2006

DISCOVERY OF LIGNIN SULFATE AS A POTENT INHIBITOR OF HSV ENTRY INTO CELLS

Jay N. Thakkar
Virginia Commonwealth University

DISCOVERY OF LIGNIN SULFATE AS A POTENT INHIBITOR OF HSV ENTRY INTO CELLS

A Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

by

JAY N. THAKKAR
BS in Pharmacy, University of Bombay, Bombay, India 2003

Director: UMESH R. DESAI
ASSOCIATE PROFESSOR, DEPARTMENT OF MEDICINAL CHEMISTRY

Virginia Commonwealth University
Richmond, Virginia
May 2006

Acknowledgements

I am extremely grateful to my advisor Dr. Umesh Desai for his continuous encouragement, comments and advice during my studies at the Virginia Commonwealth University. I would also like to thank him for the financial support provided to me during the course of my education. I would like to express my gratitude to the graduate defense committee, Dr. William Soine and Dr. Michael Hindle for their valuable comments and suggestions on my research work and on the manuscript of this thesis. I would like to thank Dr. Hindle for all his time and help in working with the mass spectrometer. I would also like to thank to Dr. Glen Kellogg for serving as the Dean's representative during my thesis defense.

I would like to thank the Department of Medicinal Chemistry for excellent education I have received during my time at Virginia Commonwealth University. I would also like to appreciate Institute of Structural Biology and Drug Discovery for the provision of the facilities they provided and all the members at the Institute who have helped me in one way or another.

I am very appreciative to Dr. Gunnarson and Arjun Raghuraman for their valuable comments and discussion during the course of this work. I would also like to thank Arjun, who has worked on the same project and has shared some of his data in here.

I would also like to thank my lab members, Dr. Mohammad Riaz, Dr. Bernhard Monien, Mohammed Rahman, Chandravel Krishnasamy, Junaid Afridi, Harminder Bhatia, Brian Henry and Abdul Quadir Khan for their positive encouragement and for creating a great working environment.

Finally, I wish to express my sincere gratitude to my parents and Berin for their constant encouragement and love.

Table of Contents

List of Tables

List of Figures

Abbreviations

HSV:	Herpes Simplex Virus
HS:	Heparan Sulfate
GAG:	glycosaminoglycans
GlcN*p*:	glucosamine
GlcA*p*:	glucuronic acid
HSPG:	Heparan Sulfate Proteoglycans
MS:	Mass Spectrometry
QS:	quercetin sulfate
MoS:	morin sulfate
HP-SEC:	High Performance Size Exclusion Chromatography
CE:	Capillary Electrophoresis
ESI-MS:	Electron Spray Ionization-Mass Spectrometry
LMWH:	Low-Molecular-Weight-Heparins
RP-HPLC:	Reverse Phase-High Performance Liquid Chromatography
DMA:	N,N-dimethyl acetamide
KOH:	potassium hydroxide
TEAST:	triethylamine-sulfur trioxide complex
PAGE:	Polyacrylamide Gel Electrophoresis
M_R:	Average Molecular Weight Range
NMWC:	Nominal Molecular Weight Cut Off
HPLC:	High Performance Liquid Chromatography
GC:	Gas Chromatography
EI:	Electron Impact
CI:	Chemical Ionization
FAB:	Fast Atom Bombardment
API:	Atmospheric Pressure Ionization
ESI:	Electrospray Ionization
APCI:	Atmospheric Pressure Chemical Ionization
LIMS:	Laser Ionization
RIMS:	Resonance Ionization
MALDI:	Matrix Assisted Laser Desorption Ionization
LC:	Liquid Chromatography
ToF-SIMS:	Time of Flight – Secondary Ionization Mass Spectrometry
TIC:	Total Ion Chromatogram

Abstract

DISCOVERY OF LIGNIN SULFATE AS A POTENT INHIBITOR OF HSV ENTRY INTO CELLS

By Jay N. Thakkar, MS

A Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

Virginia Commonwealth University, 2006

Major Director: Umesh R. Desai
Associate Professor, Department of Medicinal Chemistry

The herpes virus family consists of more than hundred members that infect organisms, of which eight, differing markedly in the biology are known to infect humans. HSV- I is the most common one, causing oral lesions and sporadic encephalitis. These infections are highly prevalent affecting at least one in three individuals in the United States.

The entry of the herpes virus into the cell is a two-step process. The initial step involves the cell surface heparan sulfate and glycoproteins in the viral envelope which enables the virus to penetrate into the cell. The second step is the fusion step. Depending on the nature of interaction and size of HS chain, a single chain may bind multiple viral ligands on a virion. There is substantial evidence showing that HS plays an important role in viral binding.

HS is a heterogeneous, linear sulfated oligosaccharide composed of alternating glucosamine and uronic acid residues, which could specify distinct receptor for various viral ligands. HS, present on most exposed cell surfaces, make an ideal snare for the capture of most herpes viruses and may facilitate subsequent interactions with other co-receptors required for entry. Number of viruses, including HSV- I, HSV- II, HIV- I and dengue virus use sites of HS as receptors for binding to cells. Recently 2000 Liu *et.al* have characterized a HS based octasaccharide that binds to HSV-I gD. The distinguished

feature in the composition of the octasaccharide is the presence of 3-O-sulfate glucosamine residue, which is an uncommon structural modification in HS. Its presence in the HSV-I gD binding sequence may confer specificity of interaction and assist HSV-I entry into the cell.

Numerous sulfated molecules have been explored as mimics of HS in the inhibition of HSV-1 entry into cells. To date, most of the sulfated molecules screened for anti-viral activity have been carbohydrates. So, we reasoned that it should be possible to mimic critical interactions of HS with one or more viral glycoprotein using synthetic, non-polysaccharide, sulfated compounds. Further, it may be possible to mimic specific sequence(s) in HS, which play a role in HSV infection, with small synthetic, sulfated, non-carbohydrate molecules. In a search for synthetic mimics of HS as inhibitors of HSV-I infection, we screened a small, synthetic, sulfated flavonoids to discover a potent inhibitory activity arising from sulfation of a macromolecule present as an impurity in a crude natural product.

The active principle was identified through an array of biophysical and chemical analyses as lignin sulfate, a heterogeneous; polydisperse network polymer composed of substituted phenylpropanoid monomers. Further, LC-MS with APCI in negative ionization mode, which have been reported in here for the first time for analysis of lignin, has been successfully used to deduce oligomeric structures present in the precursor of the active macromolecule based on the spectrum of the depolymerized lignin. This corroborates well with the structural information obtained using other analytical techniques. We hypothesize that the structural heterogeneity and polydispersity of lignin coupled with optimal combination of sulfate charge and hydrophobicity result in high potency.

Given that the native lignin is inactive, lignin sulfate discovered here provides a variety of organic scaffolds that with the critical sulfate groups in space can mimic the HSV-I gD binding sequence.

CHAPTER 1

BACKGROUND

1.1 Introduction

Herpes viruses are a leading cause of human viral disease, second only to influenza and cold viruses. They are capable of causing overt disease or remaining silent for many years only to be reactivated. The herpes family contains more than hundred viruses that infect organisms. Eight of these are known to infect human beings **(Table 1)**. In general herpes viruses cannot survive at room temperature, especially outside of the body where they can dry out. Direct contact between susceptible cells and secretions of infected cells is usually required for transmission. The herpes virus has the ability to remain latent and induce latent infections.[1,2] Latent infection occurs when the virus infects cells but remains dormant. This dormancy can last for years. For reasons that are poorly understood the dormant, inactive, infection can become active and cause illness. Reactivation of dormant viruses occurs when they start to reproduce inside cells, eventually causing them to burst and spread the virus particles to propagate infection.[3] The virus particles can also infect neurons. Breakage of latency can occur in these cells and the virus can travel back down the nerve axon, where lesions are seen.[4] This means that recurrence of infection (and therefore symptoms) can occur at the same site as the initial infection. There are several agents that seem to trigger recurrence, most of which are stress-related. It also appears

that exposure to strong sunlight and perhaps fever can lead to recurrence. These factors may cause some degree of immune suppression that leads to renewal of virus proliferation in the nerve cell.

Herpes simplex virus infects a wide range of cells in its normal life cycle, it mainly infects two distinct cell types, epithelial and neuronal. Common sites of infection for herpes simplex are the lips ("cold sores"), and the genitals ("genital herpes").

Herpes infections are highly prevalent affecting at least 1 in 3 individuals in the US. Of the eight strains known to infect humans **(Table 1)**, two types of the herpes simplex virus, HSV-1 and HSV-2 are the most common ones. The primary difference between herpes 1 and herpes 2 is their “site of preference” when establishing latency in the body.[5] Herpes simplex 1 usually establishes latency in the trigeminal ganglion, a collection of nerve cells found in close proximity to the ears. Recurring herpes outbreaks will commonly occur around the mouth or facial region. Herpes simplex 2 usually establishes latency in the sacral ganglion, a collection of nerves found at the lower base of the spine. HSV-2 recurring outbreaks will commonly occur in the genital region. Though this is the most commonly noted difference, it is not absolute. Herpes simplex 1 and herpes simplex 2 can reside in either or both parts of the body and infect orally and/or genitally.

HSV-1 is often transmitted in childhood through kissing, but can be transmitted at any age through direct skin-to-skin contact. In the US about 50% of adults carry HSV-1 antibodies by the time they are young adults and by the time they are 50 years 80-90% of Americans carry HSV-1 antibodies.[6] For both types, at least two-thirds of infected people have no symptoms, or symptoms too mild to notice such as cold sores. However, both types can recur and spread even when no symptoms are present. Occasionally, the virus spreads to the central nervous system causing meningitis or encephalitis. So far, vaccines, ILs, IFNs, therapeutic proteins, antibodies, immunomodulators and small-molecule drugs with specific or non-specific modes of action lacked either efficacy or the required safety profile to replace the nucleosidic drugs acyclovir, valacyclovir, penciclovir and famciclovir as the first choice of treatment.[7-9] All of these drugs are classified, as "Nucleoside Antimetabolites" and they are inhibitors of viral DNA polymerase. These drugs are phosphorylated by viral thymidylate kinase to the monophosphate, which is further bioactivated to the triphosphate. The triphosphate is selective inhibitor of viral DNA polymerase.

Diseases caused by the Herpes virus:

Oral herpes, Herpes keratitis, Herpes whitlow, Herpes gladiatorum, Eczema herpeticum, Genital herpes, HSV proctitis, HSV Encephalitis, HSV Meningitis, HSV infection of

neonates (This results from HSV-2 and is often fatal, although such infections are rare. Infection is especially possible if the mother is shedding virus at the time of delivery).

Table 1: Types of herpes viruses infecting humans and the site of latency

Acronym	Name	Site of Latency
HSV-1	Herpes Simplex Virus 1	Sensory Neurons
HSV-2	Herpes Simplex Virus 2	Sensory Neurons
VZV	Varicella-Zoster Virus	Sensory Neurons
CMV	Cytomegalo Virus	Monocytes, Endothelial cells
EBV	Epstein Barr Virus	B Lymphocytes
HHV-6	Human Herpes Virus 6	T cells, Monocytes
HHV-7	Human Herpes Virus 7	T cells, Monocytes
HHV-8	Kaposi Sarcoma Associated Virus	Monocytes ?

1.2 Herpes Virus Structure – General

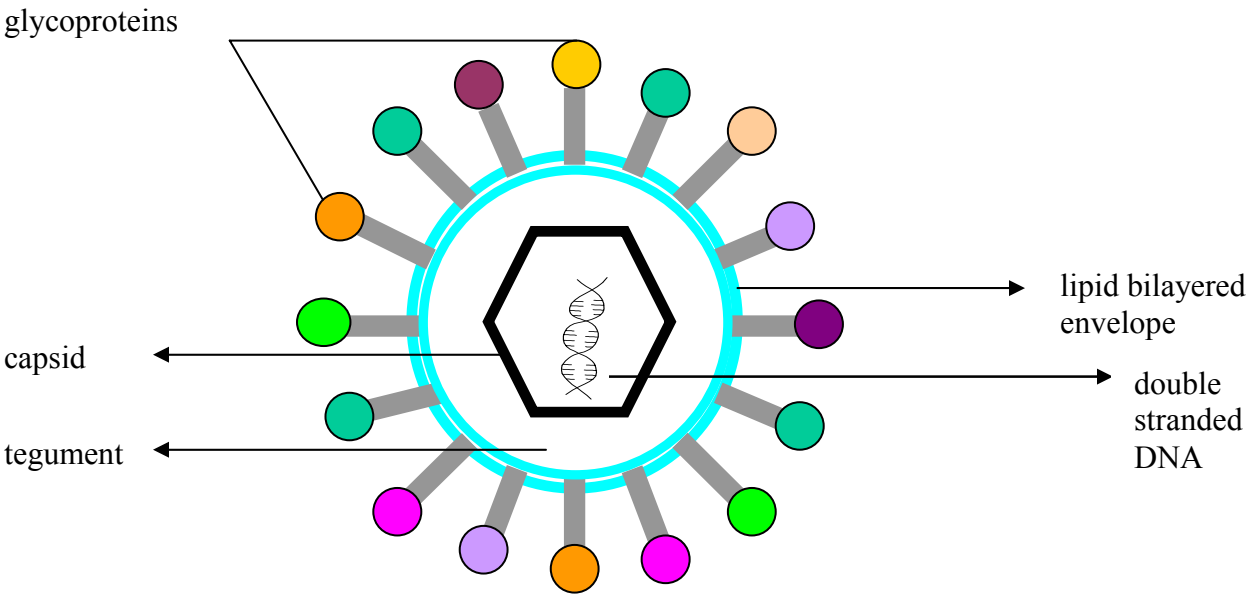


Figure 1. General structure of the herpes virus

The general structure of the herpes virus consists of the following:

Envelope

Herpes viruses are enveloped viruses with at least a dozen glycoproteins attached to the cell surface. They bud from the inner nuclear membrane, which has been modified by the insertion of herpes glycoproteins. These glycoproteins determine the cell to be infected. The viral membrane is quite fragile and a virus with a damaged envelope is not

infectious. Besides drying, the virus is also sensitive to acids, detergents and organic solvents as might be expected for a virus with a lipid envelope.

Tegument

The space between the envelope and the capsid is the amorphous tegument. This contains virally-encoded proteins, such as VP16 and *vhs* (virus host shut off), and enzymes that are involved in the initiation of replication.

Capsid

These viruses have a doughnut shaped capsomere of about 100-200 nm in diameter with an icosahedral proteinaceous nucleocapsid. The latter contains 162 capsomeres or morphological elements.

Genome

Inner core of the herpes viruses have a linear double stranded DNA viral genome encoding about 100-200 genes. Many of these proteins (about half) are not directly involved in the virus structure or controlling its replication, but function in the interaction with the host cell or the immune response of the host. There are two serotypes of herpes simplex virus, HSV-1 and HSV-2 with characteristic resemblances in their genome.

