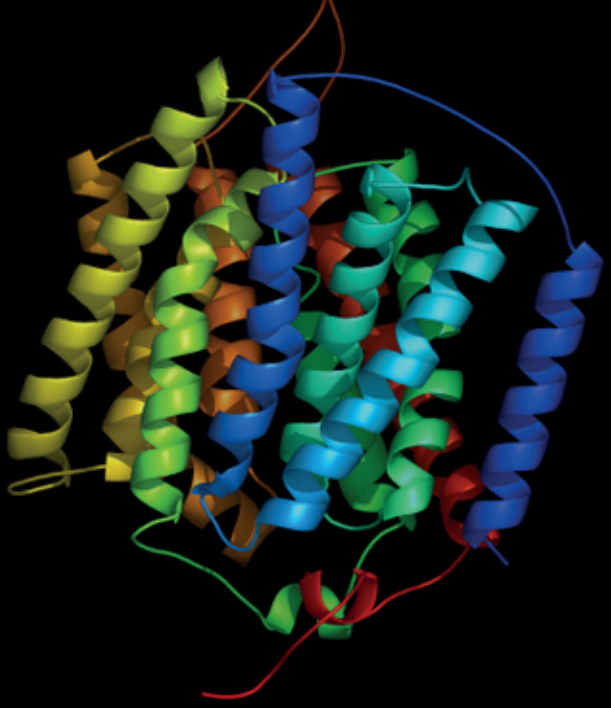
THIRD EDITION



HUMAN BLOOD GROUPS

Geoff Daniels

WILEY-BLACKWELL

Human Blood Groups

Human Blood Groups

Geoff Daniels
BSc, PhD, FRCPath
Head of Diagnostics
International Blood Group Reference Laboratory;
Senior Research Fellow
Bristol Institute for Transfusion Sciences,
NHS Blood and Transplant, Bristol, UK

Foreword by Ruth Sanger

3rd edition

WILEY-BLACKWELL
A John Wiley & Sons, Ltd., Publication

This edition first published 2013 © 1995, 2002, 2013 by Geoff Daniels

Blackwell Publishing was acquired by John Wiley & Sons in February 2007. Blackwell's publishing program has been merged with Wiley's global Scientific, Technical and Medical business to form Wiley-Blackwell.

Registered office: John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial offices: 9600 Garsington Road, Oxford, OX4 2DQ, UK
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK
111 River Street, Hoboken, NJ 07030-5774, USA

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell

Library of Congress Cataloging-in-Publication Data
Daniels, Geoff.
Human blood groups : Geoff Daniels ; foreword to first edition by Ruth Sanger. – 3rd ed.
p. ; cm.
Includes bibliographical references and index.
ISBN 978-1-4443-3324-4 (hardback : alk. paper) – ISBN 978-1-118-49354-0(epub) – ISBN 978-1-118-49359-5 (obook) – ISBN 978-1-118-49361-8 (emobi) – ISBN 978-1-118-49362-5 (epdf)
I. Title.
[DNLM: 1. Blood Group Antigens. WH 420]
612.1'1825–dc23
2012040684

A catalogue record for this book is available from the British Library.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Cover image (top right): Homology model of an Rh protein (RhD or RhCE) courtesy of Dr Nicholas Burton, University of Bristol, UK. Blood bag image: © iStockPhoto / pictorico

Cover design by Garth Stewart

Set in 9.25/11.5 pt Minion by Toppan Best-set Premedia Limited

1 2013

Contents

Foreword to 1st edition

It is a particular pleasure for me to welcome this new book on human blood groups, the more so since it emanates from the Medical Research Council's Blood Group Unit. For 25 years this Unit devoted its energies to the search for new red cell antigens and the application of those already known to various problems, particularly to human genetics. During these years Rob Race and I produced six editions of *Blood Groups in Man*.

Dr Geoff Daniels joined the Unit in 1973 on Dr Race's retirement; soon after, concurrently with the Unit's move from the Lister Institute to University College, the scope of the Unit's interest was broadened.

Having been divorced from blood groups and otherwise occupied in 12 years of retirement, I am delighted and astonished at the rapid advances made in recent years. The number of blood group loci have increased to 23 and all except one have found their chromosomal home. The biochemical backgrounds of most of the corresponding antigens are defined and hence several high and low incidence antigens gathered into systems. The molecular basis of many red cell antigens has provided an explanation for some confusing serological relationships which were observed many years before.

Dr Daniels is to be congratulated on his stamina in producing a comprehensive text and reference book on human blood groups, for which many scientists will be grateful.

Ruth Sanger
December 1994

Preface to the third edition

The primary purpose of this book, like the first two editions, is to describe human blood group antigens and their inheritance, the antibodies that define them, the structure and functions of the red cell membrane macromolecules that carry them, and the genes that encode them or control their biosynthesis. In addition, this book provides information on the clinical relevance of blood groups and on the importance of blood group antibodies in transfusion medicine in particular.

The second edition of *Human Blood Groups* was published in 2002; this new edition will appear 11 years later. There have been many new findings in the blood group world over those years. In order to prevent the book from becoming too cumbersome, my goal has been to produce a third edition roughly the same size as the first two. I have tried to do this without eliminating anything too important, although this has not been easy, with so much new material to include. Since 2002, about 69 new blood group antigens and seven new blood group systems have been identified, and all of the 38 genes representing those systems have been cloned and sequenced.

In the preface of the sixth edition of *Blood Groups in Man*, the predecessor of *Human Blood Groups*, Race and Sanger wrote, 'Here is the last edition of this book: the subject has grown to need more than our two pencils'. Well, here is the last edition of *Human Blood Groups*; the subject is rapidly growing too vast to be contained in a textbook. In the previous two editions I strove to include all fully validated blood group antigens and genetic changes associated with their expression or loss of expression. This has proved impossible and pointless in this edition so, although the genetic bases of all the important blood group polymorphisms are described, in many cases the reader is directed to web sites for a more complete list of mutations, particularly those responsible for null phenotypes. In the next few years, next-generation sequencing will become readily available and affordable, and the number of genetic variations associated with red cell change will increase exponentially.

I wish to thank again all the people who helped me produce the first two editions, in particular Patricia Tippett, Carole Green, David Anstee, and Joan Daniels. I would like to add my thanks to Dr Nicholas Burton at the University of Bristol who provided many of the protein models for this edition. Finally I would like to thank all the numerous colleagues from around the world who have provided so much of the information in this book, in published or unpublished form, over so many years.

Geoff Daniels

Some abbreviations used

ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
AET	2-aminoethylisothiourunium bromide
AIHA	Autoimmune haemolytic anaemia
bp	Base-pair
CDA	Congenital dyserythropoietic anaemia
cDNA	Complimentary DNA
CFU-E	Colony-forming unit-erythroid
Da	Daltons
DAT	Direct antiglobulin test
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
Gal	Galactose
GalNAc	*N*-acetylgalactosamine
GlcNAc	*N*-acetylglucosamine
GDP	Guanosine diphosphate
GPI	Glycosylphosphatidylinositol
GSL	Glycosphingolipid
GTA	A-transferase
GTB	B-transferase
HCF	Hydatid cyst fluid
HDFN	Haemolytic disease of the fetus and newborn
HTR	Haemolytic transfusion reaction
IAT	Indirect antiglobulin test
ISBT	International Society of Blood Transfusion (may refer to ISBT terminology)
kb	Kilo-bases
kDa	Kilo-Daltons
MAIEA	Monoclonal antibody immobilisation of erythrocyte antigens
mRNA	Messenger ribonucleic acid
MW	Molecular weight
PCR	Polymerase chain reaction
RFLP	Restriction fragment-length polymorphism
RNA	Ribonucleic acid
SDS PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SNP	Single nucleotide polymorphism

1 Human Blood Groups: Introduction

1.1 Introduction

What is the definition of a blood group? Taken literally, any variation or polymorphism detected in the blood could be considered a blood group. However, the term blood group is usually restricted to blood cell surface antigens and generally to red cell surface antigens. This book focuses on the inherited variations in human red cell membrane proteins, glycoproteins, and glycolipids. These variations are detected by alloantibodies, which occur either 'naturally', due to immunisation by ubiquitous antigens present in the environment, or as a result of alloimmunisation by human red cells, usually introduced by blood transfusion or pregnancy. Although it is possible to detect polymorphism in red cell surface proteins by other methods such as DNA sequence analysis, such variants cannot be called blood groups unless they are defined by an antibody.

