CHEMICAL REVIEWS

Review

Carbohydrate Polymers at the Center of Life's Origins: The Importance of Molecular Processivity

Robert Stern, and Mark J. Jedrzejas

Chem. Rev., **2008**, 108 (12), 5061-5085 • DOI: 10.1021/cr078240l • Publication Date (Web): 29 October 2008

Carbohydrate Polymers at the Center of Life's Origins: The Importance of Molecular Processivity

Robert Stern[†] and Mark J. Jedrzejas[*,§,‡]

Department of Pathology and Comprehensive Cancer Center, School of Medicine, University of California, San Francisco, California 94143, Microdesign Institute, 29 Kingwood Rd., Oakland, California 94619, and Center for Immunobiology and Vaccine Development, Children's Hospital Oakland Research Institute, 5700 Martin Luther King, Jr. Way, Oakland, California 94609

Received April 5, 2007

Contents

1. Introduction 5062
1.1. Foreword 5062
1.2. Early Earth 5063
1.2.1. Origins of the Earth 5063
1.2.2. Chemistry on the Early Earth 5063
1.2.3. Life on Earth May Not Be Unique 5064
1.3. Early Theories and Experiments Regarding Life on Earth 5064
1.3.1. Charles Darwin and "Warm Little Ponds" 5064
1.3.2. Stanley Miller and Electrical Discharges 5064
1.3.3. Leslie Orgel, an RNA World, and Ribozymes 5065
1.4. Beginnings of Life on Earth 5065
1.4.1. Prebiotic Earth and Its Chemistry 5065
1.4.2. Possible Icy Origins 5066
1.4.3. Synthesis of the First Polysaccharide Polymers 5066
1.5. Synthesis and Assembly of the First Cells 5068
1.5.1. Lipid Adsorption and Phase Separation 5068
1.5.2. Oil Slicks, Droplets, and the Formation of Protocellular Organelles 5068
1.5.3. First Unicellular Organisms 5068
1.5.4. Emergence of a Primitive ECM 5069
1.6. An Early Pre-DNA/Protein World and Examples of Its Current Remnants 5069
1.6.1. Ribozymes 5069
1.6.2. Primitive Viral-Like Organisms 5069
1.7. Polysaccharide Polymers Were Prerequisites for the Evolution of Life 5069
1.7.1. Polysaccharides as Food and Stored Energy Sources as Well as Structural Moieties 5069
1.7.2. Protection Against UV Radiation and Oxygen Free Radicals 5070
2. Current Polymeric Glycans 5070
2.1. Structural Polysaccharides 5071
2.1.1. Cellulose 5071
2.1.2. CT 5071
2.1.3. Bacterial Cell Wall and Surface Glycans 5071
2.1.4. Consequences of Structural Polysaccharide Evolution 5071
2.2. Storage Polysaccharides 5071
2.2.1. Starch 5071
2.2.2. Glycogen 5071
2.2.3. Pectins 5072
2.2.4. Alginate 5072
2.3. ECM and GAGs 5072
2.3.1. Heparin and HS 5072
2.3.2. Ch and ChSs 5072
2.3.3. DS 5073
2.3.4. KS 5073
2.3.5. HA 5073
2.4. Polysaccharides of Other Glycoproteins 5073
2.4.1. Branched Polysaccharides of Mucins 5073
2.4.2. Oligosaccharides of Heterogeneous Glycoproteins 5073
3. Increasing Complexity of Polysaccharide Structures and Functions 5074
3.1. Interglycosidic Bonds 5074
3.2. Polysaccharides Can Be Informational 5074
3.2.1. Fragment Size: HA Polymers Contain Size-Specific Information 5074
3.2.2. Information Coded Within Sulfation Patterns 5074
4. Polysaccharide-Active Proteins/Enzymes 5074
4.1. β-Linked Polysaccharide Synthesizing Enzymes 5075
4.1.1. Cellulose Synthases 5075
4.1.2. CTSs 5075
4.1.3. HA Synthases 5075
4.2. Polysaccharide β-Endoglycosidase Degrading Enzymes 5076
4.2.1. Cellulases 5076
4.2.2. Chitinases 5077
4.2.3. Hyals 5077
5. Modular Architecture of Carbohydrate-Active Enzymes 5078
6. Example of Molecular Mechanism on Polymeric Glycan: Mechanism of Processivity 5078
6.1. Mechanisms of Enzyme Processivity: Example of Degradation of Polysaccharides 5078
6.2. Mode of Degradation of HA and Ch by Pneumococcal Hyal 5078
6.3. Selection of the β-1,4-Glycosidic Bond for Enzymatic Catalysis 5081
6.4. Degradation of Aggregated HA and Ch by Pneumococcal Hyal 5081
6.5. General Conclusions from the Processivity of Enzymes 5082

* To whom correspondence should be addressed. Tel/Fax: +1-510-530-0211. E-mail: mjjedrzejas@yahoo.com.
[†] University of California.
[§] Microdesign Institute.
[‡] Children's Hospital Oakland Research Institute.

10.1021/cr078240l CCC: $71.00
Published on Web 10/29/2008

6.5.1. Hyal Flexibility and the Processivity Mechanism for HA Degradation 5082
6.5.2. Structural Aspects of Streptococcal Hyal Processivity 5082
7. Conclusions 5082
8. Abbreviations 5083
9. Acknowledgments 5083
10. References 5083

1. Introduction

1.1. Foreword

Life on Earth evolved over billions of years. A series of chemical reactions ultimately led to this amazing phenomenon. Simple compounds such as formaldehyde, abundant on early Earth, polymerized spontaneously. These reactions were followed by aldol condensations, leading to the production of stable saccharides. High-energy polyphosphates also existed in this prebiotic world. The production of these simple sugars and their polymerization were the initial reactions that led eventually to the synthesis of polymers with self-replicatory properties and the subsequent emergence of life. All of these were spontaneous, naturally occurring reactions. A period existed in which multiple forms of such polysaccharides were dominant on Earth, a period critical for the eventual evolution of life.

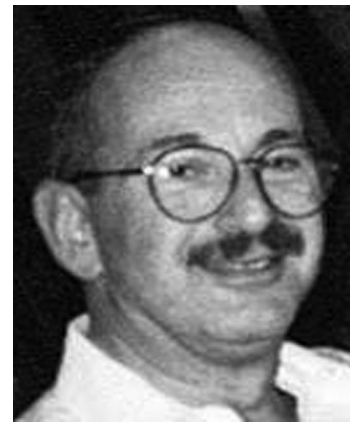
It is generally accepted that an RNA world preceded the DNA and protein world that we know today. We postulate that a carbohydrate polymer world preceded even the RNA world. In the presence of high-energy polyphosphate compounds, phospho-sugar-based nucleic acidlike analogues occurred that were capable of forming with themselves stable Watson−Crick pairs. This established pathways for the next stage of evolution, the emergence of an RNA-based world.

These polysaccharides, in addition, provided efficient sources of food and energy for subsequent life forms. However, more importantly, such simple saccharide homopolymers gave structure to organisms, as exemplified by cellulose for plants and chitin (CT) for the simplest forms of animal life. All plant life has continued to be based on cellulose structures, just as the exoskeleton of invertebrates retains chitinous structures to this day.

Intense UV (ultraviolet) radiation permeated the oxygen-poor atmosphere of early Earth. Simple sugar homopolymers such as CT protected against UV radiation. Similar glycopolymers were necessary predecessors for protecting the vital information encoded in the subsequent nucleic acid polymers, just as ribose and deoxyribose sugars protect them from UV damage today.

As atmospheric oxygen increased, so did the formation of oxygen-derived high-energy free radicals. Although the threat of UV damage diminished, protection against free radical injury became additional functions of these glycopolymers. The change in the nature of environmental stresses provided the evolutionary pressure for modifying existing polymers and for creating new polymers, as the need for different protective functions arose.

In summary, we suggest that simple molecules of early Earth condensed spontaneously to form carbohydrates and that these carbohydrates went on to form polymers, facilitated by the action of high-energy phosphates. These polysaccharides led to self-replicating polymers with high but not absolute fidelity to allow for evolutionary changes. Such self-replicating polymers presumably utilized purine- and pyri-



Robert Stern left Germany in 1938 for Seattle, Washington. He graduated from Harvard College in 1957 and obtained an M.D. degree from the University of Washington (Seattle) in 1962. While a medical student, he worked in the laboratories of Drs. Krebs and Fisher, who became Nobel laureates. He received his resident training in Pathology at the NCI and was a research scientist at the NIH for 10 years. From 1977 to 2007, he was a member of the Pathology Department at the University of California, San Francisco. He is a board-certified Anatomic Pathologist and has participated in the research, teaching, administrative, and diagnostic activities of the department. He directed the Ph.D. program in Experimental Pathology for 10 years. For the past decade, his research has focused on hyaluronan and the hyaluronidases, an outgrowth of an interest in malignancies of connective tissue, stromal−epithelial interactions in cancer, and biology of the tumor extracellular matrix. His laboratory was the first to identify the family of six hyaluronidase sequences in the human genome. These enzymes were then sequenced, expressed, and characterized. Subsequent work has identified a catabolic pathway for hyaluronan. Dr. Stern was a Fulbright scholar in Germany (1984−1985). He is a member of the editorial board of Matrix Biology and, until recently, that of the University of California Press. He is now Professor Emeritus. Since 2008, he has been teaching Pathology at Al Quds University in East Jerusalem, while also establishing a research program.

Dr. Jedrzejas received his B.A./M.S. in Physics from Jagellonian University, Krakow, Poland, in 1988, a M.S. in Chemistry in 1992 from Cleveland State University, and a Ph.D. in Structural Chemistry from Cleveland State University/Cleveland Clinic Foundation in 1993. He was appointed an Assistant Professor of Microbiology at the University of Alabama at Birmingham in 1995. In 2001, Dr. Jedrzejas moved to the Children's Hospital Oakland Research Institute as an Associate Scientist and then a Scientist. He currently directs a structural biology group at Children's and continues his studies of the structure and mechanisms of macromolecules using predominantly tools of structural biology with emphasis on X-ray crystallography. His research interests include structural aspects of Gram-positive bacterial pathogens, mainly bacteria−host interactions, with focus on *Streptococcus* species, and mechanisms of essential processes leading to the formation of bacterial spores, their germination, and outgrowth, with an emphasis on *Bacillus* and *Clostridium* species.

midine-like bases for self-replication and were precursors to the development of nucleic acid polymers.

The ribose and deoxyribose of current nucleic acids are also carbohydrate polymers. The development in the prebiotic world of self-replicating carbohydrate polymers thus led to an RNA-based and then to the DNA, protein world that we know today. However, throughout, polymer chemistry was essential to the core of life.

Here, we present details of the carbohydrate chemistry essential to the polymer reactions that led to life on Earth. We also provide an overview of the array of polysaccharides that occur in present life forms, postulating pathways that led to the increasing complexity, as life forms defied environmental pressures with ever-new survival mechanisms. To "adapt or die" is the underlying survival strategy for all life forms, as true today as it was at the dawn of life.

Finally, we explore in detail an example of chemistry critical for life, the processive chemistry of a polysaccharide. We demonstrate this using hyaluronan (HA) and its degrading enzymes, the hyaluronidases (Hyals). Such enzymes, particularly from prokaryotes, are arguably the only group of enzymes for which sufficient molecular detail is available that permits exploration of processivity and the only enzymes that can be coupled to the equally well-studied chemical properties of their polymeric HA substrates. This combination is critically important for understanding the evolution of these polysaccharide-related enzymes and the relationship to their substrates. Given the lack of comparable data, such studies on the biosynthetic enzymes would be highly speculative. For this reason, the degradative polysaccharide enzymes are the focus here.

1.2. Early Earth

1.2.1. Origins of the Earth

Our Universe is approximately 13.7 billion years old.[1] The lightest elements, hydrogen and helium, were created during the very earliest stages following the Big Bang. Additional cycles of star birth and death were required for nucleosynthesis of the elements essential for life such as carbon, nitrogen, and oxygen. However, it took several additional billions of years for sufficient amounts of these elements to accumulate. This led to a nonequilibrium type of distribution of matter, as represented by the abundance curve of the chemical elements. The distribution of the number of protons and neutrons in nuclei is governed by physical properties of their binding and leads to 92 naturally occurring stable elements (Figure 1).[2] The binding energy per nucleon reaches its maximum for nuclei with atomic numbers of ~60, reflecting the nature of their nuclear forces (Figure 1A).[2] The heavier chemical elements were created afterward by fusion within stars or supernova explosions. The six most abundant elements that resulted were helium, hydrogen, neon, oxygen, carbon, and nitrogen (Figure 1B).[3]

Radioactive decay of selected elements in rocks and meteorites leads to the conclusion that Earth was formed approximately 4.56 billion of years ago (Figure 2).[4] Its initial overall composition of elements, their type, and the relative abundance presumably reflected the composition of the entire Universe. However, the types and abundance of chemical elements in the outer crust of Earth differ significantly from that profile.

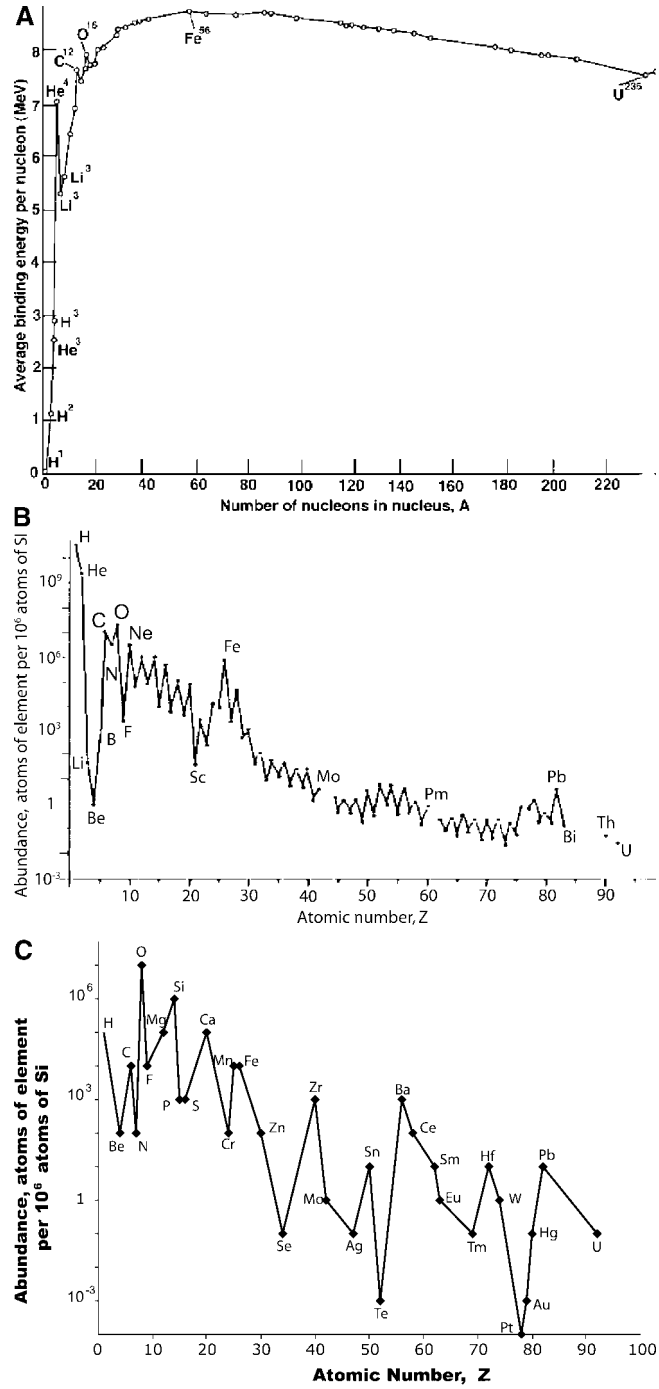


Figure 1. Chemical elements of Earth. (A) The binding energy per nucleon in nuclei of chemical elements. Both ^{58}Fe and ^{62}Ni are the more strongly bound than ^{56}Fe, with ^{62}Ni having the highest mean binding energy [2]. (B) The composition of the solar system. Abundance of the chemical elements in the solar system in terms of atoms per 10^6 of Si [3]. (C) The composition of Earth. The relative abundance of selected elements in the Earth's crust is shown.[5]

1.2.2. Chemistry on the Early Earth

The Earth is composed of an inner core of predominantly solid iron and nickel, an outer core of mostly liquid iron and nickel, a mantle composed substantially of silicates, and finally, the crust.[5,6] The inner and outer cores are the most dense portions, accounting for 70% of Earth's mass. The less dense mantle accounts for 84% of the volume, while the crust constitutes less than 0.5% of mass. The mantle also contains magnesium and iron, whereas the crust is composed

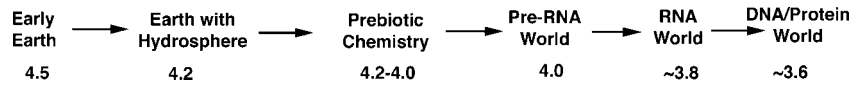


Figure 2. Selected view on the history of Earth and life on our planet. The scale underneath the pictorials is time measured in billion of years before the present. The figure is based on data from Joyce.[14]

primarily of oxygen, aluminum, silicon, iron, magnesium, calcium, sodium, and potassium.

The early Earth atmosphere was much different from that of the present atmosphere; this atmosphere contained large amounts of carbon dioxide (CO_2) and hydrogen sulfide (H_2S). The surface of the early Earth was molten, and as the surface cooled, volcanoes spilled massive amounts of carbon dioxide (CO_2), steam (H_2O), ammonia (NH_3), and methane (CH_3) into the atmosphere.[7,8] The excreted steam condensation led to the formation of a stable hydrosphere in the form of shallow seas. However, there was still no oxygen in the atmosphere at that time.

It is suspected that life began in shallow waters, close to thermal vents, a source of heat and minerals, while the remaining parts of the cooled surface remained solidly frozen. It took just over 700 million years, approximately 3.8 billion years ago, for bacterial-based life to evolve and to begin to flourish. When photosynthetic bacteria evolved and when plants evolved later, they together initiated the conversion of CO_2 from the atmosphere and H_2O into energy and oxygen. These "green" organisms and plants kept removing CO_2 from the atmosphere and, in parallel, produced oxygen, releasing it into the atmosphere. Atmospheric CO_2 became locked into sedimentary rocks in the form of carbonates and fossil fuels and also became dissolved into the oceans.[7,8]

The appearance of oxygen in the atmosphere led to reactions with ammonia, methane, and iron. Denitrifying bacteria, in turn, released nitrogen into the atmosphere. Additionally, nitrogen was released as a result of ammonia reacting with oxygen. These processes lasted for approximately 2.1 billions years. During this period, the composition of the Earth's atmosphere contained significant levels of oxygen. The ozone layer was formed subsequently, leading to the filtering out of harmful UV radiation, which further facilitated the creation of life.[8]

1.2.3. Life on Earth May Not Be Unique

1.2.3.1. A Copernican View of Life on Earth. Life on Earth is built primarily of four major chemical elements: hydrogen, oxygen, carbon, and nitrogen. All other chemical elements, although important and essential for life, contribute to less than 1% of the mass of living organisms. These four elements are part of the six most abundant elements in the Universe (Figure 1B).[3] The remaining two abundant elements of our Universe, helium and neon, are inert noble gases and are not involved in life processes. This suggests that life anywhere in our Universe, if it exists at all, is based upon the same elements and perhaps on the same chemistry. These four elements, hydrogen, oxygen, carbon, and nitrogen, are, however, not the most abundant elements of Earth, on its surface alone, or even in the oceans where presumably life began (Figure 1C).[5] It is the occurrence of these four elements that are uniquely peculiar to current life.

