The unexpected pathway to the creation of the HbA1c test and the discovery of AGE’s

A. Cerami
From the Leiden University Medical Center, Leiden, The Netherlands

Keywords: AGE’s, ageing, diabetes, diabetic complications, Hemoglobin A1c, Maillard reaction.

Introduction
In a previous study for this journal I discussed how failure in one project might give insight into another project, resulting in remarkable discoveries with significant impact, [1].

In this study, I will discuss the unpredictable road that led to the creation of the diagnostic test, HbA1c. I have found in my lifetime of scientific endeavours, that trusting the truth of clues arising from experiments, can lead to unexplored, unexpected and exciting new understanding of disease processes and treatments.

My deep motivation has always been to understand the causes of diseases and how to treat them; yet the persistent pursuit of this has led me across many different scientific disciplines. The flexibility and willingness to enter areas that are not areas of my expertise, aligned with my dogged persistence, have been and are presently crucial.

In this study, the story of the hope for defining a new treatment for patients with sickle cell disease, let to the discovery of a new diagnostic test, HbA1c, that has become the worldwide mainstay for the treatment and diagnosis of diabetes. It is especially comforting to know that using this test, many women with type 1 diabetes have delivered healthy babies.

Chemical modification of haemoglobin
This story begins at a social affair in New York for a group of haematologists. At that time I was an Assistant Professor at the Rockefeller University trying without success to isolate erythropoietin from the urine of patients with aplastic anaemia. One of the participants, Alan Schecter, mentioned that there had been a recent report of the successful treatment for sickle cell anaemia crisis in patients with large amounts of urea [2]. The rationale that was given by the investigators was that urea would disrupt the polymerized haemoglobin S in the red cells that cause entrapment in capillaries and in turn cause severe pain and tissue damage. I responded that I could not question the validity of the clinical trials, but I could not believe that one could ever give enough urea to achieve this because the amounts of urea needed to break hydrogen bonds are in the molar range. However, the possibility existed that the clinical effects could be the result of the carbamylation of HbS with cyanate because urea in solution is in equilibrium with ammonium and cyanate ions. The finding that urea solutions were in equilibrium with cyanate that could react with proteins had been first observed with the enzyme ribonuclease [3]. Working with James Manning and Martha Fedorko, we were able to demonstrate that cyanate could react with the amino terminal valines of haemoglobin S and prevent the sickling of cells following deoxygenation in vitro [4] and in vivo[5].

The question was whether this observation could be translated into a therapeutic modality for these patients. We were encouraged to go forward because urea is present in vivo and small amounts of carbamylated haemoglobin could be measured in normal people. At first, the clinical results were very encouraging [6]. The cyanate-treated patients showed less red blood cell destruction, higher haemoglobin levels and reduced incidence of crisis. However, after longer periods of time (several years) on the drug, several patients developed peripheral neuropathy and cataracts similar to that seen in diabetics. We immediately stopped administering the drug to all the patients and continued to follow the patients. Fortunately, the neuropathy and the cataracts reversed with time. It was obvious that the oral administration approach to treat the disease was not going to work. Subsequently, Deiderich et al. [7] reported that extracorporeal red cell carbamylation of patients with sickle cell disease had positive haematological and clinical responses, but the commercial development of the apparatus did not take place.
Needless to say the failure of the project was quite devastating, but gave me a very good scientific clue. As time went on I kept thinking about the fact that the neuropathologist and the ophthalmologist both noted that the cyanate-induced damages were similar to that seen in diabetes. I also remembered that there were reports [8, 9] of patients with diabetes having an increased amount of the minor haemoglobin called HbA1c. The structure was unknown, but appeared to be a sugar attached to the amino terminus of the beta chain of haemoglobin – the same place that cyanate reacts. This led me to consider that, perhaps, there was a chemically reactive compound in the blood of diabetics that could react with haemoglobin to form haemoglobin A1c and with other body proteins to explain the pathogenesis of diabetic complications. As I pondered over this possibility, I happened to visit with a friend, Scott Grundy, and told him my thoughts. During dinner, we talked about whether this could be true and how to prove it. He thought it was an interesting idea and encouraged me to follow up on this hypothesis.