The genome of herpes simplex virus encodes a number of enzymes: DNA-dependent DNA polymerase, thymidine kinase (phosphorylates thymidine and other nucleosides),

ribonucleotide reductase (converts ribonucleotides to deoxyribonucleotide) and serine-protease (converts a scaffolding protein to its final form). The genome also encodes 11 surface glycoproteins. These are involved in attachment (gB, gC, gD and gH), fusion of the viral membrane with that of the host cell (gB, gH and gL), immune escape and other functions (gC, gE and gI), required for production of virus in cell culture (gK), and unknown function (gG, gJ and gM).[10-19]

1.3 Entry of HSV-1 Into Human Cell

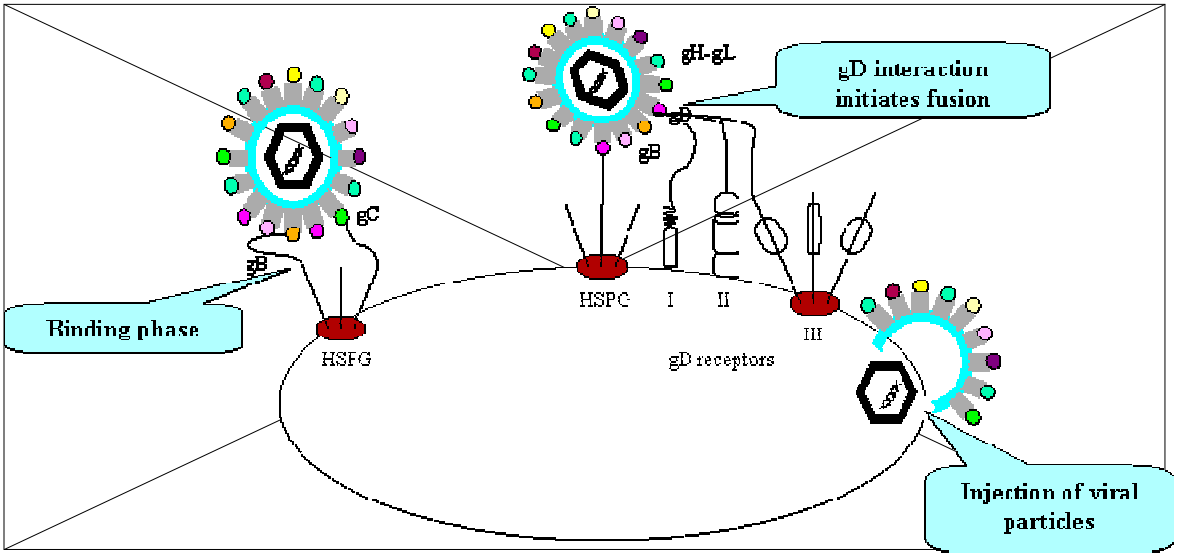


Figure 2. Pictorial representation of the mechanism of the herpes virus entering into human cell.

HSV-1 infection of cells can be divided into two phases, the binding phase and the penetration phase. The first phase involves the binding of the viral glycoprotein gC and/or gB to the glycosaminoglycans (GAGs) chains on cell surface proteoglycans that enables the virus to penetrate into the cell.[20-23] The major type of GAG chain found on the cell surface is heparan sulfate. This initial binding is not sufficient for viral penetration, which further involves fusion of the viral envelope with the cell membrane. This step involves interaction of viral glycoprotein gD with one or several receptors for gD on the cell surface, before the viral particle can fuse with the cell plasma membrane and the viral capsid is released into the cytoplasm. The fusion process is a complex set of events that involves multiple interactions between various gD receptors on the cell surface and the viral glycoprotein gD, gB and hetero-oligomers of gH and gL. Depending on the nature of interaction and size of heparan sulfate chain, a single chain may bind multiple viral ligands on a virion.[24-26] The entry receptors discovered to date fall into three categories. They inclue HVEM (herpes virus entry mediator), a member of TNF-receptor family, nectin-1 and nectin-2, members of the immunoglobin superfamily and specific site in heparan sulfate generated by certain 3-O-sulfotransferases (3-OST).[27-35]

1.4 Evidence Showing Role of Heparan Sulfate

There is substantial evidence showing that heparan sulfate plays an important role in viral binding.[36] [37]First, cells that are devoid of heparan sulfate (but not other GAGs) because of

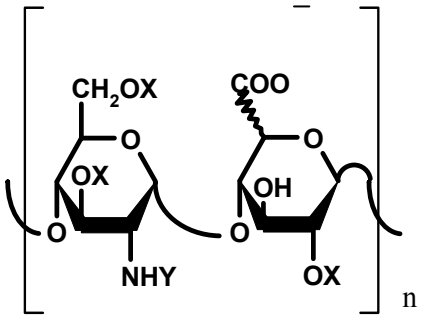
enzymatic treatment or genetic mutation are markedly less susceptible to HSV-1. These cells have greatly reduced number of receptors for viral binding. It has been shown that HSV-1 attaches weakly to heparan sulfate-deficient Chinese hamster ovary (CHO) cells.[38,39] Second, soluble heparin, which is structurally similar to heparan sulfate, inhibits binding of viruses to cells, whereas the structurally less similar GAG chondroitin sulfate fails to inhibit binding or infection.[40] Third, gC and gB bind selectively and independently to heparin-Sepharose columns under physiological conditions. Also, degree of sulfation of glycosaminoglycans is an important determinant for recognition by viral glycoprotein gB2. Heparin, which has higher degree of sulfation as compared to heparan sulfate, is a better competitive inhibitor of gB2 binding.[41] Recently, in 2002 Liu *et.al* have characterized a heparan sulfate-based octasaccharide that binds to HSV-I gD.[42] The distinguishing feature in the composition of the octasaccharide is the presence of 3-O-sulfate glucosamine residue, which is an uncommon structural modification in heparan sulfate. Its presence in the HSV-I gD binding sequence must probably confer specificity of interaction by generating a unique protein binding site, hence assisting HSV-I entry into the cell.[43] Thus, heparan sulfate is required for both viral binding and penetration – the two early events in infection.

Heparan sulfate also plays important role in attachment of other viruses such as the human immunodeficiency virus (E gp120 and gp41)[44,45], dengue virus (gC, gE and

gM)[46], foot and mouth disease virus (E VP1-4)[47,48], vaccinia virus (E A27L)[49], Sindbis virus (E2 glycoprotein)[50], respiratory syncytial virus[51,52], and Echovirus[53]. In each of these viruses, cell surface heparan sulfate is used for gaining entry into the cell.

1.5 Heparan Sulfate

Heparan sulfate is the most complex polysaccharide on the surface of the mammalian cells. Glycosaminoglycan (GAG) polysaccharides are long, polyanionic chains consisting of repeating disaccharide units.[54,55] Heparan sulfate is a GAG covalently attached to the protein core (in the cell membrane - syndecan-1, N-syndecan, fibroglycan, amphiglycan and glypican) of proteoglycans, which are expressed on nearly all cell surfaces.[56-60] Heparan sulfate is also present in the extracellular matrix (perlecan and dystroglycan).



X = H or SO_3^-
Y = H or Ac or SO_3^-

Figure 3. General structure of heparan sulfate showing the disaccharide unit, glucosamine and uronic acid residues linked in a 1-4 manner

The polysaccharide mediates the interactions between numerous different proteins. Heparan sulfate is widely expressed in human tissues and has important roles in development, differentiation and homeostasis.

Structurally, heparan sulfate is a heterogeneous linear co-polymer of glucosamine (GlcN*p*) and glucuronic acid (GlcA*p*) residues linked in a 1-4 manner **(figure 3)**, of which the GlcN*p* residues are typically acetylated at 2- position.[61] Despite this apparently simple monomeric disaccharide structure, heparan sulfate perhaps represents the most structurally complex molecule nature biosynthesizes because of two critical, essentially incomplete, structure modification steps. One, some GlcA*p* residues are epimerized to iduronic acid (IdoA*p*) and two, incomplete sulfation occurs at the 3-O-, 6-O-, and N-positions of GlcN*p* in addition to sulfation at 2-O-position of few IdoA*p* residues.[47] This primary structural diversity is further complicated by another level of complexity wherein sulfate groups may cluster in small regions and form differentially charged domains. A back-of-the envelope calculation of number of structural sequences possible with these variations, especially of the size recognized by proteins and receptors, shows millions of possibilities. The final structure of heparan sulfate depends upon the incompleteness of the reactions that occur during the biosynthetic process. The modification process is more complete in heparin where the final disaccharide IdoA*p*(2-OSO_3)-Glc$NpSO_3$(6-OSO_3) represents up to 70% of the chain, leading to a heavily *O*-sulfated polysaccharide with a

high IdoA*p*/GlcA*p* ratio. In contrast, the modifications that occur during the biosynthesis of heparan sulfate are less extensive; leading to heparan sulfate molecules characterized by lower IdoA*p* content and a lower overall degree of *O*-sulfation and resulting in high heterogeneity of distribution of the sulfate groups along the chain.[62-65] Typical concentrations of heparan sulfate proteoglycans on the cell surface are in the range of 10^5-10^6 molecules/cells as measured in various cell culture systems. The interactions are largely electrostatic in nature and depend on the distribution of hexuronic acid residues and sulfate groups along the polysaccharide chain.

1.6 Heparan Sulfate Mimics

Heparan sulfate plays a vital role in the entry of the HSV-1 into cells. Thus, structural or functional mimics of heparan sulfate are likely to competitively inhibit the entry of the virus into cells. In this context, numerous sulfated molecules have been explored as mimics of heparan sulfate in the inhibition of HSV-1 entry into cells.[66] These include heparin and its chemically modified derivatives, non-anticoagulant heparin, pentosan polysulfate, dextran sulfate, sulfated maltoheptaose, fucan sulfate, fucoidans, spirulan, sulfated galactans, and miscellaneous sulfated polysaccharides[67-80] **(figure 4)**.

These compounds compete with cell surface heparan sulfate for binding to the virus particles and therefore are most active when present during the attachment phase of the

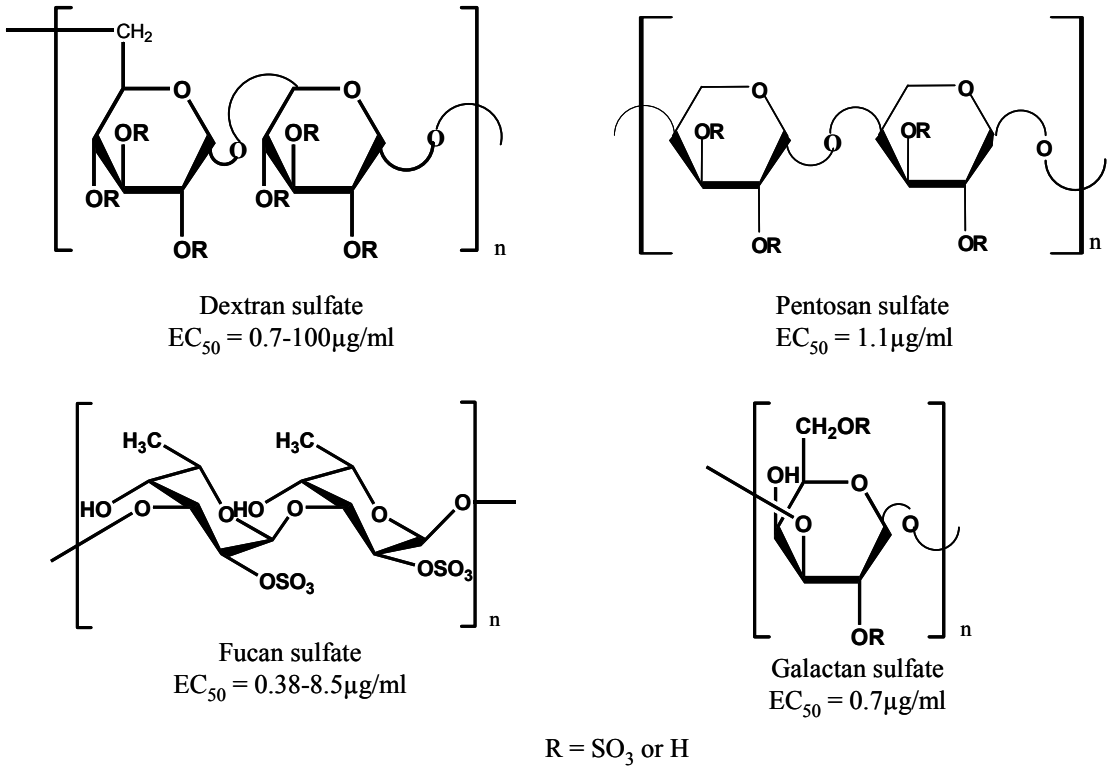


Figure 4. Some of the sulfated polysaccharides that have been tried as heparan sulfate mimics for HSV-1 inhibition activity.

viral entry. Not unexpectedly, each heparan sulfate mimic found to-date as an inhibitor of viral entry has a linear polysaccharide backbone with varying degrees of sulfation. A study of molecular weight dependence of infection inhibition activity shows that the anti-viral potency increases with increasing sulfation levels and length of chain. Yet, detailed aspects of interactions at molecular level, their nature and contributions, the order of interactions, and the step-wise molecular mechanism remains unclear.

1.7 Phenolic Compounds

1.7.1 Lignin

Lignin, after cellulose, is the second most abundant organic substance on earth. It is a major carbon sink in the biosphere accounting for about 30% of the total carbon content in the terrestrial plants. It is the most abundant natural non-carbohydrate organic compound in fibrous materials. The importance of lignin ranges from its fundamental roles in the evolution of the land plants, global carbon cycling, and plant growth and development, and its role in the abiotic and biotic stress of plants i.e: contribution to mechanical strength and protection from degradation, to the practical importance of lignin in agriculture and the utilization of plant materials.[81] The presence of lignin makes woody tissues so compact that molecules as large as proteins cannot penetrate the tissue. Thus, wood is much more resistant to microbial attack. Lignin is a highly heterogenous biomacromolecule, differing between botanical species as well as for different tissues, cells and even cell wall layers. Lignin is made up of three primary precursors, ie. coniferyl, sinapyl and *p*-coumaryl alcohols which undergo polymerization to yield lignin **(figure 5)**.[82,83] There are two different and opposing theories to date for the lignification process. The first one is the non-enzymatic random coupling of the lignin precursor units and the second one involves enzyme catalyzed dehydrogenative polymerization. *In situ*, lignin has no structural regularity. Unlike most natural polymers, such as cellulose and starch, which consist of a single intermonomeric linkage, lignin is a network polymer

made up of three major C6-C3 (phenylpropanoid) units with many carbon-to-carbon and ether (C-O-C) linkages that affects its susceptibility to chemical disruption. There are thought to be about 20 types of inter-subunit linkages in lignins, and several sites on each subunit can participate in inter-subunit bonding.[84]

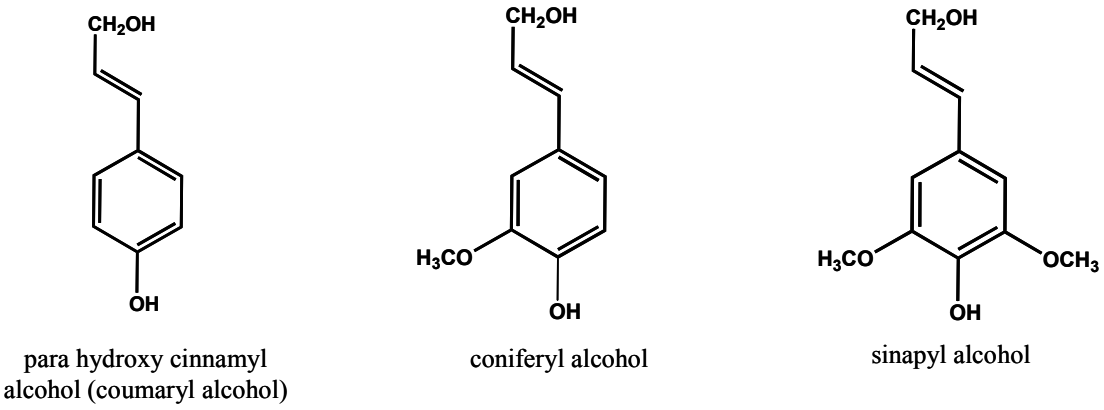


Figure 5. Monomers which serve as the precursor for lignin biosynthesis

Studies have shown that the concentration of the precursors vary in different plant species, coniferous plants have high amounts of coniferyl alcohol, deciduous plants are concentrated with sinapyl alcohol as the major unit and grasses and herb-like dicotylendons with p-coumaryl alcohol. Researchers have studied lignin for more than a century but many aspects of its biosynthesis still remain unclear. Isolation of lignin from the other components of wood without damaging its structure is still a major problem in lignin chemistry. The properties of lignin that results from its complex molecular

structure and its localization within the cell wall, has made the isolation of lignin in an unchanged form, an arduous task. Thus, isolation followed by its complete structure elucidation, has not yet been achieved. As a result, our current understanding of the structure of native lignin has been formed as a sum of the information obtained from different fields of lignin research, namely, studies concerning the elucidation of the mechanisms of lignin biosynthesis and the analytical data obtained in studies with isolated lignin specimens. The structural organization and biosynthesis of lignin are still incompletely elucidated and the current knowledge about lignin structure is fragmentary.

