The HLA system: structure and function

W F BODMER

From the Imperial Cancer Research Fund, London

SUMMARY The HLA system is the major histocompatibility system of man and was found through a search for blood group-like determinants on white blood cells that would be effective in matching for transplantation. The HLA system has its counterparts in other species of mammals, birds, and reptiles including the much studied H2 system of the mouse. The HLA system started from a series of antigens defined by a combination of relatively crude serology and genetics, supported by extensive statistical analysis. It has turned out to be a complex genetic region determining two major sets of cell surface products which mediate essential functional interactions between cells of the immune system, and so have a major role in the control of the immune response. Polymorphism in the HLA region is thus associated with a wide variety of diseases with an immune aetiology.

In this brief review I shall survey some aspects of the structure and function of the HLA system and their relation to disease susceptibility, and the normal and pathological tissue distribution of the two major classes of HLA determinants.

The HLA system

Early studies in the search for white cell blood groups, especially by Dausset, showed that there were white cell agglutinins in the serum of polytransfused subjects which were used for the definition of an antigen, but these sera turned out to be too complex to be generally useful. The further search for antigens depended on the independent discovery by Van Rood and Payne in 1958 that leucocyte agglutinins were produced by fetal-maternal stimulation. Using such sera and a statistical approach based on the analysis of 2 × 2 associations of serum reactions on a panel of cell donors, Van Rood defined a two allele system which he called group 4 in 1962. These statistical methods were further developed by Payne et al., who defined an apparently independent system of antigens which they called LA (L for leukocytes and A for the first locus). Subsequent development of the system made use of a microcytotoxicity assay dependent on complement that had been developed by Terasaki. This assay used partially purified peripheral blood lymphocytes as the target cells, and together with fetal-maternal antisera, remains the mainstay of HLA typing, although monoclonal antibodies, as will be described later, are gradually taking over. As further antigens associated with the LA and 4 groups were defined it became clear that these antigens fell into two series, each of which behaved as if they were controlled by a set of alleles at two closely linked loci, now known as the HLA-A and HLA-B loci. Subsequently, a third similar locus, HLA-C, was defined.

Following the discovery by Bach and Amos that the mixed lymphocyte culture reaction (in which the ability of lymphocytes to stimulate each other to divide is detected by the incorporation of tritiated thymidine) was controlled in association with the HLA antigens, the mixed lymphocyte culture test was used to define a new set of determinants called HLA-DW. Subsequently, it was established that these were paralleled by a set of serological determinants that could be identified predominantly on peripheral blood B lymphocytes, the HLA-DR, for D-related types. Further serological and cellular studies identified additional similar types called HLA-DQ and DP.

The development of the HLA system has been greatly enhanced by a series of international collaborative workshops started by Amos in 1964, the ninth of which took place in 1984. The proceedings of these workshops effectively define the development of the HLA system, and the latest provides a comprehensive survey of the HLA system. Further reviews of the system are available.

The loci controlling the two major sets of serologically defined antigens, the HLA-A, B, C or class I determinants, and the HLA-D region or class II determinants, are closely linked to each other at either end of the HLA region, which lies on the short arm of chromosome 6. Fig 1 gives a schematic genetic map of the HLA system. The region spans a recombination fraction of between 2 and 3%. In other words about
The HLA system: structure and function

Centromere

Fig 1 Schematic map of the HLA region.

Fig 2 Schematic structure of HLA-A, B, C molecules and their genomic organisation. The typical HLA-A, B, C gene produces a molecule with three domains outside the cell, labelled a1, a2, and a3, a transmembrane segment (tm) and a short cytoplasmic tail (c). These molecules are associated with β2-microglobulin outside the cell membrane, which associates particularly with the a3 domain. SS indicates the presence of disulphide bonds. The domains are coded for by separate exons (open boxes) labelled SS for the signal sequence, a1, a2, a3 and tm for the corresponding domains, while the cytoplasmic tail has its coding sequence split into three exons.

97% of the time alleles at either end of the system, say HLA-DP and HLA-A, will be held together on the same chromosome as they are passed on from parent to offspring.

The HLA-A, B, and C loci lie at the end of the region, distal to the centromere. The antigens they control are present on most nucleated cells in the body. Molecular data suggest that there are also at least two sets of related loci, called QA and TL by analogy with the mouse H2 system, which have similar structures but which are evolutionarily distinct and whose products have a much more restricted tissue distribution. All these products are associated with β2-microglobulin (fig 2), which is coded for by a single gene on chromosome 15. The class I region contains some 25 or more genes, about half of which, however, are not expressed.

The HLA-D region or class II products are each formed from a functional heterodimer of an α and a β chain (fig 3), but in this case both sets of chains are coded for in the HLA-D region. There are, as already mentioned, three major subsets of products, DP, DQ,
and DR and this set of genes is at the centromeric end of the region.