Blood groups were discovered at the beginning of the twentieth century when Landsteiner [1,2] noticed that plasma from some individuals agglutinated the red cells from others. For the next 45 years, only those antibodies that directly agglutinate red cells could be studied. With the development of the antiglobulin test by Coombs, Mourant, and Race [3,4] in 1945, non-agglutinating antibodies could be detected and the science of blood group serology blossomed. There are now 339 authenticated blood group antigens, 297 of which fall into one of 33 blood group systems, genetically discrete groups of antigens controlled by a single gene or cluster of two or three closely linked homologous genes (Table 1.1).

Most blood group antigens are synthesised by the red cell, but the antigens of the Lewis and Chido/Rodgers systems are adsorbed onto the red cell membrane from the plasma. Some blood group antigens are detected only on red cells; others are found throughout the body and are often called histo-blood group antigens.

Biochemical analysis of blood group antigens has shown that they fall into two main types:
1 protein determinants, which represent the primary products of blood group systems; and
2 carbohydrate determinants on glycoproteins and glycolipids, in which the products of the genes controlling antigen expression are glycosyltransferase enzymes.
Some antigens are defined by the amino acid sequence of a glycoprotein, but are dependent on the presence of carbohydrate for their recognition serologically. In this book the three-letter code for amino acids is mainly used, though the single-letter code is often employed in long sequences and in some figures. The code is provided in Table 1.2.

In recent years, molecular genetical techniques have been introduced into the study of human blood groups and now most of the genes governing blood group systems have been cloned and sequenced (Table 1.1). Many serological complexities of blood groups are now explained at the gene level by a variety of mechanisms, including point mutation, unequal crossing-over, gene conversion, and alternative RNA splicing.

Human Blood Groups, Third Edition. Geoff Daniels.

Table 1.1 Blood group systems.

No.	Name	Symbol*	No. of antigens	Associated membrane structures	CD no.	HGNC symbol(s)	Chromosome
001	ABO	ABO	4	Carbohydrate		*ABO*	9
002	MNS	MNS	46	Glycophorins, GPA, GPB	CD235 A & B	*GYPA, GYPB, GYPE*	4
003	P1PK	P1PK	3	Carbohydrate		*A4GALT*	22
004	Rh	RH	54	Rh family, RhD, RhCcEe	CD240 D & CE	*RHD, RHCE*	1
005	Lutheran	LU	20	IgSF	CD239	*BCAM*	19
006	Kell	KEL	35	Endopeptidase	CD238	*KEL*	7
007	Lewis	LE	6	Carbohydrate		*FUT3*	19
008	Duffy	FY	5	G protein-coupled SF, chemokine receptor	CD234	*DARC*	1
009	Kidd	JK	3	Urea transporter		*SLC14A1*	18
010	Diego	DI	22	Band 3, anion exchanger (AE1)	CD233	*SLC4A1*	17
011	Yt	YT	2	Acetylcholinesterase		*ACHE*	7
012	Xg	XG	2	Glycoproteins	CD99**	*XG, CD99*	X/Y
013	Scianna	SC	7	IgSF, erythroblast membrane-associated protein		*ERMAP*	1
014	Dombrock	DO	8	ADP-ribosyltransferase 4	CD297	*ART4*	12
015	Colton	CO	4	Aquaporin SF, aquaporin-1		*AQP1*	7
016	Landsteiner-Wiener	LW	3	IgSF, intercellular adhesion molecule-4	CD242	*ICAM4*	19
017	Chido/Rodgers	CH/RG	9	Complement components C4A, C4B		*C4A, C4B*	6
018	H	H	1	Carbohydrate, Type 2 H	CD173	*FUT1*	19
019	Kx	XK	1	Xk protein		*XK*	X
020	Gerbich	GE	11	Glycophorins, GPC, GPD	CD236	*GYPC*	2
021	Cromer	CROM	18	CCP SF, decay-accelerating factor	CD55	*CD55*	1
022	Knops	KN	9	CCP SF, complement regulator-1	CD35	*CR1*	1
023	Indian	IN	4	Link module SF of proteoglycans	CD44	*CD44*	11
024	Ok	OK	3	IgSF, basigin	CD147	*BSG*	19
025	Raph	RAPH	1	Tetraspanin SF	CD151	*CD151*	11
026	John Milton Hagen	JMH	6	Semaphorin SF	CD108	*SEMA7A*	15
027	I	I	1	Carbohydrate		*GCNT2*	6
028	Globoside	GLOB	1	Carbohydrate, globoside		*B3GALT3*	3
029	Gill	GIL	1	Aquaporin SF, aquaporin-3		*AQP3*	9
030	RHAG	RHAG	4	Rh family, Rh-associated glycoprotein	CD241	*RHAG*	6
031	Forssman	FORS	1	Carbohydrate, Forssman glycolipid		*GBGT1*	9
032	Junior	JR	1	ATP-binding cassette transporter ABCG2		*ABCG2*	4
033	Lan	LAN	1	ATP-binding cassette transporter ABCB6		*ABCB6*	2

HGNC, Human Genome Organisation Gene Nomenclature Committee; SF, superfamily; IgSF, immunoglobulin superfamily; CCP, complement control protein.
*ISBT gene name when in italics.
**Does not include Xg glycoprotein.

Table 1.2 The 20 common amino acids: one- and three-letter codes.

A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic acid
E	Glu	Glutamic acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

Discovery of the ABO blood groups first made blood transfusion feasible and disclosure of the Rh antigens led to the understanding, and subsequent prevention, of haemolytic disease of the fetus and newborn (HDFN). Although ABO and Rh are the most important systems in transfusion medicine, many other blood group antibodies are capable of causing a haemolytic transfusion reaction (HTR) or HDFN. Red cell groups have been important tools in forensic science, although this role was diminished with the introduction of HLA testing and has recently been displaced by DNA 'fingerprinting'. For many years blood groups were the best human genetic markers and played a major part in the mapping of the human genome.

Blood groups still have much to teach us. Because red cells are readily available and haemagglutination tests relatively easy to perform, the structure and genetics of the red cell membrane proteins and lipids are understood in great detail. With the unravelling of the complexities of blood group systems by molecular genetical techniques, much has been learnt about the mechanisms responsible for the diversification of protein structures and the nature of the human immune response to proteins of different shapes resulting from variations in amino acid sequence.

1.2 Blood group terminology

The problem of providing a logical and universally agreed nomenclature has dogged blood group serologists almost since the discovery of the ABO system. Before going any further, it is important to understand how blood groups are named and how they are categorised into systems, collections, and series.

1.2.1 An internationally agreed nomenclature

The International Society of Blood Transfusion (ISBT) Working Party on Red Cell Immunogenetics and Blood Group Terminology was set up in 1980 to establish a uniform nomenclature that is 'both eye and machine readable'. Part of the brief of the Working Party was to produce a nomenclature 'in keeping with the genetic basis of blood groups' and so a terminology based primarily around the blood group systems was devised. First the systems and the antigens they contained were numbered, then the high and low frequency antigens received numbers, and then, in 1988, collections were introduced. Numbers are never recycled: when a number is no longer appropriate it becomes obsolete.

Blood group antigens are categorised into 33 systems, seven collections, and two series. The Working Party produced a monograph in 2004 to describe the terminology [5], which was most recently updated in 2011 [6]. Details can also be found on the ISBT web site [7].

1.2.2 Antigen, phenotype, gene and genotype symbols

Every authenticated blood group antigen is given a six-digit identification number. The first three digits represent the system (001 to 033), collection (205 to 213), or series (700 for low frequency, 901 for high frequency); the second three digits identify the antigen. For example, the Lutheran system is system 005 and Lu^a, the first antigen in that system, has the number 005001. Each system also has an alphabetical symbol: that for Lutheran is LU. So Lu^a is also LU001 or, because redundant sinistral zeros may be discarded, LU1. For phenotypes, the system symbol is followed by a colon and then by a list of antigens present, each separated by a comma. If an antigen is known to be absent, its number is preceded by a minus sign. For example, Lu(a–b+) becomes LU:–1,2.