The primordial milieu in the shallow waters around the early vents contained mostly hydrogen and oxygen in the form of water. The most common elements dissolved in water are chlorine, sodium, sulfur, calcium, and potassium, but not the carbon or nitrogen that are abundant on Earth. Life clearly reflects the abundance of elements of the Universe far more closely than the composition of the Earth.[3,5] One could term this the Copernican principle of life in the Universe.[9]

1.2.3.2. Extraterrestrial Contributions. The major available elements of the Earth enter into chemical reactions with each other and into the synthesis of many compounds, including linear polymers, ring compounds such as pyranose, and their stereochemical arrangements. All of the resulting complex compounds involving those four key elements contributed to the creation of life. However, the Earth is not an isolated planet. It is and has been constantly bombarded with space dust, meteorites, and comets. Amino acids have been detected in these extraterrestrial bodies, having survived high impact temperatures. Furthermore, such high temperatures resulted in the generation of polymeric peptides. Considering these relatively common events in the first several billion years, the rich supply of extraterrestrial matter may well have contributed to the creation of the first polymers and the evolution of life on Earth.

The Murchison meteorite entered the Earth's atmosphere in 1969, and it was estimated to be more than five billion years old. It contained evidence of carbon-based compounds, which is additional evidence for extraterrestrial life.[10] The proportions of the amino acids found in the Murchison meteorite also approximated the proportions observed in the primitive atmosphere modeled in the Miller−Urey experiment described below. The extraterrestrial organic molecules could have accelerated the formation of these terrestrial organic molecules by serving as molecular templates. This meteorite also contained sugars and other hydrocarbons, including amphiphiles or their precursors, critical for the development of life.

1.3. Early Theories and Experiments Regarding Life on Earth

1.3.1. Charles Darwin and "Warm Little Ponds"

Charles Darwin first suggested that life's origins occurred around "warm little ponds". The theory of natural selection set forth by Darwin indicates that hereditary information is replicated imperfectly. Slight variations provide adaptation for survival under everchanging environmental pressures. The theory also implies that more complex forms arise from simpler ones and that all life evolved from a single simple progenitor. "From so simple a beginning, endless forms most beautiful and most wonderful have been, and are being evolved".[11]

1.3.2. Stanley Miller and Electrical Discharges

S. L. Miller performed in vitro experiments on the origin of life in 1953.[12] "The production of amino acids under primitive earth conditions" were highly regarded at the time. A circulating mixture of methane, ammonia, and hydrogen in a flask of boiling water was constantly passed over electrodes, meant to mimic lightning, for several days. A number of amino acids were detected in significant amounts,

included in particular glycine and alanine. A difficulty with these experiments was a failure to search for or to document the presence of carbohydrates. The question remains whether saccharides, amino saccharides, or polysaccharides occurred in those initial solutions. Because these were not mentioned in those original manuscripts, we can assume that they were not present or were not searched. Were carbohydrates created in those experiments—presumably yes? However, proof can now be provided only by additional experiments.

1.3.3. Leslie Orgel, an RNA World, and Ribozymes

In the late 1960s, Leslie Orgel and independently Francis Crick and Carl Woese proposed an RNA world that preceded the current DNA/protein world. In that RNA world, RNA would have catalyzed all reactions required for replication and survival. A triumph in the simulation of prebiotic chemistry was the ability to perform nonenzymatic replication of RNA in vitro.[13] In 1983, Thomas Cech and, independently, Sidney Altman discovered ribozymes, RNA molecules that had intrinsic enzymatic functions. This supported the concept that in an RNA world, these molecules carried out, in addition to replication, all functions necessary for life.

The reign of an RNA-dominated world occurred from 4.2 to about 3.6 billion years ago.[14] Ribose is the carbohydrate utilized by RNA, suggesting that polymeric glycans preceded RNA in evolution. As such, the existence of a glyco-world occurred prior to 4.2 billion years ago[15] (Figure 2). This glyco-world contributed to the evolution of a nucleic acid-based world, as a device toward a self-replicating system, independent of the clay surfaces described below.

1.4. Beginnings of Life on Earth

1.4.1. Prebiotic Earth and Its Chemistry

The concept of life originating from a "warm little pond" and electricity can be traced back to Darwin.[16] However, there are problems with this formulation. The concentrations of organic elements, even if the entire Earth biomass were to be dissolved in the ocean, are very small. Carbon, for example, would occur at a concentration of less than 0.1 mM. For chemical reactions to occur at reasonable rates, in addition to heat and electrical discharges, reagents need to be far more concentrated. Mineral-rich clay particles could concentrate the necessary reagents, serving as catalysts to lower the energy barriers so that such reactions could occur. The clay could also serve to protect chemical groups during the course of such reactions.[17] Clay-associated minerals selectively adsorb many prebiotic-like reagents and catalyze such reactions. Similarly, various carbohydrates can be synthesized by formol reactions on clay surfaces under such conditions.[18,19] Carbohydrate phosphates are also synthesized[19] under both neutral and slightly alkaline pH environments.

Simple formaldehyde, CH_2O, polymerizes spontaneously, producing small saccharides, including glyceraldehyde and dihydroxyacetone.[15] They both undergo aldol condensation, forming stable ketohexoses such as fructose. The Lobry de Bruyn rearrangement converts fructose into the aldohexoses, glucose, and mannose. In the prebiotic world, such simple chemistry made available a full set of the elementary hexoses, fructose, glucose, and mannose, from which essentially all remaining sugars can be synthesized.

These basic glycans thermodynamically prefer circular conformations, utilizing the known furanose and pyranose ring structures as opposed to their linear arrangements (Figure 3). Most biological glycans are D-enantiomers, and their ring's substituents are in a β-conformation. In subsequent evolving carbohydrate pathways, they were converted into other saccharides. One of these was galactose, utilized much later as a key recognition molecule by multicellular organisms. Runs of mannose sugars also became adopted as recognition moieties. As RNA utilizes the principal carbohydrate, ribose, in addition to purine or pyrimidine bases and phosphate, it is not surprising that polymeric glycans preceded RNA in evolution. As such, the glyco-world should be considered occurring in evolution before the presumed most widely accepted hypothesis of the RNA world[15] (Figure 2). In fact, the glyco-world might have led to the evolution of the nucleic acid-based world in an effort to devise a self-replicatory, clay mineral independent system (see below).

One of the discoveries supporting the concept of such a glyco-world is the existence of catalytic carbohydrates such as clodextrins.[20−23] These carbohydrate molecules catalyze assorted reactions using penicillin, DNA, ribonucleotides, and organophosphates as their substrates. Their catalytic properties are similar to those discovered for catalytic ribozymes, further supporting their existence prior to the RNA world.

Mineral-rich clay surfaces facilitate oligomerization of several kinds of monomers.[24] Synthesis of short polypeptides occurs under repeated cycles of wetting, drying, and heating.[25,26] Longer oligomers require some kind of template.[27−29] Their syntheses were essential in the pathway to self-replicating systems and were a prerequisite for the development of life. The synthesis of long polymers is essential to the chemistry required for the evolution of self-replicating systems. The original polymer to serve as a template for the synthesis of long oligomers may have been polyphosphate (poly P)-based.[30,31] Poly P preserved its importance for life and was present in relatively large quantities in the majority of cells (see below for more discussion) [30,32] (Table 1; see below).

The above illustrates how clay drove probiotic synthesis of simple compounds, as well as polymerization to form short and then longer oligomers, utilizing templates. At some time point, a self-replicating, polymeric system developed that made clay minerals dispensable. However, such a replicatory system needed to encompass stereochemistry of both the monomeric compounds and the linkages between them, as well as a means of storing information either in their sequence, their length, or their higher order structures. A template-driven synthesis of informational polymers would satisfy this requirement. It also provides an explanation as to why the peptide world has primarily L-amino acids and the biological polyglycans are primarily in the D-conformation (see below). The development of such information-rich replicating materials is a further step toward the emergence of life.

The presence of an ice-like well-ordered water layer on a mica surface at ambient temperatures is well-established.[33] This ordered water layer has a strong affinity for hydrophilic molecules such as the saccharide and polysaccharide sugars of early Earth and may have provided the linear template for their polymerization.

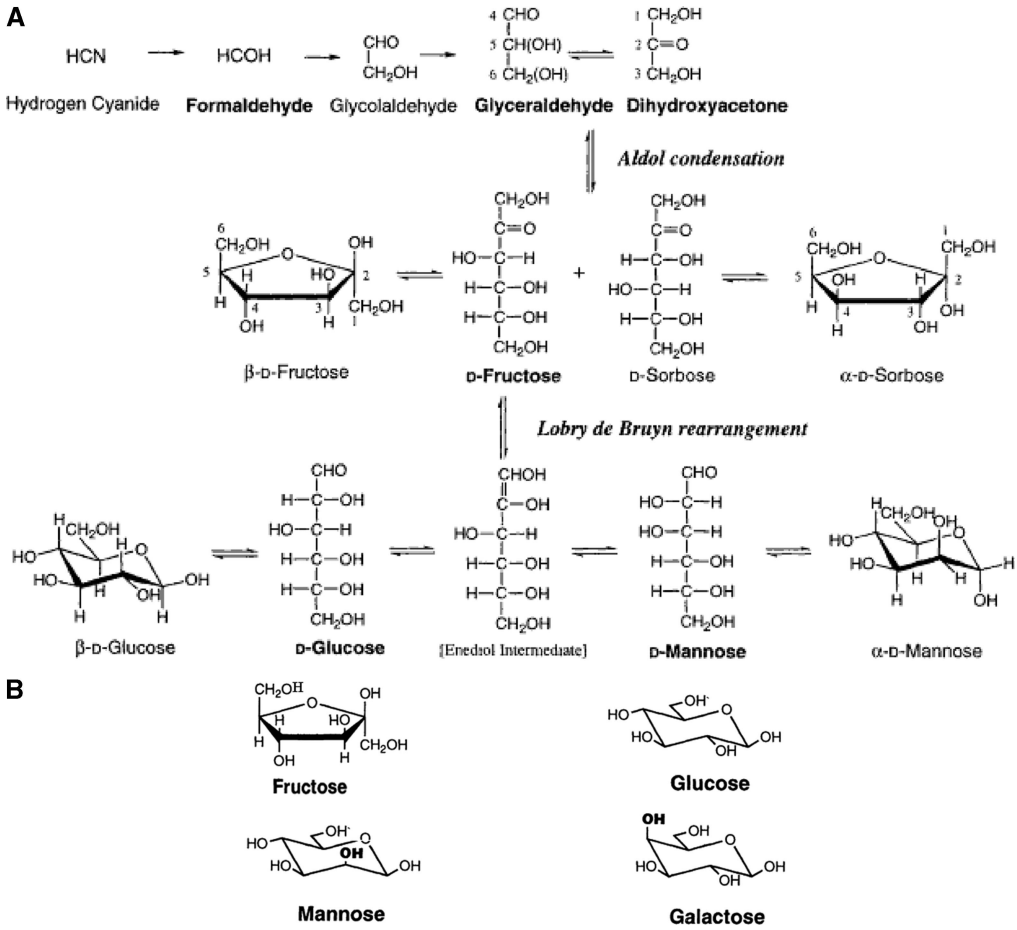


Figure 3. Synthesis and structure of basic carbohydrates. (A) The reactions to generate elementary hexoses consist of three consecutive steps: First, formol condensation generates glyceraldehyde and dihydroxyacetone from formaledehyde under weakly basic conditions; second, aldol condensation between glyceraldehyde and dihydroxyacetone generates 3,4-*trans*-ketoses (fructose and sorbose); and finally, third is Lobry deBruyn rearrangement, which converts only fructose into glucose and mannose. All hexoses are shown for simplicity in D-form. Reprinted with permission from ref 15. Copyright 1996 The University of Chicago Press. (B) Structures of the four basic hexoses. The majority of biological (keto- and aldo-) hexoses are D-enantiomers in β-conformation of their ring substituents. They utilize furanose (fructose) and pyranose (glucose, mannose, and galactose) rings in their thermodynamically most stable structures.[15]

Table 1. Levels of Poly P in Various Cells and Tissues[a]

eukaryotes	prokaryotes
fungi	bacteria
Saccharomyces cerevisiae, 120 mM	*Escherichia coli*, 0.1−50 mM
	Acinetobacter johnsonii, 200 mM
animals	archaea
rat liver, 26 mM	*Sulfolobus acidocaldaris*, 0.5−1.5 mM
cytosol, 12 mM	
nucleus, 89 mM	

[a] Poly P levels fluctuate enormously depending on the physiologic and metabolic state of the cell.[32]

1.4.2. Possible Icy Origins

Greenhouse gases were absent in the atmosphere of early Earth. With the Sun that was far less luminous at the time, ancient Earth was a permanently frozen planet except for the occasional meteor or fireball-like meteors termed bolides. Extremes of repeated freeze−thaw−heat cycles associated with such bolides were crucial for initiating reactions that contributed to life forms.[34] Hydrogen cyanide (HCN) would not have been present in sufficient concentrations to polymerize to nucleic acid bases and amino acids. Eutectic freezing, in frozen oceans, however, could generate local concentrations to facilitate such synthesis.[35,36]

The frozen oceans also concentrated substrates that facilitated catalytic reactions.[37] Sea ice has a complex multiphase structure containing mineral particles, liquid salts, gas bubbles, and ice crystals, with steep local pH and cyclic temperature gradients, ion densities, and electric potentials. Under such conditions, nonenzymatic synthesis of poly adenine (polyA) was achieved in the laboratory, reaching nucleotide lengths of 400 base pairs with predominantly 3′→5′ linkages.[37]

1.4.3. Synthesis of the First Polysaccharide Polymers

Saccharides are essential components of all living organisms and are the most abundant of all classes of biological molecules. Such carbohydrates are the most ancient biological molecules on Earth. Currently, much of the driving force of their synthesis comes from products of photosynthesis, although on the ancient Earth, simple condensations provided for their production. Now, the light-powered combination of CO_2 and water provides for the carbon hydrates from plants and several species of bacteria. The breakdown of saccharides provides the energy that currently drives nearly all biological reactions.

Polysaccharides are of variable size, have extraordinarily complex structures that are often heterogeneous, and do not have a genetic "code". Their effects are often indirect. This

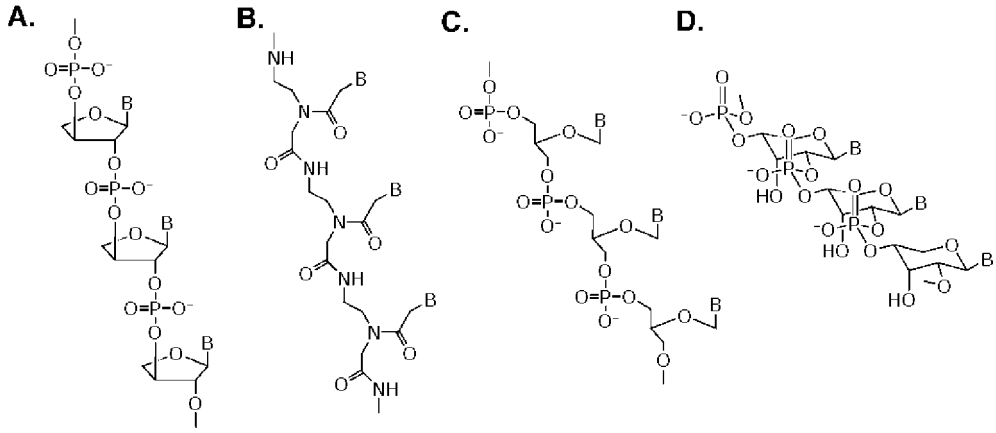


Figure 4. Types of backbones suggested for pre-RNA world. The common property seems to be the assumed existence of purine or pyrimidine bases as the main element of pairing for self-replication.[14] (A) Peptide, (B) threose, (C) glycerol, and (D) pyranose linked by phosphodiester bonds.

apparent passivity has made their study very difficult. These polymers were probably the first on Earth, and understanding their evolution is of intrinsic importance for understanding the emergence of life on this planet.

1.4.3.1. Linear Homopolymers. Clay-driven prebiotic syntheses occurred on early Earth, initially with simple compounds, followed by their oligomerization to form short oligomers, and then by longer oligomers when templates became available. At a certain point, the appropriate chemistry developed that supported a self-replicating system, which made clay minerals dispensable and facilitated the appearance of glycopolymers and RNA-like polymers. However, such a replicating system necessitated stereochemistry of both monomeric compounds and the linkages between them, as well as means of storing information, either within the monomer sequence, length of polymers, or their higher order structural features.

The template-driven synthesis of informational polymers presumably satisfied this requirement. It also explains why the peptide world is constituted primarily by L-amino acids and why carbohydrates are primarily in the D-conformation. The first successful synthesis of a template for replication of these polymers occurred entirely by chance, with one being at slight preponderance, followed by faithful accumulation of that enantiomer. A slight difference in eutectic temperatures between L- and D-amino acids, and between L- and D-sugars, particularly at solid−liquid interfaces, may also have influenced the final chirality choices. The development or evolution of such information-rich replicating material would equate to the emergence of life.

1.4.3.2. Self-Replicating Polysaccharide Polymers. In addition to current RNA/DNA self-replicating polymers, the polymeric threose, glycerol, and pyranosyl type backbones with bases attached have been suggested as constituting the pre-RNA world[14] (Figure 4). All such backbones involve phosphodiester linkages, as occur presently in nucleic acids.

Our proposal assumes that the pre-RNA world was based on simple carbohydrates that became ever more complex. Details of such an ancient carbohydrate self-replicating polymer system are not known. Perhaps it was similar to RNA and DNA utilizing poly ribose or poly deoxyribose-phosphate diester backbones. This glyco-world utilized polymeric glycan backbones with glycosidic linkages without having nucleic acid type bases attached. In support of this concept is that polymeric glycans often assume single, double, or triple helical structures, such as in the case of HA. X-ray fiber analysis of calcium HA indicates an extended left-handed single-stranded helix with three disaccharide units per turn.[38−42] The pairing of polyglycan backbones for replication is not difficult to imagine, as it would be somewhat similar or even identical to the current mode of replication of RNA and DNA that utilizes purines and pyrimidines.

Several pathways can be invoked that lead directly from carbohydrates to purines and pyrimidines. Intriguing nucleic acid analogues can be postulated containing various sugars and linkage isomers that resemble the steps that occurred during this intermediary period. Notable are α-L-threofuranosyl units joined by 3′,5′-phosphodiester linkages.[43,44] This structure forms stable Watson−Crick pairs with itself as well as with RNA. Threose is one of only two four-carbon sugars that can be joined at the 2′- and the 3′-position. From this, an easy transition to purines and pyrimidines as bases for self-replication, to an RNA world, can be envisioned. The step from carbohydrate polymers to a world in which RNA predominated, therefore, is the next transition in prebiotic evolution.