Two lines of lignin analysis are usually pursued. I) Nondestructive analysis methods employed for topochemical investigations of the presence and distribution of lignin by direct analysis of plant tissues. II) Destructive analysis methods for structural analysis of the lignin polymer at the molecular level, where the macromolecule is degraded (either thermally or chemically) to its constituent building blocks and the resulting products are analyzed by gas chromatography/liquid chromatography and/or mass spectrometry. These fragments provide structural information on the original polymer structure. Characterization of lignin is evident by elemental analysis and methoxyl group determination. Further analytical characteristics relevant to other functional groups (phenolic and aliphatic hydroxyl groups, carbonyl and carboxyl groups) signify modification of lignin structure as a result of the isolation procedures or chemical treatments. Mass spectrometry has a prominent role in the identification of the

degradation products. Using mass spectrometry, lignin species could be identified and quantified up to a MW of as high as 50,000 Da.[85,86]

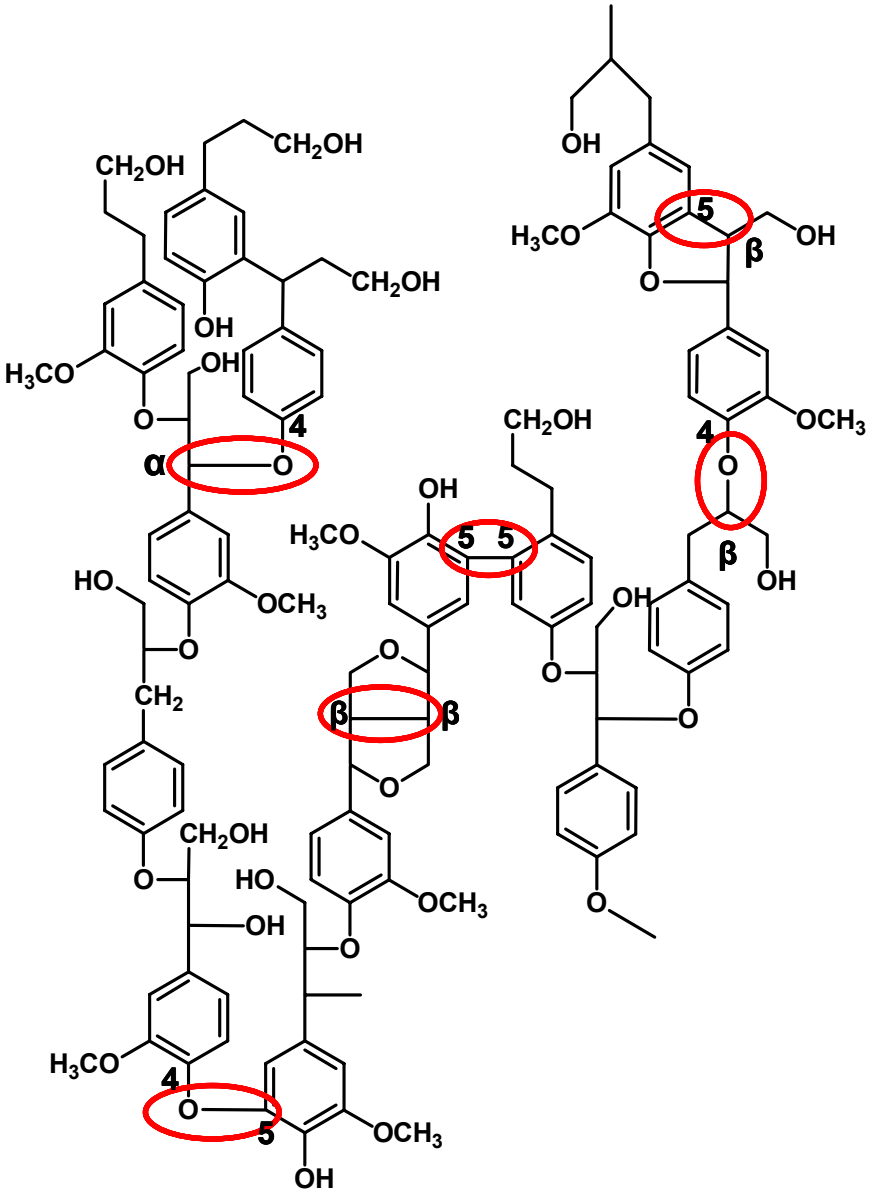


Figure 6. Proposed structure of softwood lignin (Sakakibara, A. Wood Sci. Technol. 1980, 14, 89-100)

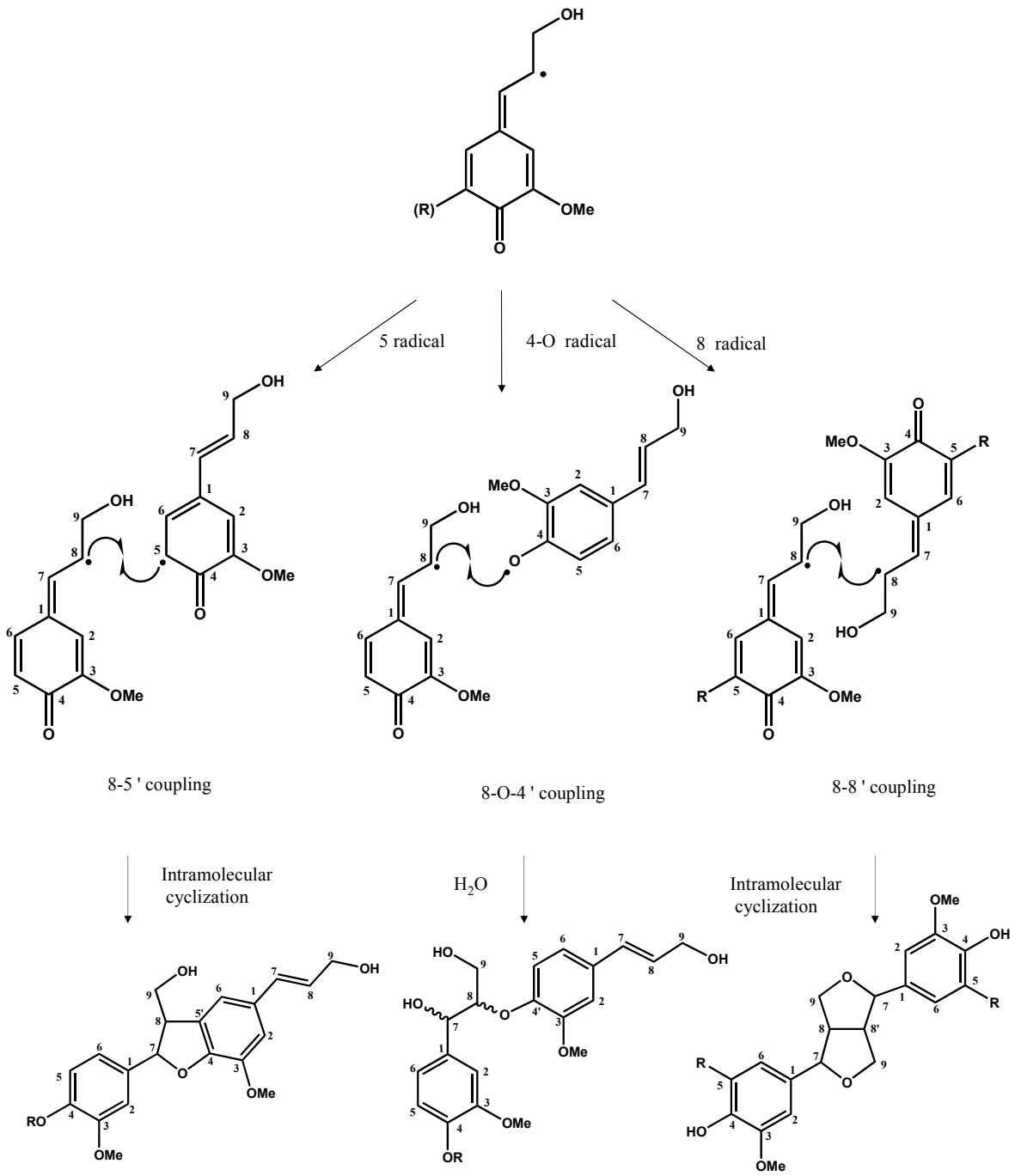
1.7.1.1 Biosynthesis of lignin

The biosynthetic pathway of lignin has been sketched many times and is still a subject of debate. The basic outlines of the lignin monomer biosynthetic pathway have been known for about 30 years, but there are still ambiguities and qualms about how lignin monomers, or monolignols, are formed and how they are further polymerized into lignins. The structure of lignins varies between different parts of the cell wall of a single cell, but in no case has a complete structure for any lignin been defined. Lignification models have been proposed and challenged many times. Many uncertainties exist about the pathway of precursor biosynthesis. The intricacy of lignins can be compared to complex carbohydrates. Such complex polysaccharides are believed to be assembled by complexes of enzymes that join particular subunits in a stereospecific mode to form oligomers, which are then amassed into a highly complex, yet ordered polymers.

The biosynthesis of lignin starts from monolignols, the lignin precursors. These monolignols include coumaryl alcohol, coniferyl alcohol and sinapyl alcohol **(figure 5)**. These precursors are transported to the cell wall, where lignin is formed through dehydrogenative polymerization by either one or combination of peroxidases, laccases, polyphenol oxidases, and coniferyl alcohol oxidase. Peroxidases use hydrogen peroxide (H_2O_2) to oxidize their substrates. Laccases (*p*-diphenol:O_2 oxidoreductases) are copper-containing, cell wall–localized glycoproteins that are encoded by multigene families in

plants. In contrast to peroxidases, laccases consume O_2 (oxygen), instead of H_2O_2 (hydrogen peroxide), to oxidize the monolignols.[87,88]

Polymerization of *p*-hydroxycinnamyl alcohols is initiated by one electron - abstracting enzymes (such as plant peroxidases) yielding phenoxy radicals, and proceeds *via* aromatic radical coupling reactions. Since these phenoxy radicals have the highest π-electron densities at the phenolic oxygen, compared to the side chain carbons, the formation of aryl ether interunit linkages (involving C4) is kinetically favored. In other words the rate of C-O bond formation is greater than the C-C bond formation thus, the final lignin structure would have greater number of ether linkages as compared to C-C linkages (such as β-5, β -1, 5-5 and β - β linkages). Although this C-C linked phenolic moiety represents a low (and variable) fraction of the total lignin, it can strongly affect the reactivity of the polymer. [89] The aromatic radical cross-coupling is imperative to the growing polymer in extending the complex three-dimensional lignin network **(figure 7)**. But, such coupling reactions are radical quenching with each coupling partner requiring a radical generating site, which presumably occurs by radical transfer from monolignols or other intermediaries. Thus, electron deficient monolignols may act as the radical shuttles to cross couple with the polymer radical. But, when the polymer is not electron-deficient, radical transfer may still occur and the monolignol will diffuse back to the peroxidase/laccase to be reoxidized. It has been proposed that redox shuttles, such as an



R = H (±) – dehydrodiconiferyl alcohols
R = , 8-5' interunit linkage

R = H (±) – erythro/threo guaiacylglycerol 8-O-4'- coniferyl alcohol ethers
R = , 8-O-4' interunit linkage

R = H (±) – pinoresinol
R = OMe, (±) - syringaresinols

Figure 7. In vitro depiction of the main free-radical coupling reactions to give corresponding dilignol dimers

Mn^{2+}/Mn^{3+} system may be involved.[90] The actual process of polymer formation, lignification, occurs without the rigid biochemical controls seen in the biosynthesis of the precursor monolignols, giving rise to a unique class of polymers.[91-93]

1.7.2 Tannins

The term 'tannin' refers to the source used to convert rawhide into durable leatherhttp://encyclopedia.laborlawtalk.com/Leather (tanning). However, the term is applied to any large polyphenolic compound containing sufficient hydroxyls and other suitable groups (such as carboxyls) to form strong complexes with proteins and other macromolecules.[94] Tannins are formed *via* the shikimic acid pathway, also known as the phenylpropanoid pathway**,** the same pathway that leads to the formation of other phenolics such as isoflavones, coumarins, lignins and aromatic aminoacids (tryptophan, phenylalanine and tyrosine).[95] Tannins occur in the vacuoles of plant cells. Their presence in trees and woody shrubs produces a bitter taste and astringency that may affect palatability and voluntary intake thus contributing to the chemical defenses that minimize damage to plants.[94,96,97] Tannins have molecular weights ranging from 500 to over 20,000. Most tannins are water soluble but some, very high molecular weight tannins are insoluble. Characteristic features of tannins are; (i) they bind and precipitates proteins, (ii) they form complexes with metal ions (iron, manganese, aluminum, calcium, *etc*.), and (iii) they posses antioxidant and radical scavenging property.[98] Tannins are found to be

present in different parts of plants such as fruit (apple, grapes), stem/bark tissues (oak, willow), leaf tissues (eucalyptus) and root tissues (mangrove). They are also found in grasses (sorghum, corn).

Tannins can be classified into hydrolysable tannins and non-hydrolysable/condensed tannins (proanthocyanidins).

Hydrolysable tannins are described as compounds containing a central core of glucose or other polyhydric alcohol esterified with gallic acid (gallotannins) or hexahydroxydiphenic acid (ellagitannins) **(figure 8)**. Hydrolysable tannins can be hydrolyzed by weak acids or weak bases to produce carbohydrate and phenolic acids. Hydrolysable tannins are usually present in low amounts in plants.

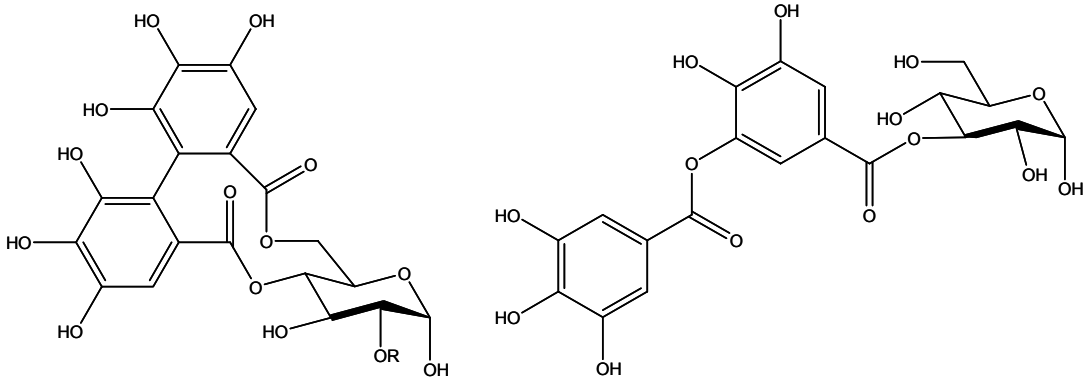


Figure 8. Structures of Hydrolysable tannins

Condensed tannins, also known as proanthocyanidins, are polyflavonoids. It consists of 2 to 50 (or more) units of mostly flavan-3-ols (catechins) and/or flavan 3,4-diols (leucoanthocyanidins) that are joined by carbon-carbon bonds, which are more resistant to breakdown by hydrolysis (figure 10). Proanthocyanidins are more widely distributed than hydrolysable tannins. While hydrolysable tannins and most condensed tannins are water soluble, some very large condensed tannins are insoluble.

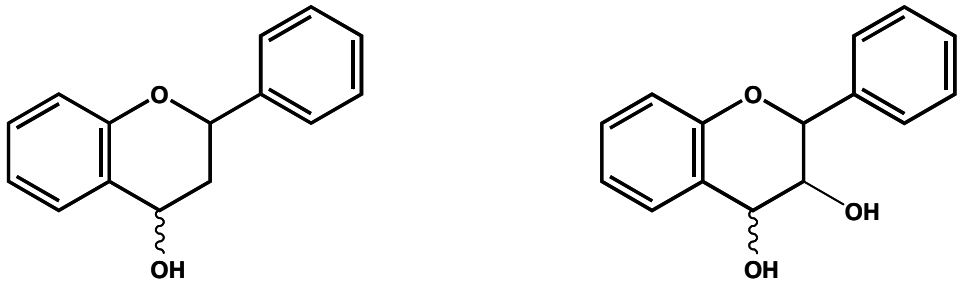
Due to the variety and complexity of tannins, several methods have been developed for their quantification from different sources. These methods can be divided into five main categories: colorimetric, volumetric, gravimetric, protein precipitation and mixed methods. None of them, however, is completely satisfactory for all purposes because each method is used for particular types of tannins and has its own advantages and drawbacks.

