Perspective

Hepatitis B virus, the vaccine, and the control of primary cancer of the liver*

Baruch S. Blumberg
Fox Chase Cancer Center, Philadelphia, PA 19111

This paper describes a project in basic research, which initially was designed to understand the role of human genetic polymorphisms in relation to inherited susceptibility to disease. At the onset, there was no obvious practical application of this project. However, building on a large body of research on hepatitis over the preceding decades, these studies resulted in the discovery of the hepatitis B virus (HBV), diagnostic methods for viral detection, and a vaccine. These applications have had a major impact on worldwide medical and public health problems. The discoveries and applications have saved many lives and also have been of significant economic value. Additional information and references are given in refs. 1–4.

Viral Hepatitis

Viral hepatitis is an inflammation of the liver caused by at least six different, mostly unrelated, viruses [hepatitis viruses A, B, C, D, E, and G (HAV, HBV, HCV, HDV, HEV, and HGV, respectively)]. It can occur in an acute form from which most patients experience a complete recovery. Acute viral hepatitis is characterized by an insidious onset, often with fever and severe malaise and loss of appetite for food, alcohol, and tobacco. Flu-like symptoms may occur early in the illness. The characteristic finding that occurs in many cases is the development of jaundice, a dramatic yellow discoloration of the skin and other surfaces. Symptoms may last for days or weeks. The acute disease usually results in complete recovery with lifelong immunity. Occasionally, acute hepatitis may advance to the fulminant phase; the patient does not recover, but develops liver failure and death may be rapid. Fortunately, this is rare. Hepatitis due to HAV and HEV is nearly always acute. Acute disease also may occur with HBV and probably HGV. HCV is usually chronic.

Some patients infected with HBV, HCV, HDV, and probably HGV may develop chronic infection. This may follow an acute attack; the virus does not resolve, but remains active or sub-active in the body for many years. More and more of the liver cells are destroyed, scarring occurs, and liver function decreases. Chronic liver disease can be life shortening. For at least two of the viruses, HBV and HCV, primary hepatocellular carcinoma (HCC) may develop, usually many years after the initial infection. The probability of HCC increases if the chronic carriers also are exposed to other agents, such as the carcinogen aflatoxin, which is produced by Aspergillus fungus that contaminates poorly stored foodstuffs. There is also epidemiological and other evidence that high body iron stores, which can be a consequence of excessive iron intake, also can increase the probability of HCC.

Chronic infection may occur without precedent acute hepatitis. Many individuals when infected with HBV, particularly in infancy and childhood, do not initially develop any symptoms of disease, but remain asymptomatic carriers of the virus for years and even decades. In some cases, the virus may be cleared, but in many the virus will remain active and/or sub-active for decades and lead to the chronic hepatitis and primary HCC already described. Chronic carriers of HBV are common (5–20%) in sub-Saharan Africa, much of Asia and the Pacific Ocean Islands, in parts of South America, and in southern and eastern Europe. There are carriers in northern Europe and in North America, but the prevalence is lower (less than 2%).

HAV and HEV are spread by fecal contamination of food and water and can occur in epidemic episodes. HBV, HCV, HDV, and HGV are spread by the transmission of blood and possibly other body fluids from one person to another. HBV is transmitted venereally as is HCV, although HCV is much less infectious by this route. HBV can be transmitted from mothers to their children; rarely before birth, more commonly at the time of birth or after birth when there is intimate interaction between mothers and their offspring. Less commonly, HCV also can be transmitted in this manner. HDV only infects persons who also are infected with HBV and is dependent on HBV for infection and, indirectly, replication.

HBV and HCV can be transmitted by the transfusion of blood from an infected to an uninfected person, or by the reuse of improperly cleaned needles or other equipment that carries infected blood. Drug abusers who share needles commonly transmit hepatitis B and C.

Viral hepatitis is one of the most common infectious diseases of humans, and HCC is one of the most common cancers of the world (ranking within the top 10). In some parts of Asia it may be the first or second most common cancer and represents a major public health problem. Treatment is difficult and often unsatisfactory. There are more than 350,000,000 carriers of HBV in the world, and about 1 million in the United States. Many of these are at risk of life shortening chronic hepatitis and cancer. It has been estimated that there are about 5,000 deaths yearly in the United States, and about 1 million worldwide due to HBV.