Devising a modern terminology for blood group alleles is more complex. One antigen, the absence of an antigen, or the weakness or absence of all antigens of a system

may be encoded by several or many alleles. Over the last few years the Working Party has been developing a new terminology for bloods group alleles. Unfortunately at the time of publication of this book, it was still incomplete, controversial, and in draft form. Consequently, it has only partially been used in this book. Basically, alleles have the system symbol followed by an asterisk followed in turn by a number or series of numbers, separated by full stops, representing the encoded antigen and the allele number. Alternatively, in some cases a letter can be used instead of a number. For example, *Lu*^{a} allele can be *LU*01* or *LU*A*. Genotypes have the symbol followed by an asterisk followed by the two alleles separated by a stroke. For example, *Lu*^{a}/*Lu*^{b} becomes *LU*01/02* or *LU*A/B*. The letters N and M represent null and mod. For example, one of the inactive *Lu*^{b} alleles responsible for a null phenotype is *LU*02N.01*, the *02* representing the *Lu*^{b} allele, even though no Lu^{b} antigen is expressed. Genes, alleles, and genotypes are italicised. For lists of blood group alleles in the ISBT and other terminologies see the ISBT and dbRBC web sites [7,8].

Symbols for all human genes are provided by the Human Genome Organisation (HUGO) Gene Nomenclature Committee (HGNC) [9]. These often differ from the ISBT symbols, as the HGNC symbols reflect the function of the gene product (Table 1.1). When referring to alleles defining blood group antigens, the ISBT gene symbol is preferred because the HGNC symbols often change with changes in the perceived functions of the gene product.

1.2.3 Blood group systems

A blood group system consists of one or more antigens, governed by a single gene or by a complex of two or more very closely linked homologous genes with virtually no recombination occurring between them. Each system is genetically discrete from every other blood group system. All of the genes representing blood group systems have been identified and sequenced.

In some systems the gene directly encodes the blood group determinant, whereas in others, where the antigen is carbohydrate in nature, the gene encodes a transferase enzyme that catalyses biosynthesis of the antigen. A, B, and H antigens, for example, may all be located on the same macromolecule, yet H-glycosyltransferase is produced by a gene on chromosome 19 while the A- and B-transferases, which require H antigen as an acceptor substrate, are products of a gene on chromosome 9. Hence H belongs to a separate blood group system from A and B (Chapter 2). Regulator genes may affect expression of antigens from more than one system: *In(Lu)* down-regulates expression of antigens from both Lutheran and P systems (Chapter 6); mutations in *RHAG* are responsible for Rh_{null} phenotype, but may also cause absence of U (MNS5) and Fy5 antigens (Chapter 5). So absence of an antigen from cells of a null-phenotype is never sufficient evidence for allocation to a system. Four systems consist of more than one gene locus: MNS has three loci; Rh, Xg, and Chido/Rodgers have two each.

Table 1.3 Blood group collections.

No.	Name	Symbol	No. of antigens	Chapter
205	Cost	COST	2	20
207	Ii	I	1	25
208	Er	ER	3	28
209		GLOB	2	4
210	(Le^{c} & Le^{d})		2	2
212	Vel	VEL	2	30
213	MNCHO	MNCHO	6	3

1.2.4 Collections

Collections were introduced into the terminology in 1988 to bring together genetically, biochemically, or serologically related sets of antigens that could not, at that time, achieve system status, usually because the gene identity was not known. Thirteen collections have been created, six of which have subsequently been declared obsolete (Table 1.3): the Gerbich (201), Cromer (202), and Indian (203) collections have now become systems; Auberger (204), Gregory (206), and Wright (211) have been incorporated into the Lutheran, Dombrock, and Diego systems, respectively.

1.2.5 Low frequency antigens, the 700 series

Red cell antigens that do not fit into any system or collection and have an incidence of less than 1% in most populations tested are given a 700 number (see Table 29.1). The 700 series currently consists of 18 antigens. Thirty-six 700 numbers are now obsolete as the corresponding antigens have found homes in systems or can no longer be defined owing to lack of reagents.

1.2.6 High frequency antigens, the 901 series

Originally antigens with a frequency greater than 99% were placed in a holding file called the 900 series, equivalent to the 700 series for low frequency antigens. With the establishment of the collections, so many of these 900 numbers became obsolete that the whole series was abandoned and the remaining high frequency antigens were relocated in a new series, the 901 series, which now contains six antigens (see Table 30.1). The 901 series antigen Jr^a and Lan became systems 32 and 33 in 2012 when their genes were identified (Chapter 27).

1.2.7 Blood group terminology used in this book

The ISBT terminology provides a uniform nomenclature for blood groups that can be continuously updated and is suitable for storage of information on computer databases. The Terminology Working Party does not expect, or even desire, that the numerical terminology be used in all circumstances, although it is important that it should be understood so that the genetically based classification is understood. In this book, the alternative, 'popular' nomenclature, recommended by the Working Party [5], will generally be used. This does not reflect a lack of confidence in the numerical terminology, but is simply because most readers will not be well acquainted with blood group numbers and will find the contents of the book easier to digest if familiar names are used. The numerical terminology will be provided throughout the book in tables and often, in parentheses, in the text.

The order of the chapters of this book is based on the order of the blood group systems, collections, and series. There are, however, a few exceptions, the most notable of which are the ABO, H, and Lewis systems, which appear together in one mega-chapter (Chapter 2), because they are so closely related, biochemically.

1.3 Chromosomal location of blood group genes

Blood groups have played an important role as human gene markers. In 1951, when the Lutheran locus was shown to be genetically linked to the locus controlling ABH secretion, blood groups were involved in the first recognised human autosomal linkage and, consequently, the first demonstration of recombination resulting from crossing-over in humans [10,11]. When, in 1968, the Duffy blood group locus was shown to be linked to an inherited visible deformity of chromosome 1, it became the first human gene locus assigned to an autosome [12]. Since all blood group system genes have now been sequenced, all have been assigned to a chromosome (Table 1.1, Figure 1.1).

1.4 DNA analysis for blood group testing

Since the discovery of blood groups in 1900, most blood group testing has been carried out by serological means. With the application of gene cloning and sequencing of blood group genes at the end of the twentieth century, however, it became possible to predict blood group phenotypes from the DNA sequence. The molecular bases for almost all of the clinically significant blood group polymorphisms have been determined, so it is possible to carry out blood grouping by DNA analysis with a high degree of accuracy.

There are three main reasons for using molecular methods, rather than serological methods, for red cell blood grouping:

1 when we need to know a blood group phenotype, but do not have a suitable red cell sample;

2 when molecular testing will provide more or better information than serological testing; and

3 when molecular testing is more efficient or more cost effective than serological testing.

1.4.1 Clinical applications of molecular blood grouping

A very important application is determination of fetal blood group in order to assess the risk of HDFN. This is a non-invasive procedure carried out on cell-free fetal DNA in the maternal plasma, which represents 3–6% of the cell-free DNA in the plasma of a pregnant woman [13]. This technology is most commonly applied to RhD typing (Section 5.7), but also to Rh C, c, and E, and K of the Kell system.

Molecular methods are routinely used for extended blood group typing (beyond ABO and RhD) on multiply transfused patients, where serological methods are unsatisfactory because of the presence of transfused red cells. These patients are usually transfusion dependent and

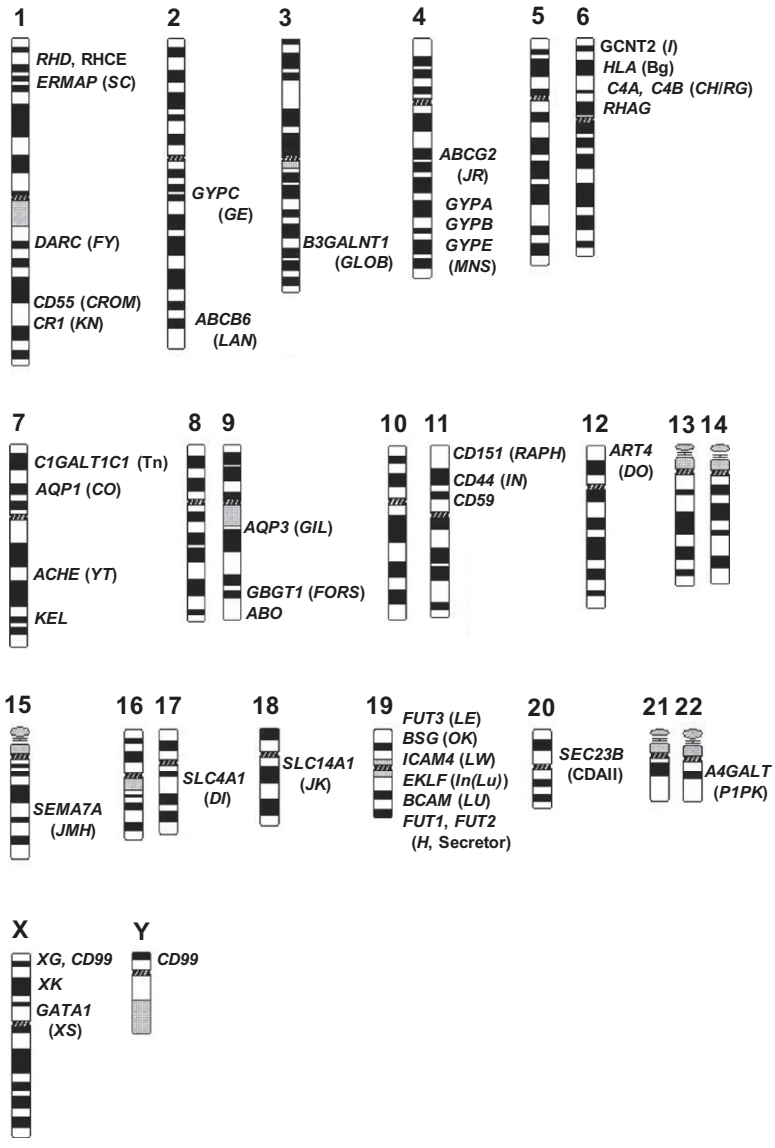


Figure 1.1 Human male chromosomes, showing location of blood group and related genes.

