, Chen JM, Chu DC. A unique 502C>T mutation in exon 7 of ABO gene associated with the B<sub>el</sub> phenotype in Taiwan. *Transfusion* 43:1254-1259, 2003.
- Loghem JJ van, Dorfmeier H, Hart M van der. Two antigens with abnormal serologic properties. *Vox Sang* 2:16-24, 1957.
- Lowe JB. Biochemistry and biosynthesis of ABH and Lewis antigens. In: Cartron JP, Rouger P, eds. Blood cell biochemistry: Molecular basis of major human blood group antigens. Plenum Press, New York, 75-115, 1995.
- Lower R. The success of the experiment of transfusing the blood of one animal into another. *Philos Trans R Soc* 1, 352, 1666.
- Martinko JM, Vincek V, Klein D, Klein J. Primate ABO glycosyltransferases: evidence for trans-species evolution. *Immunogenetics* 37:274-278, 1993.

- Mifsud NA, Watt JM, Condon JA. A novel cis-AB variant allele arising from a nucleotide substitution A796C in the B transferase gene. *Transfusion* 40:1276-1277, 2000.
- Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16:1215, 1988.
- Mollison PL, Engelfriet CP, Contreras M. Blood transfusion in clinical medicine. Blackwell Scientific, Oxford, 1993.
- Morera S, Imberty A, Aschke-Sonnenborn U. T4 phage beta-glucosyltransferase: substrate binding and proposed catalytic mechanism. *J Mol Biol* 292:717-730, 1999.
- Mourant AE, Kopec K, Domaniewska-sobczak K. The distribution of the human blood groups and other polymorphisms. Oxford university press, 1976.
- Nakamura I, Takizawa H, Nishino K. A3 phenotype with A1 gene-specified enzyme character in serum. *Exp Clin Immunogenet* 6:143-149, 1989.
- Ogasawara K, Yabe R, Uchikawa M, Saitou N, Bannai M, Nakata K, Takenaka M, Fujisawa K, Ishikawa Y, Juji T, Tokunaga K. Molecular genetic analysis of variant phenotypes of the ABO blood group system. *Blood* 88:2732-2737, 1996a.
- Ogasawara K, Bannai M, Saitou N. Extensive polymorphism of ABO blood group gene: three major lineages of the alleles for the common ABO phenotypes. *Hum Genet* 97:777-783, 1996b.
- Ogasawara K, Yabe R, Uchikawa M, Bannai M, Nakata K, Takenaka M, Takahashi Y, Juji T, Tokunaga K. Different alleles cause an imbalance in A<sub>2</sub> and A<sub>2</sub>B phenotypes of the ABO blood group. *Vox Sang* 74:242-247, 1998.
- Ogasawara K, Yabe R, Uchikawa M, Nakata K, Watanabe J, Takahashi Y. Recombination and gene conversion-like events may contribute to ABO gene diversity causing various phenotypes. *Immunogenetics* 53: 190-199, 2001.
- Olsson ML, Chester MA. A rapid and simple ABO genotype screening method using a novel B/O<sup>2</sup> versus A/O<sup>1</sup> discriminating nucleotide substitution at the ABO locus. *Vox Sang* 69:242-247, 1995.
- Olsson ML, Thuresson B, Chester MA. An A<sup>el</sup> allele-specific nucleotide insertion at the blood group ABO locus and its detection using a sequence-specific polymerase chain reaction. *Biochem Biophys Res Commun* 216:642-647, 1995.
- Olsson ML, Chester MA. Frequent occurrence of a variant O<sup>1</sup> gene at the blood group ABO locus. *Vox Sang* 70:26-30, 1996a.
- Olsson ML, Chester MA. Polymorphisms at the ABO locus in subgroup A individuals. *Transfusion* 36:309-313, 1996b.
- Olsson ML, Chester MA. Evidence for a new type of O allele at the ABO locus, due to a combination of the A<sup>2</sup> nucleotide deletion and the A<sup>el</sup> nucleotide insertion. *Vox Sang* 71:113-117, 1996c.
- Olsson ML, Guerreiro JF, Zago MA, Chester MA. Molecular analysis of the O alleles at the blood group ABO locus in populations of different ethnic origin reveals novel crossing-over events and point mutations. *Biochem Biophys Res Commun* 234:779-782, 1997.