Alternatives to this scenario can be formulated. The clutter of prebiotic chemistry, as outlined by G. F. Joyce,[44] may have been sorted out by a series of favored reactions. The synthesis of sugars from formaldehyde can be biased by starting from glycoaldehyde-phosphate, which leads to ribose 2,4-diphosphate as the predominant pentose sugar.[45] This reaction can be initiated from dilute solutions of reactants at near neutral pH when carried out in the presence of certain metal-hydroxide minerals. The polymerization of adenylate, activated as the 5′-phosphorimidazolide, yields 2′,5′-linked products in solution. However, 3′,5′-linked products predominate in the presence of a montmorillonite clay.[46]

1.4.3.3. Facilitation by Poly-Phosphates. Condensation of the first carbohydrate polymers was probably facilitated by a family of high-energy poly-P.[30,31] Such inorganic polyphosphates are families of linear phosphate molecules from ten to hundreds of residues in length. They are linked by high-energy phosphoanhydride bonds formed from inorganic phosphate by dehydration at elevated temperatures. They are derived from ancient prebiotic reactions that were crucial to evolution. Today, poly-P moieties are found in virtually every living cell, bacterial, archaeal, fungal, protozoan, plant, and animal.[32] The levels of Poly-P in various cells and tissues are provided in Table 1.

These high-energy poly-Ps resulted in the synthesis of ever longer oligomers. Poly-P can be synthesized in cycles of wet and dry in the proximity of those warm vents or "warm little ponds" adjacent to clay environments. These achieve lengths of several hundred phosphate residues, linked by high-energy phosphoanhydride bonds, as found in ATP. Poly-P can also serve as phosphate groups for life chemistry, as well as for a pre-ATP energy source. The presence of templates of certain chirality appears also to force synthesis of products that are monochiral, leading to stereospecific polymeric compounds.[47] Poly-P preserved its importance for life and is present in relatively large quantities currently in nearly all cells [30,32] (Table 1). Variants of the ATP type of high-energy phosphate bonds include the recently recognized inositol pyrophosphates.[48]

Although these polymers were formed initially by condensation, a number of poly-P polymerases are now identified. Their mutation or overexpression has facilitated greater understanding of the role of poly-P in basic biological processes. For example, poly-P can stimulate growth of breast cancer cells[49] and can function as actin-binding molecules.[50] In prebiotic evolution, these polymers may have provided a scaffold, through their high-energy phosphate bonds, for the assembly of other macromolecules.

1.5. Synthesis and Assembly of the First Cells

1.5.1. Lipid Adsorption and Phase Separation

The primordial soup may have contained a number of biopolymers that were initially in a denatured state that interacted in a nonspecific manner. The first polysaccharide polymers, by binding ions, decreased the local concentration of salts in the primordial soup. The different classes of biopolymers are essentially immiscible and form emulsions, even in aqueous solutions.[51] Synthesis of other macromolecular polymers is stimulated in such a nonmiscible polysaccharide-rich soup, resulting in phase separation and increasing concentrations within the dispersed phase particles. Cross-linking of biopolymers is stimulated by hydrophobic and electrostatic interactions and hydrogen bond formation during freezing and thawing of the mixed biopolymers. The Maillard reaction then proceeds between an amino acid and a reducing sugar, requiring the addition of heat. The carbonyl group of a sugar reacts with the amino group of an amino acid, producing N-substituted glycosylamine and water. The unstable glycosylamine undergoes Amadori rearrangement, forming ketosamines. There are several ways for the ketosamines to react further. Diacetyl, pyruvaldehyde, and other short-chain hydrolytic fission products can be formed. Such cross-linking increases the local concentrations of different biopolymers, fixes their relative positions, and makes the interactions reproducible.

1.5.2. Oil Slicks, Droplets, and the Formation of Protocellular Organelles

A location for the first simple membranes could have been an oil/water interface in the subsurface environment of the Earth. The first oil-like molecules were presumably carbohydrate-like amphiphiles, that is, short-chain fatty acids, which were in all probability available on the early Earth, can self-assemble into stable vesicles that clearly define inside and outside direction and are able to encapsulate hydrophilic solutes with catalytic activity.[52] A variety of amphiphilic compounds have the capacity to self-assemble into membranous structures in the form of bilayers. It appears that the earliest oil vesicle must have incorporated such compounds into boundary membranes.[53,54] The origins of lipids and lipidic vesicles are not known. However, considerable amounts of membrane amphiphiles were likely formed from simpler hydrocarbons, and photochemical oxidation of hydrocarbons was a likely source of such amphiphilic molecules required for the self-assembly of primary membrane structures.[54]

Other possible sources are organic material carried to the Earth's surface by meteoritic infall that served as templates or substrates for further synthesis of lipidlike substances. Samples of nonpolar substances have been extracted and analyzed from the Murchison meteorite (see above), and some of the components can produce boundary structures that resemble membranes.[53] The physicochemical conditions of the subsurface would include elevated but eventually cooling temperatures, anaerobic conditions, and protection from intense surface UV radiation.

Low molecular weight liquid hydrocarbon sources could have formed such amphiphiles and led to an oil layer covering the primeval ocean 4.0−4.4 billion years ago, preventing water from evaporating into the atmosphere.[55,56] The primordial soup was possibly an oil/water interface or emulsion on the anaerobic surface of the early Earth's oceans. In the more reducing atmosphere of 3.9 billion years ago, even more hydrocarbons as well as reactive molecules were forming. An oil layer can act as a dry solvent for reactions, where the reactive molecules can produce monomers and condensing agents.

As monomers and eventual polymers formed, they became concentrated at the oil−water interface, favoring molecular interactions, without exposure to the destructive action of UV light. Increased water leakiness of the oil layer due to accumulation of polar molecules could lead to photooxidation of liquid hydrocarbons and their emulsification at the oil−water interface. Such an environment could have been a favorable location for the assembly of the first cells on the Earth capable of growth and division.[57] Once a continuous closed membrane was formed, a central component of the first cells would have been present, a semipermeable structure permitting the in and out passage of materials. Such an open system could acquire new and evolving functions.

1.5.3. First Unicellular Organisms

Once a self-replicating polymer evolved, it had an intrinsic ability to continue to evolve, according to Darwinian principles.[58] The second requirement of continuing evolution is compartmentalization, as found inside cells. All known cells utilize lipids to define their boundaries. Lipid molecules in solution have the ability to associate and to create vesicles that grow as more lipid molecules become available.[58,59] As such vesicles grow and become too large, they then undergo division, resembling the mitotic division that occurs in cells. If the original self-replicating polymer became engulfed by such a vesicle, in principle, the first cell would have been born—the protocell, with all of the elements necessary for a potential life form.

When the interior self-replicating polymer underwent changes, dependent on the lack of strict fidelity in replication, superior cells would have been created, ready to expand and to fill a particular environmental niche. The building blocks

utilized in growth must be transported or must diffuse into the protocell or, alternatively, must be synthesized within the protocell.[58] The latter possibility caused further evolution of cells as they became more complex and more independent of the environment.

The evolution of the protocell led to their growth and adhesion of some of them into groups or colonies. Such organization provided advantage to the survival of certain early organisms. They evolved means of communicating with one another, presumably through their simple extracellular matrices (ECM), such as surface poly glycans. The colonial life within these groups would lead to specialization of function and to better utilization of environmental resources and would further endowe them with survival advantages. This led to the creation of the first multicellular organisms. It is difficult to speculate if such multicellular organisms first evolved during the pre-RNA, RNA, or during the present DNA/protein world. There are currently no examples of life based purely on such a pre-RNA and, with the exception of some RNA viruses, on an RNA world.

1.5.4. Emergence of a Primitive ECM

It is likely that the appearance of the first ECM occurred in parallel with the first multicellular organisms.[60−62] The ECM-filled spaces between cells held cells together, conferred resistance during periods of stress, and promoted communication between cells. The ECM also transferred information from the environment to such organisms, as well as providing communication between cells.[60] The prototypes for such ancient ECMs must have been profoundly different from those that exist today.

The most primitive of eukaryotes currently on Earth may have an ECM that bears some resemblance of the original ECM. Some yeast species may fulfill such a definition, with their homopolymer cuticle composed predominantly of CT or its slightly deacetylated form, chitosan. The yeast *Saccharomyces cerevisiae* is protected from environmental damage by a multilaminar ECM, the spore wall, which is assembled during spore formation in response to severe environmental stress.[63] The initial ECMs had the additional function of protecting early life forms from the severe stress of dehydration incurred during the repeating cycles of wetting and drying. The polymeric glycans of the early ECMs, because of their negative charges, retained large solvent volumes of water that prevented dehydration. HA, for example, can retain a thousand-fold volume of water compared to the initial polymer volume. Other highly negatively charged polysaccharide polymers also would have had such water-retaining abilities.

1.6. An Early Pre-DNA/Protein World and Examples of Its Current Remnants

1.6.1. Ribozymes

Did a protein or glycan world exist that overlapped with an RNA world or did it appear later, with a world dominated by DNA? If proteins did not exist, then the majority of chemical reactions were performed by enzymatically active RNA, referred to as ribozymes,[64] or were catalyzed by clay surfaces as in the early prebiotic world. The putative RNA may not have disappeared entirely but, on the contrary, appears to have left some forms of evidence. Similarly, there may be remnants of the pre-RNA world present in current life. There are no known remnants of the proposed polymeric threose, glycerol, and pyranosyl type polymeric backbones with bases attached for self-replication as it has been suggested for possible pre-RNA world[14] (Figure 4). Several examples, however, are evident for the RNA carbohydrate world.

There are at least eight natural RNA ribozymes currently known that catalyze fundamental biological processes. Among such reactions is RNA cleavage by transesterification. They carry out either their own cleavage or the cleavage of other RNAs, and they also catalyze aminotransferase activity of the ribosome. Ribozymes in the laboratory are capable of catalyzing their own synthesis. Ribozymes are true fossil remnants of very ancient life forms.

1.6.2. Primitive Viral-Like Organisms

Single-stranded RNA viruses that resembled current RNA bacteriophage and plant viroids may have been the first free-living organisms and among the first organisms to have evolved from that prebiotic, pre-RNA epoch.[65] Genomic analyses suggest that such free living RNA viruses, paradigms for all current nucleic acid polymer functions, helped shape the three major domains of life, bacteria, archaea, and eukarya.[66]

RNA is a far less stable a polymer than DNA. It is also more flexible, more versatile, and more dynamic than DNA. DNA is more impervious to change. For example, RNA is susceptible to base-catalyzed hydrolysis, while DNA is not, a property based on the mere reduction of a hydroxyl group in the 2′-position of the ribose sugar. The highly stable sequences encoded in DNA have become the memory of the organism, from one cell division to the next and from one generation to the next, while RNA is involved in the necessities of daily living. This bicameral separation of functions led to the emergence of current life forms.

1.7. Polysaccharide Polymers Were Prerequisites for the Evolution of Life

We propose that carbohydrate polymers were not only the first polymers to appear on prebiotic earth but that their evolution was probably a prerequisite for the subsequent development of life. Their structures became ever more complex, and in parallel with such development, they were sequestered, so that various functions could occur independently. Among the most important reasons for such sequestration are as food or energy sources, protection against UV radiation, structural building blocks, energy storage, and as eventual components of the ECM, in particular glycosaminoglycans (GAGs).

1.7.1. Polysaccharides as Food and Stored Energy Sources as Well as Structural Moieties

The evolution of polysaccharides involved modulation of their linkages. This evolution made it possible for polysaccharides to function as both food and energy sources as well as structural units. In plants, polyglucose can occur in the forms of either starch or cellulose, in their α- and β-linkage forms, respectively.

The α-linkages of starches are easily degraded, thus becoming a readily available food source. Currently, the saliva of vertebrates contains α-amylase enzymes that begin this catabolic process. Additional α-amylases are secreted

by the pancreas into the gut to continue the degradation of starch and other α-linked sugar polymers. Glycogen is an α-linked polymer of glucose, a major form of stored energy for most vertebrates and one that is readily assembled. Starches, similarly, are α-linked polymers that provide a stored food form for plants.

The β-linkages as found in cellulose are more stable, as would be required for structural building blocks. A sequestration of the stored food sources is a necessity and would have occurred in the process of phase separation, as described for protocellular organelles. CT as well as some of the GAGs that constitute ECM components contain such structural or durable β-linkages. Some GAGs such as chondroitin (Ch), chondroitin sulfate (ChS), and HA are composed entirely of β-linked carbohydrate polymers, while others such as heparan, heparan sulfate (HS), and keratan sulfate (KS) contain combinations of β- and α-linked sugars.

The α-linkages in some of the latter GAGs acquired their conformation not directly from biosynthetic reactions. Dermatan sulfate (DS) begins with β-linkages that becomes modified α-structures by epimerization of the 5-carbon of glucuronic acid, converting the sugar to the L-configuration. The tentative suggestion can be posited that the more structural β-linkages are more ancient and that the α-linked sugars represent a later stage of evolution. Such GAGs may reflect ancient transitions, and the fragments released by the more easily degraded linkages may then have acquired separate biological functions.

1.7.2. Protection Against UV Radiation and Oxygen Free Radicals

The β-linked carbohydrate polymers have an intrinsic ability to protect against the damage of UV radiation. Following the experimental atomic blast on Bikini atoll in the late 1940s, the only organisms that were unaffected were the insects. Beetles and other insect with their CT exoskeletons were able to survive the intense radiation.

Also, the levels of HA increase in the dermal compartment of skin in response to UV radiation and with its water of hydration accounts for the swelling associated with sunburns. This is a stress response and protects the organism from subsequent exposures. Currently, the presence of ribose sugars in RNA or deoxyribose sugars in DNA chains confines radiation damage to a small local region and prevents major energy transfer throughout the polymer.[67] It can be surmised that the evolution of carbohydrate polymers was necessary as a preliminary step to protect the integrity of subsequent nucleic acid polymers from the intense UV radiation that was still present during those early epochs.

With time, the level of oxygen increased in the Earth's atmosphere. This increasing oxygen protected early life forms from UV radiation. However, the destructive and degradative activity of high-energy reactive oxygen species increased, as the damage from UV radiation decreased. Antioxidant activities of polysaccharides became essential in protecting life forms. The degradative action of reactive oxygen species was absorbed by polysaccharides such as CT, providing a safety shield for evolving organisms.

Before the evolution of free oxygen, approximately 2.3 billion years ago, there was predominantly a reducing environment on the Earth's surface. By some genetic mechanism, organisms evolved that were able to extract four electrons from two molecules of water to form free oxygen as a metabolic waste product. This molecular oxygen was highly useful as an electron acceptor. However, it was also potentially damaging to the metabolic networks that had already evolved.[68] The triplet ground state of oxygen is highly reactive, with a propensity to generate peroxides, superoxides, and hydroxyl radicals. The polysaccharides present in cells and on cell surfaces served as protective moieties against these highly destructive metabolites of oxygen.

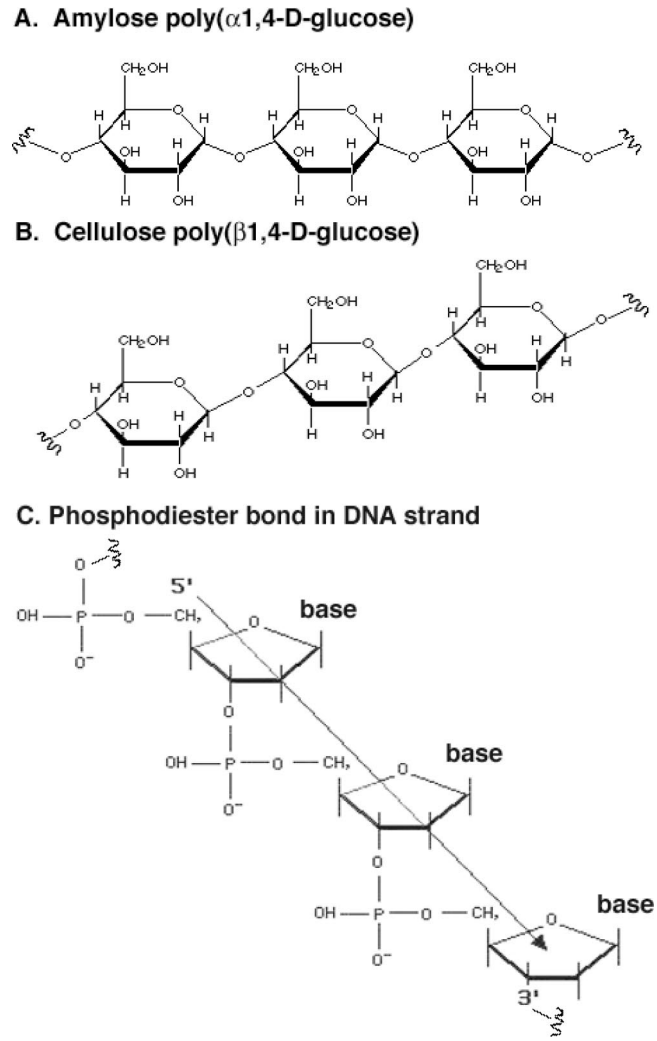


Figure 5. Examples of prevalent carbohydrate linkages. (A) α-1,4-Glycosidic linkage as in amylose, (B) β-1,4-glycosidic linkage as in cellulose, and (C) a phosphodiester bond as in nucleic acids.

2. Current Polymeric Glycans

On the prebiotic Earth, simple formol condensation accounted for the appearance of most of the glycans. The most abundant and stable glycans are the elementary keto- and aldo-hexoses, fructose, glucose, and mannose (Figure 3B). Galactose seems to have made its appearance later in evolution. The majority of biological glycans are D-enantiomers and occur in a β-conformation.[15] The β-conformation is more stable and more resistant to hydrolysis than the α-conformation (Figure 5), a relatively stable structure being one of the requirements for the process of evolution.

An overview is presented of the major polysaccharides currently present on Earth. The major polymeric glycans are, in alphabetic order, alginate, cellulose, CT, glycogen, pectin, pectinlike substances, and starch. They are briefly introduced, and their main properties are described.

2.1. Structural Polysaccharides

2.1.1. Cellulose

Cellulose, a polymeric β-1,4-linked homopolymer of D-glucose, is the building element of plants, providing shape and structure. Additionally, cellulose must have enormous weight-bearing properties, with the ability to withstand osmotic pressures as high as 20 atm between extracellular and intracellular spaces. Cellulose accounts for half the carbon in the biosphere and is the most abundant carbohydrate polymer and the most abundant polysaccharide on Earth. Cellulose is water insoluble, despite being a very hydrophilic substance. X-ray structures may explain this anomaly (for a review see Bayer et al.[69]), with parallel chains lying in alternating perpendicular patterns,[38] stabilized by intermolecular hydrogen bonds between glucose units of neighboring chains.

In the plant cell wall, cellulose fibers are cross-linked by a number of polysaccharides containing glucose and other sugars. In wood, the cross-linking agent is predominantly lignin, a polyphenol that interacts with cellulose and the hemicellulose matrix of the cell wall. It is one of the most abundant organic compounds on Earth after cellulose and CT. Lignin constitutes about one-third the dry mass of wood. Several plant genomes are now available, including rice, *Arabidopsis*, a plant in the mustard family, and most recently, a tree, the black cottonwood, a member of the poplar family. It has 45000 genes, twice the amount contained in the human genome. Of these, 93 genes of the tree genome are involved in cellulose synthesis, and 34 are involved with lignin synthesis.[70]

Cellulose occurs only once in eukaryotes, in the outer mantle of the free-swimming larval form of tunicates, the marine invertebrates also known as sea squirts. This represents probably a lateral transfer of a cellulose synthase gene. The evolutionary advantage of cellulose expression in a free-swimming larva is not known, nor why cellulose expression ceases in the adult sessile form of the sea squirt. The tunicate larva may be a remnant of a discarded experiment of nature on the pathway to chordates and not a true predecessor.