Both the HLA-D region and HLA-A, B, C sets of products function in controlling interactions between cells in the immune response and, through this, play a key part in the regulation of the immune response through interaction with the T lymphocyte antigen receptor.

The strongest evidence that the HLA system is indeed the major histocompatibility system relevant for matching for transplantation comes from the fact that kidney or bone marrow grafts exchanged between HLA identical siblings survive, with appropriate immunosuppressive treatment, almost as well as those between identical twins and far better than grafts exchanged between mismatched siblings or other relatives. Because the loci controlling all the HLA determinants are closely linked within the HLA region, on average there is a 25% chance that a pair of siblings will be HLA identical, the exceptions being due to genetic recombination occurring within the HLA region. Matching for HLA within the family, therefore, ensures that most of the time all the differences within the region are matched. Matching unrelated donors and recipients for HLA-A, B, C and D region determinants has an important effect on graft survival, but the results are much less striking than those with HLA identical siblings. This may either mean that there are other determinants controlled by the HLA region, which are not yet clearly identified, with respect to which matching for transplantation is important; or that combinations matter so that, for example, matching for either HLA-A, B, C alone or HLA-D region alone is not as good as matching for both together. Morris et al have written a recent review of the effects of HLA matching on transplantation.

In between the HLA-A, B, C and HLA-D subregions lies a series of genes for some complement components and also for the 21-hydroxylase enzyme, deficient in those with congenital adrenal hyperplasia. The complement genes in HLA are C2, the second component of the classical complement pathway, factor B (BF), the closely related product of the alternative pathway, and C4 the fourth component of the classical pathway. The C4 and 21-hydroxylase (21OH) genes are duplicated adjacent to those for C2 and BF. Recently, the genes for the tumour necrosis factor α and β chains have also been shown to be in the HLA region.

Serological and cellular analysis has shown that there are at least 18 alleles at the A locus, 41 at B, eight at C, about 20 at DR, three at DQ, and six at DP. Each allele at these loci corresponds to a type defined by the serological and cellular techniques. Each type is identified by a combination of a letter and a number, corresponding to the locus and the allele within the locus—for example, A1, B8, CW6 or DR4. (The W in CW6 stands for "workshop" to indicate, originally, a less clearly defined specificity. Now, however, the W is mainly used to distinguish the C locus products from the complement components.) The types thus fall into six different sets corresponding to the six loci, A, B, C, DR, DQ and DP, and each individual can carry up to two different types corresponding to two different alleles at any locus. Thus each individual may have up to 12 different types, two from each of the six sets corresponding to the six loci. These types can create about 80,000 million different combinations and generate, overall, an extraordinarily high level of polymorphism.

Pairs of alleles at closely linked loci within the system are often highly associated in the population due to the phenomenon of linkage disequilibrium or gametic association. Thus A1 at the HLA-A locus occurs in northern Europeans with a frequency of about 31% and B8, controlled by an allele at the closely linked HLA-B locus, with a frequency of about 21%. The combined phenotype A1 and B8 occurs with a frequency of about 17%, while if the two types were independent in the population the expected frequency would only be 0.31 × 0.21 or 6.5%. The excess of A1 B8 types over that expected if they were independent is due to gametic association or linkage disequilibrium—namely, the occurrence in the population of a relatively high frequency of chromosomes that carry the alleles A1 and B8 together on the same chromosome. In the absence of disturbing effects, such as natural selection, gametic associations decline eventually to zero at a rate (1 − r)n, where r is the recombination fraction between the loci and n the number of generations. The recombination fraction is, therefore, the major determinant of the extent of gametic association, and strong associations probably occur only for recombination fractions which are less than 0.5–0.1%. In practice strong gametic association is found between alleles at the DQ, DR, complement and HLA-B and C loci, while the average association between alleles at HLA-A and HLA-B is much less, and there is mostly no detectable association between HLA-DP alleles and other alleles of the HLA region loci. This fits in with the fact that the recombination fraction between the HLA-A and HLA-B loci is about 0.8%, while that between HLA-DP and DQ or DR is about 1.5%. The recombination fractions between the other loci must be much less, as few any recombinants have been detected between any of the HLA-DQ, DR, complement, B and C loci.

The patterns of gametic association and recombination are reflected in the distribution of recombination "hot-spots" (fig 1). These are positions where recombination is presumed to occur much
more often than elsewhere in the region, and they have been clearly identified in the mouse I region, which is the equivalent of the HLA-D region. In the HLA region recombination hot-spots are presumed to occur between DP and DQ and between A and C, and these are identified directly by relatively high recombination fractions—namely, between 0.5% and 2%, given the established molecular distance between the genes. In the case of the DQ subregion the hot-spot is identified by a relative lack of gametic association between alleles at different loci within the subregion. In general, strong gametic associations may be expected between alleles at loci within a region bounded by recombination hot-spots, but not between alleles at loci separated by a recombination hot-spot, even when both are in the HLA region.