Haemoglobin A1c

On returning to New York, I suggested to Ron Koenig, who was in the first class of students in the Rockefeller/Cornell MD/PhD programme, that we do a few simple experiments to see whether this idea was correct. The first experiment was to see whether diabetic mice also had the mouse equivalent to human HbA1c. Using the same method used for human blood, he was able to show that diabetic mice had a minor haemoglobin fraction that corresponded to that seen with haemoglobin A1c in human red cells and that diabetic animals had 2.8 times more than nondiabetic mice. The red cell is unique in that all of the haemoglobin and other proteins in the cell are produced primarily when the cell resides in the bone marrow. The question was when was HbA1c made? To address this question, radioactive iron was injected into normal mice, and a few days later, the radioactive blood cells were obtained. Aliquots of this blood were then injected into normal and diabetic mice. At intervals, small amounts of blood were removed and the amount of radioactivity in the haemoglobin A and haemoglobin A1c fractions was determined. When the ratio of the amount of radioactivity of HbA1c over that in HbA was plotted over time, it was clear to see that HbA1c increased linearly over the life of the red cell when the cells were placed in normal animals. In contrast, although the rate of HbA1c production was also linear when the cells were put into diabetic animals, the rate was 2.8 times faster (Fig. 1). In this first article, we suggested that measurement of haemoglobin A1c would be related to metabolic control, and the reactive substance that reacted with haemoglobin could be responsible for the complications of diabetes [10].

At the time this work was performed, there was an intense debate about whether the complications of diabetes were exacerbated by poor metabolic control. There were two camps of thought – the people who believed in ‘tight control’ and the other who did not believe metabolic control was important and thus promoted ‘loose control.’ We thought by measuring HbA1c in patients from their clinics that we could validate the idea that ‘tight control’ was important. We arranged to obtain blood samples from two institutions, which will go unnamed, representing each point of view. To our great surprise, the HbA1c values were between 11% and 13% compared with values of 4–6% in nondiabetics. This implied that care of patients with diabetes was far from optimum in both places in spite of what the doctors believed.
In collaboration with Joe Williamson, we measured HbA1c in a series of patients with diabetes and found that HbA1c correlated very well with fasting blood sugar and an oral glucose tolerance test [11]. At the time of this work in 1976, we thought that HbA1c would quickly replace the glucose tolerance test in clinical practice as a way to diagnose diabetes. This was not to be the case. It has taken over 30 years for the American Diabetes Association to embrace this test for diagnosis.

The Rockefeller University Hospital, where this work was carried out, was an ideal place for us to admit patients and learn how best to manage diabetes in a controlled environment. Ron Koenig [12] under the clinical direction of Charles Peterson was able to identify a number of patients who agreed to stay on the metabolic ward for extended periods of time so that we could determine how best to control blood glucose and determine its effects on HbA1c and other metabolic and clinical parameters. One of these patients was my mother who had Type 2 diabetes. She stayed in the hospital for 6 months whilst we learned how best to manage diabetes. As I write this I realize that this would probably not be allowed in most institutions today, but she was proud to participate in these studies knowing that the knowledge we gained would help others avoid the terrible complications of diabetes. After 3 weeks of the initiation of better diabetic control, there was a fall in fasting blood sugar, urinary sugar and HbA1c. Figure 2 displays the fall in HbA1c over time in a patient with Type 1 diabetes following initiation of glucose control.

The introduction of the HbA1c test to the medical community was met with considerable scepticism when it was first introduced. An example of this occurred when I suggested to the administrators at the Rockefeller that we should patent the measurement of HbA1c as a method to measure metabolic control in diabetics, I was told that this would be a waste of money because they were assured by several key diabetologists that the test would not be used because patients really did not want to know. I argued unsuccessfully that this was not the case and that a patent would promote an orderly development of new more rapid procedures, and standards so that comparison of data from different hospitals would be possible. As a result the fragmentation of the market into many providers has resulted in the absence of a primary standard that is used throughout the world. Reviewing the arguments brought up by nonbelievers of the concept of measuring control today is not very informative other than to say that it was very frustrating.

I have come to realize with age that this is a step that is necessary before adoption of new ideas.