knowledge of their blood groups means that matched blood can be provided in an attempt to save them from making multiple antibodies and, if the patient is already immunised, to facilitate antibody identification. Molecular methods can be used for determining blood group phenotypes on red cells that are DAT-positive (i.e. coated with immunoglobulin), which makes serological testing difficult. This is particularly useful in helping to identify underlying alloantibodies in patients with autoimmune haemolytic anaemia (AIHA).

There are numerous variants of D. Some result in loss of D epitopes and some in reduced expression of D; most probably involve both (Section 5.6). Individuals with some of these variant D antigens can make a form of alloanti-D that detects those epitopes lacking from their own red cells. In many cases D variants cannot be

distinguished by serological methods, so molecular methods are often used for their identification. This assists in the selection of the most appropriate red cells for transfusion in order to avoid immunisation whilst conserving D-negative blood. There are some rare D antigens, such as DEL, that are not detected by routine serological methods. Consequently, blood donors with these phenotypes would be labelled as D-negative, although evidence exists that transfusion of DEL red cells can immunise a D-negative recipient to make anti-D. As DEL and other very weak forms of D are associated with the presence of a mutated *RHD* gene, they can be detected by molecular methods. In some transfusion services all D-negative donors are tested for the presence of *RHD*, although this is still not generally considered necessary (Section 5.6.9).

Molecular tests can be used for screening for donors when serological reagents are of poor quality or in short supply. For example, anti-Do^{a} and -Do^{b} have the potential to be haemolytic, yet satisfactory reagents are not available for finding donors for a patient with one of these antibodies (Chapter 14). Some Rh variants, such as hr^{B}-negative and hr^{S}-negative, are relatively common in people of African origin but are difficult to detect serologically (Section 5.9.5). Molecular tests are often employed to assist in finding suitable blood for patients with sickle cell disease, to reduce alloimmunisation and the risks of delayed HTRs [14,15].

Molecular methods are extremely useful in the blood group reference laboratory for helping to solve serological difficult problems.

In most countries, all blood donors are tested for ABO and D, but often a proportion of the donors are also tested for additional blood group antigens, especially C, c, E, e, and K, but sometimes also C^{w}, M, S, s, Fy^{a}, Fy^{b}, Jk^{a}, and Jk^{b}. This testing is usually performed by automated serological methods, but it is likely that in the future these serological methods will be replaced by molecular methods [16–18]. Molecular typing for this purpose has already been introduced in some services [19,20]. Molecular methods are more accurate than serological methods, they are more suited to high-throughput methods, and they are either cheaper or are likely to become so in the near future. This provides justification for a switch of technologies.

1.4.2 Current and future technologies

Laboratories performing blood group testing on cell-free fetal DNA in the maternal plasma generally use real-time quantitative PCR with Taqman technology, but an alternative technology that is becoming available involves the application of matrix-assisted laser desorption/ionisation time-of-flight (MALDI TOF) mass spectrometry [21].

For other applications of molecular blood grouping, many laboratories use methods traditionally applied to single nucleotide polymorphism (SNP) testing, involving PCR with the application of restriction enzymes or PCR with allele-specific primers, followed by gel electrophoresis. Other technologies that are becoming more commonly used involve the application of allele-specific extension of primers tagged with single fluorescent nucleotides, pyrosequencing, DNA microarray technology, on chips or coloured beads coated with oligonucleotides, and MALDI TOF [18,22]. The future of molecular blood grouping and of molecular diagnostics probably lies with next generation (massively parallel) sequencing, which will be truly high-throughput [23,24]. Next generation sequencing is an extremely powerful technology that provides the capacity to sequence many regions of the genome in numerous different individuals in one run, including fetal DNA from maternal plasma [25].

1.5 Structures and functions of blood group antigens

For the half-century following Landsteiner's discovery, human blood groups were understood predominantly as patterns of inherited serological reactions. From the 1950s some structural information was obtained through biochemical analyses, firstly of the carbohydrate antigens and then of the proteins. In 1986, *GYPA*, the gene encoding the MN antigens, was cloned and this led into the molecular genetic era of blood groups. A great deal is now known about the structures of many blood group antigens, yet remarkably little is known about their functions and most of what we do know has been deduced from their structures. Functional aspects of blood group antigens are included in the appropriate chapters of this book; provided here is a synopsis of the relationship between their structures and putative functions. The subject is reviewed in [26] and computer modelling of blood group proteins, which gives detailed information about protein structure, is reviewed in [27].

1.5.1 Membrane transporters

Membrane transporters facilitate the transfer of biologically important molecules in and out of the cell. In the

red cell they are polytopic, crossing the membrane several times, with cytoplasmic N- and C-termini, and are *N*-glycosylated on one of the external loops. Band 3, the Diego blood group antigen (Chapter 10) is an anion exchanger, the Kidd glycoprotein (Chapter 9) is a urea transporter, the Colton glycoprotein is a water channel (Chapter 15), the Gill glycoprotein is a water and glycerol channel (Chapter 26), and the Lan and Junior glycoproteins are ATP-fuelled transporters of porphyrin and uric acid (Chapter 27). Band 3 is at the core of a membrane macrocomplex, which contains the Rh proteins and the Rh-associated glycoprotein, which probably function as a CO_2 channel (Chapters 5 and 10).

1.5.2 Receptors and adhesion molecules

The Duffy glycoprotein is polytopic, but has an extracellular N-terminus. It is a member of the G protein-coupled superfamily of receptors and functions as a receptor for chemokines (Chapter 8).

The glycoproteins carrying the antigens of the Lutheran (Chapter 6), LW (Chapter 16), Scianna (Chapter 13), and Ok (Chapter 22) systems are members of the immunoglobulin superfamily (IgSF). The IgSF is a large family of receptors and adhesion molecules with extracellular domains containing different numbers of repeating domains with sequence homology to immunoglobulin domains. The functions of these structures on red cells are not known, but there is evidence to suggest that the primary functional activities of the Lutheran and LW glycoproteins occur during erythropoiesis, with LW probably playing a role in stabilising the erythropoietic islands.

The Indian antigen (CD44), a member of the link module superfamily, functions as an adhesion molecule in many tissues, but its erythroid function is unknown (Chapter 21). The glycoproteins of the Xg (Chapter 12) and JMH (Chapter 24) systems also have structures that suggest they could function as receptors and adhesion molecules. The Raph antigen, a tetraspanin, may associate with integrin in red cell progenitors to generate complexes that bind the extracellular matrix (Chapter 23).

1.5.3 Complement regulatory glycoproteins

Red cells have at least three glycoproteins that function to protect the cell from destruction by autologous complement. The Cromer glycoprotein, decay-accelerating factor (Chapter 19), and the Knops glycoprotein, complement receptor-1 (CR1) (Chapter 20), belong to the complement control protein superfamily; CD59 is not polymorphic and does not have blood group activity (Chapter 19). The major function of red cell CR1 is to bind and process C3b/C4b coated immune complexes and to transport them to the liver and spleen for removal from the circulation.

1.5.4 Enzymes

Two blood group glycoproteins have enzymatic activity. The Yt glycoprotein is acetylcholinesterase, a vital enzyme in neurotransmission (Chapter 11), and the Kell glycoprotein is an endopeptidase that can cleave a biologically inactive peptide to produce the active vasoconstrictor, endothelin (Chapter 7). The red cell function for both of these enzymes is unknown. The Dombrock glycoprotein belongs to a family of ADP-ribosyltransferases, but there is no evidence that it is an active enzyme (Chapter 14).

