INTRODUCTION

This symposium about galectins has given me the opportunity to look back on the early research that led to their discovery and to take pleasure in reviewing the many fascinating developments since I moved on to other things. My interest in carbohydrate-binding proteins grew out of speculations about the possible role of protein–glycoconjugate interactions in the formation of specific synaptic connections in the brain (1), which explains how I—a psychiatrist and neuroscientist—stumbled into this field.

I first became aware of the potential biological importance of protein–glycoconjugate interactions in conversations with Victor Ginsburg, whom I met in 1961 at the National Institutes of Health in Bethesda, Maryland. I was, at the time, a postdoctoral fellow in Marshall Nirenberg’s lab and had the good fortune to participate in his studies of synthetic polynucleotides and the genetic code that soon brought him a Nobel prize. Vic, who was in the lab next door, was convinced that the next important code to solve was the one that determined how cells selectively adhere to each other to form complex tissues, and that protein–glycoconjugate interactions must play a role. He based his argument on the presence of complex glycoconjugates on cell surfaces and his hunch that their specific structures must have some biological significance. I found this idea very persuasive and tucked it away until I could explore it.

Over the next 10 years, I completed my clinical training in psychiatry, studied axoplasmic transport and the role of brain protein synthesis in memory storage,
and helped found the new School of Medicine at the University of California in San Diego. My chance to explore Vic’s idea came in 1972 when Steven Rosen joined my laboratory as a postdoctoral fellow. While in graduate school at Cornell, Steve began to study cell adhesion in cellular slime molds, and we decided to continue this project in the hope that it would teach us something about synapses in the brain. This led to a program of research that paved the way for my work with galectins. To explain the early history of galectin research, I will begin with our studies of cell adhesion in slime molds.

LECTINS IN CELLULAR SLIME MOLDS

Research on cell adhesion of the cellular slime mold, Dictyostelium discoideum, was pioneered by Gunther Gerisch in the 1960s. It was already known that slime mold cells exist as individual amebas when food is available but form multicellular aggregates when food is gone, and Gerisch had raised antisera that block aggregation (2) to search for the relevant adhesion molecules. Rosen then made antisera of his own and tested their reaction with slime mold extracts bound to sheep erythrocytes—a common immunoassay of that period. He was surprised to find that the extracts from aggregating cells agglutinated the sheep erythrocytes without the addition of antibodies, whereas the extracts from the growing cells did not.

When Rosen arrived in my laboratory, we decided to test the possibility that the extract from aggregating slime mold cells agglutinated the erythrocytes by binding to glycoconjugates on their surface, just like lectins such as concanavalin-A. This was, at the time, a wild idea because lectins were supposed to be restricted to plants. Nevertheless, it was easy to test by adding various sugars to see if they blocked agglutination. The very first experiment worked: N-acetylgalactosamine, galactose, and lactose blocked agglutination whereas other sugars did not (3).

Our first attempt to purify the active material by gel filtration on Sepharose also gave a dramatic result: none of the protein fractions that came through the column had any agglutination activity. After trying to recombine various fractions to recover the activity, we figured out that the active material must have bound to the galactose residues in the Sepharose column. Realizing that we had unwittingly used Sepharose as a matrix for affinity chromatography, we eluted the bound protein from the column with a lactose solution. When we examined the eluate by gel electrophoresis, we found a major protein that we named discoidin I, and a minor one that we named discoidin II (4, 5).

Subsequent work in my lab established that discoidin I is mainly located in the cytoplasm of aggregating slime mold cells, that it is secreted by an atypical mechanism (6), and that it plays a role in aggregation (7). When the discoidin I gene was sequenced in Rick Firtel’s lab (8) at U.C.S.D., we learned that the protein contains an arg-gly-asp (RGD) sequence similar to that found in cell adhesion proteins such as fibronectin. This suggested that the RGD in discoidin I is involved in the formation of aggregates, and we confirmed this in experiments in which we inhibited aggregation of slime mold cells with RGD-containing peptides (7).
Ironically, a functional site (RGD) that had already been discovered in studies with a human protein (fibronectin) helped us understand how aggregation works in our so-called “simple model system,” the cellular slime mold.