Hepatitis is often a disease of armies, or of people living under poor hygienic conditions in confined areas. It has been a major cause of morbidity and mortality in war, including the Vietnam conflict, where it was one of the most common causes of hospitalization. The early investigators in the field (S. Krugman, S. Sherlock, F. Deinhardt, M. Hilleman, and many others) had by the mid-1960s defined at least two forms of the disease, inferred its viral etiology, and described in general

Abbreviations: Au, Australia antigen; HBV, hepatitis B virus; anti-Au, antibody against Au, antibody against surface antigen of HBV; anti-HBs, antibody against the HBV core antigen; anti-HBs, antibody against HBsAg; HAV, hepatitis A virus; HBsAg, hepatitis B surface antigen; HCV, hepatitis C virus; HDV, hepatitis D virus; HEV, hepatitis E virus; HGV, hepatitis G virus; HCC, hepatocellular carcinoma.

*This article provided a source for a soon to be published case study entitled “The Viral Hepatitis Story” which will appear in a new series, Beyond Discovery: The Path from Research to Human Benefit (see http://www2.nas.edu/bsi). This series from the National Academy of Sciences was designed to increase public awareness of the practical benefits of basic science.
terms the epidemiology and clinical course. However, none of the viruses had been identified.

Inherited Susceptibility to Disease

One of the most intriguing questions in clinical and preventive medicine is why some people and not others succumb to disease when all are exposed to what appears to be an equal hazard. Luck may play a role, but there are also inherent and inherited differences between individuals that can account for some or all of the variation in response. I first became aware of this diversity when, as a medical student, I spent several months providing medical service and doing research at a hospital in the heavily forested area of northern Suriname in South America. We found a profound difference among different populations in the prevalence of carriers of the microfilarial form of *Wuchereria bancrofti*, the causative agent of elephantiasis in that region. Populations of heterogeneous origins whose ancestors had been brought to Suriname over several centuries to work in the sugar and other plantations lived in company-provided housing in an isolated mining community. People of African origin had a higher prevalence of microfilaria infection than did those derived from Indonesian or Chinese antecedents. These populations lived under similar environmental conditions and appeared to be equally exposed to *Culex fagitans*, the mosquito species that is the main vector for *Wuchereria* in the area.

Inherited variation and its relation to disease susceptibility can be studied by looking at those with disease and those without, and then attempting to discover the inherited biochemical differences; or, inherited biochemical variation can be identified first and then its association with disease determined. We decided to use the latter process for several reasons. In the mid-1950s, when we started our research, little was known about biochemical variation in normal individuals. Looking for the variation in the first place had the advantage that any variation that was discovered would be a contribution to the general understanding of inherited heterogeneity and would add to the accumulating catalog of these traits. In the 1950s, before the revolution in molecular biology, it was not possible to study variation at the nucleic acid level. However, phenotypes could be studied, and the gel electrophoresis technique first used in the mid-1950s made it possible to detect small differences among the blood proteins of different individuals. There was reason to believe that many of these differences would be inherited. We embarked on a project to determine the population distribution of the phenotypes of inherited systems already known (i.e., the hemoglobin-binding serum haptoglobins and the iron-binding serum transferrins) and search for serological biochemical variation that had not been previously seen. There were several scientists in the United States, Europe, and elsewhere already working in this field. It was a scientific interest group characterized by generous sharing of serum specimens, techniques, and results.

Most of this variation fell into the category of genetic polymorphisms, a term first introduced by the Oxford lepidopterist and genetic ecologist E. B. Ford. He had studied inherited variation in wing markings of butterflies and moths in relation to changes in the color background of fields and forests as a consequence of industrial contamination of the environment. His original definition is still useful:

\[ \ldots \text{the occurrence together in the same habitat of two or more discontinuous forms or phases, of a species in such proportions that the rarest of them cannot be maintained merely by current mutation.} \]

He postulated that the balance of gene frequencies was maintained by selection pressure that had different values for different phenotypes. Disease associations could account for the different selective forces, and it would be expected that the phenotype frequencies would differ greatly in populations living in different environments and subject to different disease risks.