Molecular genetics of the HLA region

All the known products of the HLA region have been cloned using standard techniques of molecular genetics, generally starting from cDNA clones through to selecting genomic clones. This has established the genomic organisation of individual genes and of the region as a whole. Sequence data show that the HLA-A, B, C and D region products are members of the immunoglobulin supergene family, sharing important though limited homology and characteristic structural features with the immunoglobulins.

There is an obvious problem in naming the HLA region products, which arises because of their having originated from studies of genetic variation, as in the case of blood groups, rather than as a protein product, like immunoglobulins and haemoglobins. When the product is studied first, the gene loci are naturally named after the product. In the case of the HLA system, however, there are names for the genes but no good names for the products. By analogy with the immunoglobulins, the HLA and other species major histocompatibility products can conveniently be called "histoglobulins." It then seems natural to call the class I products histoglobulins I or HgI, and the class II products histoglobulins II or HgII; this provisional nomenclature will be used in the rest of this paper.

Analysis of the class I or HgI products and their corresponding DNA sequence has shown them to have three protein domains outside the cell, followed by a transmembrane region, and a short intracellular cytoplasmic tail. The three main external domains have corresponding exons, as do the signal sequence and transmembrane region, while the cytoplasmic domain is distributed among three small exons (fig 2). All the HgI genes have essentially the same structure, and the major polymorphism of the HLA-A, B, C genes is found in the two domains most distal to the cell surface (α1 and α2). The membrane proximal domain (α3) is the one with the greatest similarity to immunoglobulins and is most closely associated with β2-microglobulin, itself a member of the immunoglobulin supergene family.

The class II or HgII products have a similar structure and genomic organisation to HgI (fig 3). Each of the two chains has two external domains (α1, α2 and β1, β2), coded for by separate exons. It is also the membrane proximal domains which show the greatest similarity to the immunoglobulins.

Fig 4 gives a more detailed map of the HLA-D region. From this it can be seen that in addition to the DP, DQ, and DR subregions, other genes have been found between the DP and DQ subregions, provisionally called DZα and DOβ. It is, however, not yet known to what extent they are expressed and what their functional importance might be. The DP and DQ subregions contain duplicated pairs of α and β genes, only one member of which is expressed, while the DR subregion contains two and sometimes three expressed β chain genes, but only one α chain gene. This organisation suggests an origin from a series of duplication events and some selection for diversification of the HLA-D region products. By far the most polymorphic loci are those for DQβ, DQα, and one of the DRβ chains. The other DRβ chain and the DPβ chain are also fairly polymorphic, but DQα is the only highly polymorphic α chain. This pattern of differential variability strongly implies that it is generated by natural selection. The obvious explanation for this selection is an effect of polymorphism on immune response differences to clinically important pathogens, and some form of frequency dependent selection in relation to this is still the most plausible mechanism.

Fig 4 Molecular map of the human HLA-D region. The blocks indicate the various genes in the DP, DO, DQ and DR subregions and the arrows above them the directions of transcription. ■ indicate genes known to be expressed; □ those known to be pseudogenes; and □ those which are potentially expressed, but where the expression is not definitely established (based on Trowsdale et al 1985).
for the generation of the extensive polymorphism of the HLA region.\textsuperscript{12}

Sequence comparisons between the products of the HLA region and also between them and immunoglobulin and other genes of the immunoglobulin supergene family are beginning to provide a consistent overall picture of the pattern of evolution of the region. The greatest homology between parts of the HLA molecules and immunoglobulins is about 20–25% at the amino acid level, which suggests an evolutionary separation of about 750 million years. The differences between HLA-D region and ABC products are not much less, and the approximate percentage of shared amino acid sequence between HgII $\alpha$ and $\beta$ chains is also only 30%. This suggests that a series of successive duplications first of all separated the HLA region from the immunoglobulins and other related products, and then within the HLA region separated HgI and HgII products, and subsequently within HgII $\alpha$ and $\beta$ chains. All of these events are likely to have happened more than 700 million years ago. Within the HLA-D region the differences between subregion genes, either $\alpha$ or $\beta$, are about 45%, while within a subregion the difference between the two DP $\alpha$ chains is about 25%, but between the two DQ$\alpha$ chains only about 7%. This suggests a further hierarchy of duplication events, of which the most recent, that giving rise to the division with the DQ subregion, may have occurred only some 15 million years ago. Sequence variation between alleles, especially neutral substitutions, can give some indication of their age. The data suggest that some alleles may be five or more million years old and therefore predate Homo sapiens, but postdate divergence of the hominids from the great apes. There are, however, vestiges of similarities between alleles even in man and mouse, suggesting that some allelic differences may, indeed, be much older than this but have naturally diverged during mammalian evolution.\textsuperscript{13}