1.7.2.1 Condensed tannins (proanthocyanidins)

The term “condensed” in condensed tannin is potentially confusing as the hydrolysable tannin can also undergo condensation reaction. The term, proanthocyanidins, is derived from the acid catalyzed oxidation reaction that produces red anthocyanidins upon heating proanthocyanidins in acidic alcohol solutions.[99,100] The most common anthocyanidins produced are cyanidin (flavan-3-ol, from procyanidin) and delphinidin (from

prodelphinidin).[101] Proanthocyanidin biosynthetic precursors are the leucocyanidins (flavan-3,4-diol and flavan-4-ol) **(figure 9).**[102,103]

Upon autoxidation, in the absence of heat, they form anthocyanidin and 3-deoxyanthocianidin, which, in turn, polymerize to form proanthocyanidins **(figure 10)**. These are one of the many types of secondary compounds found in plants. Condensed



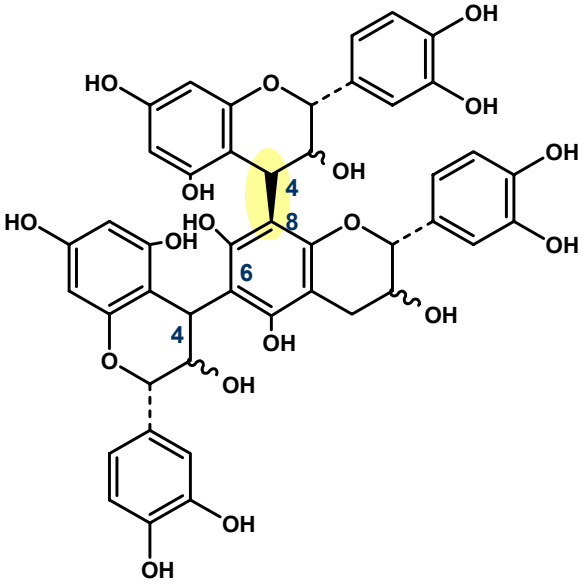
Flavan-4-ol Flavan-3,4-diol

Figure 9. Structures of precursor of proanthocyanidin biosynthesis

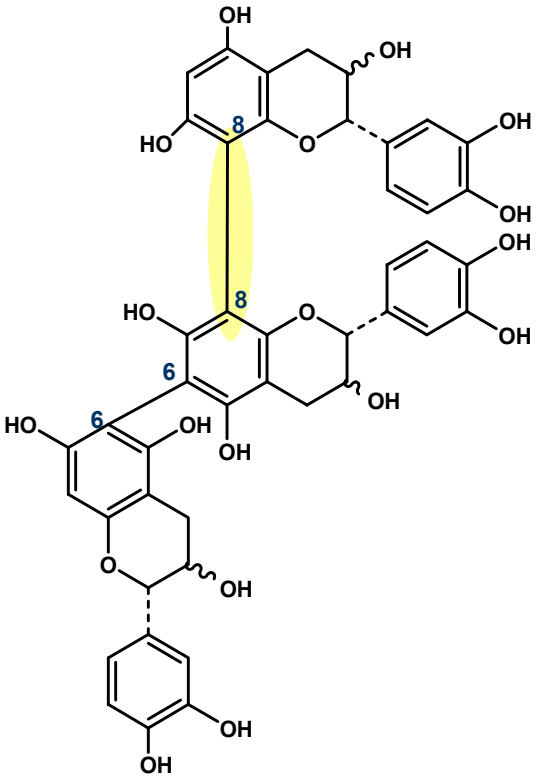
tannins are the most widely distributed in vascular plants and are made up by condensation of hydroxyflavans, leucoanthocyanidin (flavan-3,4-diol) and by catechin (flavan-3-ol). Proanthocyanidin polymers have complex structures because the flavonoid units can differ for some substituents and also because of the variability in interflavan bonding. The flavan-3-ol units are linked principally through the 4-8 and 4-6 carbon-carbon bonds known as B linkages or through 4-8 carbon-carbon linkages and 2-7 ether bonds known as A linkages **(figure 10)**, though the other linkages like the 8-8, 6-8, 6-6

are also present. High molecular weight proanthocyanidin may be completely insoluble in aqueous organic solvents. To date fifteen subclasses of proanthocyanidin have been indentified.[104]

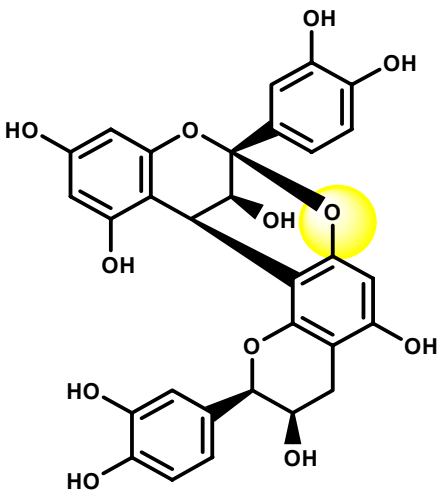
Colorimetric assays like the vanillin-HCl assay, the butanol-HCl assay, the phloroglucinol-HCl assay that are specific for proanthocyanidin can be utilized for identification of proanthocyanidin.[105]



procyanidin B - type



procyanidin B - type



procyanidin A

Fig.10 Structure of condensed tannin showing different interflavan linkages

CHAPTER 2

DISCOVERY, ISOLATION AND IDENTIFICATION OF A POTENT ANTI-HSV-1 AGENT

2.1 Introduction

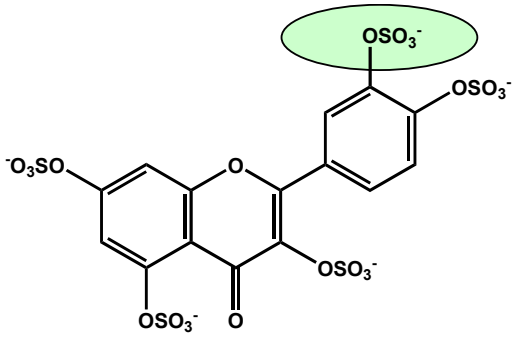
From the literature study, it is clear that heparan sulfate on the cell surface plays an imperative role in assisting the entry of the HSV-1 into cells. Therefore, designing a drug that would inhibit the entry of the virus at the very first stage and thus preventing any further proliferation seems crucial. A simple rational drug approach that comes to mind is designing heparan sulfate mimics that would competitively inhibit the entry of the virus into cells. Literature survey shows that heparan sulfate mimics have been developed to competitively inhibit viral entry. However, most heparan sulfate mimics reported so far have been polysaccharide-based sulfated molecules. We reasoned that it should be possible to efficiently inhibit HSV entry into cells using non-polysaccharide sulfated compounds, especially in light of a 1964 report by Vaheri *et al*. Towards this end, we screened a library of sulfated flavonoids, which we had synthesized earlier, to discover a high activity molecule.

2.2 Results (Part I)

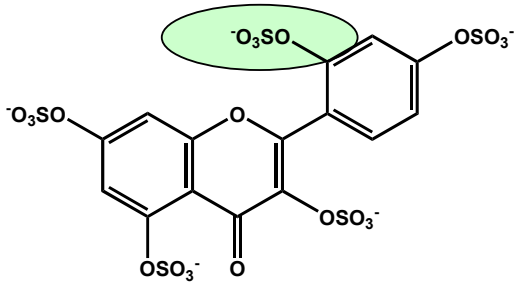
2.2.1 Discovery and Analytical Study of A Potent Anti HSV-I Inhibitor

2.2.1.1 Screening Sulfated Flavonoids

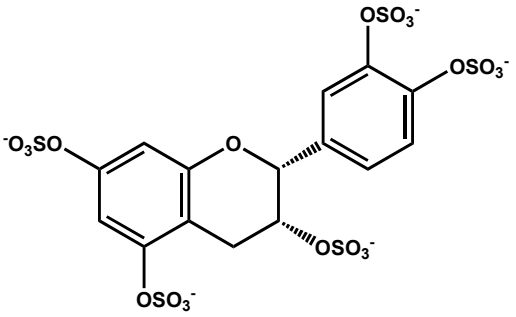
Several sulfated flavanoids were screened including (+)-catechin sulfate, (-)-epicatechin sulfate, quercetin sulfate (QS), and morin sulfate (MoS), each containing multiple sulfate groups **(figure 11)**. These sulfated flavanoids were synthesized as previously reported



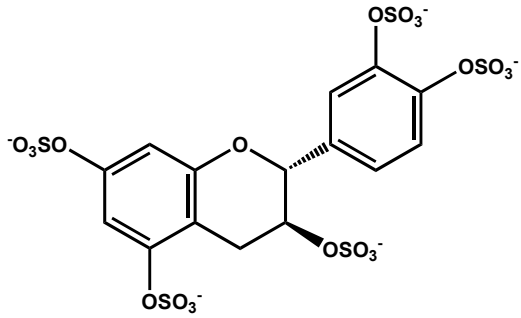
quercetin sulfate



morin sulfate



(-) – epicatechin sulfate



(+) – catechin sulfate

Figure 11. Some of the sulfated flavonoids which were screened for antiviral activity on HeLa cells

and screened for anti-HSV-1 activity. Only crude MoS reaction mixture showed inhibitiory activity. Yet, when purified MoS was tested, it lacked inhibitory activity. This gave the idea that crude MoS contained an impurity that was the seat of inhibitory activity.

2.2.1.2 Identification of Anti-HSV-1 Agent In Crude Morin Sulfate

HP-SEC, using water acetonitrile (7:3 v/v) mobile phase and an analytical column that resolves molecules up to 40,000 Da, showed two peaks at ~24 and ~26 min (peaks 2 and 3), and a broad, unsymmetrical peak between 14 and 24 min (peak 1), **(figure 12)**.

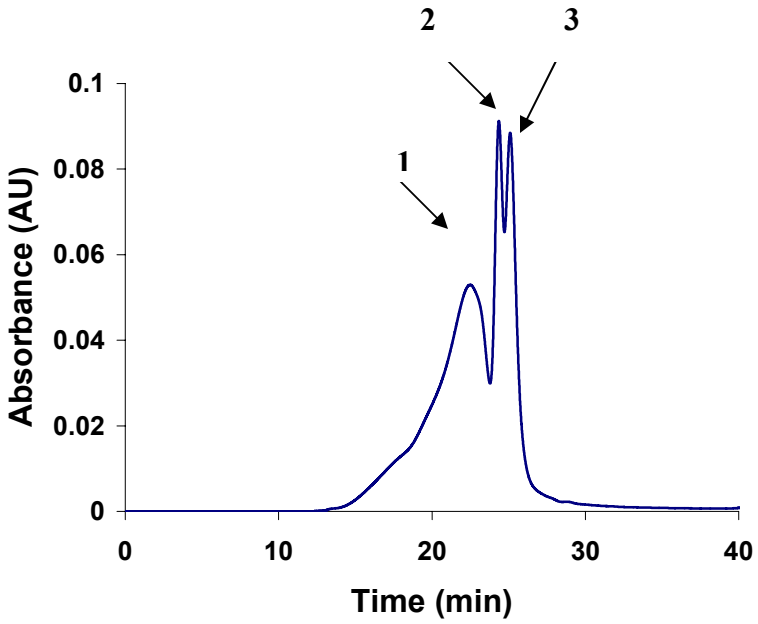


Figure 12. High performance size exclusion chromatography analysis of sulfation reaction mixture of crude morin. Peak 2 and 3 concluded to be a penta and tetra sulfate of morin based on mass spectrometry analysis

CE of the crude MoS reaction mixture under reverse polarity conditions at 25 °C in 100 mM sodium phosphate buffer, pH 2.7, showed the presence of peaks at ~15 and ~16 min, (peak 2 and 3) and a broad peak between 19 and 24 min (peak 1) **(figure 13)**.

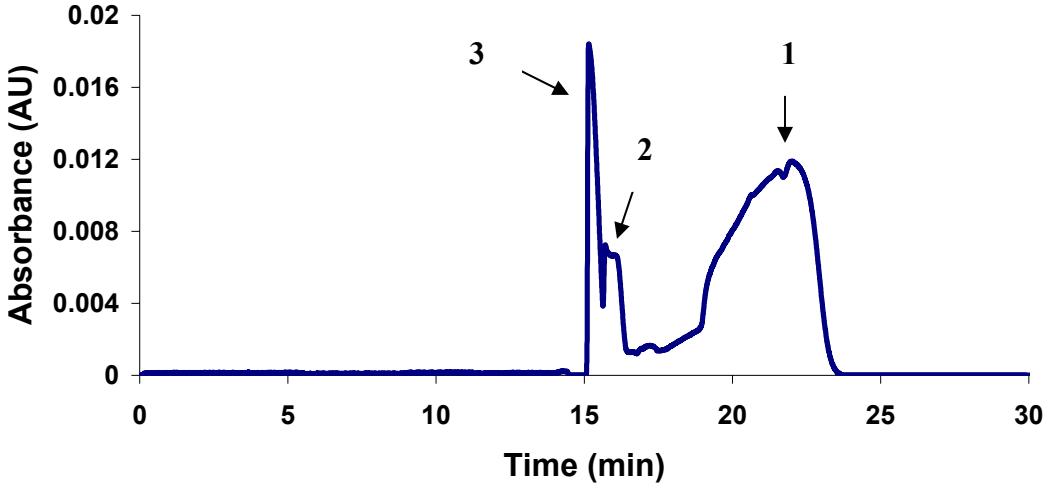


Figure 13. Capillary electrophoretic analysis of sulfation reaction of crude morin. Peak 3 and 2 concluded to be a penta and tetra sulfate of morin based on mass spectrometry analysis

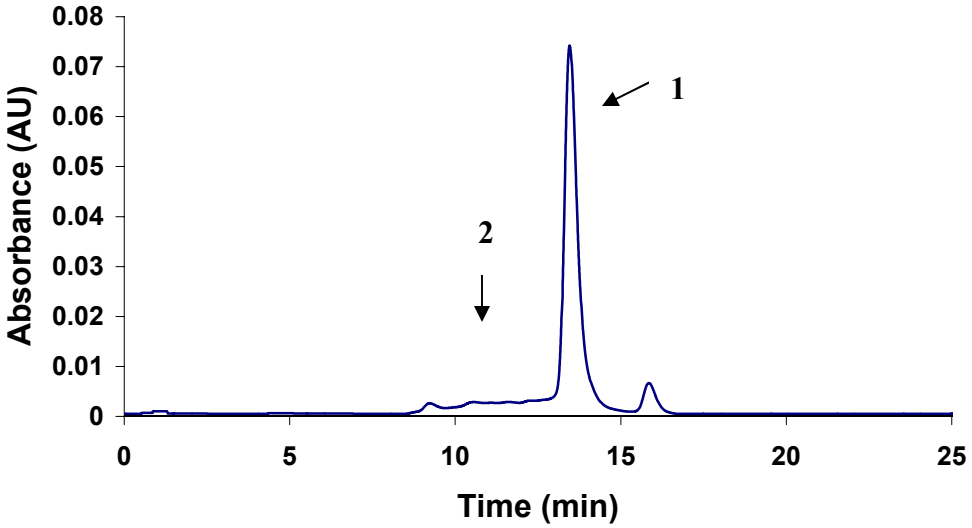


Figure 14. Reverse-phase HPLC analysis of unsulfated crude morin showing a single peak for morin

The reversed phase chromatography showed a dominant morin peak at 16.5 min (peak 1) and additional small peaks between 9 and 15 min (peak 2) **(figure 14).**

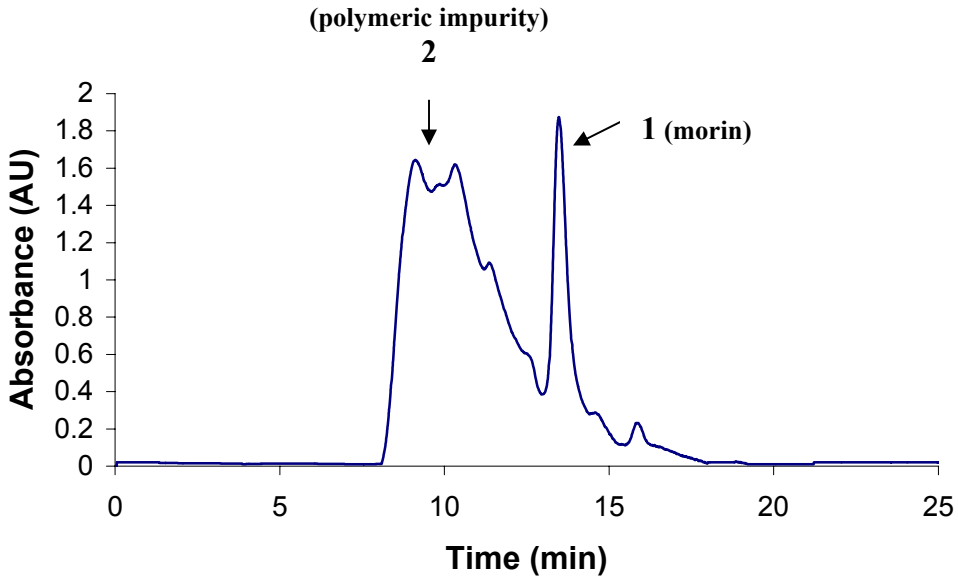


Figure 15. Reverse-phase HPLC analysis of ethyl acetate insoluble substance from crude morin (polymeric impurity of morin)

The RP-HPLC profile of the ethyl acetate insoluble fraction showed multiple unresolved peaks between 9 and 17 min (peak 2), and a morin peak at 16.5 min (peak 1) **(figure 15).**

The acid catalyzed and base catalyzed reactions of morin freed from polymer did not result in any polymer formation. No polymer was formed during the sulfation reaction of morin freed from polymer.