- Olsson ML, Chester MA. Heterogeneity of the blood group A<sup>x</sup> allele: genetic recombination of common alleles can result in the A<sub>x</sub> phenotype. *Transfus Med* 8:231-238, 1998a.
- Olsson ML, Santos SE, Guerreiro JF, Zago MA, Chester MA. Heterogeneity of the O alleles at the blood group ABO locus in Amerindians. *Vox Sang* 74:46-50, 1998b.
- Olsson ML, Hosseini-Maaf B, Hellberg Å, Chester MA. Allele-specific primer PCR across intron 6 resolves potential genotyping errors caused by recombinant hybrid alleles at the ABO locus (plenary poster). *Transfusion* 38(10S):3S, 1998c.
- Olsson ML, Chester MA. Polymorphism and recombination events at the ABO locus: a major challenge for genomic ABO blood grouping strategies. *Transfus Med* 11:295-313, 2001.
- Oriol R, Mollicone R, Coulin P, Dalix AM, Candelier JJ. Genetic regulation of the expression of ABH and Lewis antigens in tissues. *APMIS* 100:28-38, 1992.

- Oriol, R. 1995. ABO, Hh, Lewis and secretion. Serology, genetics and tissue distribution. In J. P. Cartron and P. Rouger, eds. *Blood Cell Biochemistry*. Vol. 6. New York: Plenum Press 75-115, 1995.
- Pacuszka T, Koscielak J, Seyfried H, Walewska I. Biochemical, serological and family studies in individuals with *Cis* AB phenotypes. *Vox Sang* 29:292-300, 1975.
- Pak JE, Arnoux P, Zhou S, Sivarajah P, Satkunarajah M, Xing X, Rini JM. X-ray crystal structure of leukocyte type core 2 beta1,6-N-acetylglucosaminyltransferase. Evidence for a convergence of metal ion-independent glycosyltransferase mechanism. *J Biol Chem* 281:26693-26701, 2006.
- Palcic MM, Heerze LD, Pierce M, Hindsgaul O. The use of hydrophobic synthetic glycoside as acceptors in glycosyltransferase assays. *Glycoconj J* 5:49-63, 1988.
- Patenaude SI, Seto NO, Borisova SN, Szpacenko A, Marcus SL, Palcic MM. The structural basis for specificity in human ABO(H) blood group biosynthesis. *Nat Struct Biol* 9:685-690, 2002.
- Paulson JC, Colley KJ. Glycosyltransferases. Structure, localization, and control of cell type-specific glycosylation. *J Biol Chem* 264:17615-17618, 1989.
- Persson K, Ly HD, Dieckelmann M. Crystal structure of the retaining galactosyltransferase LgtC from *Neisseria meningitidis* in complex with donor and acceptor sugar analogs. *Nat Struct Biol* 8:166-175, 2001.
- Persson M, Letts JA, Hosseini-Maaf B, Borisova SN, Evans SV, Palecic MM, Olsson ML. Structural Effects of Naturally-Occurring Human Blood Group B Galactosyltransferase Mutations Adjacent to the DxD Motif. *J Biol Chem* 2007, in press.
- Pflugrath JW. The finer things in X-ray diffraction data collection. *Acta Crystallogr D Biol Crystallogr* 55:1718-1725, 1999.
- Roubinet F, Kermarrec N, Despiau S, Apoil PA, Dugoujon JM, Blancher A. Molecular polymorphism of O alleles in five populations of different ethnic origins. *Immunogenetics* 53:95-104, 2001. Erratum in: *Immunogenetics* 53:718, 2001.
- Saitou N, Yamamoto F. Evolution of primate ABO blood group genes and their homologous genes. *Mol Biol Evol* 14:399-411, 1997.
- Schachter H, Michaels MA, Tilley CA, Crookston MC, Crookston JH. Qualitative differences in the N-acetyl-D-galactosaminyltransferases produced by human A<sup>1</sup> and A<sup>2</sup> genes. *Proc Natl Acad Sci U S A* 70:220-224, 1973.
- Schachter H, Michaels MA, Tilley CA, Crookston MC, Crookston JH. Qualitative differences in the N-acetyl-D-galactosaminyltransferases produced by human A<sup>1</sup> and A<sup>2</sup> genes. *Proc Natl Acad Sci U S A* 70:220-224, 1973.