### Immune response and disease associations

The discovery by McDevitt et al that specific immune response genes could be found in the mouse H2 region\textsuperscript{3} stimulated the search for HLA and disease associations based on immune response differences having a role in diseases with an immune or autoimmune aetiology. There is now an extraordinarily long list of HLA and disease associations.\textsuperscript{14} Table 1 gives examples of some of the more striking presently known associations. In all these cases an HLA antigen has been found much more often in patients with the disease than in appropriate controls. This indicates either a direct effect of the antigen in question, perhaps on immune response, or a susceptibility control by an allele at a closely linked locus that is in

<table>
<thead>
<tr>
<th>Disease</th>
<th>HLA Alleles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rheumatoid arthritis</td>
<td>DR4</td>
</tr>
<tr>
<td>Ankylosing spondylitis</td>
<td>~100% B27</td>
</tr>
<tr>
<td>Juvenile diabetes</td>
<td>~100% DR3 or DR4</td>
</tr>
<tr>
<td>Narcolepsy</td>
<td>~100% DR2</td>
</tr>
<tr>
<td>Graves' disease</td>
<td>DR3*</td>
</tr>
<tr>
<td>Myasthenia gravis</td>
<td>DR3</td>
</tr>
<tr>
<td>Multiple sclerosis</td>
<td>DR2</td>
</tr>
<tr>
<td>Psoriasis</td>
<td>CW6</td>
</tr>
<tr>
<td>Coeliac disease</td>
<td>DR3</td>
</tr>
<tr>
<td>Haemachromatosis</td>
<td>A3</td>
</tr>
<tr>
<td>Reactive arthritis</td>
<td>B27</td>
</tr>
</tbody>
</table>

\*The listed antigens are those with the strongest disease association.

There are relatively few published examples of a direct effect of HLA on immune response. Law et al showed that men who had undergone vasectomy and who had produced particular types of autoantibody to sperm had a significantly higher prevalence of the antigen A28 than corresponding controls,\textsuperscript{16} and this observation has been confirmed by Hancock et al.\textsuperscript{17} A recent interesting example of a difference in immune response is the suggestion that antigens of various Mycobacterium species gave a higher responsiveness in a skin test in HLA-DR4 positive subjects than in others.\textsuperscript{18} This would be exactly the type of association that might be expected to occur in selection acting on HLA variants through differential effects on resistance to important pathogens. The later price to pay for this protection is, then, the association with autoimmune diseases such as rheumatoid arthritis and IDDM.

The association between IDDM and HLA illustrates many features of the analysis of HLA effects on disease susceptibility. The initial association found by Singal and Blajchman\textsuperscript{19} was with HLA-B15, and only later was the association with the HLA-D region determinants DW3 and DW4 and then their sero-
The HLA system: structure and function

Table 2  HLA-DR genotypes of 43 diabetic patients (IDDM) (x means not 3, 4, or 1)

<table>
<thead>
<tr>
<th>DR type</th>
<th>3/x</th>
<th>3/3</th>
<th>3/4</th>
<th>4/4</th>
<th>4/1</th>
<th>4/x</th>
<th>All other</th>
</tr>
</thead>
<tbody>
<tr>
<td>No observed</td>
<td>2</td>
<td>1</td>
<td>17</td>
<td>9</td>
<td>10</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Expected</td>
<td>5.9</td>
<td>0.3</td>
<td>0.6</td>
<td>0.3</td>
<td>0.4</td>
<td>5.2</td>
<td>30.3</td>
</tr>
</tbody>
</table>

Based on data from Winearls et al 1984.23 The expected numbers are calculated from the frequencies seen in a local control population.

Table 3  B15, DR4 haplotype combinations in diabetic patients and controls

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>B15, DR4</th>
<th>B15, x</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>Controls</td>
<td>1</td>
<td>6</td>
</tr>
</tbody>
</table>

x means not DR4. Haplotype B15, DR4 means that B5 and DR4 are inherited together on the same chromosome.

Data from Winearls et al 1984.23

logical equivalents, DR3 and DR4, established.20 The distribution of HLA types in affected sibling pairs provided direct evidence for linkage between the HLA system and the gene controlling susceptibility to IDDM.21 In contrast with population association data, which depend on a strong gametic association between the antigen observed in patients (which has a higher prevalence than in controls) and a presumed allele at a closely linked disease susceptibility locus, the sibling pair family analysis is a direct search for linkage that does not depend on gametic association.22