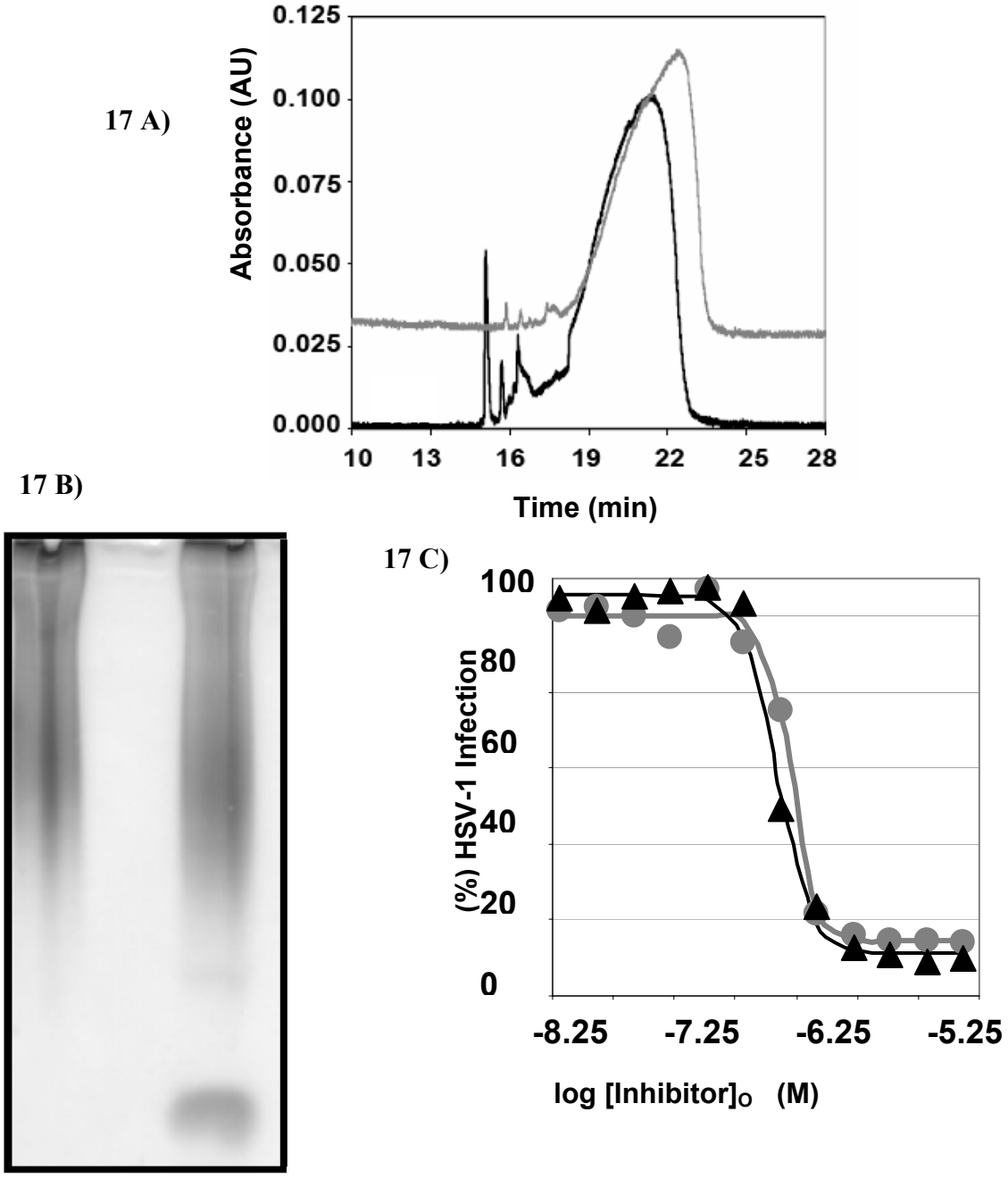


Figure 16. Equivalence of sulfated polymer obtained by isolation from crude morin sulfate reaction mixture (gray profile/lane 1) and direct sulfation of ethyl acetate/acetone insoluble substance from crude morin (black profile/lane 2) using A) capillary electrophoresis; B) polyacrylamide gel electrophoresis and C) viral inhibition studies.

2.3.1 Discussion (Part I)

On screening the library of sulfated flavonoids for anti HSV-1 inhibition only crude MoS reaction mixture showed to have the inhibition activity. Except for MoS, none of the sulfated flavonoids displayed any activity at concentrations as high as 1-2 mM. In contrast, entry of HSV-1 into HeLa cells decreased steadily from ~100% to ~15% for morin sulfate as the concentration was increased to 250 μM. This suggested that only MoS inhibited HSV-1 entry into mammalian cells.

Paradoxically, however, it suggested an exquisite specificity in this inhibition because MoS differs from QS in the position of only one -OSO3 - group, yet QS showed absence of inhibitory activity at concentrations as high as 2 mM **(see figure 11)**. To test whether such exclusive specificity has a structural basis, MoS and QS were resynthesized and rigorously purified. It is important to note at this point that the ^{1}H NMR spectrum of starting natural product morin and crude MoS reaction mixture did not show peaks other than those for morin and morin sulfate, respectively, except for a rolling baseline, thus making it difficult to assess the purity of products. MoS and QS were purified repeatedly and bioassay was performed. Repeated purification followed by bioassay showed near absence of HSV-1 infection inhibition activity in both purified MoS and QS. At the same time, the partially purified MoS reaction mixture continued to display good activity in inhibiting HSV-1 infection from their respective phenolic precursors, which are typically

isolated from plants. Thus, we reasoned that a substance other than the monomeric MoS was inhibiting HSV-1 entry into cells, which was absent in other flavonoids.

To better understand the sulfated product, high performance size exclusion chromatography (HP-SEC) and capillary electrophoresis (CE) of the MoS reaction mixture was performed[106,107]. These analytical techniques would confirm the purity of the compound based on the size of the molecule (by HP-SEC) and the charge they carry (by CE). Partially sulfate and unsulfated molecules can be distinguished by these techniques. HP-SEC, using water acetonitrile (7:3 v/v) mobile phase and an analytical column that resolves molecules up to 40,000 Da, showed two peaks at ~24 and ~26 min (peaks 2 and 3), and a broad, unsymmetrical peak between 14 and 24 min (peak 1), **(figure 12)**. Electrospray ionization mass spectroscopy (ESI-MS) of peak 2 in positive ion mode revealed a molecular ion [M + 6Na]+ at *m*/*z*) 834.75 Da corresponding to morin skeleton functionalized with five sulfate groups (MoS, not shown). Based on its elution pattern, peak 3 is likely to be tetrasulfated morin. In addition, CE of the crude MoS reaction mixture under reverse polarity conditions at 25 °C in 100 mM sodium phosphate buffer, pH 2.7, showed the presence of peaks at ~15 and ~16 min, and a broad peak between 19 and 24 min **(figure 13)**. The peak at ~15 min in CE was identified as the monomeric MoS by injecting the purified sulfated flavonoid under similar conditions, while the broad peak did not resolve even with exhaustive changes in capillary voltage, type of buffer, its pH,

and the ionic strength of buffer. Such broad peak profiles are typical of heterogeneous, polydisperse anionic polymers exemplified by full-length and low-molecular-weight heparins (LMWH), thus implying the presence of an unexpected polysulfated polymer in the MoS reaction mixture. Further, the detectability of this polymer in HP-SEC and CE, but not in ^{1}H NMR, indicated that the polymer was highly heterogeneous, and possibly polydisperse, with each chain constituting an exceedingly small proportion of the overall content. To assess whether HSV-1 infection inhibition activity arises from the polysulfated polymer present in the MoS reaction mixture, we used Sephadex G10 chromatography with 20% ethanol as eluent to separate the polymer from the monomers. The polymer, equivalent to peak 1 of **figure 12**, resolves readily from the monomeric entities peaks 2 and 3, and was obtained in ~10-14% isolated yield. When screened for HSV-1 viral infection inhibition activity, only peak 1 was found to be active, whereas peaks 2 and 3 were found to be inactive.

The reverse phase chromatography showed a dominant morin peak at 16.5 min and additional small peaks between 9 and 15 min which could possibly be arising due to the background noise, hence it was not conclusive **(figure 14)**.

HP-SEC and CE profiles confirmed that the inhibition activity was solely due to a polymeric species present in the reaction mixture and not due to MoS. The profiles

further indicated that a heterogeneous, polydisperse, polysulfated molecule present in MoS reaction mixture was the origin of HSV-1 infection inhibition activity. There were two possibilities that could have caused this polymer formation:

1) the active polymer is synthesized from the monomeric starting material morin during the sulfation reaction, or
2) it is a product of sulfation of a polymeric substance already present in the raw material.

The analysis of natural product morin by RP-HPLC using acetonitrile-water (8:2 v/v) mobile phase on a C-18 column showed a dominant peak at 16.5 min corresponding to starting material morin **(figure 14)**. However, the RP-HPLC profile showed additional small peaks between 9 and 15 min, which can merge in background noise, likely arising from small amounts of a polydisperse polymer. RP-HPLC screening of several independent sources of natural product morin showed nearly identical chromatograms, suggesting that this substance was consistently present.

The polymeric substance was isolated by exploiting its differential solubility in organic solvents. Whereas morin was found to possess good solubility in ethyl acetate or acetone, the extraneous polymer was nearly insoluble. The RP-HPLC profile of the ethyl acetate insoluble fraction showed multiple unresolved peaks between 9 and 17 min **(figure 15)**, which corresponded well with those found in crude morin. The peak shape indicates a

highly polydisperse, heterogeneous polymer. Quantitative analysis of the ethyl acetate precipitation step indicated that the polydisperse polymer was present in ~10% proportion from all three different sources of morin, an observation suggesting the likelihood of a biopolymeric nature.

To test whether the polydisperse polymer is a product of polymerization of morin, we tested the ability of morin to polymerize under acidic and basic conditions. Morin, freed from the polymer, was treated with catalytic amount boron trifluoride-diethyl ether complex or hydrochloric acid in DMA at 60 °C. Except for the presence of sulfating agent, these conditions simulate the possible generation of acid during sulfation. Likewise, morin was heated with KOH in DMA at 60 °C. Each reaction was continuously monitored on RP-HPLC using the protocol developed for separating morin and the polymer, yet no polymer formation could be detected in either of the reactions. These results indicate that morin is stable to acids and bases, and does not react readily to give a polymeric species. This conclusion is also supported by literature data suggesting that polymers of flavones, which contain a 2-en-4-one moiety, are relatively unknown.[3] In contrast, polymers of flavans, which do not contain a 2-en-4-one structure, are abundant. Further, the projected morin polymer, a flavonoid derivative, is expected to contain carbonyl groups, which were found to be absent in the isolated polymeric product

(Refer IR data on page 41 & 48). Finally, morin, freed of the polymer, on sulfation with triethylamine–sulfur trioxide complex did not yield any sulfated polymer.

To test whether the precipitated polymer is the origin of anti-HSV-1 activity, a sample of the polymer, free from monomeric morin, was prepared using centrifugal membrane filtration **(see Experimental Section)** and subjected to sulfation with triethylamine-sulfur trioxide complex (TEAST) in DMA at 65 °C. HSV-1 viral infection inhibition screening showed that this sulfated polymer was highly active **(figure 16C)**. In fact, the inhibitory activity (IC_{50}) of 6.0 μg/mL found for this polymer was essentially identical to that of the sulfated polymer obtained through purification from reaction mixture (7.5 μg/mL). In contrast, the parent ethyl acetate insoluble precipitate, the unsulfated polymer, was completely inactive. Finally, to ascertain that the sulfated polymer obtained from the ethyl acetate insoluble material was identical to the polymer purified from the MoS reaction mixture, comparative CE and PAGE techniques were used **(figure 16A and 17B)**. Both the CE and the PAGE profiles of sulfated polymer obtained from sulfation of ethyl acetate or acetone precipitate, or chromatographic purification of MoS reaction mixture, were rather similar, suggesting that both of the sulfated polymers are identical.

Table 2: Summary of the techniques and corresponding results indicate presence of a heterogeneous, polydisperse, polymeric species in crude morin

Technique	Conclusion
HP-SEC of sulfation reaction mixture	High molecular weight polydisperse molecule present
CE of sulfation reaction mixture	heterogeneous, polydisperse anionic polymer present
Reverse phase HPLC of morin	Not conclusive
Reverse phase HPLC of partially purified polymer from morin	heterogeneous, polydisperse polymer present
CE, PAGE and activity data	Both sulfated polymer obtained by isolation from morin sulfate reaction mixture and direct sulfation of ethyl acetate/acetone insoluble are the same
Acid catalyzed, base catalyzed and sulfation of pure morin reaction	Under normal sulfation reaction conditions morin cannot form a polymer

2.2 Results (Part II)

2.2.2 Elemental and Spectroscopic Analysis of Polymer

2.2.2.1 Elemental Analysis

The elemental analysis of parent un-sulfated crude polymer, devoid of monomeric morin, revealed the presence of C, H and O only, while its sulfated counterpart contained C, H, O, S and Na elements.