2.1.2. CT

CT, polymeric β-1,4-linked *N*-acetyl-D-glucosamine, is possibly the ancestral polymeric glycan, and it is arguably the most abundant glycan in animals, known mostly as the building element of exoskeletal shells of invertebrates such as crustaceans, insects, and spiders, as well as being present in the cell walls of fungi and many algae. In effect, CT may be described as cellulose with one hydroxyl group on each monomer replaced by an acetylamine group, allowing for increased hydrogen bonding between adjacent polymers. This gives the polymer increased strength.

In its unmodified form, CT is translucent, pliable, durable, and resilient. In arthropods, however, it is frequently modified, by being embedded in a hardened proteinaceous matrix, which forms much of the exoskeleton. CT is the second most abundant polysaccharide in nature, after cellulose. At least 10 gigatons of CT are synthesized and degraded each year in the biosphere, most of it in the oceans.

2.1.3. Bacterial Cell Wall and Surface Glycans

Nearly as important and ubiquitous as cellulose and CT is the building block of bacterial cell walls. Bacteria and other microbes, including fungi and algae, secrete such polysaccharides as an evolutionary adaptation to help them adhere to surfaces, as virulence factors, and to prevent them from drying out. Pathogenic bacteria, in addition, often produce a thick, mucouslike layer of polysaccharides or a capsule that cloaks antigenic proteins on bacterial surfaces that would otherwise provoke an immune response. Capsular polysaccharides are water soluble and are linear arrays of regularly repeating subunits of 1−6 monosaccharides. There is enormous structural diversity, with over 200 different polysaccharides synthesized by *E. coli* alone.

The three-dimensional (3D) structure of such bacterial cell wall glycans was until recently one of the important unsolved structural problems. Such a wall is built primarily of repeating units of β-1,4-linked *N*-acetyl-D-glucosamine (NAG) and β-1,4-*N*-acetyl-D-muramic acid (NAM). The latter has cross-linked peptide stems that provide the required structural integrity. A honeycomb helical structure with pores and channels has been documented.[71]

2.1.4. Consequences of Structural Polysaccharide Evolution

The advantage of the polysaccharide shield of Earth's organisms comes at a cost. Every cell in nature is now covered with a protective shield of polysaccharides or oligosaccharides. Some of these shields have evolved into dense glycans with a remarkable structural diversity. However, a range of pathogens have evolved in parallel and use these complex carbohydrates for binding to cells. The rapid evolutionary changes that surface carbohydrates undergo are perhaps driven in an attempt to evade pathogen binding and infection. The dynamic reciprocity between pathogens and hosts may be a major driving force in the evolution for selection and speciation.[72−74] Simultaneously, they have evolved as recognition systems, resulting ultimately in immune recognition.[75]

2.2. Storage Polysaccharides

2.2.1. Starch

Starch, a mixture of α-amylose (α-1,4-linked D-glucose) and amylopectin (α-1,6-branched every 24−39 D-glucose of "α-amylose" backbone), is an insoluble storage polysaccharide for plant cells and the main source of dietary carbohydrates. Starch is a food reserve for plants as well as animals. It is deposited in the cytoplasm of plant cells as insoluble granules of α-amylose and amylopectin. α-Amylose is a linear polysaccharide composed of thousands of α-1,4-linked glucose molecules. Amylopectin consists predominantly of α-1,4-linked glucose molecules that are branched as described above. Amylopectin may contain as many as 10^6 residues of glucose, with a molecular mass of 1.8×10^8 Da. Thus, amylopectin and HA are among the largest molecules found in nature. Although the compositions of starch and cellulose are similar, their overall structures are profoundly different.

2.2.2. Glycogen

Glycogen, the starch equivalent in animal cells, is similar to amylopectin but is branched every 8−12 glucose residues. It is the less soluble storage form of glycan for animal cells. Glycogen is the energy storage form for animal cells, comparable to starch in plant cells. It is most abundant in

hepatocytes and in skeletal muscle cells, where it is associated with a mini-organelle termed the glycogen granule.

The primary structure of glycogen resembles the amylopectin of starch but is more highly branched. Glycogen synthase activity is modulated by insulin. Glycogen phosphorylase is a complex enzyme controlled through a network of other enzymes in response to caffeine, epinephrine, and glucagon and through calcium ion flux, to muscle contraction. The highly branched structure of glycogen has many nonreducing ends, permitting mobilization in rapid response to metabolic requirements. Glycogen does not possess a reducing end. The reducing end glucose is not free but is covalently bound to the protein glycogenin in a β-linkage to a surface tyrosine residue.

2.2.3. Pectins

Pectins are a heterogeneous group of plant polysaccharides with a complex structure depending on their source. The majority of the structure consists of homopolymers of partially methylated poly-α-1,4-D-galacturonic acid residues, but there are substantial nongelling areas of alternating α-1,2-L-rhamnosyl-α-1,4-D-galacturonosyl sections containing branch points with mostly neutral side chains containing from one to 20 residues of mainly L-arabinose and D-galactose.

Of all plant cell polymers, pectins have the greatest number of functions. They make up part of the cell wall, but they also make up a layer between adjacent cell walls, that is, the middle lamella that binds cells together. Pectins also form complexes with many globular proteins.[51]

2.2.4. Alginate

Alginate, which is β-1,4-linked D-mannuronic acid, coats surfaces of many forms of brown seaweed and marine kelp and is found in certain bacteria. They are linear unbranched polymers containing β-1,4-linked D-mannuronic acid (M) and α-1,4-linked L-guluronic acid (G) residues. Although these residues are epimers, D-mannuronic acid residues are enzymatically converted to L-guluronic after polymerization and only differ at C5. Alginates are not random copolymers but, according to their source, consist of blocks of similar residues (i.e., MMMMMM, GGGGGG, or GMGMGMGM), each of which has a different conformation and different behavior.

2.3. ECM and GAGs

Vertebrates evolved GAGs originally for deposition into the extracellular space as components of the ECM. Since then, the GAGs have evolved many complex regulatory functions, too extensive to be reviewed here. They function predominantly to bind growth factors and to either hold them in abeyance or for presentation to cells. When cell-associated, GAGs constitute components of receptors, serve as adjuncts in signal transduction pathways, or are themselves signaling molecules. The GAGs, with the exception of HA, are associated with a class of glycoproteins termed proteoglycans.

2.3.1. Heparin and HS

Among early ECM developments was the appearance of GAGs. The original GAG appears to have been a heparin or heparin-like polymer.[60,76] The first of such structures occurred in the metazoan life that emerged from their unicellular choanoflagellate ancestors about 640 million years ago. Analysis of genome sequences also supports the concept that heparin-like polymers were the first ancestral GAGs (reviewed by DeAngelis[76]). Heparin is variably sulfated, with an average of 2.5 sulfates per disaccharide unit, making it one of the most negatively charged polyelectrolytes. The most common repeating disaccharide unit in heparin is composed of a 2-*O*-sulfated iduronic acid and 6-*O*- and *N*-sulfated D-glucosamine. Unlike all other GAGs, it is not associated with connective tissues or the ECM but is found in mast cells in mammalian tissues. It is best known for its ability to inhibit clot formation and may have had evolutionary significance to halt bleeding and, thus, to promote rapid recovery following injury. HS has a far more variable composition, with the most common repeating disaccharide unit of a D-glucuronic acid linked to *N*-acetyl-D-glucosamine. It is often found embedded in cell membranes and, despite its name, is less sulfated than heparin. The HS family of proteoglycans includes the syndecans,[77] perlecan,[78] glypicans,[79] and betaglycan.[80]

2.3.2. Ch and ChSs

It is speculated that the Ch, β-1,4 linked repeating units of D-glucuronic acid with a β-1,3-link to *N*-acetyl-D-galactosamine, appeared next, after the heparin and heparin sulfate, followed much later by HA, β-1,4-linked repeating units of D-glucuronic acid with a β-1,3-link to *N*-acetyl-D-glucosamine.[76] Chs have galactose in their composition instead of glucose. Galactose, originating from glucose, is a sugar that may have appeared much later than glucose, containing an epimerization at the C4 carbon.

The first organisms to produce Ch or the first to produce HA are still unknown, although it is assumed these were multicellular organisms. The analyses of genomic sequences show, for example, that the worm *Caenorhabditis elegans* (Nematoda) and the fruit fly *Drosophila melanogaster* (Insecta) both have heparin, HS, and Ch. *C. elegans* has predominantly the unsulfated form of ChS. Neither of these organisms, however, contain HA.[81] Therefore, it is suggested that the second GAG polymer to evolve was Ch. It also appears that sulfation of GAGs occurred later in evolution. The evolution of Ch prior to HA seems to be in conflict with the appearance of galactose, one of the building blocks of Ch. It occurs later in evolution than the three original hexoses: fructose, glucose, and mannose. By deductive reasoning, one can postulate that the polymeric glycans built from the three original hexoses would have evolved earlier than those containing galactose. Therefore, HA should have preceded Ch in evolution. However, this appears not to be the case. The resolution of the conundrum must await further investigation.

It is also possible that the glycosyl transferases that synthesize ChS evolved from the HS transferases as their mechanism appears similar.[76] HA synthesis (as described below) appears to have evolved from chitin synthases (CTSs) to establish a completely different mechanism.

Chondroitin-4-sulfate (C4S) is a major constituent of cartilage and occurs as a component of ChS-rich proteoglycans such as aggrecan and versican. These ChS-rich proteoglycans often bind HA through their protein core proteins (PGCP) and are known together as the lectican family of proteoglycans. Aggrecan contains a great number of KS chains attached to its core protein, while versican has none. Why such major differences should occur is not totally clear, but KS in aggrecan contributes much more charge density to the macromolecule, which in turn gives cartilage greater

resistance to compression under load. Versican, on the other hand, has a similar domain volume but much less charge density. Therefore, it can compress much more, which is more suitable for tissues such as arteries that undergo large changes in volume.

The HA chains form the center of many associated ChS-proteoglycans in a bottlebrush configuration. Chondroitin-6-sulfate (C6S) is sulfated at the 6-position of *N*-acetyl-D-galactosamine, instead of the 4-position and is scattered throughout a number of tissues. Most tissue ChS, however, contains both 4- and 6-sulfate groups but in much different proportions.

2.3.3. DS

DS is so named because of its prevalence in mammalian skin. It is in reality a variant of C4S, with an inversion about the configuration of C5 of β-D-glucuronate to form α-L-iduronate. This results from epimerization of Ch in its passage through the Golgi apparatus. It is also referred to as ChS B. It is found in skin but also in blood vessels, heart valves, tendons, and lungs. DSs have roles in coagulation, cardiovascular disease, carcinogenesis, infection, wound repair, and fibrosis. DS accumulates abnormally in several of the mucopolysaccharidoses.

2.3.4. KS

KS is the most heterogeneous of the GAGs. It is a linear polymer that consists of a repeating β-1,3-linked disaccharides of D-galactose linked by the 1,4-linkage to *N*-acetyl-D-glucosamine. Its sulfate content is highly variable and contains small amounts of mannose, sialic acid, and fucose. Various forms of KS are particularly prominent in the cornea, cartilage, and bone. KS-containing molecules have also been identified in numerous epithelial and neural tissues in which KS expression occurs in embryonic development and in wound healing. Evidence supports functional roles of KS in receptors on cell surfaces, in cellular recognition of multiple ligands, as well as being key to axonal guidance, cell motility, and in embryonic implantation.

2.3.5. HA

HA appears to have evolved from Ch because of increasing environmental pressures, resulting in substituting every *N*-acetyl-D-galactosamine with an *N*-acetyl-D-glucosamine. However, as described above, it cannot be rejected that HA might have evolved from HS, the original GAG.[76] The hydroxyl group at the C4 position in galactosamine is frequently used for immune recognition in eukaryotes. Epimerization of that hydroxyl group provides the stealthlike qualities of HA and provides an ability to avoid immune recognition. HA may have emerged when organisms reached a critical size, when embryonic cells or stem cells were required to travel some distance in the process of development, or when special compartmentalized regions of the organism were required for the maintenance of stem cell-like cells. The emergence of the specialized stem cell "niche" may correspond with the evolution of HA.[82,83] The HA pathway is often commandeered by cancer cells for their own survival, migration, and proliferation. It is tempting to predict that the presence of HA correlates with the first occurrence of malignancy in the evolution of organisms.

By contrast, except for the capsule of some bacteria, no HA occurs in invertebrates. The HA capsules of these bacteria function as virulence factors, promoting adhesion to host vertebrate tissues. The ability to synthesize HA may represent a lateral transfer of genes from hosts to the bacteria. On the other hand, HA is universally present in vertebrates.[76] The enzymes for HA catabolism do occur, however, in invertebrates. Hyals (haluronidases) occur in bacteria, bacteriophage, bacteria, fungi, parasites such as hookworms, nematodes, annelids such as leeches, insects, particular insect venoms as found in wasps and bees, and in crustaceans.[84]

Lack of reliable data makes it difficult to resolve the conundrum regarding the consecutive origin of the various GAGs. Much data support the concept that HA is a late arrival on the evolutionary scene. All other GAGs are synthesized in the Golgi apparatus and are covalently bound, using a common motif, to proteins. The proteoglycan core protein (PGCP) to which the GAGs are attached are together termed "proteoglycans". HA, on the other hand, is synthesized in vertebrate cells in an entirely different manner. The HA synthases (HAS) enzymes are embedded on the cytoplasmic surface the plasma membrane. The growing HA chain is extruded through the membrane into the extracellular space as it is being synthesized. The enormous size of the HA polymer together with its solvent volume of hydration, if it were retained within the cell, would create great havoc. HA is the only GAG not associated with a PGCP that is not sulfated (other than Ch) and that is not synthesized in the endoplasmic reticulum (ER) and Golgi apparatus.[85]

2.4. Polysaccharides of Other Glycoproteins

2.4.1. Branched Polysaccharides of Mucins

Mucins are a family of heavily glycosylated proteins, with a mass of between 1 and 10 million Da, some membrane-bound and others that are secreted onto mucosal surfaces, predominantly in the lungs and gastrointestinal tract, and are present in high concentrations in saliva. They are a major component of the protective Biofilm barrier on all mucosal surfaces. The dense "sugar coating" provides the capacity for an extensive volume of water of hydration. Their predominant function may be to bind and inactivate microorganisms. Their production in response to malignancies remains enigmatic. The apomucin genes constitute a wide gene family. There are at least 19 such genes in the human. A large central region of multiple tandem repeats contain sequences, 50% of which are serine and threonine, for the attachment of *O*-linked oligosaccharides. *N*-Linked oligosaccharides are also present at much lower concentrations.

2.4.2. Oligosaccharides of Heterogeneous Glycoproteins

Polylactosamines contain repeats of the *N*-acetyllactosamine units and are a fundamental structure of glycans carried on *N*- and *O*-glycans of glycoproteins [86,87] and glycolipids.[87] Polylactosamine oligosaccharide synthesis and its subsequent addition onto glycoproteins is an immune regulatory factor presumably suppressing excessive responses during immune reactions.[88] Similarly, the addition of polylactosamines on to CD44, the cell surface receptor for HA, appears to be involved in the cleaning up apoptotic granulosa cells.[89]

Also, another oligosaccharide, polysialic acid, a polymer of sialic acid, is known to be covalently linked to cell

adhesion molecules.[90] Polysialic acid is an unusual post-translational modification that occurs on neural cell adhesion molecules (NCAM), for example.[91] Polysialic acid is considerably anionic. This strong negative charge gives this modification the ability to change the proteins surface charge and binding ability. In the context, polysialation of NCAM prevents its ability to bind to NCAM's on the adjacent membranes. Similarly to polylactosamines, polysialic acid plays regulatory roles in biology. For example, polysialic acid is involved in immunogenicity and also in neurogenes, in fertilization, and early development. The functional significance of chain length of polysialic acid has been shown in terms of its stage-dependent, tissue-specific, and carrier protein-specific expression.

Inhibition of cell adhesion through repulsive interactions between polyanionic molecules such as polylactosamines and/or polysialic acid on the cell surface has also been described.[90] Thus, it appears that polysaccharide polymers have had a long evolutionary history in relation to control of cell functions. The ability of the carbohydrate chain length to be endowed with specific functions appears to be a newly emerging theme in biology.[92]

Other polymeric and polymeric-like substances such as polyprenols, dolichols (Dols), and dolichol-phosphates (Dol-P) have many important functions,[93] most of which are restricted to membranous environments. For example, Dols and Dol-P are carriers of lipid oligosaccharyl moieties and their synthetic precursors. These include Dol-P-Mann, the initial precursor, all occurring within bilayer type membranes. They are used to glycosylate nascent polypeptides as they cross ER membranes. Both *N*-linked and *O*-linked protein glycosylation is involved.[93−96] The compromised function of Dols[97] and the protein glycosylation that they facilitate result in human disease.[98] In addition, Dols take part in CT synthesis as carriers of the product of CT polymer synthesis.[99]

3. Increasing Complexity of Polysaccharide Structures and Functions

3.1. Interglycosidic Bonds

Some GAGs, such as Ch, ChS, KS, and HA, contain only β-bonds. However, a mixture of α- and β-bonds are contained within other GAGs, such as heparin, HS, and DS. The evolutionary advantage of such hybrid structures is not known. A difference in cleavage susceptibility has been outlined previously. The greater ease of cleavage of the α type bond supports a hypothesis that GAG fragments might have biological activity. HA fragments are known to encompass a variety of biological activities and are referred to as an "information-rich system".[39] Whether GAG fragments may also have biological activities has not been investigated. Additional evidence comes from the observation that DS is a derivative of ChS, with the inversion about the configuration of C5 of β-D-glucuronate to form α-L-iduronate. This occurs in the passage through the Golgi apparatus. Nature has gone out of its way to change a β-glycosidic bond to an α-bond. This may constitute a change in 3D structure, but it also suggests that a change in the catabolic pathway is occurring, with additional putative functions conferred onto the resulting cleavage products. The other major intermolecular bond for polysaccharides is the phospho-diester link.

3.2. Polysaccharides Can Be Informational

The ease of formol type synthesis of glycans on primordial Earth suggests that such a glyco-rich world did exist. This world is currently evident in a number of polymeric glycans that are, undoubtedly, information-rich systems. The "remnant" glycans from the early world (i.e., glycans currently present) are essential for recognition and in ligand−cell, cell−cell, and cell−organelle signaling type of communication. While DNA/protein information is encoded predominantly in linear sequences, polymeric glycans code information in their sequence, by way of their branching patterns, in postsynthetic modifications, such as sulfation, phosphorylation, and epimerization patterns, and finally in higher order secondary and tertiary structural properties. Some carbohydrates utilize the length of their polyglycan chains as the method for information coding, especially those carbohydrates that are the simple polymers such as poly-β-glucose (cellulose), poly-β-*N*-acetylglucosamine (CT), poly-β-(glucuronic acid-*N*-acetylgalactosamine) or Ch, and poly-β-(glucuronic acid-*N*-acetylglucosamine) or HA.[100]

3.2.1. Fragment Size: HA Polymers Contain Size-Specific Information

HA is an information-rich system.[100] The HA fragment sizes can themselves be informational. An ambiguity exists here in distinguishing between HA specific fragment size that is informational because of an ability to interact with other cellular components and information secondary to the physical-chemical configuration specific for that sized fragment. Shape, aggregation, and other manifestations of higher order structure can be informational and are also responsive to environmental condition. In all probability, it is a combination of these two phenomena that results in information transfer coded within the size of this polysaccharide.