**FINDING GALECTINS**

Encouraged by the discovery of the discoidins, Tom Nowak, a graduate student, began looking for agglutinins of erythrocytes in extracts of vertebrate tissues. He started with extracts of mouse brains because of the hope that a lectin like discoidin might play a role in synapse formation. He also tested extracts of embryonic chick muscle because of the possibility that mammalian lectins might be made at a specific developmental stage as they are in slime molds. In initial experiments, Tom found some agglutination activity in both the brain and muscle extracts. But it was not blocked by sugars.

Scientists at the Weizmann Institute had better luck. Also interested in the possible role of lectins in synapses, and encouraged by our work with slime molds (3, 4), Vivian Teichberg and his colleagues published a classic paper in 1975 about a lactose-inhibiting hemagglutinin in extracts of the electric organ of an electric eel. When they applied the extracts to a modified Sepharose column and eluted it with beta-galactosides, they obtained a pure protein that they named electrolectin (9)—the first purified galectin. They also found a similar lectin activity in extracts of several vertebrate tissues including chick muscle, which shares some properties with the electric organ.

Why did Teichberg et al. succeed while we had failed? The answer is that their buffers contained a reducing agent that is essential for maintaining the lectin activity, whereas ours did not. When we subsequently prepared chick muscle extracts in buffers that contained dithiothreitol or beta-mercaptoethanol, we too found a lactose-inhibitable lectin. Furthermore, when we studied chick muscle extracts prepared at various stages of embryogenesis we found that, like discoidin, the chick lectin was developmentally regulated (10): it was virtually undetectable in 8-day-old embryos but abundant by day 16.

Stimulated by the pioneering work of Teichberg et al., several groups began purifying related lectins from other vertebrate tissues. The first, which we now know as galectin-1, was purified in 1976 in Stuart Kornfeld’s laboratory from calf heart and lung (11). This was followed in 1977 by the purification of the lectin from embryonic chick muscle by Halina Den and David Malinzack (12), and in my lab (13). Like electrolectin, those from the other sources were dimers with subunit molecular weights of about 15,000; and in each case, activity was only retained in buffers with sulfhydryl-containing compounds. This led Kurt Drickamer (14) to call these lectins “S-type” (for sulfhydryl-dependent type), a term that was later abandoned in favor of the term galectins when it became clear that other members of this family retain activity in the absence of reducing agents.

The first evidence of other members of this family came in 1980 from Eric Beyer, a student in my laboratory, who found an abundant lectin in chick intestine with a subunit molecular weight of 14,000 and other differences from the muscle lectin.
To distinguish them, we named the one from muscle chicken-lactose-lectin I (CLL-I), and the one from intestine chicken-lactose-lectin II (CL-II) (15).

Further evidence that there are multiple members of this family came in the 1980s from work on mammals. First, a group led by John Wang showed that cultured mammalian fibroblasts have three lectins and went on to purify the one now known as galectin-3 from mouse lung (16, 17). Then, Robert Cerra, Michael Gitt, and I purified three lectins from rat lung, now known to be galectin-1, -3, and -5 (18, 19). Later, Hakon Leffler, Frank Masiarz, and I found yet another member in extracts of rat intestine, now known to be galectin-4 (20).

Subsequent discovery of new galectins was greatly facilitated by cloning and sequencing of cDNA and genomic DNA. In the first attempt that Michael Gitt and I made to clone the cDNA for galectin-1 from a human hepatoma cDNA expression library, we found a related cDNA for a hitherto unknown galectin, now called galectin-2 (21). Many of the other galectins discussed in this symposium were initially identified on the basis of nucleotide sequencing and were later confirmed as galectins by demonstrating that the purified proteins bind beta-galactosides. The availability of cDNAs for galectins also made possible the synthesis of large quantities of recombinant protein for crystallography. In fact, the conserved structure of the carbohydrate-binding domain of galectins was first identified in Jim Rini’s lab in 1993 (22, 23) by studying crystals of recombinant galectin-2 that we made with Gitt’s cDNA.