Table 2 shows the distribution of genotypes at the HLA-DR locus for 43 patients with IDDM, based on studies of their families.23 Patients all have either DR3 or DR4, and there is a striking excess of DR3/DR4 and DR4/DR1 heterozygotes, which agrees with earlier data, at least for DR3/DR4.21 Another feature of the data (table 3) is the fact that the association between B15 and DR4 is much stronger in patients with IDDM than it is in a control population. This observation accounts for the fact that the first association between HLA and IDDM was with B15. This was initially puzzling in relation to the subsequently established association with DR4, as there was no strong association between B15 and DR4 in the normal population. The data clearly indicate that the primary effect of the HLA system on IDDM is not due to DR4 alone. The susceptibility could either be due to an allele at another locus closely linked to both the DR and B loci, which shows strong gametic association with both B15 and DR4, or it could be due to DR4 together with an allele at another locus that is associated with B15. The heterozygote effect, as pointed out by Svejgaard et al24 and Bodmer25 can be explained by an effect of a particular combination of α and β chains that is only found, say, in DR3/DR4 heterozygotes. As DQα is the only polymorphic α chain locus, this would imply that the combinations must be DQα/β combinations,23 25 which is consistent with the population association data on both B15 and DR4. Attempts are now being made to refine the data on HLA and disease associations by studying restriction fragment length polymorphisms that can be identified using clones of the various HLA region genes.26 Using this approach, Bell et al27 found suggestive evidence for the role of HLA-DQβ chain polymorphism in the association between DR3 and myasthenia gravis.

The associations between HLA and cancers are much less clear cut than those with diseases with a definite immune aetiology.28 The most obvious possibility concerns effects of HLA variation on immune response to viruses which are aetologically associated with cancer. Suggestive associations have been found between HLA-DR5 and the acquired immune deficiency syndrome (AIDS) virus, human immunodeficiency virus (HIV)29 30 and between Burkitt's lymphoma, the original source of the Epstein-Barr virus, and HLA-DR7.31 A weak population association between HLA-B locus antigens and Hodgkin's disease has been confirmed by a sibpair family analysis on patients with Hodgkin's disease, showing a severe distortion of the Mendelian HLA segregation expected in the absence of an HLA effect.28 It is important to emphasise that only a very small proportion, perhaps up to 3%, of cases of Hodgkin's disease occur in families with two or more affected persons. It is within such families that HLA typing can establish whether the affected pairs of siblings with Hodgkin's disease are identical, share only one HLA chromosome but not the other, or have neither chromosome in common. The expected incidence of these three classes based on the assumption of Mendelian segregation and in the absence of an HLA effect is 1:2:1. It is the departure from this expectation that shows that there may be a gene in the HLA region which confers susceptibility to Hodgkin's disease. The data suggest that it may be possible to find an HLA variant that is highly associated with Hodgkin's disease in the population, and HLA-DP is an obvious possibility.32 Hodgkin's disease is interesting in that it shows how a gene, which may only confer a small chance, say 1 in 20, of getting the disease, may nevertheless, if there is virtually no chance of the disease developing without the gene, produce an unequivocal inherited susceptibility, even when there is no significant familial clustering of the disease.33

The effects of the HLA system may not be directly on disease susceptibility itself but on the pattern of
expression of disease. Thus Oliver et al.\textsuperscript{34} suggest that the extent to which germ cell testicular tumours metastasise may depend on the patient’s HLA type. Patients who were DR7 showed a higher incidence of stage IV disease, with haematogenous spread to lung, liver, brain or bone than those with other types. Further data, both from families and population studies, are needed to substantiate this interesting possibility of a genetic influence on the pattern of tumour metastasis.

The bases for the effects of the HLA system on susceptibility to disease need to be established by appropriate functional studies. Transfection of cloned genes into a cell in which their expression can readily be detected now provides enormous scope for investigating the relation between function and structure. The cloned genes can be changed in defined positions and the effects of this on expression investigated, both at the serological level and in terms of function—for example, in antigen presentation. Thus in our laboratory DP cosmid clones have been transfected into mouse L cells and expressed both in terms of recognition by monoclonal antibodies, and functionally, through presentation of antigen.\textsuperscript{35} The specific sequence at the DNA and protein level that has a role in an immune response underlying an HLA and disease association must eventually be identified. This may mean reconstructing in vitro the basis for an autoimmune attack using appropriate T cell clones, which either recognise the autodeterminant in cellular destruction, or, mediated through T cells, generate a damaging autoimmune response. Manipulation of cloned genes expressed in suitable cells will provide the basis for testing whether, for example, a given sequence suspected of underlying an HLA and disease association is critical for the presentation of a particular antigen. The triggers for the development of autoimmunity, be they viruses, other infections, or exposure to antigens in the environment through diet or in other ways, also need to be identified.\textsuperscript{32,36}

Monoclonal antibodies and the distribution of HLA determinants in normal and pathological tissues