3.2.2. Information Coded Within Sulfation Patterns

Other GAG chains such as heparin, HS, and ChS are also informational as HA is, but in addition, they code information using other properties such as their sulfation patterns. Anticoagulant effects of heparin and HS occur, for example, as a result of specific sulfation patterns.[101] Also, recent exploration of ChS chains reveals that highly specific sulfation patterns can either stimulate or inhibit neurite outgrowth.[102] Although GAGs contribute to diverse physiological processes, an understanding of their molecular mechanisms has been hampered by the inability to access homogeneous GAG structures. Well-defined ChS oligosaccharides have now been generated using a convergent, synthetic approach that permitted placement of sulfate groups at precise positions along the carbohydrate chain. Using such defined structures, specific sulfation motifs were shown to function as molecular recognition elements for growth factors and modulation of neuronal growth. These results indicate that there is a "sulfation code" whereby GAGs contain functional information in a sequence-specific manner analogous to that of DNA, RNA, and proteins. However, a thorough investigation, to date, has not been performed.

4. Polysaccharide-Active Proteins/Enzymes

With the emergence of the DNA/proteins world, numerous enzyme proteins evolved to synthesize, modify, and degrade the carbohydrates necessary for the existence of current life.

Examples of such proteins with enzymatic activities that act upon carbohydrates are provided below.

4.1. β-Linked Polysaccharide Synthesizing Enzymes

The genes that synthesize the β-linked polysaccharide polymers, CT, cellulose, and HA, have sequence homologies, suggesting that they are all descendents of an ancient single β-polysaccharide synthesizing gene.[103]

4.1.1. Cellulose Synthases

The completed genome sequence of rice (www.prl.ms-u.edu/walton/CSL_updates.htm) has provided insight into the cellulose synthase superfamily (for a review, see Keegstra and Walton[104]). Over 40 cellulose synthase type genes are present in rice. This superfamily is comprised of a number of subfamilies, some of which are found in all plants, some form cellulose, others hemicellulose, and others whose functions are uncertain.

Hemicelluloses themselves are a complex family each with different additional sugars attached requiring specific enzyme activities. Hemicellulose contains a variety of different sugar monomers, in contrast to cellulose, which contains only glucose. Hemicelluloses contain most of the D-pentose sugars and small amounts of L-sugars as well. Xylose is always the sugar monomer present in the largest amount, with mannose and galactose occurring at lower concentrations. Hemicelluloses are imbedded in the cell walls of plants, in chains that bind pectin to cellulose to form a network of cross-linked fibers.

4.1.2. CTSs

Although there are gigatons of CT synthesized and degraded yearly in the biosphere, little is known about the enzymes involved. Most of the fundamental information has been derived from common yeast, *Saccharomyces cerevisiae*. The major CTS activity is not required for synthesis of the chitinous primary septum. The mechanism of in vivo synthesis of CT has been clarified by cloning the structural gene for CTS 2, a relatively minor activity in yeast. Disruption of this gene results in loss of septa and in growth arrest, establishing that the gene product is essential for both septum formation and cell division.

4.1.3. HA Synthases

4.1.3.1. Origins and Prokaryotic Enzymes. Sequence homologies suggest that HA synthases (HAS) evolved originally from CTSs.[105] All organisms known to date that produce HA utilize HAS genes. These organisms usually have 3−4 isoforms of such genes. [76,106] There are two major groups in which studies of HASs have been carried out:, the bacterial enzymes, for example, from *Streptococcus pyogenes*,[107−109] and the vertebrate enzymes.[106,110−112] They are likely to be modular, and their functional properties are well-studied.[105,109] No 3D structures are available for *S. pyogenes* and for human HAS enzymes. It is likely that three gene duplication events occurred at sometime in their evolution.[106] The existence of an original ancestral HAS gene can be predicted, although the isoform closest to the ancestral one has not yet been identified.

The identification of this gene or further evidence of HAS genes evolving from CTS could facilitate resolution of the order of evolutionary sequence for HA and CT. Indeed, nodulation protein C (NodC) *Rhizobium* sp. N33 (NodC) is an *N*-acetyl-glucosaminyl transferase synthesizing CT backbone in that organism. NodC, CTS, adds UDP-*N*-acetyl D-glucosamine at the reducing end of growing CT chains.[113−115] Another protein differentially expressed in gastrulation (DG42) in *Xenopus laevis* (African clawed frog) was identified by homology to NodC and has HAS1 synthase activity.[106,116] In zebra fish, DG42 is expressed during early embryogenesis.[117] The DG42/HAS1 proteins are able to synthesize HA or CT, dependent upon conditions. For example, in mouse, HAS1 or DG42 can express either HA or CT in vitro.[118] This demonstrates that either polymers, HA or CT, can be synthesized by the same enzyme,[114,119,120] reflecting perhaps a weakness of the putative allosteric substrate effect upon the enzyme that normally would facilitate catalysis of alternating substrates during the synthetic process.

Not surprisingly, both group of enzymes belong to the same group of glycosyltransferases enzymes in Carbohydrate-Active EnZymes (CAZy) database, http://afmb.cnrs-mrs.fr/CAZY/, glycosyltransferase family 2 (GT-2).[121] Examples from this family of enzymes include cellulose synthase (EC 2.4.1.12), CTSs (EC 2.4.1.16), dolichyl-phosphate-D-mannosyl transferase (EC 2.4.1.83), dolichyl-phosphate-glucosyltransferase (EC 2.4.1.117), *N*-acetylglucosaminyltransferase (EC 2.4.1.-), *N*-acetylgalactosaminyltransferase (EC 2.4.1.-), HAS (EC 2.4.1.-), chitin oligosaccharide synthase (EC 2.4.1.-), and -1,3-glucan synthase (EC 2.4.1.34). Each of these enzymes have similar functions and a similar inverting mechanism of action.[122]

There is only one enzyme from the entire GT-2 family for which there is structural information, SpsA from *Bacillus subtilis*.[123−125] Remarkably, SpsA's precise function, substrates, and products are unknown.[126] A modeled structure of dolichyl-phosphate-D-mannosyltransferase (EC 2.4.1.83) has recently become available, shedding considerable light on structural and mechanistic properties of this family of enzymes.[126] CTSes (EC 2.4.1.16) transfer UDP *N*-acetyl-glucosaminyl groups to CT to facilitate the growth of CT chains using β-1,4-linkages with UDP as an additional product. HAS enzymes (EC 2.4.1.212) utilize UDP-*N*-acetyl-D-glucosamine and UDP-D-glucuronate, adding these glycans in a strict alternating manner to the growing HA chain, one sugar unit at a time, using β-1,4- and β-1,3-glycosidic linkages. UDP and increasing lengths of HA chains are generated as products.

It is not surprising, therefore, that CTSs are able to accept D-glucuronate in addition to *N*-acetyl-D-glucosamine as substrates. HAS enzymes presumably evolved to accept both types of substrate. Perhaps, once D-glucuronate is being added to the growing HA chain, its fit in the HAS active site that favors the 1,3-linkage instead of 1,4-, as in the case of *N*-acetyl-D-glucosamine. If a HAS such as DG42 is provided only *N*-acetyl-D-glucosamine residues, in its UDP form, it produces 1,4-linked linear polymers of CT only.[106,120,127,128] Therefore, it is tempting to suggest that the type of the glycosidic bond formed is primarily a function of the residue being added to the growing chain. If *N*-acetyl-D-glucosamine is added, β-1,4-linkages are formed, and if D-glucuronate is added, β-1,3-linkages are produced for at least some, if not all, HAS and CT synthases.

It is, however, a mystery how the enzyme understands which residue to add in order to synthesize an HA polymer

with strictly alternating glycan residues. It is conceivable that the active site changes its conformation by an allosteric-like mechanism, responding to the terminal residue of the growing chain. Amino acid substitutions that would endow the enzyme proteins with such allosteric mechanisms could be considered a further evolutionary step. The presence of one sugar at the growing end of the polymer confers a change in the protein so that the second sugar becomes the preferred substrate for the next step in chain growth. Such allosteric modifications induced by two alternating substrates may mark the transition from the homopolymer to the alternating polymer.

The question then arises as to what the advantage might be, from the evolutionary perspective, of alternating β-linked polymers, such as Ch and HA, over a β-linked homopolymer, such as CT. Complex 2- and 3D structures that are themselves informational may provide the answers, structures that are not yet available for the strict homopolymer enzymes. Alternatively, HA chains bound to HA synthase enzymes or CT that remained electrostatically bound to CTSs may have provided early forms of glycoproteins, with functions profoundly different from the unadorned proteins.

Ch thus appears to have evolved separately from either CT or HA. The enzymes synthesizing Ch, the Ch synthases, utilizing a similar inverting mechanism, appear to be different from either HAS or CTSs. Ch synthases belongs to CAZY family 31 of GT, GT-31. This entire family does not have any representative with a known 3D structure. In addition, it has a molecular mass much larger than that associated with either CTS or HAS. One might suspect that once galactose evolved from glucose, an already existing CTS enzyme would have been utilized for the evolution of CH. However, major differences between HAS and CTS on one side and CHS on the other preclude such a possibility.

4.1.3.2. Eukaryotic Hyaluronan Synthases (HASs). The enzymes that synthesize HA are the HASs, HAS-1, -2, and -3. They were not described until the 1990s. These multipass transmembrane proteins lose activity when solubilized and have thus been difficult to study. They polymerize the HA chain on the intracellular membrane surface, in contrast with all other GAGs. The latter are synthesized in the rough ER attached to core proteins that together generate the PG end products. During synthesis, the growing HA polymers are extruded through the plasma membrane onto the pericellular surface sheath termed the glycocalyx or into the ECM. The process of membrane extrusion uses the multidrug resistance system[129] with homology to the bacterial ABC (ATP-binding cassette) transporter system.[130] However, this is not universally accepted, and other mechanisms of HA extrusion from the cell are invoked.

Extrusion of the growing HA chain extracellularly through the plasma membrane permits unrestrained growth of the polymer, so that it can reach 1000−10000 kDa. Synthesis of such an enormous polymer, containing up to 25000 disaccharide units, could not be possible intracellularly. Nor could the attendant high viscosity be tolerated within the cell. The three-member HAS isoenzyme family, localized to three separate chromosomes, was identified in both human and mouse genomes (for reviews, see refs 105, 110, and 131). Sequence data indicate that there are seven transmembrane regions and that a central cytoplasmic domain contains consensus sequences that are substrates for phosphorylation by protein kinase C.[132]

The catalytic rate and mode of regulation for each isozyme is different.[110] HAS1 is the least active and drives the synthesis of high molecular weight HA, suggesting low constitutive levels of synthesis. HAS2 is more catalytically active and also generates a high molecular weight form of HA.[133] This may be the HAS enzyme that responds to stress-induced increases in synthesis, as found in shock, septicemia, inflammation, massive wounding, after major blood loss, and in burn patients. HAS2 is also implicated in developmental and repair processes involving tissue expansion and growth. HAS3 is the most active HAS enzyme and drives the synthesis of large amounts of lower molecular weight HA chains.[133] The products of HAS3 may provide the pericellular glycocalyx and the HA that interacts with cell surface receptors. Such shorter HA chains may trigger cascades of signal transduction events and major changes in cellular behavior. However, the context and conditions of these experiments need to be considered in such evaluation. For example, overexpression of HAS3 led to extended coats of HA chains on microvilli-like structures, and such HA would appear to be rather large all things considered and contradict the above statement regarding small HA size.[134,135]

The HAS proteins may be part of a larger protein complex,[136,137] components of the hyalosome that regulate enzyme activity and coordinate interactions with other cellular components. The hyalosome is a putative miniorganelle within cells that contains a complex of synthetic and degradative activities for the regulation of HA deposition, together with regulatory proteins and hyaladherins that respond to the metabolic schemes of the cell.[138,139]

4.2. Polysaccharide β-Endoglycosidase Degrading Enzymes

Enzymes that degrade the more resistant carbohydrate polymer structures evolved independently from those that catabolize polymers such as starch and sugar. Despite the existence of such degradative enzymes, remains of ancient creatures are still available, hundreds of millions of years later. Many are hydrolases, and the dehydrated states of their substrates have prevented their enzymatic action.

4.2.1. Cellulases

As much as 10^{15} kg of cellulose are degraded annually by the cellulase class of enzymes.[140] They are found in bacteria, some of which are in the digestive tracts of herbivores, since vertebrates do not possess such an activity. Fungi and termites also contain the enzymes and are largely responsible for the removal of decaying trees, plants, and of course attacking houses. The degradation of cellulose, the cleavage of the β-1,4-glycosidic linkage of cellulose, is a slow and complex process because of its tight packing and extensive hydrogen bonding, even after a great number of glycosidic bonds have been cleaved.

Three general types of enzymes make up the cellulase enzyme complex. Endocellulases cleave internal bonds to disrupt the crystalline structure to expose individual cellulase polysaccharide chains. Exocellulases cleave 2−4 saccharides from the ends of the exposed chains resulting in the tetrasaccharides or disaccharide such as cellobiose. Cellobiases or β-glucosidases hydrolyze endocellulase products into individual monosaccharides. Within the cellulases, there are progressive and nonprogressive types. Progressive cellulases continue to cleave a single strand of cellulose, while

nonprogressive cellulases interact once and then disengage and engage another cellulose strand in a random manner.

There is now intense interest in plant cellulases. Genetic engineering of plants will provide biofuels as an energy source of the future that will no longer be dependent on oil. Plant cellulases are not as thoroughly studied, despite their potential importance. They are also critical for the modulation of plant structure that occurs with growth and age. It may also be possible to formulate a cellulosome that parallels the glycogen granule and hyaluronasome miniorganelle described earlier.

4.2.2. Chitinases

The origin of CT, β-1,4-linked *N*-acetyl-D-glucosamine, is placed in metazoan life. Many invertebrates contain CT, but it is entirely absent in vertebrates. Chitinases are found in many bacteria, in yeast, in fungi, and in plants. Some plant chitinases are members of the pathogenesis-related (PR) inducible proteins. Expression is mediated by the NPR1 gene and the salicylic acid pathway.[141]

The chitinase enzymes that degrade CT and its partially deacetylated form, chitosan, are very potent. Yet, they are a relatively neglected family of enzymes. This is remarkable considering that they are among the key enzymes of marine biology. An example of the potency of chitinases is found in the biology of krill. These tiny shrimplike crustaceans, congregating in large, dense, swarmlike masses in the oceans, are the major food source of baleen whales, sea birds, and other predators. When a spoonful of krill is warmed to room temperature, they expire, rapidly autodigest, and the solution becomes water clear, the CT exoskeletons having become degraded by powerful chitinases. The turnover of the CTs of crustaceans, including shrimp, molluscs, and other sea creatures amounts to billions of tons per year. Yet, we know virtually nothing about these enzymes, so critical for marine biology and for the health of the seas and, ultimately, of the planet.

Chitinases, omnipresent throughout nature, are also produced by vertebrates in which they play important roles in defense against CT-containing pathogens and in food processing.[142] Chitinases occur in vertebrate monocytes and macrophages,[143] having evolved possibly as a protection against protozoan parasites. Curiously, a thousand-fold increased chitinase activity occurs in plasma from patients with Gaucher's disease, the most common of the lysosomal storage diseases, and decreases upon therapeutic intervention.[144,145] Chitinase is an excellent marker for Gaucher patients and is now widely used clinically. Many other members of the chitinase family exist in mammals,[144] including an acid-active chitinase involved in the pathogenesis of asthma.[142] These enzymes, although initially to be an ancient relic of evolution, are important in many aspects of current biology.

Cholera, caused by the Gram-negative bacillus *Vibrio cholerae*, must have exerted enormous survival pressure on our ancient ancestors. Colonization and survival in the human gut are dependent on CT surfaces, probably zooplanktons, and a chitinase has been identified as a human survival factor.[146] Remarkable also is the structural similarity of plant chitinase and the lysozymes from animals and phage.[147]

Lysozyme, the first enzyme to be characterized crystallographically, breaks down the bacterial cell wall NAG-NAM peptidoglycans. Lysozymes have an evolutionary relationship to the chitinases, both catalyzing the hydrolysis of β-1,4-glycosidic bonds. While lysozymes break down NAG-NAMs primarily, some are also able to degrade CTs. Similarly, some chitinases have lysozyme activity. These enzymes are related, and the superfamily may have a common evolutionary ancestor. The physiological ability of the enzyme, discovered by Alexander Fleming,[148] to dissolve bacterial cell walls, is secondary to its glycosidase activity. The enzyme functions as an *N*-acetylmuramidase. It also has the ability to cleave CT. Of note is that lysozyme is capable of catalyzing a transglycosylation reaction, as are the some of the vertebrate Hyals.

4.2.3. Hyals

There are two classes of Hyals with major difference in the mechanism of action. Most invertebrate Hyals are eliminases that function by β-elimination of the glycosidic linkage. There is no homology with the vertebrate enzymes that use substrate hydrolysis for the cleavage reaction. There are a number of exceptions, the most notable being the Hyal of bee venom, which has sequence homology with the vertebrate enzymes. This may be a case of lateral gene transfer. Additionally, while the majority of these enzymes are endo-β-*N*-acetylglucosaminidases, the leech and crustacean enzymes are endo-β-glucuronidases. To date, there have been no sequence analyses of the latter enzymes, and their mechanisms of actions are unknown.

4.2.3.1. Prokaryotic Hyals. A wealth of sequence, structural, and mechanistic information has become available during the past decade, facilitating formulation of precise mechanisms for these enzymes. These function by β-elimination with the introduction of an unsaturated bond.[39] The mechanism involves an acid/base type of proton acceptance and donation. Most bacterial and bacteriophage Hyals degrade through an initial nonprocessive endolytic bite followed by exolytic processive degradation. This generates unsaturated disaccharides of HA as products of exhaustive degradation. Most of these enzymes, like their eukaryotic counterparts, are also able to degrade Ch and ChS substrates. This reaction proceeds though the nonprocessive endolytic method, with generation of similar unsaturated Ch and ChS disaccharides.

4.2.3.2. Eukaryotic Hyals. The eukaryotic class of Hyals is endo-β-*N*-acetyl-hexosaminidases that employ substrate hydrolysis as their mechanism of action. They also have intrinsic transglycosylation activity, with the ability to cross-link chains of HA, presumably when they are partially digested and in the oligosaccharide size range.[149] These enzymes also have the ability to hydrolyze Ch and ChS and, therefore, have the potential ability to form cross-linked hybrid chains of HA and Ch or HA and ChS. The vertebrate Hyals degrade their substrates through endolytic nonprocessive reactions and generate predominantly tetrasaccharides.

Isolation and characterization of Hyals from vertebrate somatic tissues were accomplished only recently.[150] They occur in exceedingly small amounts, with very high but unstable specific activities. They require the constant presence of protease inhibitors and detergents during the purification procedures to maintain activity but are stable, once purified.