- Schenkel-Brunner H, Chester MA, Watkins WM. Alpha-L-fucosyltransferases in human serum from donors of different ABO, secretor and Lewis blood-group phenotypes. *Eur J Biochem* 30:269-277, 1972.
- Schenkel-Brunner H. In human blood groups. Springer-Verlag, New York, 2000
- Seltsam A, Hallensleben M, Eiz-Vesper B, Lenhard V, Heymann G, Blasczyk R. A weak blood group A phenotype caused by a new mutation at the ABO locus. *Transfusion* 42:294-301, 2002.
- Seltsam A, Hallensleben M, Kollmann A, Blasczyk R. The nature of diversity and diversification at the ABO locus. *Blood* 102:3035-3042, 2003.
- Seltsam A, Blasczyk R. Missense mutations outside the catalytic domain of the ABO glycosyltransferase can cause weak blood group A and B phenotypes. *Transfusion* 45:1663-1669, 2005.
- Seto NO, Compston CA, Evans SV, Bundle DR, Narang SA, Palcic MM. Donor substrate specificity of recombinant human blood group A, B and hybrid A/B glycosyltransferases expressed in *Escherichia coli*. *Eur J Biochem* 259:770-775, 1999.
- Seyfried H, Walewska I, Werbinska B. Unusual inheritance of ABO group in family with weak B antigens. *Vox Sang* 9:268-277, 1964.
- Shapiro M. The ABO, MN, P and Rh blood group systems in the South African Bantu, a genetic study. *S Afr Med J* 25:165-170, 1951.
- Smith GR. Hotspots of homologous recombination. *Experientia* 50:234-241, 1994.
- Spriger GF, Horton RS and Forbes M. Origin of anti-human blood group B agglutinins in white Leghorn chicks. *J Exp Med*. 110:221-244, 1959.

- Stahl F, Myers R. Old and new concepts for the role of chi in bacterial recombination. *J Hered* 86:327-329, 1995.
- Storry JR, Olsson ML. Genetic basis of blood group diversity. *Br J Haematol* 126:759-771, 2004.
- Suzuki K, Iwata M, Tsuji H, Takagi T, Tamura A, Ishimoto G, Ito S, Matsui K, Miyazaki T. A de novo recombination in the ABO blood group gene and evidence for the occurrence of recombination products. *Hum Genet* 99:454-461, 1997.
- Svensson L, Rydberg L, Hellberg A, Gilliver LG, Olsson ML, Henry SM. Novel glycolipid variations revealed by monoclonal antibody immunochemical analysis of weak ABO subgroups of A. *Vox Sang* 89:27-38, 2005.
- The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50:760-763, 1994.
- Thuresson B, Dykes J, Olofsson T, Olsson ML. The number of repeats in the CBF/NF-Y enhancer VNTR region does not influence the level of ABO transcripts in blood and erythropoietic culture. *Vox Sang* 91:135, 2006.
- Unligil UM, Zhou S, Yuwaraj S, Sarkar M, Schachter H, Rini JM. X-ray crystal structure of rabbit N-acetylglucosaminyltransferase I: catalytic mechanism and a new protein superfamily. *EMBO J* 19:5269-5280, 2000.
- Watkins WM. Biochemistry and genetics of the ABO, Lewis and P blood group systems. In: Harris H, Hirschhorn K, (eds.) *Advances in Human Genetics*. Plenum Press, New York, USA, Vol. 10:1-136, 1980.
- Watkins WM. Blood-Group Substances. *Science* 152:172-181, 1966.
- Yamamoto F. Cloning the ABH genes. *Transfusion* 30:671-672, 1990.
- Yamamoto F, Hakomori S. Sugar-nucleotide donor specificity of histo-blood group A and B transferases is based on amino acid substitutions. *J Biol Chem* 265:19257-19262, 1990.
- Yamamoto F, Clausen H, White T, Marken J, Hakomori S. Molecular genetic basis of the histo-blood group ABO system. *Nature* 345:229-233, 1990a.
- Yamamoto F, Marken J, Tsuji T, White T, Clausen H, Hakomori S. Cloning and characterization of DNA complementary to human UDP-GalNAc:Fuc-alpha1-2Gal alpha1-3GalNAc transferase (histo-blood group A transferase) mRNA. *J Biol Chem* 265:1146-1151, 1990b.