The project initially was inductive in nature; that is, we collected data on the distribution of polymorphisms with the expectation that the observations would lead to the formulation of hypotheses. Research of this character is currently difficult to fund, because granting committees usually want to see a well formulated hypothesis to guide research. However, inductive research has the advantage of generating data unrestricted by previous notions of outcome and can result in totally new ideas. This phase of the work was done at the National Institutes of Health in Bethesda.

The execution of the project required collecting blood specimens (often as part of public health surveys), determining the distribution of known polymorphisms, and discovering hitherto unknown systems. We undertook field studies in northern Spain, Nigeria, Alaska, the Marshall Islands in the Central Pacific, northern Europe, among Native American and other populations in North and South America and eventually, many other places. Additions were made to the list of polymorphic systems found in blood and to their distribution in different populations and geographic locations.

In the summer of 1960, in collaboration with A. C. Allison (then at the Medical Research Council Laboratory at Mill Hill, London), we introduced a method for finding antigenic serum polymorphisms. By 1960 it was known that there were many serum protein polymorphisms. It was inferred that some of these could be antigenic, and if a patient received many transfusions it was likely that he or she would have been exposed to a protein that had not been acquired or inherited. This could result in the development of antibodies in the transfused person. These antibodies then could be used to identify the antigenic variant that had not been inherited by the patient, but had been inherited by the blood donor or others in the population.

Using the immunological technique of double diffusion in agar gel, an antibody that reacted with the low-density lipoproteins of human sera was identified. As predicted, it identified a complex polymorphic system, which we initially called the “Ag system.” We found an association with diabetes, but because of the difficulty of standardizing the human antisera this line of research was not pursued. Subsequently, K. Berg in Norway developed a series of rabbit antisera against lipoproteins that determined a different complex antigenic polymorphism of the serum lipoproteins that has revealed important disease associations.

**Australia Antigen and HBV**

The hypothesis that new polymorphic systems could be discovered by the use of blood from transfused persons had proven to be productive. We continued the search for additional antibodies and antigens. In 1963, we found an isoprecipitin that reacted with an antigen that appeared to be different from the lipoprotein Ag system we had initially discovered. It had fewer or no lipid-staining characteristics, and the antigen had a much different distribution than the Ag antigens. The new antigen was very rare in Western populations, but common among Australian Aborigines, Micronesians, Vietnamese, and Taiwanese. We made the curious and interesting observation that it was also common in patients in the United States with leukemia. This “new” antigen-antibody reaction had been initially seen by H. Alter, who was then working in my laboratory at the National Institutes of Health in Bethesda and subsequently has become one of the leading hepatitis researchers.

The next year my laboratory moved to the Institute for Cancer Research at the Fox Chase Cancer Center in Phila-
We referred to the serum antigen as “Australia antigen” (Au), because the earliest studies were done on the serum of an Australian Aborigine. We formulated several hypothesis based on the association of Au with leukemia:

(i) Individuals with Au have an increased susceptibility to leukemia, and this susceptibility is inherited.

(ii) Leukemia causes Au.

(iii) Au is related to the virus that had been postulated to be the cause of leukemia.

Family studies of carriers were consistent with a genetic hypothesis. However, we also considered the possibility of infection. During the course of the next few months, a series of observations inclined our thinking to the possibility that Au was related to a hepatitis virus. We found Au in the blood of a patient who had developed hepatitis and in the blood of several people who had received blood transfusions. These findings established a mind set, and we started to collect sera from patients with hepatitis to test the hypothesis that Au was associated with this disease. However, the event that triggered a much more intensive testing of the Au/hepatitis hypothesis occurred in one of our own patients.

We had embarked on the testing of the hypothesis that certain individuals were susceptible both to the development of Au and leukemia. A corollary of this hypothesis is that individuals who are at risk for leukemia would have a greater probability of also having Au. Children with Down syndrome, mental retardation associated with trisomy of chromosome 21, are at much greater risk than other children of developing a form of leukemia. We tested the blood of Down syndrome patients in a large public institution for the mentally retarded and found that our hypothesis was amply supported. About 30% of Down syndrome patients had Au compared with about 5% in other mentally retarded patients in the same institution. The support of our hypothesis was very encouraging, and it also allowed us to observe persons with Au over an extended time. We found that usually individuals who were negative when first tested remained negative and those who were positive remained positive. On one occasion, contrary to expectation, a patient who previously had been negative was found to be positive on a subsequent testing. He developed hepatitis coincident with the appearance of Au in his blood. This observation accelerated our focus on the hepatitis hypothesis, and we soon determined that there was a significant association of Au with acute and chronic hepatitis.