One of the first groups of clinicians to embrace the concept of monitoring HbA1c was obstetricians who cared for pregnant Type 1 diabetic patients. At that time, the incidence of birth defects in children born to these mothers was >20%. By promoting tight control of blood glucose through use of HbA1c measurements, it was possible to reduce the incidence of birth defects to that found in nondiabetics [13, 14]. Another group that very quickly began to measure HbA1c were insurance companies. One day soon after the publication of the New England Journal article referred to above, I received a call from a commercial laboratory asking for technical help on measuring HbA1c. Out of curiosity I asked who would be using so many measurements. He then replied that insurance companies were routinely measuring HbA1c on the blood samples that they obtained when people were applying for insurance to detect patients with diabetes. I was horrified that a test developed to help diabetics could potentially be used against them.

There is no question that the DCCT in type 1 diabetics and the UKPD trial in type 2 diabetics convincingly showed that tight control as assessed by HbA1c could prevent the complications of diabetes. The DCCT, which was carried out by the NIH from 1983 to 1993, showed that the onset and progression of diabetic complications of the eye, kidney and nerve damage were significantly reduced in the patients with Type 1 diabetes.
who had stricter glycaemic control as evidenced by HbA1c measurements [15]. A follow on study to the DCCT, called the EDIC, continued to follow these patients and showed that the risk of having a heart attack, stroke or death from cardiovascular disease was less in the patient group with lower HbA1c [16]. Similar results were observed in the UKPDS with Type 2 diabetic patients. This trial was initiated by Bob Turner [17] in 1977 and ran until 1997. At the 10 year mark, there was a significant reduction in death, and diabetic eye and kidney disease. These studies are milestones in the field of diabetes. The evidence for the importance of tight control is overwhelming. As a reflective note I refer you to Fig. 3, which is a composite of these two studies in terms of the HbA1c values observed between the different degrees of control and the type of diabetes, as well as the range seen with nondiabetic people. As I mentioned earlier when we first studied HbA1c, the values we observed were often in the range of 10–18%. I presume that this difference reflects the increased knowledge of the importance of glucose control to patients or at least those that volunteered for clinical studies. The second observation is that relatively small differences in HbA1c between the groups over a period of 10 years results in significant differences in the development of complications. These two studies were critical in pointing to the importance of glucose control in patients with diabetes and resolving once and for all that tight control was important in the development of the complications of diabetes.

Pathogenesis of diabetic complications and ageing

In addition to HbA1c being a biological marker, we believed that the chemistry of the nonenzymatic reaction of glucose with other body proteins was responsible for the complications of diabetes and possibly some aspects of ageing in general. But first let me review the chemistry of the reaction of glucose with the amino terminal valine of haemoglobin and the epsilon amino groups of lysines in proteins (Fig. 4). The reaction of the amino terminal valine of haemoglobin with glucose occurs so readily because this amino group has a much lower pK. At neutral pH, approximately 50% of these amino groups have no charge. This compares with about 1% of the epsilon amino groups of lysine. Noncharged amino groups can react with the carbonyl of glucose to form a Schiff base. Thus, the amino terminus has a much higher probability of forming a Schiff base. This Schiff base can undergo a rearrangement to form what is called an Amadori product by chemists. The structure of HbA1c was found to be Haemoglobin A with an Amadori product attached to the amino terminus of the beta chain [18, 19]. This relatively stable product can rearrange slowly to release glucose and mannose. This back reaction accounts for the inability of HbA1c to accurately reflect an integral beyond approximately 1 month [20].

Reasoning that a similar reaction of glucose with other body proteins might be occurring, we began by studying the lens. The lens, like the erythrocyte, is not dependent on insulin for glucose transport and develops cataracts in rats and humans with diabetes. The incubation of lens crystalline proteins with glucose in vitro led to Amadori products and an increased susceptibility to form disulphide linked high-molecular weight aggregates that were similar to that seen in the cataracts of diabetic rats [21]. These results suggested that cataracts in patients with diabetes could be the result of glucose modification of lens proteins.

As we studied the chemistry of the reaction of glucose with proteins, we were introduced to the food chemistry literature. In the early part of the 20th century, the French physician/scientist, L.C. Maillard noted that reducing sugars like glucose could react with proteins and peptides when heated, to form a complex mixture of brown pigments now called Maillard products. Food chemists for the rest of the century studied this reaction because of its importance in cooking and storage of food stuffs. The initial Amadori product can undergo many rearrangements to form an extremely diverse series of compounds. For example, more than 900 chemical entities have been identified after the roasting of coffee beans. Most of us have carried out the Maillard reaction by putting sugar on the surface of a duck or ham before cooking and enjoying

Fig. 3 The HbA1c values of the DCCT (Type 1 diabetics) and the UKPDS (Type 2 diabetics) for the intensively treated patients compared with a comparisional group of patients are shown over a 9-year period of time.
the wonderful, tasty Maillard products. In addition to forming during cooking, the Maillard products can form even in the deep freeze. Thus, chickens stored in the freezer will become tough as a result of protein crosslinking if they are stored beyond 1 year.