The Hyals constitute an enzyme protein family with a high degree of sequence homology. In the human genome, there are six genes tightly clustered at two chromosomal locations with Hyal-like sequences. The three genes, *Hyal-1*, *Hyal-2*, and *Hyal-3*, coding for Hyal-1, Hyal-2, and Hyal-3, are on

chromosome 3p21.3. They are organized in a complex and overlapping manner in an area densely packed with transcribed genes.[151] An example of this complex packing is a sequence coding for an *N*-acetyltransferase partially imbedded in an intron of Hyal-3. Polycistronic transcription and tissue coexpression of sets or cassettes of these genes occur, suggesting that some unknown but probably important physiology is occurring here.

The three genes *Hyal-4*, *Phyal1*, and *Spam1* (Sperm Adhesion Molecule1) are clustered in a similar fashion on chromosome 7q31.3, coding, respectively, for Hyal-4, a pseudogene transcribed but not translated in the human, PHYAL1, and PH-20. This chromosomal pattern is highly suggestive of an original sequence that underwent two ancient gene duplications, followed by en masse block duplication of the resulting three genes to constitute the final six sequences.

5. Modular Architecture of Carbohydrate-Active Enzymes

The enzymes synthesizing or degrading polysaccharides (polymeric glycans) have usually a modular architecture. The reason being the inaccessible nature of glycosidic linkages for degradation by enzymes. This is caused by close interactions of such glycans and the hidden nature of their glycosidic bonds. Such modules, in addition to the major catalytic portion (also a module), are often carbohydrate binding moieties (CBM). These CBMs promote enzyme concentration at the site of action and facilitate association of the catalytic portion with the substrate, for example, by spreading physically the associated polymeric substrate, making them available for enzymatic catalysis.[152]

The CAZy database divides carbohydrate enzymes, that is, those that degrade, modify, or create glycosidic bonds, into five groups: glycosidases and transglycosidases, glycosyltransferases, polysaccharide lyases, carbohydrate esterases, and carbohydrate-binding modules.[153,154] Here, we focus on selected aspects of these enzymes such as processive mechanism of synthesis and degradation of polymeric glycans. Many enzymes synthesizing glycans are glycosyltransferases, whereas degrading enzymes are either hydrolases or lyases; therefore, we focus on analysis of properties and mechanisms of the last two groups.

6. Example of Molecular Mechanism on Polymeric Glycan: Mechanism of Processivity

Polymeric carbohydrate needs specialized chemistry to act in the most efficient manner. To take advantage of elongated/polymeric structure, a repetitive mechanism of chemistry on such polymers has evolved. It is highly efficient and faster than just the sum of individual reactions. A specific processive molecular mechanism evolved to facilitate such improved efficient chemistry on elongated, linear sugar polymers.

Once protocells evolved and become more complex, they likely acquired specialized chemistry leading to the formation, modification, and even the degradation of self-replicating polymers. Synthesis would involve replication, transcription, and, in the DNA/protein world, translation of these polymers. Other polymers or oligomers that appeared with time, not necessarily self-replicatory ones, would also benefit from such specialized chemistry. Polymer specialized chemistry would be based on repetition of the same chemical tasks over and over again in an ever faster and more efficient manner. Such chemistry was therefore based on reactions that were sequential, in seriatim, a process referred to as processivity. The cells evolved to include processive replication, transcription, translation, simile synthesis, modification, and degradation of polymers.

6.1. Mechanisms of Enzyme Processivity: Example of Degradation of Polysaccharides

The X-ray crystal structures of bacterial Hyal illustrate a truncated two-domain molecule, the catalytic domain having an α_5/α_5 barrel structure, termed the α-domain. The other or β-domain is comprised mainly of a layered β-sandwich (Figure 6A). The catalytic cleft, transversing the α-domain has a positive charge, termed the positive patch. The active site of the enzyme is located within this barrel domain, on one wall of the cleft area[41,155−157] (Figure 6B, C).

The full-length enzyme contains additional residues at the N-terminal arranged in two domains[158] that are not of catalytic importance. The first of these domains is a substrate binding domain that enhance the Hyal affinity for its two substrates, HA and Ch (for a full discussion, see refs 158 and 159). The second and smaller domain acts as a spacer, to distance the catalytic from the substrate-binding domain at the extreme N terminus.[158] The β-domain functions only to support catalysis, as a modulator of substrate access to the catalytic cleft and its consequent release (see below). Catalysis is performed solely by residues of the α-domain. The catalytic residues are Asn349, His399, and Tyr409 (Figure 6C). The amino acid residues that position the substrate precisely for catalysis are Trp291, Trp292, and Phe343 and are referred to as the hydrophobic patch, whereas residues implicated in product release are Glu388, Asp398, and Thr400 [41,42,160−164] (Figure 6C).

6.2. Mode of Degradation of HA and Ch by Pneumococcal Hyal

The structural differences between Hyal and chondroitinase (Ch'ase) correlate with their modes of action, as revealed by kinetic analysis using a variety of substrates.[162,165,166] Cleavage of HA, the predominant substrate for pneumococcal Hyal, proceeds via an initial endolytic reaction, followed by rapid exolytic and processive activity. This degradative reaction produces unsaturated disaccharides as end products.[41,155,156,167] In contrast, the Ch'ase of *F. heparinum* cleaves unsulfated Ch and ChS, and to a lesser extent DS, in a nonprocessive manner producing a mixture of unsaturated oligosaccharide products containing even numbers of saccharides.[166] In the cases of 4- and 6-ChS, but not DS, cleavage continues, eventually forming disaccharides.[166] This occurs in a random, endolytic manner.[165]

Both HA and Ch, ChS have somewhat similar 3D structures. They assume a 2-fold or a similar helical structure,[40,168−170] whether in crystals, in solution, or in complexes with proteins. Modeling studies demonstrate that the extensive, unimpeded cleft in the Ch'ase enzyme forms a binding site that can easily accommodate long polymeric Ch oligosaccharides, thus facilitating endolytic degradation.[40]

Comparison of Hyal and Ch'ase structures reveals dramatic structural differences consisting primarily of a block at one end of the Hyal binding site cleft. This change prevents long oligosaccharides from being easily accommodated in the more favorable 2-fold helical conformation. The steric restriction occurs in the region containing the binding site

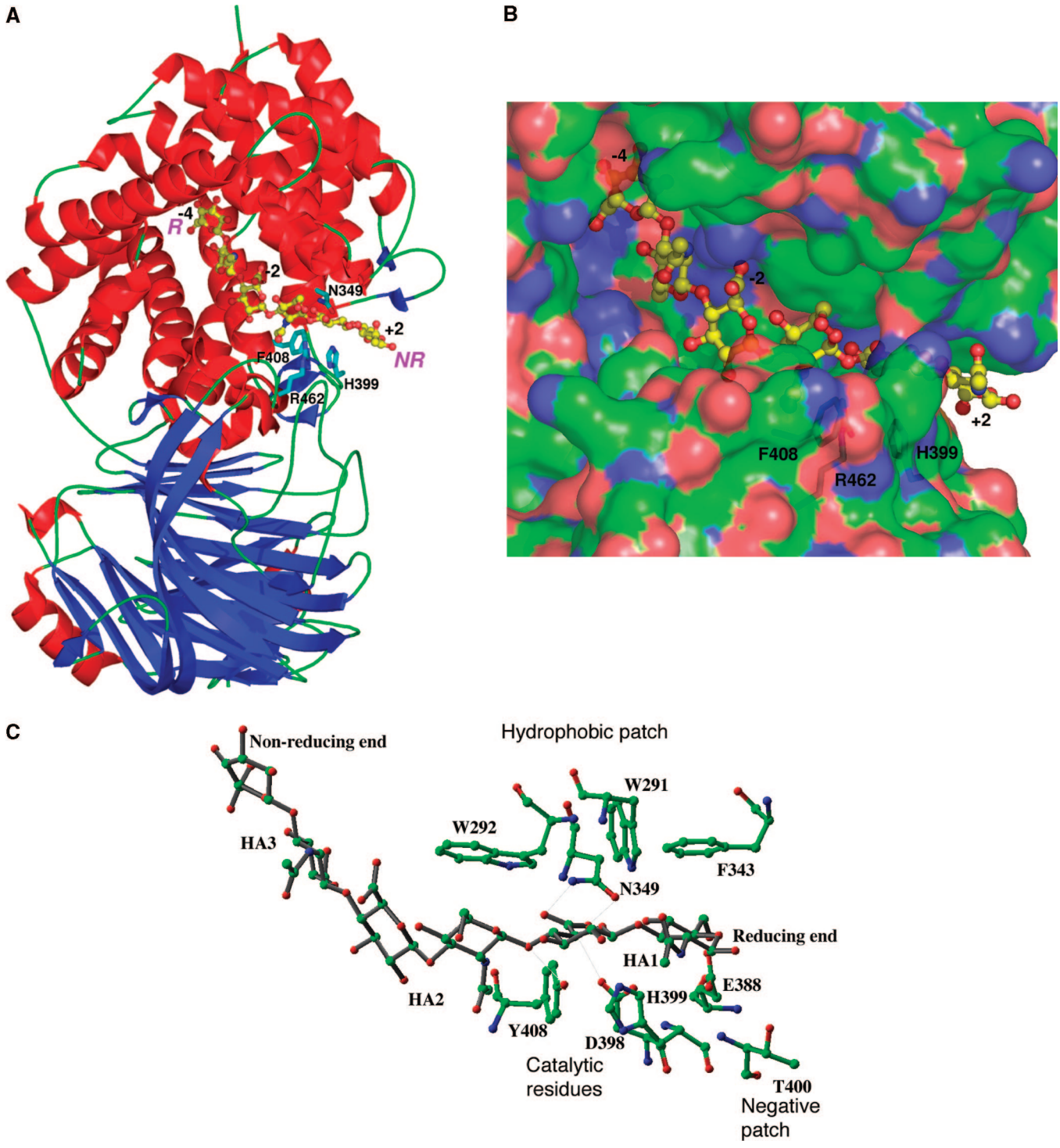


Figure 6. Structure of *S. pneumoniae* HA lyase. Panels A and B were adopted with permission from ref 40. Copyright 2003 American Society for Biochemistry and Molecular Biology. (A) Overall view of *S. pneumoniae* Hyal bound to HA hexasaccharide.[41] The ligand is shown as ball-and-stick, and some catalytic site residues mentioned in the text are drawn as sticks. Some binding subsites (nomenclature of Davies et al.[184]) are labeled, as are the reducing (R) and nonreducing (NR) ends of the substrate chain. (B) Close-up view of *S. pneumoniae* Hyal bound to HA hexasaccharide. The orientation is similar as in panel A. A semitransparent molecular surface is shown. (C) Catalytic center of the enzyme. The residues directly involved in mechanism of action, Asn349, His399, and Tyr408 (catalytic residues), positioning of the substrate Trp291, Trp292, and Phe343 (hydrophobic patch), and release of the product Glu388, Asp398, and Thr400 (negative patch), are shown together with the HA hexasaccharide substrate (HA1−3 are consecutive HA disaccharide from the reducing end). The most essential enzyme−substrate interactions for the catalytic process are shown as lines.[41,161]

for the nonreducing end of the oligosaccharide. Hyal can make initial endolytic cuts in larger substrates, albeit more slowly than exolytic cuts. Modeling demonstrates that this can be achieved but only when the substrate is deformed away from a straight, elongated 2-fold helical conformation. In such a case, a longer oligosaccharide can bind and be endolytically cleaved.[40] It is apparent that HA has sufficient intrinsic flexibility, so that binding to Hyal occurs in the middle of long polysaccharide chains. However, the distortion introduced, relative to the ideal 2-fold helical HA conformation, leads to suboptimal binding, particularly at the −3 to −4 subsite relative to the cleaved β-1,4 linkage

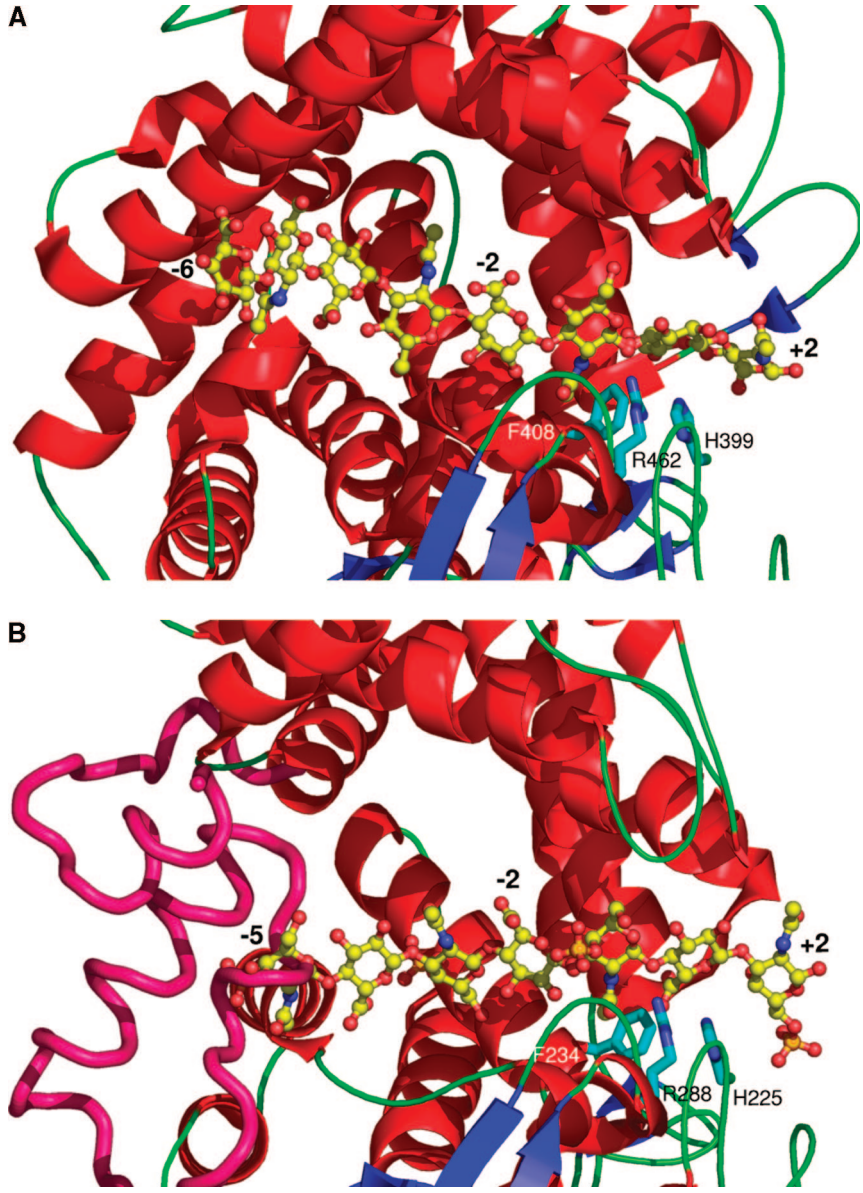


Figure 7. Binding of longer oligosaccharides in the entire length of the active site clefts of lyases. Binding modes in subsites −1, +1, and +2 (A) and subsites −2, −1, +1, and +2 (B) are those observed crystallographically.[41,170] The remaining oligosaccharides are modeled, as detailed in the text. Key catalytic residues are labeled in both cases. In panel B, the two large insertions in Hyal, relative to Ch'ase, which prevent extended binding to the former in the favorable 2-fold helical conformation, are shown as magenta loops. Reprinted with permission from ref 40. Copyright 2003 American Society for Biochemistry and Molecular Biology. (A) *S. pneumoniae* HA lyase and (B) *F. heparinum* Ch AC lyase.

(Figure 7A). This Hyal property exhibits the enzyme's unique feature that only binding to subsites −2 to +2, relative to the cleaved β-1,4-linkage, is essential for rapid exolytic cleavage (Figures 6C and 7A).

Optimal binding occurs, as observed crystallographically, with multiple hydrogen bonding, electrostatic, and hydrophobic interactions.[41,164] In contrast, to cleave endolytically in the middle of the chain, Hyal, even with a steric blockage at one end of the substrate binding cleft, must accommodate the oligosaccharide chain in the region beyond subsite −2. This interaction can be obtained satisfactorily because of the flexibility of HA and presumably Ch. However, this seems to involve the energetic cost of less than ideal interactions at subsite −2, as well as energetic costs associated with bending of the substrate as well as flexibility of the enzyme.

The mechanism by which Hyal achieves processive cleavage of HA, but not Ch, has been studied through structural and flexibility analyses.[41,155,156,163,167,171] Some structural features, such as extensive use of hydrophobic interactions with substrate interactions that generally lack directionality, are also observed with other processive type enzymes.[163,172] Lack of directionality in the interactions will aid in the movement of substrates along the active site cleft. Twisting motions between the two sides of the cleft and the two principal domains, revealed by simulated dynamics, and later by molecular dynamics, and confirmed by crystal structures of pneumococcal Hyal in new conformation also appear to play a role.[41,164,171] Given the overall structural similarity of Hyal and Ch'ase enzymes, one can expect similar dynamic behaviors of these two enzymes,[41,164,173] as well as conserved hydrophobic interactions with their respective substrates. Ch'ase, at least from *F. heparinum*, does not exhibit such processive substrate cleavage reactions[165] (Figure 7B). It appears that added substrate bulk, due to the irregular sulfation pattern of CH, is the cause of such nonprocessivity. Both enzymes, Hyal and Ch'ase, share

a constriction in the active site cleft near the catalytic residues (Figure 7). For processivity to occur, the chain undergoing cleavage must remain bound to the enzyme. Thus, after cleavage of the substrate, the unsaturated disaccharide product placed in subsites +1 and +2 must dissociate, and the remaining chain with its reducing end in subsite −1 must move by one disaccharide to occupy once more subsites −2 to +2. This translation places the next glycosidic bond for cleavage in the catalytic site (Figure 7).

As well as this significant movement, the approximate 2-fold symmetry of the bound substrate helix means that a rotation of around 180° about the helical axis must occur. It seems that these reorientations occur with greater difficulty for a sulfated substrate than for unsulfated HA. This must not be the only factor determining the inability of Ch'ase to carry out processive cleavage, since Ch'ase action on unsulfated Ch is also nonprocessive.[165] Similarly, bacterial Hyal action upon Ch, including the unsulfated form, is a nonprocessive endolytic reaction. It appears that in addition to Ch sulfation, the altered anomeric configuration at the C4 carbon of the *N*-acetyl-D-galactosamine, plays significant role for both enzymes in hampering processivity.

In addition, even with significant primary, secondary, and tertiary structures, Ch'ase appears to be designed not to carry out processive cleavage. In such an endolytic mode, both products must dissociate before the next productive substrate binding can occur. Binding in the entire active site cleft that leads to endolytic cleavage would presumably be tighter than binding of HA for processive degradation, irrespective of length, with the reducing end in subsite +2.