Our colleague K. Okochi, then in Tokyo, had identified Au and anti-Au shortly after our discovery, and he soon confirmed our findings. Investigators in the United States and Italy also provided substantial data on the association between clinical hepatitis and the presence of Au in the blood. M. Bayer and B. Werner in the Philadelphia laboratory, using electron microscopy, identified particles with the appearance of a virus in the serum of individuals with Au. These were subsequently shown to be particles that contained only the surface antigen of the virus and were neither infectious nor pathogenic. The whole virus particle was identified sometime later by the British scientist D. S. Dane and his colleagues. I will return to this later in the discussion of the hepatitis vaccine.

I. Millman and V. Coyne, in our laboratory, were able to localize Au in the liver cells of patients with hepatitis. They also were able to sustain the growth of the virus for weeks and months in liver cells obtained from patients with liver disease associated with Au, although this observation was difficult to confirm.

We distributed sera containing the antibody or the antigen to other investigators. The immunological method used for the detection of Au, double diffusion in gel, was easy to duplicate, and the equipment to do the experiments consisted only of glass slides and agar gel. Hence it was easy for others to repeat the observations even in laboratories with a minimum of laboratory equipment. There were numerous confirmations of the initial studies. In retrospect, the wide distribution of the reagents had a major impact on the speedy acceptance of the identification of HBV.

Based on the contributions of many clinical and laboratory investigators whose work preceded our own, it was known that there were at least two hepatitis viruses; HAV or infectious hepatitis, transmitted by the fecal-oral route, and HBV or serum hepatitis, transmitted by blood and serum. The virus associated with Au was most similar to the HBV, and the designation HBV became accepted. Australia antigen, the surface antigen of HBV, was termed HBsAg.

What happened to our interest in genetic polymorphism, leukemia, and Down syndrome after the main thrust of the research tracked the hepatitis virus? We subsequently realized that the virus had been transmitted to the leukemia patients by transfusion of blood contaminated with HBV, and that the high prevalence in the Down syndrome population was a consequence of crowding in large institutions. But, the answer was not that simple. Leukemia patients were more likely than other transfused patients to become carriers of HBV; Down syndrome patients were more likely than patients with a different diagnosis within the same institution to become carriers. We concluded that there were specific immunologic and genetic characteristics of these diseases that indicated a specific affinity with HBV. We tested the hypothesis that the virus thought to cause certain forms of human leukemia might be related to HBV, although this line of research was not continued. Additional studies, some of which are still in progress, were consistent with a genetic susceptibility to persistent infection with HBV, which is part of a complex interaction of polymorphic systems. Hence, the research on genetic polymorphisms related to differences in disease susceptibility was central to the discovery of HBV.

The Control of Post-Transfusion Hepatitis B

The clinical implications of the findings were quickly realized. In the United States in the 1960s a large percentage of blood donor units were obtained from paid donors, although this was not the case in many countries in Europe, in Canada, and elsewhere. The incidence of post-transfusion hepatitis, both clinical and occult, was enormous. In some studies, post-transfusion hepatitis occurred in up to 50% of patients who were receiving large numbers of transfusions for extensive surgical treatments such as open-heart surgery. There were casual observations that recipients of donor blood containing Au developed hepatitis, but there were no systematic observations on the extra risk imposed. Okochi undertook such a study in Tokyo and reported a significantly increased risk for hepatitis in recipients of the positive donor units.

We initiated a study at Philadelphia General Hospital, where the incidence of post-transfusion hepatitis was high, to determine the efficacy of a blood screening program. The results were very convincing. Before the screening program the incidence had been 17.9%, and after instituting it had been reduced by two-thirds to 5.9%. In retrospect the residual cases were a consequence of the low sensitivity of the immunodiffusion method and cases due to non-B hepatitis. It is now known that most of these cases were due to HCV. The results indicate that, at this time, most of the post-transfusion hepatitis was a consequence of hepatitis B.