Although Maillard in his articles suggested that this chemistry would be important in biology and medicine, we could not find any papers demonstrating that it did occur \textit{in vivo}. Vincent Monnier and I decided to look for Maillard products in biological samples. One of the characteristics of a small subset of Maillard products is that they have a fluorescence with a characteristic excitation and emission spectra. When we examined human cataracts for this signature fluorescence, we observed spectra that were the same as bovine lens crystallins that had been incubated with glucose [22]. Originally, we called this assortment of glucose rearrangement products as Maillard products as the food chemists had done. This proved to be confusing because many of the Maillard products studied by food chemists were formed under harsh conditions and temperatures. To distinguish those products formed under physiological conditions of temperature, pH, osmolarity etc., we coined the acronym AGE-advanced glycation end products. A subsequent study of dura collagen isolated from people aged 20–95 showed a linear increase in the amount of AGE’s with age and that diabetic individuals had an amount of AGE’s that was equal to someone twice their chronological age [23]; for example, Fig. 5 shows the accumulation of fluorescent AGE’s as a function of age and diabetes in the aorta.

Since these initial articles, there have been over 5000 articles cited in Pubmed that have examined AGE’s in various tissues as a function of diabetes or ageing. Figure 6 is a cartoon drawing of the human body where AGE’s have been suggested to be involved in the development of the complications of diabetes or ageing.

Another important bio-molecule that can react with reducing sugars is DNA. We were prompted to investigate this possibility because we knew that high blood glucose during pregnancy could cause damage to the foetus during the first trimester of pregnancy. A number of molecules that interfere with DNA metabolism are known to be teratogenic. The addition of glucose-6-phosphate to DNA solutions led to the formation of AGE’s that had the same fluorescence spectra noted above for proteins and caused breaking of the DNA strands [24]. Subsequently, Annette Lee and Rick Bucala [25] showed that there was a twofold increase in the mutation rate of mouse foetuses, which had the LacI gene, developing in a diabetic mother compared with a nondiabetic mother. Ageing animals also accumulate mutations over time. A linear increase in the number of mutations in the LacI gene was observed...
from birth to 24 months that we proposed was because of the reaction of AGE’s with DNA. The potential impact of DNA damage from AGE’s is in need of further work.

Several unexpected results occurred during the years that we worked on the chemistry of the Maillard reaction. The first occurred early in our work when we tried to develop an antibody assay for HbA1c. Utilizing this antibody, we observed that the antibody-derived data reflected the data obtained with the values obtained by chromatography. However, there was a subset of people who had lower values with the antibody compared with the chromatographic method. Investigation of the history of these patients revealed that this subset of people had a history of alcohol abuse. We reasoned that acetaldehyde formed in vivo by the oxidation of ethanol could be reacting with the Amadori product of HbA1c in such a way that the antibody could no longer interact. What we were able to show was that the incubation of acetaldehyde with red cells led to stable adduct formation in a time-dependent manner [26]. Further studies showed that the measurement of this adduct could estimate alcohol consumption for a previous 2–3-week period. Although we thought it would be a useful assay, several attempts to raise the money to pursue this were met with scepticism and not awarded. We reluctantly abandoned further work for nearly 20 years. At a laboratory meeting, we discussed in theory the possibility that acetaldehyde could form an adduct between the OH group of the Amadori product and the amino group of valine of haemoglobin. The importance of this proposed structure was that it could explain our previous data. More importantly it could explain how this adduct could stabilize and protect the Amadori product from undergoing further rearrangements to form AGE’s, in vitro and in vivo (Fig. 4). In the lapse of two decades, the development of chemical methods, such as isolation methods, and mass spectroscopy had improved greatly. In short order, we were able to prove that the structure that we had hypothesized was correct. We proposed in this article that the ‘red wine’ effect for preventing atherosclerosis was the result of acetaldehyde preventing AGE formation [27].