1.5.5 Structural components

The shape and integrity of the red cell is maintained by the cytoskeleton, a network of glycoproteins beneath the plasma membrane. At least two blood group glycoproteins anchor the membrane to its skeleton: band 3, the Diego antigen (Chapter 10), and glycophorin C and its isoform glycophorin D, the Gerbich blood group antigens (Chapter 18). Mutations in the genes encoding these proteins can result in abnormally shaped red cells. In addition, there is evidence that glycoproteins of the Lutheran (Chapter 6), Kx (Chapter 7), and RHAG (Chapter 5) systems interact with the cytoskeleton and their absence is associated with some degree of abnormal red cell morphology.

1.5.6 Components of the glycocalyx

Glycophorin A, the MN antigen (Chapter 3), band 3 are the two most abundant glycoproteins of the red cell surface. The *N*-glycans of band 3, together with those of the glucose transporter, provide the majority of red cell ABH antigens, which are also expressed on other glycoproteins and on glycolipids (Chapter 2). The extracellular domains of glycophorin A and other glycophorin molecules are heavily *O*-glycosylated. Carbohydrate at the red cell surface constitutes the glycocalyx, or cell coat, an extracellular matrix of carbohydrate that protects the cell from mechanical damage and microbial attack.

1.5.7 What is the biological significance of blood group polymorphism?

Very little is known about the biological significance of the polymorphisms that make blood groups alloantigenic. In any polymorphism one of the alleles is likely to

have, or at least to have had in the past, a selective advantage in order to achieve a significant frequency in a large population, though genetic drift and founder effects may also have played a part [28]. Glycoproteins and glycolipids carrying blood group activity are often exploited by pathogenic micro-organisms as receptors for attachment to the cells and subsequent invasion; surviving malaria possibly being the most significant force affecting blood group expression. In some cases, however, selection may have nothing to do with red cells; the target for the parasite could be other cells that carry the protein. It is likely that most blood group polymorphism is a relic of the selective balances that can result from mutations making cell surface structures less suitable as pathogen receptors and resultant adaptation of the parasite in response to these selective pressures. It is important to remember that whilst blood group polymorphism undoubtedly arose from the effects of selective pressures, these factors may have disappeared long ago, so that little hope remains of ever identifying them. To quote Darwin (*The Origin of Species*, 1859), 'The chief part of the organisation of any living creature is due to inheritance; and consequently, though each being assuredly is well fitted for its place in nature, many structures have now no very close and direct relations to present habits of life'.

References

1 Landsteiner K. Zur Kenntnis der antifermentativen, lytischen und agglutinietenden Wirkungen des Blutserums und der Lymphe. *Zbl Bakt* 1900;27:357–366.

2 Landsteiner K. Über Agglutinationserscheinungen normalen menschlichen Blutes. *Wien Klein Wochenschr* 1901;14:1132–1134.

3 Coombs RRA, Mourant AE, Race RR. Detection of weak and 'incomplete' Rh agglutinins: a new test. *Lancet* 1945; ii:15.

4 Coombs RRA, Mourant AE, Race RR. A new test for detection of weak and 'incomplete' Rh agglutinins. *Br J Exp Path* 1945;26:255–266.

5 Daniels GL and members of the Committee on Terminology for Red Cell Surface Antigens. Blood group terminology 2004. *Vox Sang* 2004;87:304–316.

6 Storry JR and members of the ISBT Working Party on red cell immunogenetics and blood group terminology: Berlin report. *Vox Sang* 2011;101:77–82.

7 The International Society of Blood Transfusion Red Cell Immunogenetics and Blood Group Terminology Working Party. http://www.isbtweb.org/working-parties/red-cell-immunogenetics-and-terminology (last accessed 5 October 2012).

8 Blood Group Antigen Gene Mutation Database (dbRBC). http://www.ncbi.nlm.nih.gov/projects/gv/mhc/xslcgi.cgi?cmd=bgmut/home (last accessed 5 October 2012).

9 HUGO Gene Nomenclature Committee. http://www.genenames.org (last accessed 5 October 2012).

10 Mohr J. A search for linkage between the Lutheran blood group and other hereditary characters. *Acta Path Microbiol Scand* 1951;28:207–210.

11 Mohr J. Estimation of linkage between the Lutheran and the Lewis blood groups. *Acta Path Microbiol Scand* 1951;29:339–344.

12 Donahue RP, Bias WB, Renwick JH, McKusick VA. Probable assignment of the Duffy blood group locus to chromosome 1 in man. *Proc Natl Acad Sci USA* 1968;61:949–955.

13 Daniels G, Finning K, Martin P, Massey E. Non-invasive prenatal diagnosis of fetal blood group phenotypes: current practice and future prospects. *Prenat Diagn* 2009;29:101–107.

14 Pham B-N, Peyrard T, Juszczak G, *et al.* Analysis of RhCE variants among 806 individuals in France: consideration for transfusion safety, with emphasis on patients with sickle cell disease. *Transfusion* 2011;51:1249–1260.

15 Wilkinson K, Harris S, Gaur P, *et al.* Molecular typing augments serologic testing and allows for enhanced matching of red blood cell for transfusion in patients with sickle cell disease. *Transfusion* 2012;52:381–388.

16 Avent ND. Large-scale blood group genotyping: clinical implications. *Br J Haematol* 2008;144:3–13.

17 Anstee DJ. Red cell genotyping and the future of pretransfusion testing. *Blood* 2009;114:248–256.

18 Veldhuisen B, van der Schoot CE, de Haas M. Blood group genotyping: from patient to high-throughput donor screening. *Vox Sang* 2009;97:198–206.

19 Perreault J, Lavoie J, Painchaud P, *et al.* Set-up and routine use of a database of 10 555 genotyped blood donors to facilitate the screening of compatible blood components for alloimmunized patients. *Vox Sang* 2009;87:61–68.

20 Jungbauer C, Hobel CM, Schwartz DWM, Mayr WR. High-throughput multiplex PCR genotyping for 35 red blood cell antigens in blood donors. *Vox Sang* 2011;102:234–242.

21 Bombard AT, Akolekar R, Farkas DH, *et al.* Fetal RHD genotype detection from circulating cell-free fetal DNA in maternal plasma in non-sensitised RhD negative women. *Prenat Diagn* 2011;31:802–808.

22 Monteiro F, Tavares G, Ferreira M, *et al.* Technologies involved in molecular blood group genotyping. *ISBT Sci Ser* 2011;6:1–6.

23 ten Bosch JR, Grody WW. Keeping up with the next generation. Massively parallel sequencing in clinical diagnosis. *J Molec Diagn* 2008;10:484–492.

24 Su Z, Ning B, Fang H, *et al.* Next-generation sequencing and its applications in molecular diagnosis. *Expert Rev Mol Diagn* 2011;11:333–343.
25 Liao GJW, Lun FMF, Zheng YWL, *et al.* Targeted massively parallel sequencing of maternal plasma DNA permits efficient and unbiased detection of fetal alleles. *Clin Chem* 2011;57:92–101.
26 Daniels G. Functions of red cell surface proteins. *Vox Sang* 2007;93:331–340.
27 Burton NM, Daniels G. Structural modelling of red cell surface proteins. *Vox Sang* 2011;100:129–139.
28 Anstee DJ. The relationship between blood groups and disease. *Blood* 2010;115:4635–464

2 ABO, H, and Lewis Systems

Part 1: History and introduction

Described in this chapter are three blood group systems, ABO, H, and Lewis (Table 2.1), although Lewis is really an 'adopted' blood group system because the antigens are not intrinsic to the red cells, but introduced into the membrane from the plasma. These three systems are genetically discrete, but are discussed in the same chapter because they are phenotypically and biochemically closely related. A complex interaction of genes at several loci controls the expression of ABO, H, Lewis, and other related antigens on red cells and in secretions.

The science of immunohaematology came into existence in 1900 when Landsteiner [1] reported that, 'The serum of healthy humans not only has an agglutinating effect on animal blood corpuscles, but also on human blood corpuscles from different individuals'. The following year Landsteiner [2] showed that by mixing together sera and red cells from different people three groups, A, B, and C (later called O), could be recognised. In group A, the serum agglutinated group B, but not A or C cells; in group B, the serum agglutinated A, but not B or C cells; and in group C (O), the cells were not agglutinated by any serum, and the serum appeared to contain a mixture of two agglutinins capable of agglutinating A and B cells. Decastello and Stürli [3] added a fourth group (AB), in which the cells are agglutinated by sera of all other groups and the serum contains neither agglutinin. Healthy adults always have A or B agglutinins in their serum if they lack the corresponding agglutinogen from their red cells (Table 2.2).