2.2.2.2 IR Analysis

The IR spectrum of the sulfated polymer shows a broad band in the region 3200-3600 cm^{-1} and bands at 1610, 1500, 840, and 760 cm^{-1} **(figure 17)**

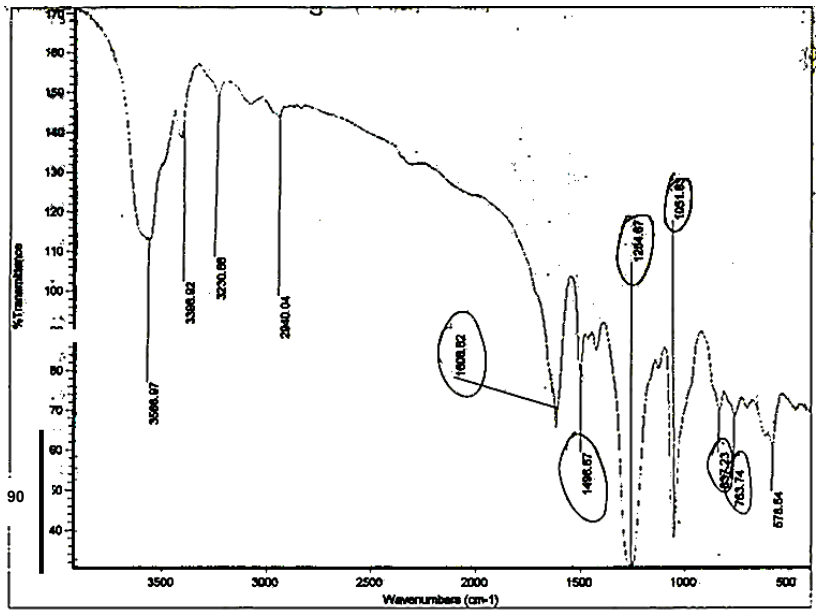


Figure 17: IR spectrum of the sulfated polymer showing presence of aromatic region, presence of acidic and hydroxyl groups but absence of carbonyl groups

2.2.2.3 UV-Visible Analysis

The UV spectrum of the parent, unsulfated polymer in ethanol showed absorbance at 280 and 330 (sh) nm **(figure 18)**. The band at 280 nm underwent bathochromic shift of 26 nm in 1 N NaOH. Sulfation of the polymer resulted in hypsochromic shift of 12 nm in the band at 280 nm.

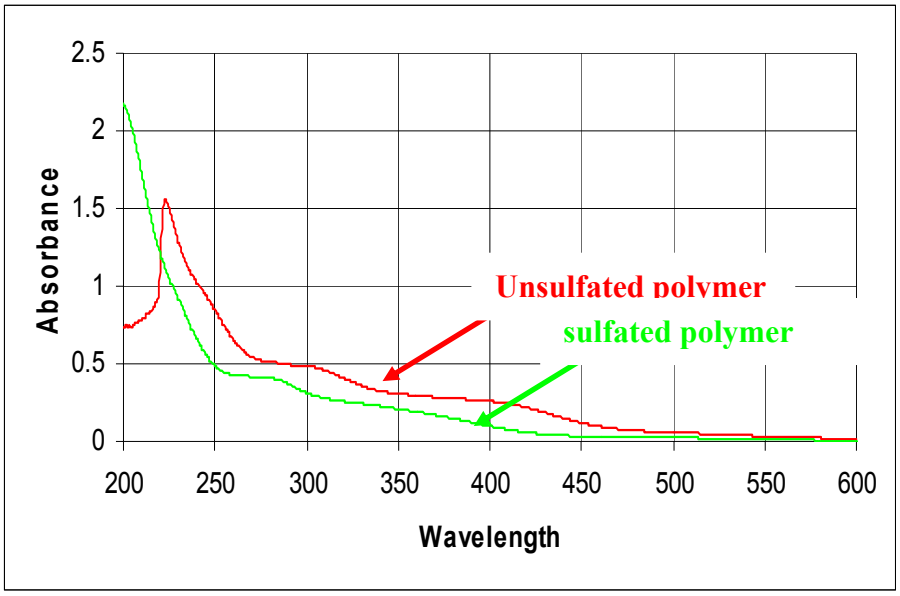


Figure 18. UV spectrum of the unsulfated and sulfated polymer

2.2.2.4 ^{1}H NMR

The ^{1}H NMR of 1.9 kDa sulfated polymer fraction in $^{2}H_2O$ at 25 °C showed several broad signals corresponding to significant polydispersity and heterogeneity in the sample **(figure 19)**. The spectrum indicates the presence of aromatic protons between 6 and 8 δ,

methine protons attached to multiple electron-withdrawing groups (OC*H*-Ph) at 5 δ, aromatic methoxys (ArOC*H*) at 4.2 δ, aliphatic methoxys (ROC*H*) at 3.6 δ, benzylic methylenes (ArCH) at 2.4 δ, and aliphatic methyls (RCH_3) and methylenes (RCH_2) at 0.9 and 1.2 δ.

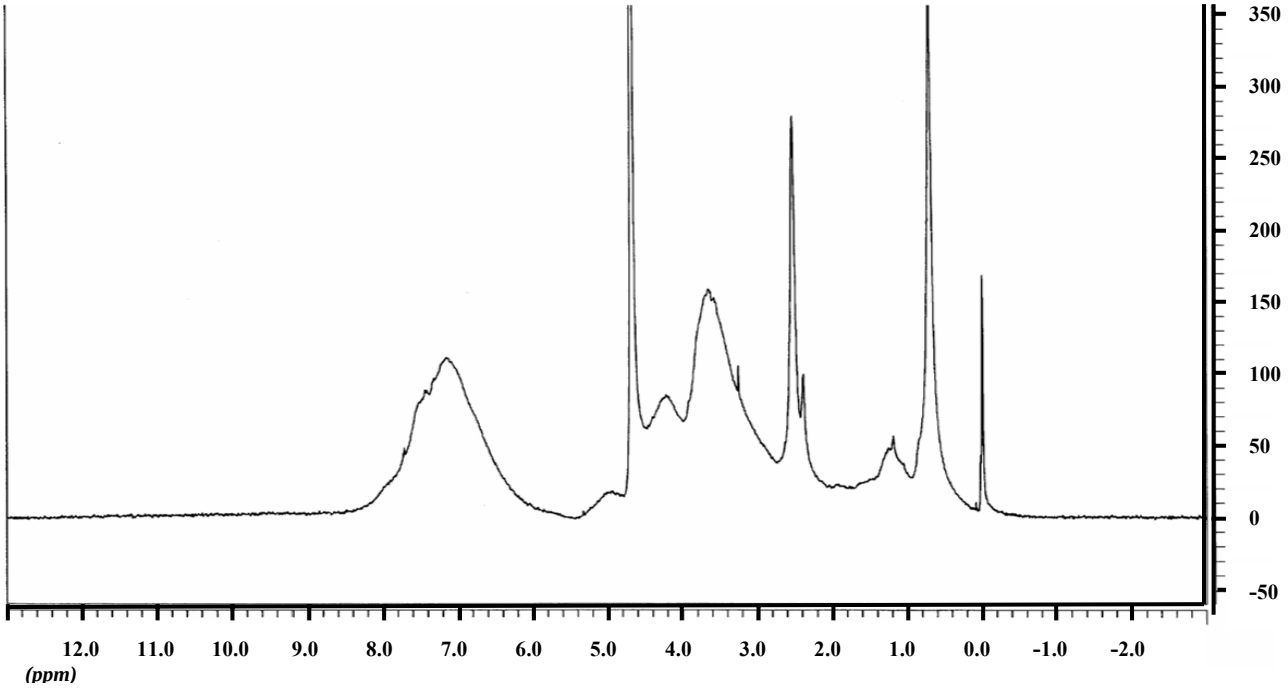


Figure 19. ^{1}H NMR of the sulfated polymer

2.2.2.5 ^{13}C NMR

The ^{13}C NMR spectrum of the sulfated polymer, recorded with a sample (80 mg/mL) at 60 °C and a line broadening factor of 50 Hz, showed broad peaks at 10 and 56 ppm

corresponding to -CH_3/CH_2 and -OCH_3 groups, respectively, in addition to peaks for aromatic carbons in the range 115-160 ppm **(figure 20)**. A broad signal at 48 ppm is also observed. This peak is unusual with very few organic groups resonating at this position and has been assigned to the HO*C*H- groups, which typically resonate at ~50 ppm considering that the three-dimensional structure of the molecule may introduce special shielding environments that cause this 1-5 ppm shift **(figure 20).**

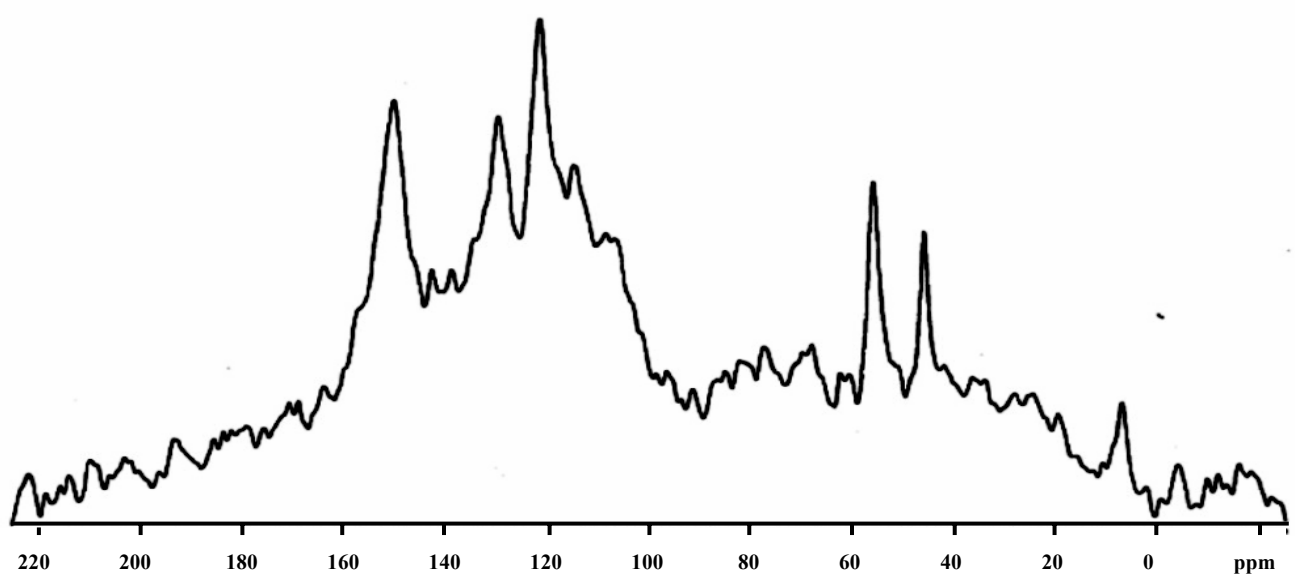


Figure 20. ^{13}C NMR of sulfated polymer

2.3.2 Discussion (Part II)

Elemental analysis is an experiment that determines the amount (typically a weight percent) of an element in a compound. Just as there are many different elements, there are

many different experiments for determining elemental composition. The most common type of elemental analysis is for carbon, hydrogen, and nitrogen (C, H and N analysis). This type of analysis is especially useful for organic compounds (compounds containing carbon-carbon bonds). The elemental analysis of a compound is particularly useful in determining the empirical formula of the compound. The empirical formula is the formula for a compound that contains the smallest set integer ratios for the elements in the compound that gives the correct elemental composition by mass.

The elemental analysis of parent un-sulfated crude polymer, devoid of monomeric morin, revealed the presence of C, H and O only, while its sulfated derivative contained C, H, O, S and Na elements, suggesting that the polymer does not contain N.

To assist in structure elucidation of both the unsulfated, native polymer and its sulfated counterpart, a sulfated polymeric fraction of M_R 1900 Da was isolated from the heterogeneous, polydisperse preparation using a combination of centrifugal filtration through molecular membrane (NMWC 5K) and preparative size-exclusion chromatography. The elemental analysis of parent unsulfated crude polymer, devoid of monomeric morin, revealed the presence of C, H, and O only, while its sulfated derivative contained C, H, O, S and Na elements, suggesting that the polymer does not contain N.

The IR spectrum of the sulfated polymer shows a broad band in the region 3200-3600 cm^{-1} corresponding to -OH stretch and bands at 1610, 1500, 840, and 760 cm^{-1} showing the presence of an aromatic structure, but more importantly displaying absence of certain groups, such as the carbonyl and triple bond.

The UV spectrum of the parent, unsulfated polymer in ethanol showed absorbance at 280 and 330 (sh) nm, indicating the presence of aromatic structure **(figure 18)**. The band at 280 nm underwent bathochromic shift of 26 nm in 1 N NaOH, which is characteristic of an acidic ionizable group such as a phenolic hydroxyl or carboxylic acid. Sulfation of the polymer resulted in hypsochromic shift of 12 nm in the band at 280. This change in λMAX to lower wavelengths by 6-10 nm is indicative of *O*-sulfation reaction. Finally, addition of alkali to the sulfated polymer solution resulted in this band at 268 nm red shifting by 12 nm, indicating that the sulfated polymer retains some underivatized phenolic groups.

The ^{1}H NMR of 1.9 kDa sulfated polymer fraction in $2H_2O$ at 25 °C showed several broad signals corresponding to significant polydispersity and heterogeneity in the sample **(figure 19)**. The spectrum indicates the presence of aromatic protons between 6 and 8 δ, methine protons attached to multiple electron-withdrawing groups (OC*H*-Ph) at 5 δ,

aromatic methoxys (ArOC*H*) at 4.2 δ, aliphatic methoxys (ROC*H*) at 3.6 δ, benzylic methylenes (ArCH) at 2.4 δ, and aliphatic methyls (RCH_3) and methylenes (RCH_2) at 0.9 and 1.2 δ. These broad resonances sharpen at higher temperatures, possibly because of enhanced flexibility, but do not resolve into sharp peaks for a more definite interpretation. The signal at 0.9 δ suggests the presence of methyl groups in a rather hydrophobic environment of an anionic polymer. The observation that broad resonances dominate the ^{1}H NMR spectrum of a sulfated polymer with a relatively low *M*R of 1900 Da (even at elevated temperature) suggests the presence of a highly networked polymer. More importantly, it indicates that the monomers making up the network are small (~100-200 Da).

The ^{13}C NMR spectrum of the sulfated polymer, recorded with a sample (80 mg/mL) at 60 °C and a line broadening factor of 50 Hz, showed broad peaks at 10 and 56 ppm corresponding to -*C*H$_3$/*C*H$_2$ and -O*C*H$_3$ groups, respectively, in addition to peaks for aromatic carbons in the range 115-160 ppm **(figure 20)**. A broad signal at 48 ppm is also observed. This peak is unusual with very few organic groups resonating at this position and has been assigned to the HO*C*H- groups, which typically resonate at ~50 ppm considering that the three-dimensional structure of the molecule may introduce special shielding environments that cause this 1-5 ppm shift. This spectrum correlates well with the ^{1}H NMR spectrum discussed above. Also, the ^{13}C NMR spectrum shows absence of

resonances in the region 180-200 ppm that correspond to carbonyl carbons, further supporting the FT-IR data.

Table 3: Summary of the techniques and corresponding results indicate presence of a heterogeneous, polydisperse, polyphenolic polymeric species in the starting material morin

Analytical Technique	Conclusions
Elemental Analysis	C, H, O present; N absent
IR	Ar & -OH present; C=O & triple bond absent
UV	ArOH, -COOH present
^{1}H NMR	Ar-OCH_3, R-OCH_3, Ar-CH, R-CH, R-CH_3
^{13}C NMR	Ar, OCH_3, ROCH-, CH_3, CH_2 present C=O absent

2.2 Results (Part III)

2.2.3 Polymer Is A Lignin Derivative

2.2.3.1 Acid butanol test

The acid butanol test, which is specific for condensed tannin failed to develop pink coloration for our heterogeneous, polydispersed polymer.

2.2.3.2 Phloroglucinol – HCl Test

The phloroglucinol HCl reaction used for detecting condensed tannins did not yield any change in the parent polymer.

2.2.3.3 Test for tannin/lignin

The test carried out by using the LaMotte Octet Comparator kit showed light blue coloration development on reacting with the polymer.

2.3.3 Discussion (Part III)

The polymer isolated from morin sample is unlikely to have been formed synthetically from morin, and is more likely to be a structurally distinct class of natural product which is present in the starting material that may have squeezed during the extraction of morin from a plant or tree. The anti HSV-1 sulfated polymer is a sulfated derivative of a polymer present in raw material, morin.

The biophysical properties of the polymeric natural product indicate that the polymer is polyphenol-based. Natural polyphenolic polymers with characteristics matching the ones found by biophysical analysis include condensed tannins and lignins.

