Vox Sanguinis (2004) **87 (Suppl. 1)**, S58–S62
© 2004 Blackwell Publishing

ORIGINAL PAPER ES05.01

The complexities of the Rh system

M. L. Scott
Bristol Institute for Transfusion Sciences, National Blood Service, UK

Short history of the Rh blood-group system

In 1939, Levine & Stetson first described the case of a mother, after giving birth to a stillborn child, having a haemolytic transfusion reaction following transfusion her husband's blood. Her serum agglutinated her husband's red cells and those of 80% of ABO-compatible donors. In 1940, Landsteiner & Wiener injected Rhesus-monkey red blood cells into rabbits. The rabbit serum agglutinated Rhesus-monkey red cells, and also 85% of human red cells. They called this antibody anti-Rh, and in 1941 Levine & Stetson's human antibody was shown to have the same pattern of reactivity as the rabbit anti-Rh antibody. However, by 1942, Fisk & Foord had demonstrated a difference between rabbit and human anti-Rh: red cells from all newborn babies, whether Rh positive or negative as defined by human anti-Rh were positive with the rabbit anti-Rh. In 1963, Levine *et al.* finally proved that human and rabbit anti-Rh did not react with the same antigen. However, owing to common usage of the term, 'Rh' was kept as the title for the human antibodies. The rabbit 'anti-Rh' was then called anti-LW in honour of Landsteiner & Wiener. Even today people often make the mistake of referring to the 'Rhesus' blood group system; this is incorrect, as the blood-group antigen has nothing to do with Rhesus monkeys, and should be referred to as the 'Rh' blood group system.

In 1943, Wiener demonstrated the presence of six alleles in the Rh system with three different antisera, and since then the system has been shown to have many further variants and complexities. In 1953, it was discovered that some RhD-positive individuals can make anti-D, and this led to the hypothesis that the RhD antigen is a complex antigen comprised of a series of epitopes. A lack of part of the antigen means that an individual can make antibody to this part.

In 1942, Levine confirmed that Rh incompatibility between mother and fetus was the major cause of haemolytic disease of the newborn (HDN). Various studies showed that, depending on the population, 3–25% lack the Rh antigen D, and that RhD-negative individuals readily make anti-D. RhD is thus of critical importance in blood transfusion, and all transfusions are routinely matched for RhD compatibility. In emergency situations RhD-negative blood is given, particularly if the recipient is a female of child-bearing age, to avoid immunization. RhD-negative women can be treated with prophylactic anti-D to prevent HDN. Although the mechanism of action for this process is not fully understood, it is apparent that rapid clearance of fetal red cells from the maternal circulation by passively administered anti-D prevents the mother making immune anti-D. A very successful programme of routine postdelivery administration of anti-D to RhD-negative mothers bearing RhD-positive babies has greatly reduced the incidence of HDN. However, some fetal cells can also enter the maternal bloodstream during pregnancy, and cause immunization. To prevent this cause of HDN, routine antenatal administration of anti-D is now being adopted in many countries.

Correspondence: M. L. Scott, Bristol Institute for Transfusion Sciences, National Blood Service, Southmead Road, Bristol BS10 5ND UK
E-mail: marion.scott@nbs.nhs.uk

Notation

Two notations were developed to describe the Rh blood groups. Fisher & Race proposed a system based on a model of three closely linked loci, *C*, *D* and *E*. They proposed that there were two antithetical antigens produced by the C and *E* loci, producing C, c, E and e antigens, but only one antigen, D, produced by the *D* locus. They proposed that as the allele of *D* produces no antigen, the symbol d would represent the lack of D.

Wiener proposed an alternative system, based on his theory that there were multiple alleles at a single locus, with each agglutinogen composed of several serological determinants – for example the agglutinogen produced by *R′* expresses at least three different blood factors, Rh_0, rh′ and h″ (D, C and E in Fisher–Race terminology).