6.3. Selection of the β-1,4-Glycosidic Bond for Enzymatic Catalysis

The structure of the hexasaccharide HA complex with *S. pneumoniae* or *S. agalactiae* Hyal[41,164] was utilized to elucidate the mechanism of this enzyme's preference for β-1,4-cleavage, rather than the β-1,3-bond. Several factors contribute to this selection process. First, the *N*-acetyl-β-D-glucosamine portion of the HA substrate takes part in the binding interaction, primarily through hydrogen bonds with enzyme residues once the substrate is bound in the proper catalytic position for degrading the β-1,4-bond. Second, because of the different chemical nature of the β-D-glucuronic acid, such hydrogen bonds could not form if this sugar were placed at that same site. Third, placement of HA in the cleft in such a way as to situate the β-1,3-bond close to the catalytic residues would induce steric clashes between substrate and enzyme. In addition, no charge compensation is available for the carboxylate group of the glucuronate. Indeed, the closest residue to this carboxylate is a negatively charged Asp residue. Therefore, cleavage of the β-1,3-glycosidic bond is not possible by the bacterial Hyals. The geometry of both the HA substrate and the enzyme's cleft is responsible for the exclusivity of the β-1,4-glycosidic bond. A similar rationale is responsible for the selection of the same glycosidic linkage in the Ch substrate as well.[40] Sequence and structural analyses of the leech and crustacean Hyals that are endo-β-glucuronidases that cleave exclusively the β-1,3-bond are not yet available but will be of intrinsic interest. It is also not clear what the necessity or the evolutionary advantage would be for such a Hyal activity.

6.4. Degradation of Aggregated HA and Ch by Pneumococcal Hyal

Degradation of HA or Ch by bacterial Hyals, including the Hyal of *S. pneumoniae*, is a complex phenomenon. The catalytic mechanism is a multistep process[161] that can proceed in either a processive or a nonprocessive manner. The degradation of HA by *S. pneumoniae* Hyal in a processive manner[41,155,156,162−164,167] is initiated by a random endolytic "initial bite" resulting in cleavage of the polymer into two parts. This is followed by processive, exolytic cleavage of one HA disaccharide at a time until the entire remaining chain is degraded. Even tetrasaccharides can be cleaved in this manner. The presence of different sulfation patterns in naturally occurring Ch and the different epimerization of C4 carbon of *N*-acetyl-β-D-galactosamine prevents processive cleavage as delineated above for HA. The structural evidence correlates perfectly with the biochemical data. The Hyal of *S. pneumoniae* Hyal can degrade Ch only at the β-1,4-position when the disaccharide on the nonreducing side is either unsulfated or 6-sulfated. 4-Sulfation is not tolerated on the nonreducing side, although it is accepted on the reducing side of the bond to be cleaved. 2-Sulfated Ch is not cleaved.[40] The reason for this specificity is directly related to the steric clashes between enzyme and substrate. The degradation of Ch precedes, as a consequence, by an endolytic "random bite", nonprocessive mechanism.

HA and Ch possess specific, well-defined 3D structures that are dependent on the environment. Nuclear magnetic resonance (NMR) studies indicate that, in the presence of NaCl or divalent cations Ca^{2+} or Mg^{2+}, high molecular mass HA (over 300 kDa) assumes a 2-fold helix structure, as seen in crystalline structures, both HA/CH alone, or in complex with proteins. These are further arranged into antiparallel β-sheet structures that are stabilized by H-bonds between chains.[174,175] This structural property, for HA at least, is still under discussion, and others believe in a more highly dynamic and flexible HA structure.[176] However, the β-sheet structure of HA/Ch appears to be consistent with biochemical data on HA degradation by bacterial Hyals and provides possible explanations for such data. Assuming, however, that HA is aggregated under physiological conditions in the form of β-sheets, the initial degradation of the polymer appears to proceed by random endolytic cleavage, presumably only at sites where chains expose their β-1,4-linkages in a proper conformation. Because of the presumed 2-fold helical conformation of HA in these β-sheets, the next β-1,4-linkage is rotated by ∼180° and not likely to be accessible for catalysis. As the size of HA chains decreases, the ability to aggregate also decreases. At molecular size below 300 kDa, electron microscopy-rotary shadowing confirms decreased aggregation.[177] Light scattering provides evidence that HA chains below ∼50 disaccharides (∼20 kDa) do not aggregate in salt solutions.[178]

However, we also note that recent NMR structural work in aqueous solution has suggested that an exchange between hydrogen bonds and water molecules is another explanation that leads to HA chains being highly flexible in a solution, and no β-sheet structures are formed.[179,180] This results in a locally dynamic HA structure in solution that is on average a contracted 4-fold helix.[181] Similarly, additional rheological and scattering studies have also suggested that HA chains are semiflexible in solution.[182]

These still largely arguable properties indicate that at some size below 300 kDa HA is degraded by bacterial Hyals by

a processive mechanism because of decreased aggregation. At greater than 300 kDa, the mechanism is nonprocessive and endolytic. In this model, as the average size of HA chain decreases, the processive mechanism takes over, leading to exponential HA degradation and the exponential formation of HA disaccharides (Jedrzejas, M. J. Unpublished results). Such behavior is consistent with the size-dependent tertiary structures of HA in solution containing β-sheets at high molecular mass and the lack of aggregation below 20 kDa.

The proposed structure of Ch is similar to that of the HA 2-fold helices.[183] Indeed, Chs assume a 2-fold helical structure in D_2O solution with various patterns of aggregation dependent on patterns of sulfation.[183] The 6-sulfated Ch, C6S, but not the 4-sulfated molecules can also form duplex structures. The unsulfated Ch forms similar higher aggregated structures similar HA.[183] Therefore, degradation of these polymeric unsulfated and sulfated Ch molecules, just as for high molecular weight HA, proceeds via an endolytic, "random bite" type mechanism. In this model, because of selectivity of Hyal for preferred Ch sulfation patterns, endolytic cleavage predominates until the entire polymer is degraded. This differs from the situation with the HA substrate in which exolytic cleavage takes over as substrate size decreases. The final degradation product of sulfated Ch is also a disaccharide unit, with the limitation of selectivity due to sulfation patterns, as described above.

Interestingly, the extreme N-terminal portion of *S. pneumoniae* Hyal, not visualized crystallographically because of enzyme instability, bears homology to carbohydrate-binding domains of other carbohydrate active enzymes.[158] Presumably this domain facilitates (i) colocalization of enzyme and substrate, thereby enhancing Hyal's catalytic efficiency, (ii) disruption of the higher-order aggregated substrate conformations by spreading individual helical strands apart to make the degradative process possible, (iii) feeding of the HA/CH chains to the catalytic domain after their binding by CBM (carbohydrate-binding modules/motif) containing domain, and helping to retain HA binding to the enzyme after catalysis, to allow for the translational repositioning for the next processive cleavage step.[40] Hyal would then be able to degrade the large HA chains primarily through processive degradation, bypassing the initial nonprocessive action pattern.[158]

6.5. General Conclusions from the Processivity of Enzymes

6.5.1. Hyal Flexibility and the Processivity Mechanism for HA Degradation

The main feature of the Hyal catalytic domain is the presence of a long, large cleft where substrates bind and undergo catalysis. This domain has great flexibility that facilitates degradation. The motions involved in this flexibility includes (i) a rotation/twisting motion of the sides of the cleft that yields a ~10 Å movement along the cleft axis, (ii) an opening/closing of the width of the substrate-binding cleft, accompanied by the movement of catalytic amino acid residues and those responsible for the release of degradation products. In terms of the entire enzyme, a third movement (iii) involves a shift in position of the β-domain, resulting in an effective opening and closing of the access/entrance to the cleft.[41,164] This motion depicts the functional role of the β-domain as the modulator of access to the enzyme's cleft. These motions are directly related to function: catalysis, endolytic "random bite" degradation, and exolytic cleavage in a processive manner of one disaccharide at a time in the reducing to nonreducing direction.[40]

These individual modes of collective enzyme−substrate fluctuations interplay in a complex manner and do not correspond to specific functional tasks. The α-domain, containing the positive patch, is responsible for the majority of protein−ligand contacts. The twist motion (i) provides exposure of the positively charged cleft to the enzyme's environment and to the substrate. Hence, this twisting motion facilitates movement of the substrate along the cleft in the reducing direction by one disaccharide, to reposition it in the catalytic site, in preparation for further cleavage. The shift of hydrophobic patch residues Trp371, Trp372, and Phe423 is by an average distance of 11.7 Å, which matches closely the distance between HA/CH disaccharide units, ~11 Å. Another movement (ii) illustrates two distinct factors contributing to a change in characteristics favorable for the binding of the negatively charged substrate. First, the positive patch on the α-domain becomes more exposed, and the residues from the active site, the hydrophobic and the negative patches, remain inaccessible and are not influenced by this motion. Second, the overall reduction of the negative field is observed around the C-terminal β-domain. Finally, the opening/closing of the access to the substrate binding cleft (iii) brings a catalytic histidine residue closer to the HA substrate's β-1,4-glycosidic bond. Each of these steps facilitates the complex mechanism of Hyal action.

6.5.2. Structural Aspects of Streptococcal Hyal Processivity

The consequent likely relatively weaker binding of the substrate due to the flexibility of Hyals, as compared with an enzyme with a rigid structure, is presumably relevant physiologically. As the enzyme degrades HA in a processive mode of action, a weaker binding of substrate might allow for easier shifting of the enzyme with respect to the substrate (sliding/threading mechanism of processivity) along the cleft length.[41,164] For the processivity to take place, the substrate binding needs to be relatively weak to allow for the sliding of the enzyme along the substrate toward its nonreducing end to bring the catalytic part of the cleft over the β-1,4-glycosidic linkage to be cleaved.[41,172]

7. Conclusions

It was intended here to examine both broad principles as well as specific details on the possible evolution of life, concentrating on the importance of carbohydrate polymers in that process. The contention is that the synthesis of these polysaccharides was an essential first step in the multiple steps that generated the earliest life forms.

The subtext here has been that a natural order of events is inherent in that process and that a rerun of the tape would provide similar results. It is suspected that life has evolved on several occasions in billion and trillion year time scales and will continue to do so. The same basic elements, the basic metabolism, and the same basic energy sources and requirements are intrinsic to the system. The strongest argument for this hypothesis is that the key reactive elements that went into the creation of life, hydrogen, oxygen, carbon, and nitrogen, are the most abundant elements that make up the Universe. However, these are profoundly different from the materials that make up the composition of the Earth and

other planets. It is suspected, therefore, that if the tape is run again, the results will be remarkably similar.

The most amazing varieties of "endless forms most beautiful and most wonderful", as Darwin noted, have a remarkable similarity in underlying principles. There may indeed be a single tape.

8. Abbreviations

3D	three-dimensional
CBM	carbohydrate-binding moieties
Ch	chondroitin
Ch'ase	chondroitinase
ChS	chondroitin sulfate
C4S	chondroitin-4-sulfate
C6S	chondroitin-6-sulfate
CT	chitin
CTS	chitin synthase
Dol	dolichol
DS	dermatan sulfate
ECM	extracellular matrix
ER	endoplasmic reticulum
GAGs	glycosaminoglycans
HA	hyaluronan also termed hyaluronic acid
HAS	HA synthase
HS	heparan sulfate
Hyal	hyaluronidase
KS	keratan sulfate
NCAM	neural cell adhesion molecule
NMR	nuclear magnetic resonance
NodC	nodulation protein C
PGCP	proteoglycan core protein
polyA	poly adenine

9. Acknowledgments

We acknowledge the assistance and many helpful discussions with present and past members of our laboratories, as well as our colleagues and collaborators.

10. References

(1) Hinshaw, G.; Weiland, J. L.; Hill, R. S.; Odegard, N.; Larson, D.; Bennett, C. L.; Dunkley, J.; Gold, B.; Greason, M. R.; Jarosik, N.; Komatsu, E.; Nolta, M. R.; Page, L.; Spergel, D. N.; Wollack, E.; Halpern, M.; Kogut, A.; Limon, M.; Meyer, S. S.; Tucker, G. S.; Wright, E. L. *Astrophysics* **2008**, in press.
(2) Fewell, M. P. *Am. J. Phys.* **1995**, *63*, 653.
(3) Anders, E.; Ebihara, M. *Geochim. Cosmochim. Acta* **1982**, *46*, 1363.
(4) Baker, J.; Bizzarro, M.; Wittig, N.; Connelly, J.; Haack, H. *Nature* **2005**, *436*, 1127.
(5) Lide, D. R., Ed. *CRC Handbook of Chemistry and Physics. Section 14: Geophysics, Astronomy, and Acoustics; Abundance of Elements in the Earth's Crust and in the Sea*, 85th ed.; CRC Press: Boca Raton, FL, 2005.
(6) Bossuyt, X.; Blanckaert, N *Biochem. J.* **1993**, *296*, 627.
(7) Lunine, J. I. *Evolution of Habitable World*; Cambridge University Press: Canmbridge, MA, 1999.
(8) Zimmer, C. *Evolution: The Triumph of an Idea*; Harper Collins Publishers: New York, NY, 2001.
(9) Gonzalez, G.; Richards, J. *The Privileged Planet: How Our Place in the Cosmos Is Designed for Discovery*; Regnery Publishing, Inc.: Washington, DC, 2004.
(10) Kvenvolden, K. A.; Lawless, J. G.; Ponnamperuma, C. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 486.
(11) Darwin, C. *On the Origin of Species*; John Murray Press: London, 1859.
(12) Miller, S. L. *Science* **1953**, *117*, 528.
(13) Orgel, L. E. *Cold Spring Harb. Symp. Quant. Biol.* **1987**, *52*, xxx.
(14) Joyce, G. F. *Nature* **2002**, *418*, 214.
(15) Hirabayashi, J. *Q. Rev. Biol.* **1996**, *71*, 365.
(16) Schwartz, A. W. *Chem. Biol.* **1996**, *3*, 515.
(17) Bernal, J. D. *The Physical Basis of Life*; Routledge & Kegan Paul: London, 1951.
(18) Miller, S. L. *Cold Spring Harb. Symp. Quant. Biol.* **1987**, *52*, 17.
(19) Stanek, J.; Cerby, M.; Kocourek, J.; Pacak, J. In *The Monosaccharides*; Stanek, J., Ed.; Academic Press: New York, 1963.
(20) Breslow, R.; Huang, D. L. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 4080.
(21) de Figueiredo, P.; Terra, B.; Anand, J. K.; Hikita, T.; Sadilek, M.; Monks, D. E.; Lenskiy, A.; Hakomori, S.; Nester, E. W. *Extremophiles* **2007**, *11*, 133.
(22) Spies, M. A.; Schowen, R. L. *J. Am. Chem. Soc.* **2002**, *124*, 14049.
(23) Ishiwata, S.; Kamiya, M. *Chemosphere* **1999**, *39*, 1595.
(24) Laszlo, P. *Science* **1987**, *235*, 1472.
(25) Lahav, N.; White, D.; Chang, S. *Science* **1978**, *201*, 67.
(26) Lahav, N.; White, D. H. *J. Mol. Evol.* **1980**, *16*, 11.
(27) Acevedo, O. L.; Orgel, L. E. *Nature* **1986**, *321*, 790.
(28) Acevedo, O. L.; Orgel, L. E. *J. Mol. Biol.* **1987**, *197*, 187.
(29) Ferris, J. P.; Hill, A. R., Jr.; Liu, R.; Orgel, L. E. *Nature* **1996**, *381*, 59.
(30) Kornberg, A. *J. Bacteriol.* **1995**, *177*, 491.
(31) Brown, M. R.; Kornberg, A. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 16085.
(32) Kornberg, A.; Rao, N. N.; Ault-Riche, D. *Annu. Rev. Biochem.* **1999**, *68*, 89.
(33) Spagnoli, C.; Loos, K.; Ulman, A.; Cowman, M. K. *J. Am. Chem. Soc.* **2003**, *125*, 7124.
(34) Bada, J. L.; Bigham, C.; Miller, S. L. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 1248.
(35) Miyakawa, S.; Cleaves, H. J.; Miller, S. L. *Origins Life Evol. Biospheres* **2002**, *32*, 209.
(36) Miyakawa, S.; Yamanashi, H.; Kobayashi, K.; Cleaves, H. J.; Miller, S. L. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 14628.
(37) Trinks, H.; Schroder, W.; Biebricher, C. K. *Origins Life Evol. Biospheres* **2005**, *35*, 429.
(38) Voet, D.; Voet, J. G. *Biochemistry*; John Wiley & Sons, Inc.: New York, 2004.
(39) Stern, R.; Jedrzejas, M. J. *Chem. Rev.* **2006**, *106*, 818.
(40) Rigden, D. J.; Jedrzejas, M. J. *J. Biol. Chem.* **2003**, *278*, 50596.
(41) Jedrzejas, M. J.; Mello, L. V.; de Groot, B. L.; Li, S. *J. Biol. Chem.* **2002**, *277*, 28287.
(42) Li, S.; Jedrzejas, M. J. *J. Biol. Chem.* **2001**, *276*, 41407.
(43) Schoening, K. U. *Science* **2000**, *290*, 1347.
(44) Joyce, G. F. *Nature* **2002**, *420*, 278.
(45) Mueller, D. *Helv. Chim. Acta* **1990**, *73*, 1410.
(46) Ferris, J. P.; Ertem, G. *Science* **1992**, *257*, 1387.
(47) Joyce, G. F.; Visser, G. M.; van Boeckel, C. A.; van Boom, J. H.; Orgel, L. E.; van Westrenen, J. *Nature* **1984**, *310*, 602.
(48) Saiardi, A.; Bhandari, R.; Resnick, A. C.; Snowman, A. M.; Snyder, S. H. *Science* **2004**, *306*, 2101.
(49) Wang, L.; Fraley, C. D.; Faridi, J.; Kornberg, A.; Roth, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 11249.
(50) Chaussepied, P.; Mornet, D.; Kassab, R. *Biochemistry* **1986**, *25*, 6426.
(51) Tolstoguzov, V. *Origins Life Evol. Biospheres* **2004**, *34*, 571.
(52) Monnard, P. A.; Deamer, D. W. *Anat. Rec.* **2002**, *268*, 196.
(53) Deamer, D. W. *Origins Life Evol. Biospheres* **1986**, *17*, 3.
(54) Deamer, D.; Kuzina, S. I.; Mikhailov, A. I.; Maslikova, E. I.; Seleznev, S. A. *J. Evol. Biochem. Physiol.* **1991**, *27*, 212.
(55) Trevors, J. T. *Cell Biol. Int.* **2003**, *27*, 451.
(56) Trevors, J. T. *Riv. Biol.* **2004**, *97*, 79.
(57) Nilson, F. P. *Origins Life Evol. Biospheres* **2002**, *32*, 247.
(58) Szostak, J. W.; Bartel, D. P.; Luisi, P. L. *Nature* **2001**, *409*, 387.
(59) Szostak, J. W. *Harvey Lect.* **1997**, *93*, 95.
(60) Har-el, R.; Tanzer, M. L. *FASEB J.* **1993**, *7*, 1115.
(61) Gross, J. In *Biology of Invertebrate and Lower Vertebrate Collagens*; Bairati, A., Garrone, A., Eds.; Plenum Press: New York, 1985.
(62) Mathews, M. B. In *Biology of Invertebrate and Lower Vertebrate Collagens*; Bairati, A., Garrone, A., Eds.; Plenum Press: New York, 1985.
(63) Coluccio, A.; Bogengruber, E.; Conrad, M. N.; Dresser, M. E.; Briza, P.; Neiman, A. M. *Eukaryotic Cell* **2004**, *3*, 1464.
(64) Doudna, J. A.; Cech, T. R. *Nature* **2002**, *418*, 222.
(65) Gora-Sochacka, A. *Acta Biochim. Pol.* **2004**, *51*, 587.
(66) Zimmer, C. *Science* **2006**, *312*, 870.
(67) Benstein, S. L.; Kempner, E. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 6410.
(68) Falkowski, P. G. *Science* **2006**, *311*, 1724.
(69) Bayer, E. A.; Chanzy, H.; Lamed, R.; Shoham, Y. *Curr. Opin. Struct. Biol.* **1998**, *8*, 548.
(70) Tuskan, G. A.; Difazio, S.; Jansson, S.; Bohlmann, J.; Grigoriev, I.; Hellsten, U.; Putnam, N.; Ralph, S.; Rombauts, S.; Salamov, A.; Schein, J.; Sterck, L.; Aerts, A.; Bhalerao, R. R.; Bhalerao, R. P.; Blaudez, D.; Boerjan, W.; Brun, A.; Brunner, A.; Busov, V.; Campbell, M.; Carlson, J.; Chalot, M.; Chapman, J.; Chen, G. L.; Cooper, D.; Coutinho, P. M.; Couturier, J.; Covert, S.; Cronk, Q.; Cunningham, R.; Davis, J.; Degroeve, S.; Dejardin, A.; Depamphilis, C.; Detter, J.; Dirks, B.; Dubchak, I.; Duplessis, S.; Ehlting, J.; Ellis,

B.; Gendler, K.; Goodstein, D.; Gribskov, M.; Grimwood, J.; Groover, A.; Gunter, L.; Hamberger, B.; Heinze, B.; Helariutta, Y.; Henrissat, B.; Holligan, D.; Holt, R.; Huang, W.; Islam-Faridi, N.; Jones, S.; Jones-Rhoades, M.; Jorgensen, R.; Joshi, C.; Kangasjarvi, J.; Karlsson, J.; Kelleher, C.; Kirkpatrick, R.; Kirst, M.; Kohler, A.; Kalluri, U.; Larimer, F.; Leebens-Mack, J.; Leple, J. C.; Locascio, P.; Lou, Y.; Lucas, S.; Martin, F.; Montanini, B.; Napoli, C.; Nelson, D. R.; Nelson, C.; Nieminen, K.; Nilsson, O.; Pereda, V.; Peter, G.; Philippe, R.; Pilate, G.; Poliakov, A.; Razumovskaya, J.; Richardson, P.; Rinaldi, C.; Ritland, K.; Rouze, P.; Ryaboy, D.; Schmutz, J.; Schrader, J.; Segerman, B.; Shin, H.; Siddiqui, A.; Sterky, F.; Terry, A.; Tsai, C. J.; Uberbacher, E.; Unneberg, P. *Science* **2006**, *313*, 1596.