NAMING GALECTINS

As this information accumulated, it became clear that a standard nomenclature was needed. In 1994, I contacted Kurt Drickamer, who agreed that “S-type lectins” was no longer appropriate. I then contacted the other main workers in the field and we all decided to give up the various names we were using in favor of the general term galectin. In the joint statement we published in Cell (24), we indicated that membership in the galectin family required “fulfillment of two criteria: affinity for beta-galactosides and significant sequence similarity in the carbohydrate-binding site, the relevant amino acid residues of which have been determined by X-ray crystallography.” For mammalian galectins, numbers were assigned for those then known and continued to be assigned for new ones (e.g., (25, 26)) as they were discovered: there were 4 when we agreed on these criteria and now there is a pretty good agreement that there are at least 15. Because Gitt and I had already worked with a nomenclature committee to name the galectin-1 gene LGALS-1 and the galectin-2 gene LGALS-2 (27–29), the designation LGALS for “lectin, galactoside-binding, soluble” was used for all the others as they were identified.

LOOKING BACKWARD AND FORWARD

When Steve Rosen and I discovered discoidin I in Dictyostelium discoideum, we took this as a support for the general idea that protein–carbohydrate binding plays a role in cell–cell and cell–matrix interactions during tissue development not only in slime
molds but also in vertebrates. Looking back at discoidin I, it is remarkable how many molecular and biological properties it shares with galectins such as galectin-1: (a) they both bind galactosides but with relatively low affinity; (b) they are both developmentally regulated, often rising to very high levels at a particular stage of development; (c) they are both concentrated in the cytosol and may constitute as much as 1% of the soluble cytosolic proteins; (d) like other cytosolic proteins they both lack a signal peptide; (e) despite the absence of a signal peptide they both are secreted onto the cell surface and into the extracellular matrix by atypical secretory mechanisms (6, 30–33); (f) they both influence cell–matrix interactions and cell migration. These striking similarities suggest that soluble lectins (34) fill a special biological niche that has remained important through vast periods of cellular evolution (35, 36).

Another remarkable thing about discoidin I is that since 1993 many human proteins that contain discoidin domains have been identified. The first to be discovered (37), named discoidin domain receptor-1 (DDR-1), is an integral membrane protein with an extracellular N-terminal domain that resembles discoidin I. It also has a tyrosine kinase domain that is activated when specific ligands such as collagen bind the discoidin domain, and this interaction activates intracellular signaling pathways (38, 39).

The discovery of discoidin domain proteins is of personal interest to me because several have important roles in the brain. For example, DDR-1 has been implicated in axon migration (40), and neurexin, another discoidin domain protein, has been found in synapses where it interacts with neurologins (41). Even more tantalizing is the finding that patients with rare forms of autism have mutations in neurexins and neurologins (41). Having been teased by some of my psychiatric colleagues when I devoted so much time to discoidin I, I find it amusing that this slime mold protein is related to a human one that appears to play a role in a mental disorder.

Also of interest to psychiatrists and neuroscientists is the identification of galectin-1 and -3 in specific populations of neurons (42–44). But most galectin research is now focused elsewhere, and their role in the nervous system is not yet well established.

In the 10 years since I left galectin research, and turned my attention to psychiatric genetics and psychopharmacology, there has been a great deal of progress, as described in the course of this volume. One idea that continues to guide this work is that galectins influence cell adhesion, cell signaling, and other functions by interaction with glycoconjugates on and around cells. But, as you will see, there is a growing awareness that these proteins also have other domains (45), have important effects inside the cells that make them, and have been adapted for functions that no one imagined at the time they were discovered.

REFERENCES