Epstein and Ottenberg [4] suggested that blood groups may be inherited and in 1910 von Dungern and Hirschfeld [5] confirmed that the inheritance of the A and B antigens obeyed Mendel's laws, with the presence of A or B being dominant over their absence. Bernstein [6,7] showed that only three alleles at one locus were necessary to explain ABO inheritance (Table 2.2).

Some group A people produce an antibody that agglutinates the red cells of most other A individuals. Thus A was subdivided into A_1 and A_2, and the three allele theory of Bernstein was extended to four alleles: A^1, A^2, *B* and *O* [8] (Section 2.4). Many rare subgroups of A and B have now been identified (Sections 2.7 and 2.8).

The structure and biosynthesis of the ABO, H, and Lewis antigens is well understood, thanks mainly to

Human Blood Groups, Third Edition. Geoff Daniels.

Table 2.1 Numerical notation for the ABO, Lewis, and H systems, and for Le^c and Le^d.

ABO (system 001)		Lewis (system 007)		H (system 018)		Collection 210	
ABO1	A	LE1	Le^a	H1	H	210001	Le^c
ABO2	B	LE2	Le^b			210002	Le^d
ABO3	A,B	LE3	Le^{ab}				
ABO4	A1	LE4	Le^{bH}				
		LE5	Ale^b				
		LE6	BLe^b				

Obsolete: ABO5, previously H.

Table 2.2 The ABO system at its simplest level.

ABO group	Antigens on red cells	Antibodies in serum	Genotype
O	None	Anti-A,B	*O/O*
A	A	Anti-B	*A/A* or *A/O*
B	B	Anti-A	*B/B* or *B/O*
AB	A and B	None	*A/B*

the pioneering work in the 1950s of Morgan and Watkins [9,10] and of Kabat [11]. A and B red cell antigens are carbohydrate determinants of glycoproteins and glycolipids and are distinguished by the nature of an immunodominant terminal monosaccharide: *N*-acetylgalactosamine (GalNAc) in group A and galactose (Gal) in group B. The *A* and *B* genes encode glycosyltransferases that catalyse the transfer of the appropriate immunodominant sugar from a nucleotide donor to an acceptor substrate, the H antigen. The *O* allele produces no active transferase (Sections 2.2 and 2.3). The sequences of the *A* and *B* alleles demonstrate that A- and B-glycosyltransferases (GTA and GTB) differ by four amino acid residues; the most common *O* allele contains a nucleotide deletion and encodes a truncated protein.

There are a multitude of *ABO* alleles, many of which affect phenotype, and at least two different terminologies. In this chapter the original terminology (e.g. A^1, A^2, O^1) will be used, with the dbRBC terminology often provided in parentheses.

H antigen is synthesised by a fucosyltransferase produced by *FUT1*, a gene independent of *ABO*. Very rare individuals lacking *FUT1* have no H antigen on their red cells and, consequently, are unable to produce A or B antigens, even when the enzyme products of the *A* or *B* genes are present (Section 2.12).

H antigen is present in body secretions of about 80% of Caucasians. The presence of H in secretions is governed by *FUT2*, another fucosyltransferase that is closely linked to *FUT1*. Individuals who secrete H also secrete A or B antigens if they have the appropriate *ABO* alleles. Non-secretors of H secrete neither A nor B, even when those antigens are expressed on their red cells (Section 2.6).

The first two examples of anti-Lewis, later to be called anti-Le^a, were described by Mourant [12] in 1946. These antibodies agglutinated the red cells of about 25% of English people. Andresen [13] found an antibody, later to become anti-Le^b, that defined a determinant only present on Le(a–) cells of adults. Six percent of group O adults lacked both antigens. Although Le^a and Le^b are not synthesised by red cells, but are acquired from the plasma, they are considered blood group antigens because they were first recognised on red cells. The terminology Le^a and Le^b is misleading as these antigens are not the products of alleles.

The Lewis gene (*FUT3*) encodes a fucosyltransferase that catalyses the addition of a fucose residue to H antigen in secretions to produce Le^b antigen or, if no H is present (non-secretors), to the precursor of H to produce Le^a. Consequently, as these structures are acquired from the plasma by the red cell membrane, red cells of most H secretors are Le(a–b+) and those of most H non-secretors are Le(a+b–). The Lewis-transferase can also convert A to ALe^b and B to BLe^b. About 6% of white people and 25% of black people are homozygous for a silent gene at the

FUT3 locus and, as they do not produce the Lewis enzyme, have Le(a–b–) red cells and lack Lewis substances in their secretions (Sections 2.3 and 2.15). In East Asia the red cell phenotype Le(a+b+) is common, caused by a weak secretor allele (Section 2.6.3).

The antigens Le^c and Le^d represent precursors of the Lewis antigens and are present in increased quantity in the plasma of Le(a–b–) individuals. Le^c is detected on the red cells of Le(a–b–) non-secretors of H and Le^d is detected on the red cells of Le(a–b–) secretors of H. Le^x and Le^y antigens, isomers of Le^a and Le^b, are not present in substantial quantities on red cells (Section 2.18.2).

ABH and Lewis antigens are often referred to as histo-blood group antigens [14] because they are ubiquitous structures occurring on the surface of endothelial cells and most epithelial cells. The precise nature of the histo-blood group antigens expressed varies between tissues within the same individual because of the intricacy of the gene interactions involved (Section 2.19).

ABO is on chromosome 9; *FUT1*, *FUT2*, and *FUT3* are on chromosome 19 (Sections 2.3.1, 2.3.2.4, and 2.3.5).

Part 2: Biochemistry, inheritance, and biosynthesis of the ABH and Lewis antigens

2.2 Structure of ABH, Lewis, and related antigens

ABH and Lewis antigens are carbohydrate structures. These oligosaccharide chains are generally conjugated with polypeptides to form glycoproteins or with ceramide to form glycosphingolipids. Oligosaccharides are synthesised in a stepwise fashion, the addition of each monosaccharide being catalysed by a specific glycosyltransferase. The oligosaccharide moieties responsible for expression of ABH, Lewis, and related antigens are shown in Table 2.3 and abbreviations for monosaccharides are given in Table 2.4. The biosynthesis of these structures is described in Section 2.3 and represented diagrammatically in Figure 2.1. There is a vast literature on the biochemistry of these blood group antigens and only some of the relevant references can be given in this chapter. The following reviews are recommended: [10,14–27].

2.2.1 Glycoconjugates expressing ABH and Lewis antigens

Two major classes of carbohydrate chains on glycoproteins express ABH antigens:

1 *N*-glycans, highly branched structures attached to the amide nitrogen of asparagine through GlcNAc; and
2 *O*-glycans, simple or complex structures attached to the hydroxyl oxygen of serine or threonine through GalNAc.

Glycosphingolipids consist of carbohydrate chains attached to ceramide. They are classified as lacto-series, globo-series, or ganglio-series according to the nature of the carbohydrate chain. Glycosphingolipid-borne ABH and Lewis antigens are present predominantly on glycolipids of the lacto-series, although ABH antigens have also been detected on globo-series and ganglio-series glycolipids. The carbohydrate chains of most ABH-bearing glycoproteins and of lacto-series glycolipids are based on a poly-*N*-acetyllactosamine structure; that is, they are extended by repeating Galβ1→4GlcNAcβ1→3 disaccharides (see Table 2.5 for examples).

On red cells, most ABH antigens are on the single, highly branched, poly-*N*-acetyllactosaminyl *N*-glycans of the anion exchange protein, band 3, and the glucose transport protein, band 4.5 [28]. There are about 1 million monomers of band 3 protein and half a million monomers of band 4.5 protein per red cell [29]. The other major red cell glycoprotein, glycophorin A, carries very low levels of ABH activity on both *O*- and *N*-glycans (Sections 3.2.1 and 3.2.2) and ABH determinants have also been detected on the Rh-associated glycoprotein [30]. Lewis antigens on red cells are not expressed on glycoproteins; they are not intrinsic to red cells, but are acquired from the plasma.

Glycolipids play a minor role in red cell ABH expression compared with glycoproteins. Red cell glycosphingolipids of the poly-*N*-acetyllactosaminyl type that express ABH antigens may have relatively simple linear or branched carbohydrate chains [15] (Table 2.5) or may be highly complex, branched structures called polyglycosylceramides, with up to 60 carbohydrate residues per molecule [31].