In literature there few tests, which are specific for condensed tannins; like the acid butanol test and the phloroglucinol test.[109]

Acid butanol test: The butanol-HCl reaction is specific to the proanthocyanidins and the rarely occurring flavan-3,4-diols. Under appropriate conditions, virtually all condensed tannins can be degraded to yield anthocyanidins that can be spectrophotometrically determined. This method estimates all condensed tannins in terms of a single chemical product that is common to a wide range of parent compounds. In this method, the formal subunit of the condensed tannins, the flavanoid, is determined after chemical modification to yield a chromophore (the anthocyanidin), which forms a complex with a ferric reagent. Complexation results in the development of a distinctive pink color, which can be quantified providing a convenient colorimetric method for estimating total condensed tannins. In other terms the assay uses an acid-catalyzed oxidative depolymerization of condensed tannins (proanthocyanidins) to yield red anthocyanidins. As the resulting anthocyanidins, mostly being either cyanidin or delphinidin, depend on the catechin units of the condensed tannin, this assay has been used for identification of the polyflavan structure. The assay's greatest strength lies in the qualitative confirmation of a polymeric interflavan structure. Indeed, it has efficiently been applied as one element in structure elucidation of polymeric proanthocyanidins.

Phloroglucinol HCl test: This is a mild acid catalyzed scission, of condensed tannins where the generated extender flavan carbocations are captured with nucleophilic phloroglucinol. This reaction results in the cleavage of the terminal flavanoid units. The generally accepted view of tannin formation has involved nucleophilic attack by the aromatic A ring on a carbocation or, as more recent evidence suggests, a quinone-methide intermediate, either of which may be derived from a flavan-3,4-diol. In the presence of both the resorcinol- and phloroglucinol-type nucleophiles, the latter has considerably stronger nucleophilic centres [S] and hence would be preferentially involved in condensation with reactive intermediates to form higher oligomers.

Both the acid-butanol and the phloroglucinol tests failed for the parent unsulfated polymer, indicating the absence of condensed tannin-type structure. The above results suggest that our parent, unsulfated polydisperse polymer was not tannin, but most probably a lignin.

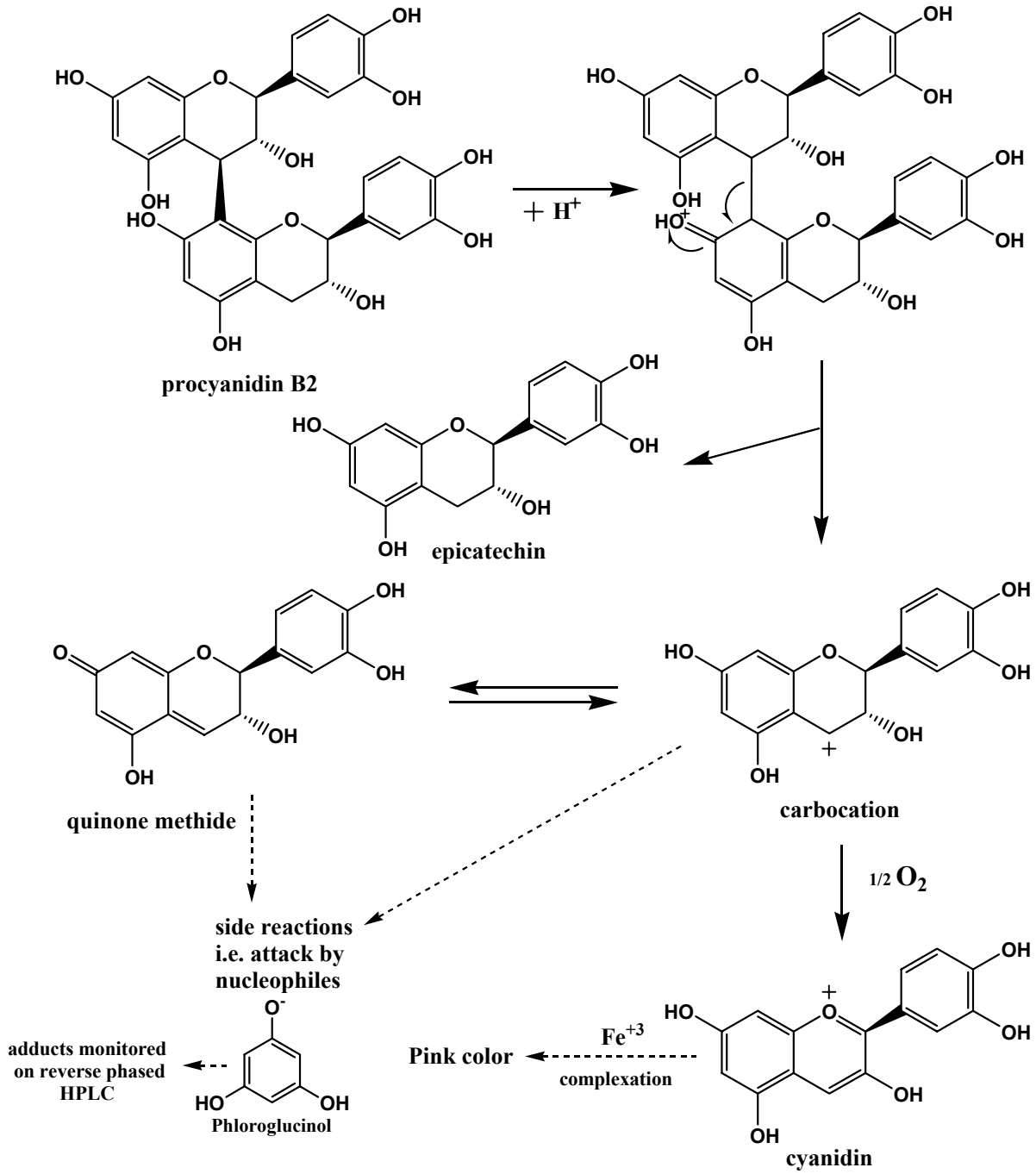


Figure 21. Reaction mechanism explaining the acid-catalyzed cleavage of proanthocyanidins

To ascertain the polymer was a lignin type compound, LaMotte Octet Comparator was used. The test is not specific for tannin or lignin and the result may be reported in more general terms such as “tannin-like” or “lignin-like” compounds. Lignin produces about one-third the color that is given by tannin or tannic acid. In this test complex polymeric ions formed from phosphomolybdic and phosphotungstic heteropoly acids are reduced by phenolic (and other reducing) compounds to complex molybdenum-tungsten blue. Thus, the color yield in the assay depends on the redox potential of the reference standard used as well as the phenolic compounds and other interfering compounds such as aromatic amines, carbohydrates. Consequently, quantitative results have to be interpreted carefully and should be seen as total polyphenol estimate rather than total polyphenol content, particularly when comparing results from different sources or from different assays. Lignin develops one-third the color that is developed by tannins, which was found to be true when we ran a test for our polymer against standard tannin.

Table 4: Summary of the techniques and corresponding results indicate presence of a heterogeneous, polydisperse, polyphenolic lignin class of polymeric species in the starting material morin

Technique	Conclusions
Acid butanol test	No pink coloration, condensed tannin absent
Phloroglucinol test	No change observed on the reverse phase HPLC, condensed tannin absent
Test for tannin-like/lignin-like	Polymer developed one-third the coloration, tannin-like compound absent

So, it was most likely that the heterogeneous, polydispersed polymer appeared to be a lignin class of compound.

Lignin is a plant cell-wall material, which is also a common source for commercial morin. In contrast, proanthocyanidins are typically isolated from grapes and wines. Second, the observation that a broad ^{1}H NMR resonance suggests a highly networked polymer arising from small monomeric units with numerous different types of intermonomer linkages. Varying self- and intermonomer combinations of these small C6-C3 units result in a highly complex three-dimensional network-type structure. The spectroscopic data, discussed above, obtained on our polydisperse polymer, both sulfated and unsulfated, are consistent with a lignin structure. Finally, the ^{13}C spectra of purified

low molecular weight lignosulfonate obtained from pine wood chip were essentially equivalent **(figure 23)** to that determined for our polymer. In addition to this, the ^{1}H NMR spectra of the brownstock residual lignin corroborates well with the spectra of our polymer **(figure 22)**. ^{13}C NMR spectroscopy has been shown to be of significant potential in providing detailed structural information for lignins. In particular, the advent of multidimensional NMR techniques has extended the prospect of lignin structural analysis considerably. In addition, ^{13}C NMR has been indispensable in the quantitative determination of the amounts of different structural units in lignin. Whereas the broad proton NMR signals that occur over a narrow frequency range render limited quantitative information, ^{13}C NMR spectroscopy provides an elegant alternative, mainly due to its significantly larger chemical shift dispersion.

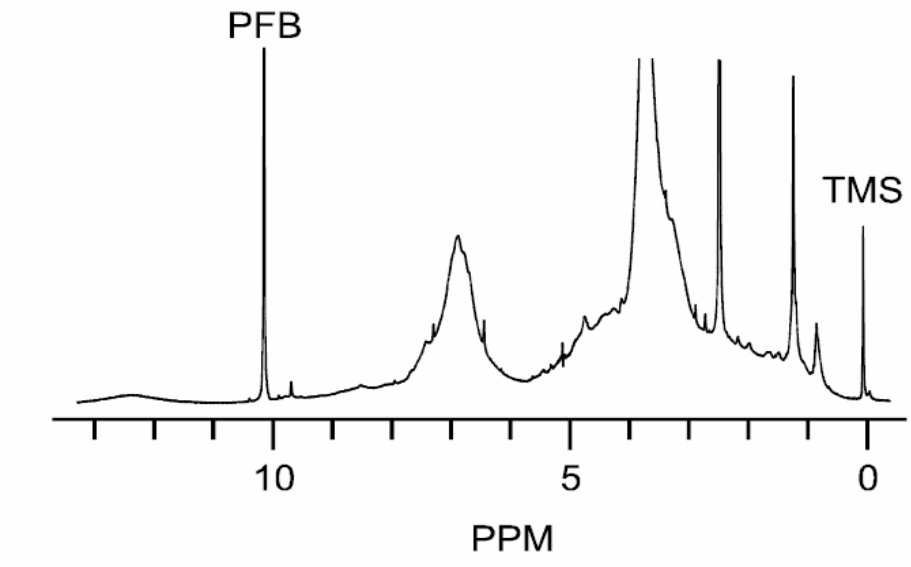


Figure 22. ^{1}H NMR of brownstock residual lignin (Moe, S. T.; Ragauskas, A. J.)[110]

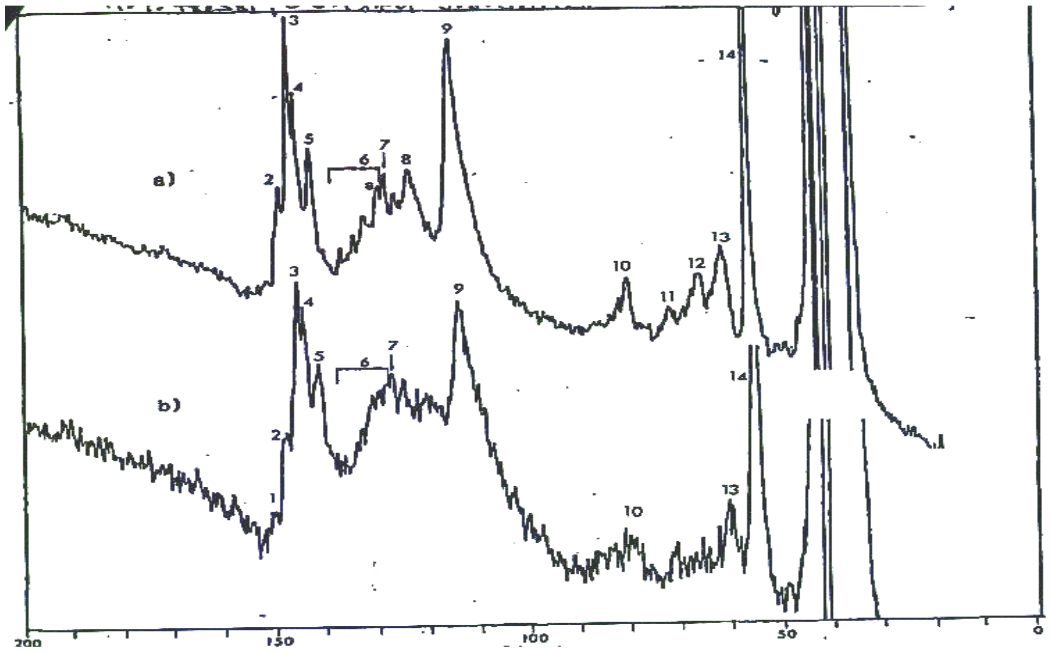


Figure 23. ^{13}C NMR of low molecular weight lignosulfonate obtained from pine wood chip (Hassi et. al)[111]

Conclusions

We were able to isolate the active polymer from morin by using the differential solubility of morin and the polymer in organic solvent like ethyl acetate and acetone. Using chemical and biophysical analysis we were able to conclude that the polymer was a highly networked, heterogeneous, polyphenol based molecule belonging to lignin class of compound, which on sulfation shows the anti viral activity.

2.4 Chemicals

Morin was obtained from Indofine (Somerville, NJ), Fluka (Milwaukee, WI), and Aldrich (Milwaukee, WI). (+)-catechin, (-)-epicatechin, quercetin, were purchased from Aldrich (Milwaukee, WI). *N*,*N*-Dimethyl-*m*-phenylenediamine dihydrochloride, *N*,*N*-dimethyl-*p*-phenylenediamine monohydrochloride, sulfur trioxide-triethylamine complex, and NP-40 were from Sigma-Aldrich (Milwaukee, WI). *β*-Galactosidase substrate and *o*-nitrophenyl *β*-D-galactopyranoside (ONPG) were from Pierce (Rockford, IL). High purity water, obtained from NERL Diagnostics (RI), was used in all experiments.

2.5 Methods

2.5.1 Sulfation of flavonoids[112-114]

All flavonoids (1 g, 3.3 mmoles) were sulfated using triethylamine-sulfur trioxide complex (0.9 g, 4.95 mmoles) at 65 °C in dimethylacetamide (50 mL) for 3h. The reaction mixture was poured into acetone (250 mL) containing 5 mL triethylamine and left undisturbed at 4 ^{o}C for 24 h. The crude oil formed at the bottom was washed with acetone and suspended in 30% (w/v) sodium acetate (20 mL) to exchange the triethylamine cation for sodium. The suspension was added to ethanol (100 mL) to precipitate the sodium salt of sulfated flavonoids.

2.5.2 Cells and Viruses

Dr. Patricia Spear (Northwestern University) provided HeLa cells and the reporter viruses listed here. HSV-1 virus strain carrying the *lac*Z gene of E. *coli* and capable of expressing β-galactosidase as a reporter of entry included HSV-1(KOS) gL86, and HSV-1(KOS)-tk12. The experiments shown here were done with HSV-1(KOS)gL86 alone. Briefly, in this assay, a mutant strain of HSV-1 that contains the *lacZ* gene is used so as to enable a spectrophotometric determination of infection. The method is simpler and less tedious than a plaque formation assay used earlier. This HSV-1 viral infection inhibition assay involves the exposure of a constant dose of virus to HeLa cells, which internalize the virus, in the presence and absence of sulfated inhibitors. Following incubation for 6 h at 37 °C, the internalized viral particles are quantified using the *α*-galactosidase activity of the viral genome, which decreases in a sigmoidal manner **(equation 1)** as the concentration of the sulfated inhibitor increases.

2.5.3 HP-SEC

High performance size exclusion chromatography (HP-SEC) of the sulfation reaction of morin was carried out on a Shimadzu chromatography system composed of LC10Ai pumps and a SPD-10A VP UV-vis detector controlled by a SCL-10A VP system controller connected to a computer. The reaction mixture was analyzed using ASAHIPAK GS 520 HQ (Phenomenex, Torrance, CA, 7.6 mm i.d. × 300 mm). A guard

column (Phenomenex) was used to remove particulates, if present. The mobile phase consisted of acetonitrile-water (7:3 v/v) at the constant flow rate of 0.4 mL/min. Detection was performed at 266nm.