The possibility of drastically reducing post-transfusion hepatitis galvanized the medical, health care, and commercial
medical communities. There were frequent meetings of government and professional societies, such as the American Association of Blood Banks. Eventually legislation and regulations required the testing of blood donor units for HBsAg. There was a scramble by commercial companies to produce kits of increasing sensitivity to tap the enormous market for the testing the more than 10 million units of donor blood collected each year. We had introduced, and Fox Chase Cancer Center had patented, a sensitive radioimmunoassay for HBsAg, and this technique and similar ones were commercialized. With the improvements in the tests and the nearly universal requirement of donor blood screening, post-transfusion hepatitis due to HBV became a rarity although cases of hepatitis due to other viruses continued to be a problem until detection methods for HCV became available in the 1990s. There are still other causes of post-transfusion hepatitis (hepatitis G for example), but the problem is far less grievous than it was before the introduction of the screening tests.

**Invention of the Hepatitis B Vaccine**

It is likely that the most important outcome of the research on Au and HBV has been the invention, development, and application of the vaccine against HBV. This came about as a consequence of a series of events that coalesced in 1969. The data confirming the hypothesis that Au was a component of HBV were accumulating. Because I had experience in preventive medicine, primarily because of my work in Suriname and elsewhere in the tropics, the advantages of a vaccine were obvious. The experience of the South American jungles, where large segments of the population were devastated by infectious diseases (malaria, tuberculosis, yaws, diarrhea-causing bacteria, elephantiasis, and many others) provided the motivation to find a vaccine that would prevent large-scale infection.

The arrival of Millman in our laboratory in June 1967 was a major factor in the invention of the vaccine. He was trained as a microbiologist and had developed a purified pertussis vaccine while working for Merck in its vaccine facility near Philadelphia. His reaction to the data were influenced by his experience with other infectious agents. We had calculated the amount of Au in the serum of carriers and estimated that in some it amounted to about 1% of the serum proteins. His immediate response was that if this was all virus it would be incompatible with the life of the carrier. This became of considerable importance later when we devised the vaccine.

Sometime in 1969, Timothy Talbot, the director of the Institute for Cancer Research, told us that the United States government, which funded most of the basic research in medical science, would expect scientists to generate their own funds to replace a percentage of government grants. The inference was that commercial exploitation and patenting of discoveries would be encouraged. We decided to apply for a patent for a HBV vaccine. We knew that the Au particles, which we had isolated, and which Bayer and Werner had visualized in the electron microscope, did not contain nucleic acid. We inferred that these were nonreplicating, incomplete forms of the virus made up entirely of surface antigen, and that there must be additional particles, which contained nucleic acid, that were replicating, infectious, and pathogenic. The surface antigen particles were highly resistant to proteases, which allowed their purification from the serum proteins in the blood. In animal experiments Millman and London found that partially purified Au particles that presumably also contained the whole virus particles, which we had not yet visualized, could be transmitted by inoculation into experimental animals. The fully purified particles from which the whole virus had been removed were not infectious. The implication was that we could separate the noninfectious particles containing only the surface antigen from the pathogenic whole virus particles.

Based on these findings we submitted a patent application in 1969 for the extraction of the surface antigen particles from the serum of human carriers of the virus. The particles were purified by centrifugation, enzyme digestion, column gel filtration, differential density centrifugation in sucrose and cesium chloride, and dialysis. Afterward, standard adjuvants and preservatives were added. This was a unique method for producing a vaccine, which had not previously been used. It took some time before the concept was accepted by virologists and vaccine manufacturers who were more accustomed to dealing with vaccines produced by attenuation of viruses, or the use of killed viruses produced in tissue culture, or related viruses that were nonpathogenic but protective (i.e., smallpox). However, by 1971, we were able to interest Merck, which had considerable experience with vaccines. During the next few years, a series of human and primate observations by scientists including Hilleman (who was responsible for vaccines at Merck), S. Krugman, R. Purcell, P. Maupas, and others provided additional support for the vaccine. In 1980, the results of the first extensive field trial were published by W. Szmuness and his colleagues in New York City. They showed that the vaccine was highly effective (over 90% were protected) and that no untoward side effects were observed. The appropriate FDA approvals were obtained, and by 1982 the vaccine was available for general use. Soon afterward, recombinant hepatitis B vaccine was developed in several laboratories, and enormous quantities of the vaccine have been produced by this method. HBV vaccine is the first, and so far, only vaccine to be produced commercially by recombinant methods. Vaccine still is produced from the blood of carriers in China, Korea, and elsewhere, and at present, remains a major source of supply.