All the early work establishing the structures of the ABH and Lewis determinants was carried out on body secretions, especially the pathological fluid from human ovarian cysts, an abundant source of soluble A, B, and H substances [32]. ABH and Lewis antigens in secretions are glycoproteins; oligosaccharide chains attached to mucin by *O*-glycosidic linkage to serine or threonine (for reviews see [9,10]). These macromolecules have molecular weights varying from 2×10^5 to several millions. In milk and urine, free oligosaccharides with ABH and Lewis activity are also found [33,34]. ABH and Lewis determinants are present in plasma on glycosphingolipids, some

Table 2.3 Structures of A, B, H, Lewis, and related antigens (for abbreviations see Table 2.4).

Type 1		*Type 2*	
Precursor (Le^c)	Galβ1→3GlcNAcβ1→R †	**Precursor**	Galβ1→4GlcNAcβ1→R
H (Le^d)	Galβ1→3GlcNAcβ1→R † 2 ↑ Fucα1	**H** (CD173)	Galβ1→4GlcNAcβ1→R * 2 ↑ Fucα1
A	GalNAcα1→3Galβ1→3GlcNAcβ1→R † 2 ↑ Fucα1	**A**	GalNAcα1→3Galβ1→4GlcNAcβ1→R * 2 ↑ Fucα1
B	Galα1→3Galβ1→3GlcNAcβ1→R † 2 ↑ Fucα1	**B**	Galα1→3Galβ1→4GlcNAcβ1→R * 2 ↑ Fucα1
Le^a	Galβ1→3GlcNAcβ1→R † 4 ↑ Fucα1	**Le^x**	Galβ1→4GlcNAcβ1→R 3 ↑ Fucα1
Le^b	Galβ1→3GlcNAcβ1→R † 2 4 ↑ ↑ Fucα1 Fucα1	**Le^y**	Galβ1→4GlcNAcβ1→R 2 3 ↑ ↑ Fucα1 Fucα1
ALe^b	GalNAcα1→3Galβ1→3GlcNAcβ1→R † 2 4 ↑ ↑ Fucα1 Fucα1	**ALe^y**	GalNAcα1→3Galβ1→4GlcNAcβ1→R 2 3 ↑ ↑ Fucα1 Fucα1
BLe^b	Galα1→3Galβ1→3GlcNAcβ1→R † 2 4 ↑ ↑ Fucα1 Fucα1	**BLe^y**	Galα1→3Galβ1→4GlcNAcβ1→R 2 3 ↑ ↑ Fucα1 Fucα1
sialyl-Le^a	Galβ1→3GlcNAcβ1→R 3 4 ↑ ↑ NeuAcα2 Fucα1	**sialyl-Le^x**	Galβ1→4GlcNAcβ1→R 3 3 ↑ ↑ NeuAcα2 Fucα1

of which may become incorporated into the red cell membrane (Section 2.15.4).

2.2.2 Carbohydrate determinants

Expression of H, A, and B antigens is dependent on the presence of specific monosaccharides attached to various precursor disaccharides at the non-reducing end of a carbohydrate chain. There are at least five precursor disaccharides, also called peripheral core structures (reviewed in [14,18,21,23]):

Type 1	Galβ1→3GlcNAcβ1→R
Type 2	Galβ1→4GlcNAcβ1→R
Type 3	Galβ1→3GalNAcα1→R
Type 4	Galβ1→3GalNAcβ1→R
Type 6	Galβ1→4Glcβ1→R.

(Type 5 has only been chemically synthesised.)

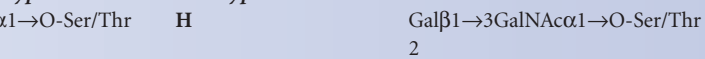
H-active structures have Fuc α-linked to C-2 of the terminal Gal [35,36]; A- and B-active structures have GalNAc and Gal, respectively, attached in α-linkage to

Table 2.3 (*Continued*)

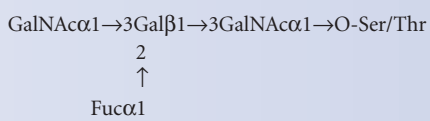
Type 3: O-linked mucin type

Precursor (T antigen) Galβ1→3GalNAcα1→O-Ser/Thr

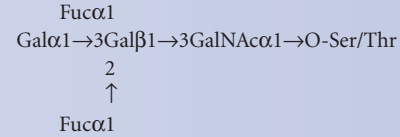
H



A

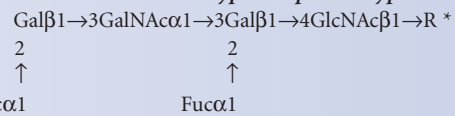


B

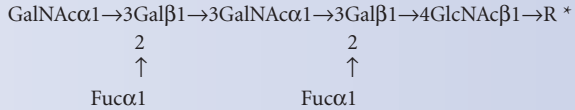


Type 3: repetitive type

H

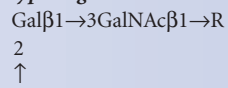


A

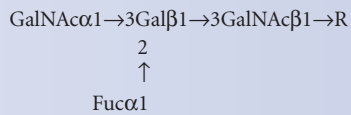


Type 4: globo-series

Globo-H



Globo-A



*Intrinsic to red cells and detected in significant quantity on red cells of individuals of appropriate genotype.
†Adsorbed onto red cells from plasma in individuals of appropriate genotype.

Table 2.4 Some abbreviations for monosaccharides and the structures they are linked to.

Gal	D-galactose	Cer	Ceramide
GalNAc	*N*-acetyl-D-galactosamine	Asp	Asparagine
GlcNAc	*N*-acetyl-D-glucosamine	Ser	Serine
Fuc	L-Fucose	Thr	Threonine
NeuAc	Sialic acid (*N*-acetylneuraminic acid)		
Man	Mannose		
Glc	Glucose	R	Remainder of molecule

C-3 of this α1→2 fucosylated Gal residue (Table 2.3). Although Fuc does not represent the whole H determinant, it is the H immunodominant sugar because its loss results in loss of H activity. Likewise GalNAc and Gal are the A and B immunodominant sugars, respectively.

Le^a and Le^b antigens are expressed when Fuc is attached to the GlcNAc residue of the Type 1 precursor and Type 1 H, respectively [37–40]. Le^x and Le^y are the Type 2 isomers of Le^a and Le^b [36,39,41,42]. Fuc is linked α1→4 to the GlcNAc residue of a Type 1 chain in Le^a and Le^b and α1→3 to the GlcNAc of a Type 2 chain in Le^x and Le^y. Le^x and Le^y are not present in significant quantities on red cells [43]. The monofucosylated Le^a and Le^x

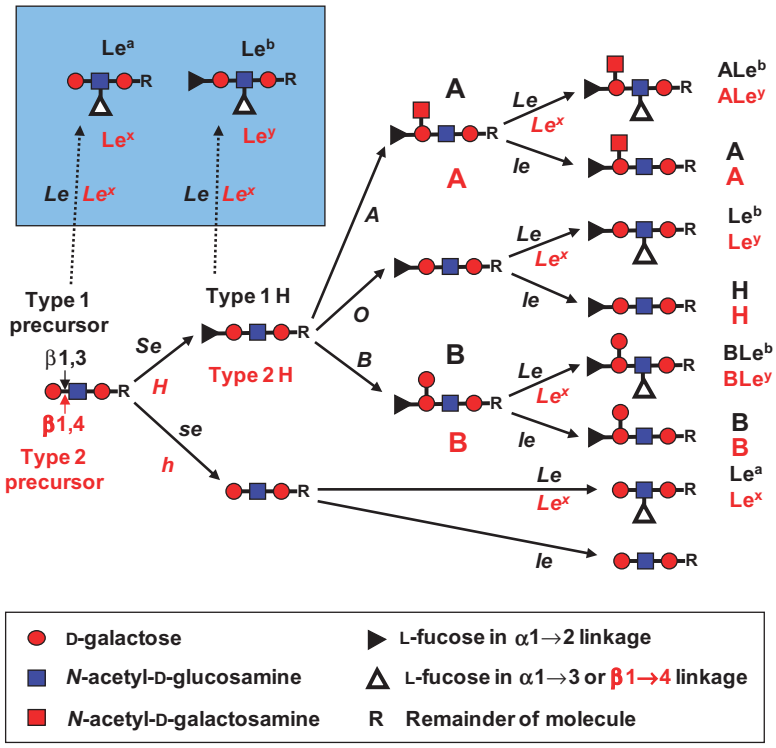


Figure 2.1 Diagram representing the biosynthetic pathways of ABH, Lewis, Le^x, and Le^y antigens derived from Type 1 and Type 2 core chains. Genes controlling steps in the pathway are shown in *italics* and the gene products are listed in Table 2.6. Type 1 and Type 2 precursors differ in the nature of the linkage between the non-reducing terminal Gal and GlcNAc: β1→3 in Type 1 and β1→4 in Type 2. Type 1 and Type 2 structures and the genes acting on them are shown in black and red, respectively. Dashed lines show how Le^a (Le^x) and Le^b (Le^y), produced from the precursor and H structures respectively, are not substrates for the H, Se, or ABO transferases and remain unconverted.

structures may be sialylated at the C-3 of Gal [44–46] (Table 2.3).