For routine serology, both systems are in common usage. The CDE nomenclature of Fisher–Race is used when the full phenotype is to be written, but the symbols of Wiener are often used as 'shorthand' for the gene complexes. For example, Dce/dce in Fisher–Race notation is normally referred to as R_1r in routine communications.

In 1986, based on an analysis of a lot of accumulated serological data, Patricia Tippett proposed a modification of the Fisher–Race theory, with only two structural Rh loci, *D* encoding the D antigen, and *CE* encoding the C, c, E and e antigens. Later molecular genetic analysis (see below) has

Table 1 Rh notations

CDE Fisher/Race	Rh-Hr Wiener shorthand	Numerical Rosenfield
DCe	R_1	RH 1, 2, 5
dce	r	RH 4, 5
DcE	R_2	RH 1, 3, 4
Dce	R_0	RH 1, 4, 5
dcE	r″	RH 3, 4
dCe	r′	RH 2, 5
DCE	R_z	RH 1, 2, 3
dCE	r_y	RH 2, 3

Table 2 Rh Cc Ee antigens: specific residues

Antigen	Amino-acid position				
	16	60	68	103[a]	226[b]
ce	Trp	Leu	Asn	Pro	Ala
Ce	Cys	Ile	Ser	Ser	Ala
cE	Trp	Leu	Asn	Pro	Pro
CE	Cys	Ile	Ser	Ser	Pro
D	Trp	Ile	Ser	Ser	Ala

[a]Cc depends on residue 103.

[b]Ee depends on residue 226.

proved this model to be correct, but the 'shorthand' Wiener terms are still in common usage.

Rosenfield *et al.* also introduced a numerical terminology, to make digital data storage easier. This simply labelled the D, C, E, c and e antigens 1, 2, 3, 4 and 5, respectively. This numerical notation is now the basis for the International Society for Blood Transfusion terminology.

Table 1 shows the relationship between the three terminologies for the 8 known Rh gene complexes. When representing the probable genotype of an individual, it is common to use the Wiener shorthand; DCe/DCe is commonly referred to as R_1R_1; DcE/DcE as R_2R_2; DCe/dce as R_1r, etc. The easiest way to remember these is that if D is present there is a capital R, if D is absent there is a small r. 1 or ′ means that C is present, 2 or ″ means that E is present. 0 means that neither C or E are present, and Z or y means that both are present. This complex nomenclature makes it difficult for non-serologists to understand the Rh blood-group system!

Rh molecular biology

In 1990, the *RHD* and *RHCE* genes were cloned; the two genes are highly homologous (93·8%), each consisting of 10 exons, with only a predicted difference in 36 of the 417 encoded amino acids. The genes are on the short arm of chromosome 1. They are separated by about 30kb, and they are in tail-to-tail configuration. In 1991, it was shown that in whites, most RhD-negative individuals completely lack the *RHD* gene. Later work showed that sometimes an *RHD* gene is present in RhD-negative individuals, but it has some form of mutation that gives rise to a stop sequence, which prevents expression of the protein. In black Africans it is common to find an *RHD* pseudogene (*RHDψ*) that contains a 37-bp duplication of the last 19 nucleotides of intron 3 and the first 18 nucleotides of exon 4; there is also a nonsense mutation in exon 6, which ensures that no RhD protein is present in the red-cell membrane. In 1996, expression of the *RHD* and *RHCE* cDNA in erythroleukaemia cells showed definitively that RhD and Rh C, c, E and e antigens are encoded by two genes and are carried on two separate proteins. Expression of C, c, E and e antigens was shown to be dependent on five amino acids of the CE protein (see Table 2).

Hydropathy plots of the predicted amino-acid sequences indicate that the RhD and RhCE proteins both have 12 hydrophobic transmembrane domains, six extracellular loops, and cytoplasmic N and C termini. Only nine of the RhD-specific residues are predicted to be exposed at the extracellular surface of the red cell; one in extracellular loop 1, three in extracellular loop 3, two in extracellular loop 4 and three in extrcellular loop 6 (Fig. 1).