- Yamamoto F, McNeill PD, Hakomori S. Human histo-blood group A<sub>2</sub> transferase coded by A<sup>2</sup> allele, one of the A subtypes, is characterized by a single base deletion in the coding sequence, which results in an additional domain at the carboxyl terminal. *Biochem Biophys Res Commun* 187:366-374, 1992.
- Yamamoto F, McNeill PD, Yamamoto M, Hakomori S, Harris T, Judd WJ, Davenport RD. Molecular genetic analysis of the ABO blood group system: 1. Weak subgroups: A<sup>3</sup> and B<sup>3</sup> alleles. *Vox Sang* 64:116-119, 1993a.
- Yamamoto F, McNeill PD, Yamamoto M, Hakomori S, Harris T. Molecular genetic analysis of the ABO blood group system: 3. A<sup>x</sup> and B(A) alleles. *Vox Sang* 64:171-174, 1993b.
- Yamamoto F, McNeill PD, Kominato Y, Yamamoto M, Hakomori S, Ishimoto S, Nishida S, Shima M, Fujimura Y. Molecular genetic analysis of the ABO blood group system: 2. cis-AB alleles. *Vox Sang* 64:120-123, 1993c.
- Yamamoto F, McNeill PD, Yamamoto M, Hakomori S, Bromilow IM, Duguid JK. Molecular genetic analysis of the ABO blood group system: 4. Another type of O allele. *Vox Sang* 64:175-178, 1993d.
- Yamamoto F, McNeill PD, Hakomori S. Genomic organization of human histo-blood group ABO genes. *Glycobiology* 5:51-58, 1995.
- Yamamoto F, McNeill PD. Amino acid residue at codon 268 determines both activity and nucleotide-sugar donor substrate specificity of human histo-blood group A and B transferases. In vitro mutagenesis study. *J Biol Chem* 271:10515-10520, 1996.
- Yamamoto F & Yamamoto M. Molecular genetic basis of porcine histo-blood group AO system. *Blood* 15;97:3308-3310, 2001.

- Yamamoto M, Lin XH, Kominato Y, Hata Y, Noda R, Saitou N, Yamamoto F. Murine equivalent of the human histo-blood group *ABO* gene is a *cis-AB* gene and encodes a glycosyltransferase with both A and B transferase activity. *J Biol Chem* 276:13701-13708, 2001.
- Yamamoto F. ABO blood group system-ABH oligosaccharide antigens, anti-A and anti-B, A and B glycosyltransferases, and *ABO* genes. *Immunohematol* 20:3-22, 2004.
- Yazer MH, Palcic MM. The importance of disordered loops in ABO glycosyltransferases. *Transfus Med Rev* 19:210-216, 2005.
- Yazer MH, McGuirt J, Hellberg Å, Hosseini-Maaf B, Hult A, Cortese-Hassett A, Triulzi DJ, Olsson ML. Heterozygosity for the non-deletional *O<sup>2</sup>* allele does not cause discrepancies in automated blood donor ABO grouping (abstract). *Blood* 108:957, 2006.
- Yip SP, Choy WL, Chan CW, Choi CH. The absence of a B allele in acquired B blood group phenotype confirmed by a DNA based genotyping method. *J Clin Pathol* 49:180-181, 1996.
- Yip SP. Single-tube multiplex PCR-SSCP analysis distinguishes 7 common ABO alleles and readily identifies new alleles. *Blood* 95:1487-1492, 2000.
- Yip SP. Sequence variation at the human ABO locus. *Ann Hum Genet* 66:1-27, 2002.
- Yu LC, Lee HL, Chan YS, Lin M. The molecular basis for the B(A) allele: an amino acid alteration in the human histoblood group B alpha-(1,3)-galactosyltransferase increases its intrinsic alpha-(1,3)-N-acetylgalactosaminyltransferase activity. *Biochem Biophys Res Commun*. 262:487-493, 1999.
- Yu LC, Chang CY, Twu YC, Lin M. Human histo-blood group ABO glycosyltransferase genes: different enhancer structures with different transcriptional activities. *Biochem Biophys Res Commun* 273:459-66, 2000.