Type 1 ABH and Lewis structures are present in secretions, plasma, and endodermally derived tissues [21]. They are not synthesised by red cells, but are incorporated into the red cell membrane from the plasma [47]. Lewis antigens (Le^a and Le^b) are only present on Type 1 structures. Elongated carbohydrate chains with Type 1 peripheral structures are generally extended by repeating poly-*N*-acetyllactosamine disaccharides with the Type 2 (β1→4) linkage [48] (Table 2.5). Extended Type 1 structures with Le^a and Le^b activity have been detected in plasma, particularly in persons with Le(a+b+) red cells [49,50].

Antigens on Type 2 chains represent the major ABH-active oligosaccharides on red cells and are also detected in secretions and various ectodermally or mesodermally derived tissues [15,21]. Type 2 structures in secretions are probably more often difucosylated (Le^y, ALe^y, BLe^y) than monofucosylated (H, A, B) [51,52].

There are two forms of Type 3 ABH antigens, the *O*-linked mucin type and the repetitive A-associated type. In the *O*-linked mucin type the precursor exists as a disaccharide linked directly, by *O*-glycosidic bond, to a serine or threonine residue of mucin [53]. This precursor represents the T cryptantigen (see Section 3.17.2), but is not usually expressed because it is masked by substitution with sialic acid residues or other sugars. Type 3 ABH antigens of the *O*-linked mucin type are not found on red cells [54]. Repetitive Type 3 chains are present on red cell glycolipids and secreted mucins from group A individuals. They are restricted to group A because they are biosynthesised by the addition of Gal in β1→3 linkage to the terminal GalNAc of an A-active Type 2 chain followed by

Table 2.5 Examples of H-active glycoconjugates with Type 2 precursor chains (for abbreviations see Table 2.4).

Glycosphingolipid (simple linear)

Fucα1→2Galβ1→4GlcNAcβ1→3Galβ1→4Glcβ1→Cer

Glycosphingolipid (branched)

Fucα1→2Galβ1→4GlcNAcβ1→3(Galβ1→4GlcNAcβ1→3)$_n$Galβ1→4Glcβ1→Cer
6
↑
Fucα1→2Galβ1→4GlcNAcβ1

N-*linked glycoprotein*

Fucα1→2Galβ1→4GlcNAcβ1
↓
6
Fucα1→2Galβ1→4GlcNAcβ1→3(Galβ1→4GlcNAcβ)$_n$1→2Man
Man-GlcNAc-GlcNAc-Asn
Fucα1→2Galβ1→4GlcNAcβ1→3(Galβ1→4GlcNAcβ)$_n$1→2Man
6
↑
Fucα1→2Galβ1→4GlcNAcβ1

O-*linked glycoprotein (complex mucin type)*

Fucα1→2Galβ1→4GlcNAcβ1
↓
6
Fucα1→2Galβ1→4GlcNAcβ1→3(Galβ1→4GlcNAcβ1→3)$_n$GalNAcα1→Ser/Thr
6
↑
Fucα1→2Galβ1→4GlcNAcβ1

n, 0–5 or more.

the fucosylation of that Gal to form Type 3 H [43,54–56] (Figure 2.2). Repetitive Type 3 chains are only present on group A cells because they are produced by the addition of Gal to the terminal GalNAc of a Type 2 A chain.

Type 4 ABH structures are only located on glycolipids. Type 4 precursor chain of the globo-series results from the addition of terminal Gal to globoside [57] (P antigen, see Chapter 4). Type 4 globo-H and globo-A have been detected in small quantities on red cells [57,58], but are more abundant in kidney [59]; Type 4 globo-B has only been found, in minute quantities, in kidney [60]. Kidney from a group A person with the p phenotype, which prevents extension of the globo-series structures, lacked Type 4 A [61] (see Chapter 4).

Type 6 chains have been found as free oligosaccharides in milk and urine [33,34].

The internal carbohydrate chains express I and i antigens. In fetal cells linear chains predominate and i antigen is expressed, whereas in adult glycoproteins and glycolipids there is branching of the inner core chains and I antigen is expressed (see Chapter 25).

2.3 Biosynthesis, inheritance, and molecular genetics

The carbohydrate antigens of the ABO, H, and Lewis blood group systems are not the primary products of the genes governing their expression. Carbohydrate chains are built up by the sequential addition of monosaccharides, each extension of the chain being catalysed by a specific glycosyltransferase. These enzymes catalyse the

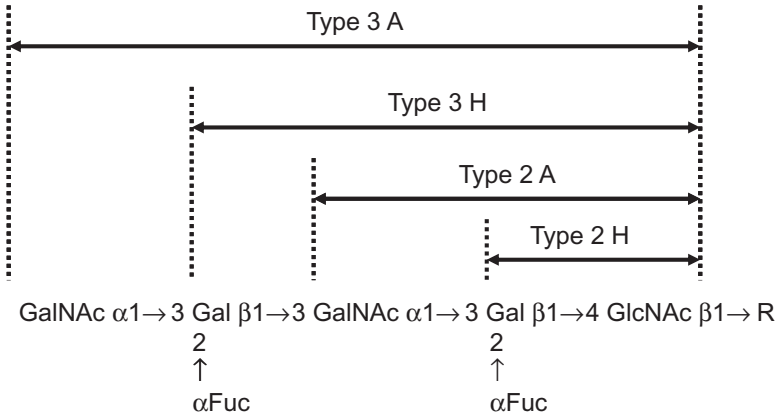


Figure 2.2 Diagram showing how a repetitive Type 3 A chain is built up from a Type 2 H chain. From right to left, Type 2 H is converted to Type 2 A in group A people. Type 2 A may be converted to Type 3 H. Type 3 H is then further converted to Type 3 A.

Table 2.6 Some ABH-related blood group genes and the glycosyltransferases they produce.

Locus	Allele	Transferase	
FUT1 (*H*)	*H*	α1,2-L-fucosyltransferase	EC 2.4.1.69
	h	None	
FUT2 (*SE*)	*Se*	α1,2-L-fucosyltransferase	EC 2.4.1.69
	se	None	
ABO	*A*	α1,3-*N*-acetyl-D-galactosaminyltransferase	EC 2.4.1.40
	B	α1,3-D-galactosyltransferase	EC 2.4.1.37
	O	None	
FUT3 (*LE*)	*Le*	α1,3/4-L-fucosyltransferase	EC 2.4.1.65
	le	None	

transfer of a monosaccharide from its nucleotide donor and its attachment, in a specific glycosidic linkage, to its acceptor substrate. Glycosyltransferases represent the primary products of the *ABO*, *FUT1* (H), *FUT2* (secretor), and *FUT3* (Lewis) genes (Table 2.6).

At least 100 glycosyltransferases are required for synthesis of the known human carbohydrates. The genes producing most of them have been identified and sequenced, including those for the ABO, H, and Lewis blood groups, and for secretion of H. The gene products are trans-membrane proteins of the Golgi apparatus. They share a common domain structure comprising a short N-terminal cytoplasmic tail, a 16–20 amino acid membrane-spanning domain, and an extended stem region followed by a large C-terminal catalytic domain. Soluble glycosyltransferases present in secretions may result from the release of membrane-bound enzymes by endogenous proteases or they may lack the membrane-spanning domain as a result of mRNA translation-initiation at an alternative site (reviewed in [62,63]).

The regulatory mechanisms required to assure that carbohydrate chains with the appropriate sequences are produced are complex. They involve the presence or absence of certain enzymes according to the genes expressed in various tissues and at different stages of development, and according to the genotype of the individual. Competition between different transferases for the same donor or acceptor substrate is also important in determining the carbohydrate chain produced (reviewed in [16]).