(71) Meroueh, S. O.; Bencze, K. Z.; Hesek, D.; Lee, M.; Fisher, J. F.; Stemmler, T. L.; Mobashery, S. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 4404.
(72) Karlsson, K. A. *Adv. Exp. Med. Biol.* **2001**, *491*, 431.
(73) Paulson, J. C.; Blixt, O.; Collins, B. E. *Nat. Chem. Biol.* **2006**, *2*, 238.
(74) Rybak, M. J. *Am. J. Infect. Control* **2006**, *34*, S38.
(75) Buzas, E. I.; Gyorgy, B.; Pasztoi, M.; Jelinek, I.; Falus, A.; Gabius, H. J. *Autoimmunity* **2006**, *39*, 691.
(76) DeAngelis, P. L. *Anat. Rec.* **2002**, *268*, 317.
(77) Carey, D. J. *Biochem. J.* **1997**, *327*, 1.
(78) Iozzo, R. V. *Matrix Biol.* **1994**, *14*, 203.
(79) Filmus, J. *Glycobiology* **2001**, *11*, 19R.
(80) Cheifetz, S.; Massague, J. *J. Biol. Chem.* **1989**, *264*, 12025.
(81) Toyoda, H.; Kinoshita-Toyoda, A.; Selleck, S. B. *J. Biol. Chem.* **2000**, *275*, 2269.
(82) Haylock, D. N.; Nilsson, S. K. *Regener. Med.* **2006**, *1*, 437.
(83) Khaldoyanidi, S. *Cell Stem Cell* **2008**, *2*, 198.
(84) Frost, G. I.; Csoka, A. B.; Wong, T.; Stern, R.; Csoka, T. B. *Biochem. Biophys. Res. Commun.* **1997**, *236*, 10.
(85) Prehm, P. *Biochem. J.* **1984**, *220*, 597.
(86) Fukuda, M.; Carlsson, S. R.; Klock, J. C.; Dell, A. *J. Biol. Chem.* **1986**, *261*, 12796.
(87) Fukuda, M.; Bothner, B.; Ramsamooj, P.; Dell, A.; Tiller, P. R.; Varki, A.; Klock, J. C. *J. Biol. Chem.* **1985**, *260*, 12957.
(88) Togayachi, A.; Kozono, Y.; Ishida, H.; Abe, S.; Suzuki, N.; Tsunoda, Y.; Hagiwara, K.; Kuno, A.; Ohkura, T.; Sato, N.; Sato, T.; Hirabayashi, J.; Ikehara, Y.; Tachibana, K.; Narimatsu, H. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 15829.
(89) Miyake, Y.; Matsumoto, H.; Yokoo, M.; Miyazawa, K.; Kimura, N.; Tunjung, W. A.; Shimizu, T.; Sasada, H.; Aso, H.; Yamaguchi, T.; Sato, E. *Biol. Reprod.* **2006**, *74*, 501.
(90) Dityatev, A.; El-Husseini, A. *Molecular Mechanisms of Synaptogenesis*; Springer: New York, NY, 2006.
(91) Colley, K. J. *Neurochem. Res.* **2008**, in press.
(92) Ong, E.; Nakayama, J.; Angata, K.; Reyes, L.; Katsuyama, T.; Arai, Y.; Fukuda, M. *Glycobiology* **1998**, *8*, 415.
(93) Xing, J.; Forsee, W. T.; Lamani, E.; Maltsev, S. D.; Danilov, L. L.; Shibaev, V. N.; Schutzbach, J. S.; Cheung, H. C.; Jedrzejas, M. J. *Biochemistry* **2000**, *39*, 7886.
(94) Parodi, A. J. *Annu. Rev. Biochem.* **2000**, *69*, 69.
(95) Sosinska, G. *Postepy Biochem.* **2003**, *49*, 107.
(96) Helenius, J.; Aebi, M. *Semin. Cell Dev. Biol.* **2002**, *13*, 171.
(97) Bergamini, E.; Bizzarri, R.; Cavallini, G.; Cerbai, B.; Chiellini, E.; Donati, A.; Gori, Z.; Manfrini, A.; Parentini, I.; Signori, F.; Tamburini, I. *J. Alzheimers Dis.* **2004**, *6*, 129.
(98) Koscielak, J. *Acta Biochim. Pol.* **1999**, *46*, 727.
(99) Palli, S. R.; Retnakaran, A. *Exs* **1999**, *87*, 85.
(100) Stern, R.; Asari, A. A.; Sugahara, K. N. *Eur. J. Cell Biol.* **2006**, *85*, 699.
(101) Shworak, N. W.; Shirakawa, M.; Colliec-Jouault, S.; Liu, J.; Mulligan, R. C.; Birinyi, L. K.; Rosenberg, R. D. *J. Biol. Chem.* **1994**, *269*, 24941.
(102) Gama, C. I.; Tully, S. E.; Sotogaku, N.; Clark, P. M.; Rawat, M.; Vaidehi, N.; Goddard, W. A., 3rd; Nishi, A.; Hsieh-Wilson, L. C. *Nat. Chem. Biol.* **2006**, *2*, 467.
(103) Lee, J. Y.; Spicer, A. P. *Curr. Opin. Cell Biol.* **2000**, *12*, 581.
(104) Keegstra, K.; Walton, J. *Science* **2006**, *311*, 1872.
(105) Weigel, P. H.; Hascall, V. C.; Tammi, M. *J. Biol. Chem.* **1997**, *272*, 13997.
(106) Spicer, A. P.; McDonald, J. A. *J. Biol. Chem.* **1998**, *273*, 1923.
(107) Heldermon, C.; DeAngelis, P. L.; Weigel, P. H. *J. Biol. Chem.* **2001**, *276*, 2037.
(108) Weigel, P. H. *IUBMB Life* **2002**, *54*, 201.
(109) DeAngelis, P. L. *Cell. Mol. Life Sci.* **1999**, *56*, 670.
(110) Itano, N.; Sawai, T.; Yoshida, M.; Lenas, P.; Yamada, Y.; Imagawa, M.; Shinomura, T.; Hamaguchi, M.; Yoshida, Y.; Ohnuki, Y.; Miyauchi, S.; Spicer, A. P.; McDonald, J. A.; Kimata, K. *J. Biol. Chem.* **1999**, *274*, 25085.
(111) Spicer, A. P.; Nguyen, T. K. *Biochem. Soc. Trans.* **1999**, *27*, 109.
(112) Monslow, J.; Williams, J. D.; Norton, N.; Guy, C. A.; Price, I. K.; Coleman, S. L.; Williams, N. M.; Buckland, P. R.; Spicer, A. P.; Topley, N.; Davies, M.; Bowen, T. *Int. J. Biochem. Cell Biol.* **2003**, *35*, 1272.
(113) Kamst, E.; Pilling, J.; Raamsdonk, L. M.; Lugtenberg, B. J.; Spaink, H. P. *J. Bacteriol.* **1997**, *179*, 2103.
(114) Kamst, E.; Bakkers, J.; Quaedvlieg, N. E.; Pilling, J.; Kijne, J. W.; Lugtenberg, B. J.; Spaink, H. P. *Biochemistry* **1999**, *38*, 4045.
(115) van der Holst, P. P.; Schlaman, H. R.; Spaink, H. P. *Curr. Opin. Struct. Biol.* **2001**, *11*, 608.
(116) Bakkers, J.; Kijne, J. W.; Spaink, H. P. *Exs* **1999**, *87*, 71.
(117) Semino, C. E.; Allende, M. L. *Int. J. Dev. Biol.* **2000**, *44*, 183.
(118) Yoshida, M.; Itano, N.; Yamada, Y.; Kimata, K. *J. Biol. Chem.* **2000**, *275*, 497.
(119) Semino, C. E.; Specht, C. A.; Raimondi, A.; Robbins, P. W. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 4548.
(120) DeAngelis, P. L.; Achyuthan, A. M. *J. Biol. Chem.* **1996**, *271*, 23657.
(121) Coutinho, P. M.; Deleury, E.; Davies, G. J.; Henrissat, B. *J. Mol. Biol.* **2003**, *328*, 307.
(122) Jedrzejas, M. J.; Stern, R. *Proteins: Struct., Funct., Bioinf.* **2005**, *61*, 227.
(123) Charnock, S. J.; Davies, G. J. *Biochemistry* **1999**, *38*, 6380.
(124) Hammerschmidt, S.; Talay, S. R.; Brandtzaeg, P.; Chhatwal, G. S. *Mol. Microbiol.* **1997**, *25*, 1113.
(125) Tarbouriech, N.; Charnock, S. J.; Davies, G. J. *J. Mol. Biol.* **2001**, *314*, 655.
(126) Lamani, E.; Mewbourne, R. B.; Fletcher, D. S.; Maltsev, S. D.; Danilov, L. L.; Veselovsky, V. V.; Lozanova, A. V.; Grigorieva, N. Y.; Pinsker, O. A.; Xing, J.; Forsee, W. T.; Cheung, H. C.; Schutzbach, J. S.; Shibaev, V. N.; Jedrzejas, M. J. *Glycobiology* **2006**, *16*, 666.
(127) Pummill, P. E.; Achyuthan, A. M.; DeAngelis, P. L. *J. Biol. Chem.* **1998**, *273*, 4976.
(128) Varki, A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 4523.
(129) Prehm, P.; Schumacher, U. *Biochem. Pharmacol.* **2004**, *68*, 1401.
(130) Ouskova, G.; Spellerberg, B.; Prehm, P. *Glycobiology* **2004**, *14*, 931.
(131) Itano, N.; Sawai, T.; Atsumi, F.; Miyaishi, O.; Taniguchi, S.; Kannagi, R.; Hamaguchi, M.; Kimata, K. *J. Biol. Chem.* **2004**, *279*, 18679.
(132) Itano, N.; Kimata, K. *Biochem. Biophys. Res. Commun.* **1996**, *222*, 816.
(133) Baggenstoss, B. A.; Weigel, P. H. *Anal. Biochem.* **2006**, *352*, 243.
(134) Kultti, A.; Rilla, K.; Tiihonen, R.; Spicer, A. P.; Tammi, R. H.; Tammi, M. I. *J. Biol. Chem.* **2006**, *281*, 15821.
(135) Itano, N.; Kimata, K. *IUBMB Life* **2002**, *54*, 195.
(136) Mian, N. *Biochem. J.* **1986**, *237*, 333.
(137) Mian, N. *Biochem. J.* **1986**, *237*, 343.
(138) Stern, R. *Glycobiology* **2003**, *13*, 105R.
(139) Stern, R. *Eur. J. Cell Biol.* **2004**, *83*, 317.
(140) Chapin, F. S.; Matson, P. A.; Mooney, H. A. *Principles of Terrestrial Ecosystem Ecology*; SpringerVerlag: New York, 2002.
(141) Kasprzewska, A. *Cell. Mol. Biol. Lett.* **2003**, *8*, 809.
(142) Couzin, J. *Science* **2004**, *304*, 1577.
(143) Renkema, G. H.; Boot, R. G.; Strijland, A.; Donker-Koopman, W. E.; van den Berg, M.; Muijsers, A. O.; Aerts, J. M. *Eur. J. Biochem.* **1997**, *244*, 279.
(144) Bussink, A. P.; van Eijk, M.; Renkema, G. H.; Aerts, J. M.; Boot, R. G. *Int. Rev. Cytol.* **2006**, *252*, 71.
(145) Barone, R.; Sotgiu, S.; Musumeci, S. *Clin. Lab.* **2007**, *53*, 321.
(146) Kirn, T. J.; Jude, B. A.; Taylor, R. K. *Nature* **2005**, *438*, 863.
(147) Holm, L.; Sander, C. *FEBS Lett.* **1994**, *340*, 129.
(148) Fleming, A. *Proc. R. Soc. (London)* **1922**, *93*, 306.
(149) Hoffman, P.; Meyer, K.; Linker, A. *J. Biol. Chem.* **1956**, *219*, 653.
(150) Kreil, G. *Protein Sci.* **1995**, *4*, 1666.
(151) Shuttleworth, T. L.; Wilson, M. D.; Wicklow, B. A.; Wilkins, J. A.; Triggs-Raine, B. L. *J. Biol. Chem.* **2002**, *277*, 23008.
(152) Boraston, A. B.; Bolam, D. N.; Gilbert, H. J.; Davies, G. J. *Biochem. J.* **2004**, *382*, 769.
(153) Coutinho, P. M.; Henrissat, B.; AFMB−CNRS−Universités Aix-Marseille I & II, 2008.
(154) Coutinho, P. M.; Henrissat, B. In *Recent Advances in Carbohydrate Bioengineering*; Gilbert, H. J., Davies, G., Henrissat, B., Svensson, B., Eds.; The Royal Society of Chemistry: Cambridge, United Kingdom, 1999.
(155) Jedrzejas, M. J. In *Science of Hayluronan Today*; Hascall, V. C., Yanagishita, M., Eds.; Glycoforum, www.glycoforum.gr.jp/science/hyaluronan, 2002.
(156) Jedrzejas, M. J. In *Recent Research Developments in Biophysics and Biochemistry*; Pandali, S. G., Ed.; Research Signpost: Trivandrum, India, 2002; Vol. 2.
(157) Li, S.; Taylor, K. B.; Kelly, S. J.; Jedrzejas, M. J. *J. Biol. Chem.* **2001**, *276*, 15125.
(158) Rigden, D. J.; Jedrzejas, M. J. *Proteins: Struct., Funct., Bioinf.* **2003**, *52*, 203.

(159) Rigden, D. J.; Galperin, M. Y.; Jedrzejas, M. J. *Crit. Rev. Biochem. Mol. Biol.* **2003**, *38*, 143.
(160) Ponnuraj, K.; Jedrzejas, M. J. *J. Mol. Biol.* **2000**, *299*, 885.
(161) Li, S.; Kelly, S. J.; Lamani, E.; Ferraroni, M.; Jedrzejas, M. J. *EMBO J.* **2000**, *19*, 1228.
(162) Kelly, S. J.; Taylor, K. B.; Li, S.; Jedrzejas, M. J. *Glycobiology* **2001**, *11*, 297.
(163) Nukui, M.; Taylor, K. B.; McPherson, D. T.; Shigenaga, M. K.; Jedrzejas, M. J. *J. Biol. Chem.* **2003**, *278*, 3079.
(164) Mello, L. V.; de Groot, B. L.; Li, S.; Jedrzejas, M. J. *J. Biol. Chem.* **2002**, *277*, 36678.
(165) Jandik, K. A.; Gu, K.; Linhardt, R. J. *Glycobiology* **1994**, *4*, 289.
(166) Gu, K.; Linhardt, R. J.; Laliberte, M.; Zimmermann, J. *Biochem. J.* **1995**, *312*, 569.
(167) Jedrzejas, M. J. *Front. Biosci.* **2004**, *9*, 891.
(168) Winter, W. T.; Arnott, S.; Isaac, D. H.; Atkins, E. D. *J. Mol. Biol.* **1978**, *125*, 1.
(169) Scott, J. E.; Heatley, F.; Wood, B. *Biochemistry* **1995**, *34*, 15467.
(170) Huang, W.; Boju, L.; Tkalec, L.; Su, H.; Yang, H. O.; Gunay, N. S.; Linhardt, R. J.; Kim, Y. S.; Matte, A.; Cygler, M. *Biochemistry* **2001**, *40*, 2359.
(171) Rigden, D. J.; Littlejohn, J. E.; Joshi, H. V.; de Groot, B. L.; Jedrzejas, M. J. *J. Mol. Biol.* **2006**, *358*, 1165.
(172) Breyer, W. A.; Matthews, B. W. *Protein Sci.* **2001**, *10*, 1699.
(173) Mello, L. V.; van Aalten, D. M.; Findlay, J. B. *Biochemistry* **1998**, *37*, 3137.
(174) Scott, J. E.; Heatley, F. *Biomacromolecules* **2002**, *3*, 547.
(175) Scott, J. E.; Heatley, F. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 4850.
(176) Day, A. J.; Sheehan, J. K. *Curr. Opin. Struct. Biol.* **2001**, *11*, 617.
(177) Scott, J. E.; Thomlinson, A. M. *J. Anal.* **1998**, *192*, 391.
(178) Turner, R. E.; Lin, P. Y.; Cowman, M. K. *Arch. Biochem. Biophys.* **1988**, *265*, 484.
(179) Cowman, M. K.; Matsuoka, S. *Carbohydr. Res.* **2005**, *340*, 791.
(180) Gribbon, P.; Heng, B. C.; Hardingham, T. E. *Biochem. J.* **2000**, *350*, 329.
(181) Almond, A.; Deangelis, P. L.; Blundell, C. D. *J. Mol. Biol.* **2006**, *358*, 1256.
(182) Almond, A. *Cell. Mol. Life Sci.* **2007**, *64*, 1591.
(183) Scott, J. E.; Chen, Y.; Brass, A. *Eur. J. Biochem.* **1992**, *209*, 675.
(184) Davies, G. J.; Wilson, K. S.; Henrissat, B. *Biochem. J.* **1997**, *321*, 557.

CR078240L