Another surprising result occurred when we were studying AGE’s in the arteries of diabetic individuals. Utilizing a polyclonal antibody that recognized AGE’s, we could find extensive accumulation of AGE’s in arteries of patients with diabetes that had limbs amputated because of critical limb ischaemia (CLI). The unusual finding was that a number of nondiabetic patients with CLI also had extensive amounts of AGE’s in their vessels. In discussing how this could be, Carla Cerami pointed out that the nondiabetic had been or were heavy smokers. A possible explanation was that highly reactive AGE’s are present in tobacco smoke that were being inhaled and getting absorbed through the lungs and eventually travel via the blood throughout the body. The preparation of tobacco utilizes classic Maillard chemistry. The leaves of the plant, which are laden with glucose, are heated in sheds to eventually form tobacco. A simple experi-

Fig. 5 Some of the AGE’s that accumulate over time in tissues have a characteristic fluorescent spectra. The amount of fluorescent AGE’s in the aorta increases linearly over time in nondiabetics. Diabetics have increased amounts of these AGE’s compared with age matched nondiabetics.

Fig. 6 AGE’s have been implicated in the pathogenesis of microvascular and macrovascular complications of diabetes and ageing.
ment revealed that passing cigarette smoke through a protein solution led to the accumulation of Maillard products on the proteins. Animals exposed to cigarette smoke had their fur turn yellow-brown but also accumulated AGE’s on serum and body proteins. As many of the consequences of smoking are a form of accelerated ageing, we put forth the concept that the accelerated ageing of smokers could be the result of the inhalation of AGEs in the smoke [28]. We eventually discovered ways to create cigarette filters that would absorb most of these reactive products. Although we thought that the world would embrace such an invention that did not turn out to be the case, and like many good ideas, the time was not right. The technology sits on the shelf of undeveloped products. As cigarette consumption in the world continues to be >20% of the adult population, in spite of campaigns to reduce smoking, I believe that the scientific and medical community should devise ways to decrease the damage incurred by a habit that is ingrained.

Inhibiting and reversing AGEs

Over the years, we attempted to pharmacologically intervene with AGE formation or AGE crosslinking. The first molecule, aminoguanidine, had an amino group attached to positively charged guanidine which is always charged at neutral pH. This inhibits the free amino group from becoming protonated and thus allows it to be readily able to react with carbonyl groups. Animal studies of aminoguanidine with diabetic animals showed that we could inhibit AGE formation and pathological changes in the kidney [29]. A clinical trial in diabetic patients with impaired renal function was performed over two to 4 years. Whilst there were positive clinical effects observed in the patients receiving the drug, the drug treatment groups did not reach statistical significance over the placebo group for the primary end-point of doubling of serum creatinine [30]. Needless to say this was a disappointment. The second molecule ALT711 was a thiazolium compound that could break AGE crosslinks in vitro and in vivo [31]. Diabetic rats, and aged dogs and monkeys showed a significant decrease in stiffness of the heart and blood vessels [32, 33]. However, a clinical study in patients with heart failure failed to show an improvement in outcome [34]. Development of both compounds has been halted.

Lessons learned?

What are the lessons to be learned from these clinical translational studies? The first obvious answer is that clinical trials are difficult and expensive to carry out, especially if the trials have to be carried out for long periods of time. The second is that we need better ways to monitor early trials of diabetic complications with biomarkers relevant to the disease process, so that the amount of drug and the dosing schedule can be ascertained in early phase 2 studies. A good example of the type of biomarker that I am referring to is HbA1c. Over the past four decades, HbA1c has become the gold standard for the pharmaceutical industry and the regulatory agencies for allowing new agents to be approved to treat hyperglycaemia. Although we have come a long way in being able to treat diabetes, these patients still go on to develop retinopathy, nephropathy, neuropathy and cardiovascular disease. New agents are needed more than ever to treat the complications of diabetes because of the epidemic of Type 2 diabetes in the world today. A century ago Maillard suggested that the reaction of glucose with proteins in the body could cause disease; hopefully, 100 years from now we will know how to decrease the damage caused by glucose. In closing, it is important to reflect that glucose, which is essential for life, is also the harbinger of ageing and death.

Conflict of interest statement

No conflict of interest was declared.

Acknowledgement

I wish to thank Michael Brines and Ann Dunne for their help in the preparation of this manuscript.

References