There is also another homologous gene, *RHAG*, which codes for another similar protein,the Rh-associated glycoprotein. Studies have shown that although this glycoprotein does not carry any residues that relate to any Rh blood-group polymorphism, it is highly associated with the RhD and RhCE proteins in the red cell membrane, and appears to be involved in chaperoning these proteins into the membrane. Changes to this gene can also therefore effect expression of the Rh antigens.

Variants of RhD

The RhD antigen can vary in both the quantity of antigen expressed and the qualitative nature of the antigen. Many complex serological observations over the years have now been explained by more recent studies on the molecular biology of the system.

Quantitative variants

Enhanced D antigens

Cells of the type D – have enhanced expression of the antigen (110000-202000 sites per cell, compared with 10000-30000 in normal phenotypes). These cells totally lack the RhCE protein, and it is most likely that normally the RhD and RhCE proteins compete for binding to the Rh-associated glycoprotein, thus absence of the RhCE proteins will result in more D

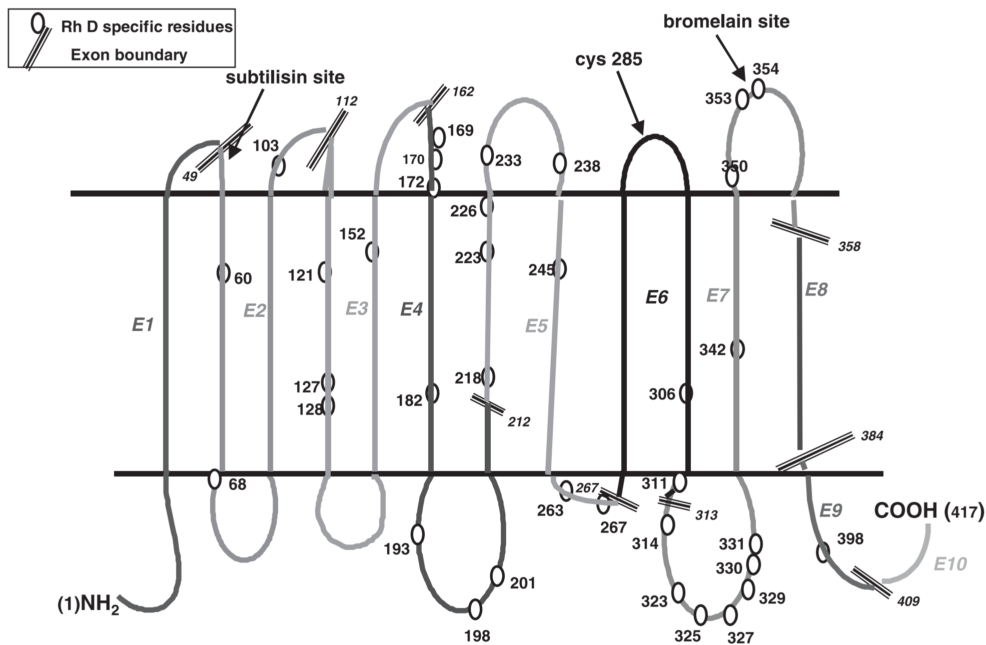

Fig. 1 Rh model.

antigen being expressed. Cells of these types can be readily agglutinated directly by IgG anti-D, whereas cells with normal numbers of the D antigen will only be agglutinated by IgG anti-D if the cells are enzyme-treated, or an antiglobulin reagent is used. Some other D variants (e.g. DIIIC, DIVa) also show elevated expression of the D antigen. Again, this may be due to increased binding to the Rh-associated glycoprotein.