**The International Hepatitis B Vaccination Program**

The major methods for the transmission of HBV are:

(i) From a carrier mother to a child, at, or soon after the time of birth, and, in some cases, probably before birth.

(ii) From a carrier mother and/or carrier siblings in the first few years of life.

(iii) Venereal.

(iv) By transfusion or injection of blood from a carrier to another person.

In many cases the mechanism for transmission cannot be determined.
People who are infected with HBV in infancy or childhood have a greater risk of becoming carriers. The frequency of transmission and the prevalence of carriers can be greatly decreased by the vaccination of children soon after birth, preferably within the first few months. Universal childhood vaccination has been adopted in more than 70 countries. Programs are still required in other countries, particularly in Africa and the Indian sub-continent where the need is great but the resources are inadequate.

The vaccine is highly effective in decreasing the prevalence of carriers. Several reports were presented at the IX Triennial International Symposium on Viral Hepatitis and Liver Disease in Rome in 1996. In Taiwan, where universal childhood vaccination has been used for more than 10 years, the prevalence of HBV in the 1- to 2-year age group decreased from 10.7% in 1984 to 1.5% in 1989, and the decrease since that time has been even greater. A vaccination program was undertaken between 1981 and 1983 in Native Americans in Alaska. Acute hepatitis B infection dropped from 215 cases per 100,000 before the vaccination program, to between 7 and 14 cases in 1993. In 1995, no cases were reported.

A vaccination program was started in 1983 in Afragola, a community in southern Italy near Naples, selected because of its high rates of HBV carriers and morbidity and mortality from liver disease. There were major decreases in HBV prevalence in the vaccinated population; for example, the prevalence of HBsAg in males aged 5–10 dropped from 11.9% in 1978 to 1.6% in 1989 after the vaccination program was in place for about 5 years. There were similar decreases in other age groups. Of particular interest was the drop in the prevalence of HBsAg and anti-HBc (an indicator of viral infection) in the unvaccinated population. HBsAg dropped from 13.4% in 1978 to 7.3% in 1989. This implies that the reduction of the carrier prevalence in the vaccinated group has an indirect effect on unvaccinated carriers and susceptibles, and suggests that there may be an amplification characteristic to the vaccination program. There are other examples in the People’s Republic of China, Korea, Gambia, and elsewhere that demonstrate the considerable success of the vaccination programs. These programs have raised the question of the eradication of HBV, particularly if it becomes possible to treat carriers and to extend the vaccination programs. This possibility already had been considered at the International Conference on Prospects for Eradication of HBV in Geneva in 1989.

There are indications that the incidence of HCC, much of which is due to chronic infection with HBV, also is decreasing. At the 1996 meeting in Rome it was reported that the incidence of HCC has significantly decreased among children in the vaccinated population from Taiwan. In a study from Korea, the results indicate that not only the impacted vaccinated group, but also several unvaccinated cohorts, had a decrease in the incidence of HCC. This has happened after only 10 years of the vaccination program. There is good reason to be hopeful that cancer of the liver will decrease significantly in many parts of the world where HBV is the main cause of the cancer. M. Kane, who directs the hepatitis activities for the World Health Organization, has said that the prevention of HCC is now considered to be the second most important cancer control program, exceeded only by the campaign against cigarette smoking.

Conclusion

An apparently esoteric research project on inherited biochemical and immunologic variation resulted, unexpectedly, in the identification of a virus and the development of diagnostic methods and a preventative vaccine that have been of enormous medical and public health importance. At the inception of the project it would have been impossible to foretell that it would lead to large decreases in morbidity and mortality, and that the results would be of economic benefit.

Funding for basic research is based on the faith that it will, in due course, lead to an outcome that benefits society. Many practical applications of science have derived from this belief, and many more will in the future. Scientists have the responsibility to inform those who provide funds for research about how this belief has been vindicated.

This work was supported by U.S. State Public Health Service Grant CA-06927 from the National Institutes of Health and by an appropriation from the Commonwealth of Pennsylvania.