Because of their specificity, unlimited availability, and their potentially extraordinarily high reactivity, monoclonal antibodies are making a major contribution to the analysis of the HLA system. One of the earliest monoclonal antibodies against any human determinant was W6/32 which reacts with all HLA-A, B, and C molecules.\textsuperscript{37} This and other similar antibodies, including also some against β2-microglobulin, have been widely used for studies on the tissue distribution of the HLA-A, B, C or Hgl determinants.\textsuperscript{38} Monoclonal antibodies against HLA-D region or HgII determinants have made an important contribution to the analysis of HLA-D region serology and biochemistry, including the sorting out of the three sets of products DP, DQ, and DR.\textsuperscript{39,40} In addition to the monomorphic antibodies, there are now many polymorphic antibodies for both HLA-A, B, C\textsuperscript{41} and HLA-D region\textsuperscript{42} specificities. Monoclonal antibodies to the D region specificities are also helping define new determinants, and aiding the serological identification of differences previously only defined by cellular techniques. New approaches to the production of specific HLA monoclonal antibodies, as well as the development of more sensitive techniques based on enzyme linked immunosorbent assay (ELISA),\textsuperscript{43} should eventually make HLA typing using these approaches the routine standard. Thus mouse L cells transfected with human HLA-DP genes have been used to make a DP specific polymorphic monoclonal antibody in C3H mice. As the L cells are derived from C3H mice, HLA specificities expressed on their surface should be the only determinant, apart from a possible tumour antigen, that the C3H mouse can react to.

For some time, red blood cells seemed to be the only major human normal tissue that did not express HLA-A, B, C or Hgl determinants. Subsequently, especially following the use of monoclonal antibodies, sperm and most trophoblasts were also found to lack expression of Hgl determinants on their surface.\textsuperscript{35} This absence of expression by sperm has been the subject of some controversy but has recently been clearly confirmed by Kuhlmann et al.\textsuperscript{45} Its functional importance is not yet clear, though it may simply represent a gradual slowing down and stopping of turnover of the HLA-A, B, C gene products as the sperm mature. This is also assumed to be the case for red blood cells as reticulocytes have been reported still to express Hgl determinants.

Immune T cell attack depends on the presence of Hgl histoglobulins on the surface of the target cells.\textsuperscript{46} The lack of HLA-A, B, and C determinants on the trophoblast, which carries paternal as well as maternal genes (and so should express antigens foreign to the maternal environment) therefore protects the trophoblasts from cellular immune attack against any antigen. This explains in simple terms why the fetus survives as an allograft—one of the classical conundrums of immunology.\textsuperscript{47,48} Avoidance of rejection of the fetus is most likely to be due to protection against cell mediated immune attack; HLA and other antibodies are commonly produced as a result of fetal-maternal stimulation and, in most cases, have no untoward consequences for the fetus. Some apparent expression of HLA-A, B, C related determinants on extra villous trophoblasts has been described by Redman et al.\textsuperscript{49} but the importance of this is unclear.

A detailed survey of the distribution of HLA-A, B,
The HLA system: structure and function

C antigens on normal human organs has been reported by Daar et al.\textsuperscript{50} The survey shows that in addition to red cells, sperm, and trophoblast, certain other cells, such as hepatocytes and neurones, have either very low or absent expression of Hgl. The lack of HLA-A, B, C antigens from sperm may protect them from cellular immune attack in the female, before fertilisation, and so may have a functional basis that is similar to the lack of these determinants from trophoblasts.\textsuperscript{51} There is no obvious similar rationale for the low expression or lack of these determinants on other tissues, though it may have implications for transplantation—for example of the liver—or using neurones.

The lack of Hgl determinants on choriocarcinomas\textsuperscript{32,53} presumably explains why this tumour (which, since it carries genes introduced from the male as does the trophoblast, should express foreign antigens), is unaffected by HLA incompatibility.\textsuperscript{54} The choriocarcinoma is protected from T cell immune attack just like its normal progenitor the trophoblast. These cells, interestingly, switch off HLA-A, B, C gene expression in coordination, while still synthesising and secreting β2-microglobulin.

The first human tumour cell line anomalously shown to lack completely surface expression of HLA-A, B, C or Hgl products was the Daudi cell line derived from a Burkitt's lymphoma. Genetic experiments with Daudi showed that its primary genetic defect was the expression of β2-microglobulin,\textsuperscript{55} and subsequent biochemical studies confirmed that HLA-A, B, and C products were synthesised, but in the absence of β2-microglobulin were not functionally expressed on the cell surface.\textsuperscript{52} The lack of expression of Hgl in a tumour derived cell line Daudi, which arose from a cell type that normally expresses their determinants on the surface, was suggested by Arce-Gomez et al.\textsuperscript{55} to be a change that could have been selected for during tumour progression through an advantage of resistance to T cell immune attack. Such resistance is presumably only an advantage to tumours that express tumour specific antigens in a form that can induce T cell immunity. In the case of the Burkitt lymphoma derived cell line the Epstein-Barr virus determinants would be a natural target for such immunity. The fact that only few Burkitt's lymphomas seem to have deranged expression of HLA-A, B, C or β2-microglobulin determinants\textsuperscript{52} may be explained by the observation of Rooney et al.\textsuperscript{56} that many Burkitt's lymphomas may lack the Epstein-Barr virus determinants recognised by cytotoxic T cells, and so may have used this alternative route to escape from host T cell immunity directed against the tumour.