Weak D

Weak D is generally considered as a qualitatively normal D antigen, but expressed at low levels. Some types of weak D show normal inheritance patterns, whereas others are caused by the Ceppellini effect, where D expression is diminished when the haplotype encoding D is partnered by a haplotype encoding C but no D. The two types can be distinguished by family studies. Weak D is normally defined as a type where the red cells are not agglutinated by an IgM anti-D, but require the use of an IgG anti-D and an antiglobulin reagent. As the avidity of IgM typing reagents has increased with the careful selection of monoclonal antibodies, fewer samples are now classified as weak D. A large number of weak D samples has been analysed by DNA sequencing of the *RHD* gene, and in all cases nucleotide changes have been found that encode amino-acid substitutions in the RhD protein. At least 21 different genetic types of weak D have been found. Most are point mutations, and they are commonly clustered in four areas: amino acid residues 2-13, 149, 179-225 and 269-397 (Fig. 1). All of the amino-acid substitutions associated with weak D are in the predicted membrane-spanning or cytoplasmic portions of the RhD proteins; none is predicted to be extracellular. The mechanism of the low expression of weak D is not clear, but presumably the weak D mutations in some way compromise the efficient expression of the RhD protein in the membrane, again perhaps by influencing the interaction of the RhD protein with the Rh-associated glycoprotein.

D_{el}

D_{el} is a very weak form of D found in the Far East, and can only be reliably detected by absorption and elution tests. Again, several different molecular changes have been discovered in the *RHD* gene. It is not clear how these mutations affect the very reduced levels of RhD protein expression.

Rh_{null} and Rh_{mod}

In these very rare phenotypes, there is no expression or very little expression of any Rh antigens on the red cell. This is caused by mutations in the *RHAG* gene. Two molecular types of Rh_{null} have been described, each of which have termination mutations in the *RHAG* gene. In an Rh_{mod} type, a missense mutation was found. This clearly shows the importance of the Rh-associated glycoprotein for correct and efficient expression of the Rh antigens.

Qualitative variants of RhD

It has been known since 1953 that some RhD-positive individuals can make anti-D. Their red cells were considered to

lack part of the normal D antigen, such that they had made anti-D to the missing part on challenge with the whole D antigen, thus they were called 'partial D'. In 1962, Tippett and Sanger studied the reaction patterns of cells and serum from D-positive people who had made anti-D. They observed a limited number of reaction patterns, and divided them into six categories. These categories of partial D were designated with roman numerals. Subsequently, most categories were subdivided, and another category, category VII was added. Categories are denoted by 'D' followed by roman numerals, e.g. DVI is partial D, category VI. In many cases, novel low-incidence antigens are associated with types of partial D antigen. All DIVa are Go(a)+ all DVa are D^w+ all DVII are Tar+ and all DFR are FPTT+.

Since this time, more partial-D individuals who have made anti-D have been identified, but it was not possible to classify them into the category system, because the materials were no longer available for the necessary cross-testing. The successful production of a large number of monoclonal antibodies has provided more reliable tools for the characterization of partial D phenotypes. Each partial D has a different pattern of reactivity with panels of different monoclonal antibodies. This has been presumed to be due to the different monoclonal antibodies recognizing different epitopes of the D antigen. Partial D antigens will not react with monoclonal antibodies that are specific for an epitope that they lack. This has been extensively studied in three international workshops, resulting in an estimation of 30 different epitopes of the RhD antigen. New partial-D antigens that cannot be assigned to the categories are denoted as 'D' followed by 2-3 Arabic capital letters, e.g. DFR, DBT.

Routine RhD typing with two different monoclonal antibodies has revealed that such D variants are more common than had been previously thought, as many discrepant typings were found to be due to D variants, who were previously undetected because they had not made anti-D. Many new D variants are still being reported.

The molecular basis of most of the partial-D and D-variant types has now been investigated. They fall into three types: substitution of part of the *RHD* gene with *RHC*; deletion of part of RHD; or single amino-acid substitutions in parts of the *RHD* gene that code for parts of the D protein that are extracellular (Fig. 2).

Comparison of the serological reaction profile of partial-D variants with their predicted amino-acid sequence indicates that many of the epitopes defined by monoclonal antibodies require interactions between the different extracellular loops of the protein. This has been confirmed by site-directed mutagenesis experiments, where D-specific bases were sequentially inserted into an *RHCE* gene, and the protein expressed in erythroleukaemia cells and tested for reactivity with monoclonal antibodies. This work has allowed the construction of an epitope 'map' of the RhD antigen.

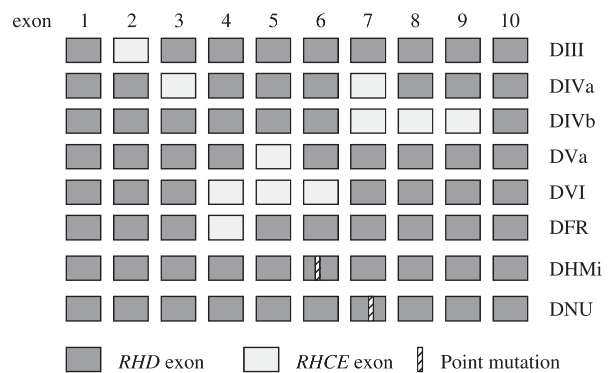

Fig. 2 Examples of D variant genes.

The high serological complexity of the RhD antigen, demonstrated by the presence of many naturally occurring variants and the differing fine specificities of monoclonal anti-D, can be explained by changes in nine critical extracellular amino acids of the protein, present on four extracellular loops. As with most antibody-antigen interactions, 2–4 residues of the paratope and epitope are required to interact to give rise to demonstrable binding, and these residues can be either on one loop, or scattered across two, three or four loops that are close together in the tertiary formation of the RhD protein. Given this combinatorial picture, it is not suprising that 30 epitopes have been defined serologically; if three residues are critical for every binding, then there are theoretically 84 possible combinations. Constraint in the numbers is probably dependent on the span of the antibody paratope and the configuration of the extracellular loops in the membrane. Such is the nature of fine serological specificity that changes in just one residue of the antigen or antibody can determine whether binding occurs. Creation of new epitopes (such as the low-incidence antigens associated with some partial D) can be caused by changes to single residues.

Variants of Rh C, c, E and e

It is now becoming clear that a similar complex range of variants exist for the other Rh antigens. RhE variants have been classified into seven types on the basis of testing 19 partial E variant samples with 23 monoclonal antibodies, and different patterns of reactivity are becoming apparent with anti-e, anti-C and anti-c complex of the blood-group systems. Where the molecular basis has been investigated, there is a similar picture of gene exchange with the RHD gene, exon deletion, and single base-pair mutations.

Conclusions

The Rh system has homologous, closely linked genes: *RHD* producing the RhD antigen and *RHCE* producing the Cc and Ee antigens. Each antigen consists of a mosaic of epitopes,

which are recognized by different antibodies, monoclonal and polyclonal. Exchanges between the two genes have produced a series of variants, which are recognized in different ways by the different epitope-specific antibodies. Insertion of the Rh proteins into the red cell membrane is highly influenced by another protein, the Rh-associated glycoprotein, such that changes in the *RHAG* gene can influence the expression of both of the Rh proteins. Both RhD and RhCE proteins are polytopic proteins with 12 membrane-spanning domains and six extracellular loops. Most of the epitopes rely on interactions between these loops for antibodies to bind. An understanding of the molecular basis of the Rh blood-group system has assisted our interpretation of the complex array of serological phenomena associated with Rh.

Acknowledgement

This research was supported by the National Blood Service.

Suggested reading

Daniels G: *Human Blood Groups*, Chapter 5. Oxford, UK: Blackwell Publishing Ltd, 2002.

Flegel W: Available at: http://www.uni.ulm.de/~wflegel/RH/

Liu W, Avent ND, Jones JW, Scott ML, Voak D: Molecular configuration of Rh D epitopes as defined by site directed mutagenesis and expression of mutant Rh constructs in K562 erythroleukaemia cells. *Blood* 1999; **94**:3986-96

Scott ML: Section 1A: Rh serology co-ordinator's report. *Transfus Clin Biol* 2002; **9**:23-9