These ideas naturally extend to the question of whether carcinomas may be found that lack Hgl expression and, if so, whether this represents an escape from immune response to tumour specific antigens selected for during tumour progression. In initial studies a surprisingly high proportion of carcinoma derived cell lines were shown to have low or absent Hgl expression, suggesting that these might correspond to clinical situations in which tumours had a specific antigen that generated a T cell immune attack.\textsuperscript{38,52} Other examples of this phenomenon have subsequently been described.\textsuperscript{50,57}

A most interesting ramification of these ideas is the observation that certain adenovirus strains switch off Hgl antigen expression in the cells which they infect or transform.\textsuperscript{58,59} Though there is some controversy about the extent to which this reduced expression favours tumour growth, it seems most likely that these adenoviruses have evolved a mechanism that allows them to hide inside a cell and so escape T cell immune attack with respect to virally induced or expressed determinants.

These observations on variations of Hgl expression on tumours suggest the possibility of a correlation with prognosis. Some unpublished observations from our laboratory support this idea in the case of colorectal carcinomas. Thus in a screen of sections from just over 50 cases derived from a typical sample of colorectal carcinomas seen at St Mark's Hospital two were found with complete lack of reactivity with the HLA-A, B, C monomorphic antibody, W6/32. (Richman P, personal communication). In contrast, we have found that only three of eight colorectal carcinoma derived cell lines express detectable amounts of Hgl on their surface. There is therefore a suggestion that those that do express are either more differentiated or less “aggressive” tumours in terms of their growth characteristics and consequently prognosis (Spurr N, Makin C, and Bodmer W, unpublished observations). The clear implication is that the cell lines are derived from tumours with a more aggressive phenotype and therefore with a poorer prognosis, which is consistent with the relatively low yield of tumours that give rise to growing cell lines. It would then be those tumours with a poorer prognosis which tend more often to lack Hgl surface expression. Perhaps as the tumours progress, they are more likely to develop changes on their cell surface associated with their malignancy that may be recognised as tumour specific antigens by the immune system. As a result, they are then subject to stronger selection by the development of immunity to T cell attack through reduced or absent functional surface expression of HLA-A, B, C determinants.

The cellular and serological basis on which HLA-D region or HgII determinants were defined suggested a limited tissue distribution, specifically on peripheral blood B lymphocytes and also monocytes or macro-
phages. An early cellular study also showed the presence of these molecules on the surface of endothelial cells.\textsuperscript{60} This distribution was confirmed by initial studies with monoclonal antibodies, which in addition hinted at their unexpected presence in other non-lymphoid tissues, in particular the kidney.\textsuperscript{38} Systematic surveys of the distribution of HLA-D region HgII determinants subsequently showed their widespread, though not universal, presence on endothelial cells and an occasional patchy distribution on epithelial cells, including on colorectal carcinomas.\textsuperscript{61,62}

The clue to the probable clinical importance of this distribution of HgII determinants came first from the observations of Lampert \textit{et al}\textsuperscript{63} and Mason \textit{et al}\textsuperscript{64} that graft versus host disease induced the expression of HgII products in adjacent epithelial tissues. This was followed by the observation of Bottazo \textit{et al}\textsuperscript{65} that HgII determinants were expressed on thyroid follicular cells in patients with Graves' disease. As a result of this came a general indication that autoimmune reactions are associated with induction of expression of HgD products on affected epithelial cells. They pointed out that such cells expressing HLA-D are able to present antigen to lymphocytes, and that this may be a contributory factor to the development of autoimmunity.\textsuperscript{32} At about the same time that these observations were made it was shown by Wallach \textit{et al}\textsuperscript{66} and subsequently by many others that \(\gamma\)-interferon could induce the expression of HLA-D region determinants on a wide variety of cell types, including in particular, epithelial cells. As \(\gamma\)-interferon is released by T cells during immune stimulation, this can explain nearly all the occasions when epithelial cells have been observed to express HgII determinants. This expression may be an important adaptation that enhances T cell response to viral determinants expressed on infected epithelial cells. The occasional triggering of autoimmunity may be an unfortunate side effect of this mechanism for enhancing the body's immune response to viral infections.

The first observation of unexpected expression of HgII determinants on tumours was the description given by Winchester \textit{et al}\textsuperscript{67} of their presence on certain human melanomas, an observation that was subsequently confirmed by many others. There is no evidence that the normal melanocyte expresses HgII (Bobrow L, personal communication). Perhaps the expression of HgII on melanomas is induced initially by \(\gamma\)-interferon, or some other similar differentiation inducer, and is associated with the expression of other cell surface molecules, such as growth factor receptors. This is supported by the indication that \(\gamma\)-interferon may induce the expression of the surface receptor for tumour necrosis factor.\textsuperscript{68} Surface expression of a growth factor receptor may then provide a strong selective advantage for tumour progression, and the expression of HgII products may simply be a by-product of this change. Interestingly, Ruiter \textit{et al}\textsuperscript{69} have observed that melanomas also often lack expression of HgI determinants, and they have suggested that there is a better prognosis associated with melanomas that have a high expression of HgI and lack HgII than those that express HgII and have a low or absent expression of HgI. The changed expression of the HLA-A, B, C or HgI products may be connected with escape from T cell immunity, as seems likely to be the case for other tumours.

Monoclonal antibodies which react differentially with the HLA-DR, DQ, and DR products have been used to show that in leukaemias and lymphomas at least coordinate expression of all three sets of products does not always occur.\textsuperscript{70-72} These observations have recently been extended to colorectal carcinomas for which Ghosh \textit{et al}\textsuperscript{73} have shown that DR followed by DP, is most abundantly expressed, DQ being present in much smaller quantities. These authors state that HgII expression did not correlate with the Dukes stage of the tumours, though all their seven Dukes' A expressed HgII, while only 16 of 25 of the remainder did so. Observations in our laboratory suggest that increasing expression of HgII determinants on colorectal adenocarcinomas may be associated with a tendency to be better differentiated (Richman P, personal communication). These patterns of differential expression of the HLA-D subregion products suggest some differences in function. As evolution is a somewhat messy patching up process, however, a clear cut functional distinction between these products should not necessarily be expected.\textsuperscript{32}

Monoclonal antibodies are proving to be essential for the clear cut characterisation of HLA products in normal tissues, in tumours, and in other disease states. Variations in expression may have important functional implications and justify investigation of HLA expression as a part of normal histopathological investigation. For this purpose, it is extremely valuable to have antibodies that are effective in routinely fixed wax embedded sections. Results with an antibody against the HLA-DR subregion \(\alpha\) chain, 1B5, have been described by Epenotos \textit{et al}\textsuperscript{74} and show the value of such a reagent for routine immunohistology. It should now be possible to make a wider range of these reagents that will help in the investigation of the expression of HLA region determinants, especially in different sorts of tumours and their metastases.

**Conclusion**

The HLA region constitutes about 1/1000th of the human genome, corresponding to a size of about 3 \(\times\)
The HLA system: structure and function

10^6 base pairs. The total coding sequences within the region, which include 6 D region α chains, 10 β chains, about 25 class I chains, and the genes for C4, factor B, C2 and 21-hydroxylase code (including pseudogenes) for a total of about 14 500 amino acids. The proportion of the DNA sequence which codes, or which potentially codes for a protein—the coding ratio—is therefore 3 × 14 500/3 × 10^6, or about 1:70. This is similar to coding ratios found for other known genes, but nevertheless leaves room for further functionally expressed genes which could perhaps have a role in disease associations,11 and the tumour necrosis factor genes are a possible example. A complete characterisation of the set of expressed genes in the HLA region would immediately limit the possible explanations for the effects of HLA variation on a wide range of inherited susceptibilities to some major chronic disease.

The structures and functions of molecules are inextricably interrelated. Thus if we are to understand the function of the HLA products at a molecular level it is essential to establish their three dimensional structure using the techniques of x-ray crystallography, and through this and other approaches work out the functional basis for their interaction with the T cell receptor and other key molecules with a role in the immune response. Such structural knowledge, together with the power of recombinant DNA and cellular techniques, should enable us to unravel the role of the HLA genes in transplant rejection, immune response, disease associations, and tumour progression.

References

6 Spies T, Morton CC, Nedosposav SA, Fiers W, Pious D, Strominger JL. Genes for the tumour necrosis factor α and β are linked to the human major histocompatibility complex. Proc Natl Acad Sci USA 1986;83:6999–702.
8 Möller G. Immunological reviews No 84. Copenhagen: Munksgaard, 1985a.
29 Pollack MS, SAFAI B, DUPONT B. HLA-DR5 and DR2 are susceptibility factors for acquired immunodeficiency syndrome with Kapoai’s sarcoma in different ethnic subpopulations. Disease Markers 1983;1:135–9.
32 Bodmer WF. HLA today. Hum Immunol (In press).
35 Austin P, Trowsdale J, Rudd C, Bodmer W, Feldmann M,


Requests for reprints to: Sir Walter Bodmer, FRS, ICRF Laboratories, PO Box 123, Lincoln's Inn Fields, London WC2A 3PX, England.