2.5.4 Capillary Electrophoresis

Capillary electrophoresis of morin sulfate, sulfation reaction of morin and unknown sulfated polymer was carried out under reverse polarity conditions using a Beckmann PACE/MDQ unit with an integrated power supply.[115] An uncoated fused silica capillary of 50 μm internal diameter and 32.5 cm effective length to the detector window was used. The electrophoresis run buffer consisted of 100 mM sodium phosphate, pH 2.7, filtered through 0.22 μm filter. The capillary was activated by washing at 20 psi for 2 min each with 0.1 M phosphoric acid, 0.5 M sodium hydroxide and water, followed by rinsing with the electrophoresis run buffer. Samples were typically injected under a pressure of 0.5 psi for 10 sec and detected spectrophotometrically using a 254 nm filter. Electrophoresis was performed at a constant voltage of 5 kV and at 25 °C.

2.5.5 Isolation of polymer

Morin (5 g) was stirred in 400 mL of ethyl acetate at room temperature for 2 h, following which the precipitate so remaining was filtered on Whatman filter paper (No. 1). The precipitate so obtained was dissolved in 70% ethanol – water mixture and the solution

filtered through Amicon centrifugal concentrator with membrane filter 5K. The retentate on the membrane was evaporated to give the parent polymer.

2.5.6 Reversed phase HPLC of unsulfated

The analysis of morin, ethyl acetate or acetone insoluble substance (polymeric contaminant of morin), phloroglucinol adducts, or acid butanol reaction products was performed on a Shimadzu VP system. Typically samples (0.1 mg/mL) were analyzed using YMCTM ODS-AQ S-5 120A (Waters, Milford, MA) 4.6 × 250 mm column in analytical mode. A guard column of the same material was used to remove any extraneous particulate matter. The mobile phase consisted of acetonitrile – water mixture (either 1:1 or 7:3 v/v) at a constant flow rate of 0.1–0.5 mL/min. Detection was performed at 279 nm.

2.5.7 Sulfation of pure morin

Refer earlier method for sulfation of flavonoids **(page 57)**.

2.5.8 Acid-catalyzed reaction of pure morin

Morin, freed from the polymer, was treated with catalytic amount boron trifluoride-diethyl ether complex or hydrochloric acid in DMA at 60 °C. Except for the presence of sulfating agent, these conditions simulate the possible generation of acid during sulfation.

2.5.9 Base catalyzed reaction of pure morin

Morin was heated with KOH in DMA at 60 °C. Except for the presence of sulfating agent, these conditions simulate the possible generation of acid during sulfation.

2.5.10 Sulfation of the polymeric contaminant

The parent polymeric sample (0.5 g) was sulfated using triethylamine-sulfur trioxide complex (3.2 g) at 65 ^{o}C in DMA for 3 h in the presence of molecular sieves. Following the reaction, the mixture was poured in acetone containing 0.5–2 mL triethylamine and the solution left undisturbed at 4 ^{o}C for 24 h. A crude oil formed at the bottom. This oil was collected, washed with chilled acetone, and suspended in 30% sodium acetate. After about 2 h, the suspension was added to 50 mL ethanol to precipitate the sodium salt of the sulfated polymer (0.2 g).

2.5.11 HSV-1 inhibiton activity study

Assays for infection of cells were based on quantitation of β-galactosidase expressed by the mutant HSV-1 viral genome containing the *lacZ* gene. HeLa cells were grown in 96-well tissue culture dishes ((2-4) × 104 cells/well), washed after 16 h of growth, and exposed to 10 plaque forming units (PFU)/cell of HSV-1 virus in 50 μL of phosphate-buffered saline (PBS) containing glucose and 1% calf serum (PBS-G-CS) for 6 h at 37

°C. To test for inhibitory activity, the sulfated compounds were simultaneously added to this 50 μL medium in varying amounts ranging from 0.2 μL to 1.6 ng. Following incubation, the cells were solubilized in 100 μL of PBS containing 0.5% NP-40 and 10 mM ONPG. The initial rate of hydrolysis of the substrate was monitored spectrophotometrically (Spectra MAX 190, Molecular Devices) at 410 nm, which corresponds to the concentration of the β-galactosidase within HeLa K-1 cell. The initial rate of hydrolysis of the substrate in the absence of any added sulfated molecule formed the control and assigned a value of 100% HSV-1 infection. Assays were performed in duplicate and analyzed using equation 1 to obtain IC_{50} values. For comparison of IC_{50} values for the sulfated polymer obtained by two methods, purified either from sulfation reaction of natural product morin or from sulfation of purified polymer from morin, after ascertaining that the electrophoretic profiles are similar.

$$F = F_{MIN} - \frac{F_{MAX} - F_{MIN}}{1 + 10^{(\log[I]_o - \log IC_{50}) \times b}} \qquad \textbf{…… (equation 1)}$$

2.5.12 PAGE

Sulfated polymer obtained by two methods, purified either from sulfation reaction of natural product morin or from sulfation of purified polymer from morin, was analyzed using a protocol commonly for proteins. The electrophoresis run buffer was 100 mM

Tris, and 100 mM boric acid buffer, pH 8.3, containing 2 mM EDTA. Sulfated samples were analyzed using gel concentrations ranging from 6% to 18%. No stacking gel was used. The samples, 30 μg in 4 μL of sulfated polymer fractions, were loaded onto the resolving gel using 40% glycerol. Electrophoresis was performed at a constant current. The gels were stained for 10 min using the high iron diamine stain followed by destaining with water for 24 h. Densitometric analysis was performed using BioRad VersaDoc 4000 ChemImager equipped with a 610 nm filter against a transwhite background. Relative front (*Rf*) of bands, defined as the ratio of band migration distance to the length of gel, were calculated using BioRad Quantity 1 software. The mobility of a band was then calculated by multiplying its *Rf* value by gel length and factoring in the run time.

2.5.13 Elemental Analysis

Elemental analysis of the sulfated and unsulfated polymer fractions was obtained from Atlantic Microlabs (Norcross, GA).

2.5.14 IR

IR was performed on the unsulfated polymer. Samples were prepared by drying it overnight at 50^{o}C. Potassium bromide, mortar pestle and spatula all were dried overnight at 50^{o}C to remove any moisture present. Sample was grounded with potassium bromide (1:50) to form a homogenous powder. Manual press was used to make the pellets. A

blank was recorded for potassium bromide devoid of the sample and the mixture (sample and KBr) was recorded. The blank was subtracted from the sample to for the final spectrum.

2.5.15 UV-Visible

UV was performed on the sulfated and unsulfated polymer samples. Samples were analyzed using Shimadzu UV-1610PC UV-Visible Spectrophotometer. The samples were scanned from λmax 200nm to 600nm. A blank for sulfated and unsulfated in form of water and methanol respectively was used.

2.5.16 ^{1}H NMR

Low molecular weight sulfated polymer sample for NMR analysis was prepared as follows. The filtrate from the 5 kDa molecular weight filter was lyophilized and the solid loaded on a Sephadex G-25 column. Elution with water separated the polymer (MR = 1.9 kDa) from the monomer, morin sulfate. The polymer sample was lyophilized twice from $^{2}H_2O$ to yield the sample for NMR analysis. ^{1}H NMR spectra were recorded on a 500 MHz Oxford-Varian spectrometer in $^{2}H_2O$ at 25, 35 or 50 °C,

2.5.17 ^{13}Carbon NMR

^{13}C NMR spectra were recorded on a 300 MHz Varian-Gemini spectrometer at 60 °C. For ^{13}C NMR analysis nearly 80 mg of the sulfated polymer was dissolved in 0.5 mL of $^{2}H_2O$ and the signal acquired for 20,272 scans with an acquisition time of 1.7 sec and a pulse delay of 2 sec. A line-broadening factor of 50 Hz was used to extract the ^{13}C NMR spectrum of the polymer from noise.

2.5.18 Acid-butanol test

The acid-butanol test for condensed tannins essentially followed the previously developed protocol. Briefly, two reagents were prepared. Reagent A consisted of butanol – HCl mixture (95:5 v/v) and was prepared by mixing 950 mL of n-butanol with 50 mL concentrated HCl. Reagent B consisted of 2% ferric ammonium sulfate in 2 N HCl, which was prepared by diluting 16.6 mL of conc. HCl to 100 mL with distilled water and then dissolving in it 2g ferric ammonium sulfate. A small amount of the polymeric sample (1-5 mg) was dissolved in 0.5 mL aqueous acetone (70%) so that the absorbance at 550 nm is less than 0.6. To this solution was added 3.0 mL of reagent A and 0.1 mL of reagent B. The tube was vigorously shaken on a vortexer, tightly covered and heated at 100 ^{o}C for 1 h. The absorbance of the tube at 550 nm was then recorded (If the extract has condensed tannins, a pink color develops). The tannin content is directly correlated to the absorbance at 550 nm.

2.5.19 Phloroglucinol test

The phloroglucinol test for condensed tannins followed the previously developed protocol.[116] Briefly, a solution of 0.1 N HCl in methanol containing 50 g/L phloroglucinol and 10 g/L ascorbic acid was prepared. To this solution (20 mL), the ethyl acetate or acetone insoluble substance (100 mg) was added and heated at 50 ^{o}C for 20 min. The final solution was brought to room temperature and combined with 100 mL of 40 mM aqueous sodium acetate. The phloroglucinol adducts, if any, were analyzed by RP-HPLC as described earlier.

2.5.20 Test for lignin-like/tannin-like compounds

Test for lignin/tannin was performed using LaMotte Octet Comparator kit. Test tube (0230) was filled to 5 mL line with sample dissolved in methanol. To this 3 drops of Reagent #1 (7833) was added capped and mixed. Using a 1 mL pipette (0354) from the kit, 1 mL of Reagent #2 (7834) was added to the sample solution capped and mixed. The solution mixture was allowed to stand for 30 minutes. Finally the sample colors were matched to the standard color. A blank was also run simultaneously.

CHAPTER 3

LIQUID CHROMATOGRAPHY-MASS SPECTROMETRIC (LC-MS) CHARARTERIZATION OF LIGNIN

3.1 Introduction

3.1.1 General Principles

The four basic parts of any mass spectrometer are: a sample introduction device; an ionization source; a mass analyzer; and an ion detector. Combining these components a mass spectrometer determines the molecular weight of chemical compounds by ionizing, separating, and measuring molecular ions according to their mass-to-charge ratio (m/z). The ions are generated in the ionization source by inducing either the loss or the gain of a charge (e.g. electron ejection, protonation, or deprotonation). The ions thus formed in the gas phase can be electrostatically directed into a mass analyzer, where they would be separated according to mass and finally detected. The result of ionization, ion separation, and detection is a mass spectrum that can provide molecular weight or even structural information. Each of these mass spectrometry components exists in many forms and is effectively combined to produce a wide variety of mass spectrometers with specialized characteristics.

3.1.2 Components of Mass Spectrometer[117-122]

3.1.2.1 Sample Introduction

The sample inlet is the interface between the sample and the mass spectrometer. The sample can be introduced directly into the ionization source, or can be inserted via some

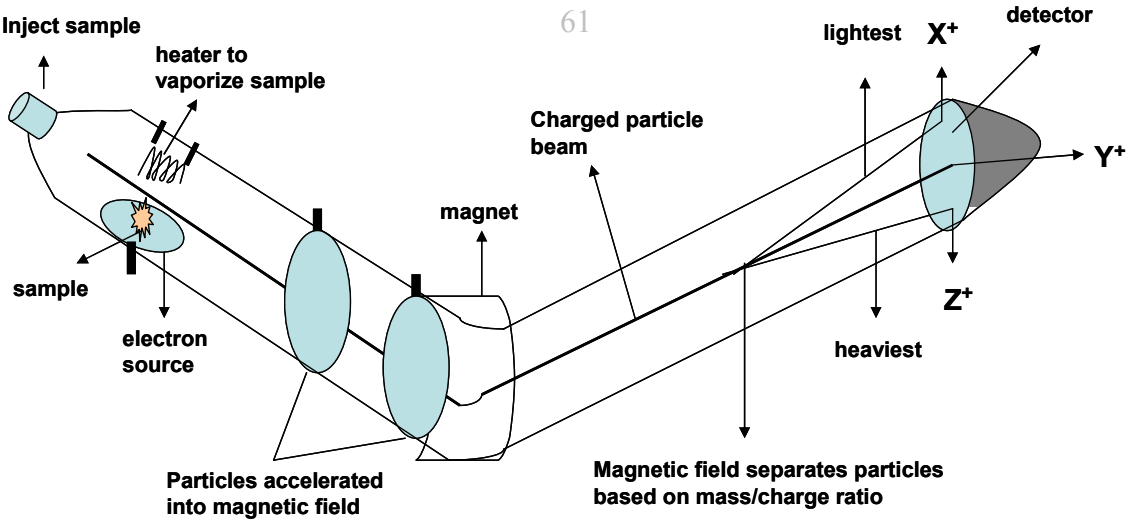


Figure 24. Basic principle of working of mass spectrometer

type of chromatography route to the ionization source. One approach to introduce a sample is by placing the sample on a probe, which is then inserted, usually through a vacuum lock, into the ionization region of the mass spectrometer. The sample can then be heated to assist thermal desorption or undergo any number of high-energy desorption processes used to achieve vaporization and ionization. Capillary infusion is often used

because it can efficiently introduce small quantities of a sample into a mass spectrometer without destroying the vacuum.

The other method of sample introduction usually involves the mass spectrometer being coupled directly to a high pressure liquid chromatography (HPLC), gas chromatography (GC) or capillary electrophoresis (CE) separation column, and hence the sample is separated into a series of components which then enter the mass spectrometer sequentially for individual analysis.

3.1.2.2 IONIZATION

Ionization is the act of placing a charge on a neutral molecule.

Some of the common ionization techniques include the following:

Electron Impact	(EI)
Chemical Ionization	(CI)
Fast Atom Bombardment	(FAB)
Atmospheric Pressure Ionization	(API)
Electrospray Ionization	(ESI)
Atmospheric Pressure Chemical Ionization	(APCI)
Laser Ionization	(LIMS)
Resonance Ionization	(RIMS)
Matrix Assisted Laser Desorption Ionization	(MALDI)
Collision Induced Dissociation	(CID)

Electron Ionization (Electron Bombardment and Electron Impact)

In this technique the sample must be delivered as a gas, which is usually accomplished by heating the sample to vaporize it off of the probe. Once in the gas phase, the compound passes into an electron ionization region where it interacts with a beam of electrons of nearly homogeneous energy (70 electron volts), typically causing electron ejection and some degree of fragmentation. Electron ionization is most useful for compounds below a

molecular weight of 400 Daltons since larger molecules tend to thermally degrade during vaporization. The principal problems associated with it include excessive fragmentation in the ionization source, the non-volatility of large molecules and thermal decomposition during vaporization.

Chemical Ionization (CI)

CI is applied to samples similar to those analyzed by EI and is primarily used to enhance the abundance of the molecular ion. CI uses ion-molecule reactions in the gas phase to produce ions from the sample molecule. The CI process is initiated with a reagent gas such as methane, isobutane, or ammonia, which is initially ionized by EI. A high gas pressure in the ionization source results in the propagation of ion-molecule reactions between the reagent gas ions and reagent gas neutrals, some of these ions can react with the analyte molecules to produce ions.

Fast Atom Bombardment (FAB)

The FAB ionization technique is a soft ionization method that typically requires the use of a direct insertion probe for sample introduction and a high-energy beam of xenon atoms or cesium ions to sputter the sample and matrix from the probe's surface. The matrix, such as m-nitrobenzyl alcohol, is used to dissolve the sample and facilitate desorption as well as ionization. The FAB matrix is a nonvolatile liquid material that

serves to continuously replenish the surface with new sample as the incident ion beam bombards this surface.

Atmospheric Pressure Ionization (API)

API is used in conjunction with LC/MS techniques. The ions are formed at atmospheric pressure. There are two common types of atmospheric pressure ionization: ESI and APCI.

Electrospray Ionization (ESI)

In an electrospray interface, a high potential (3.5kV) is applied between a solvent (usually an aqueous or aqueous/organic) emerging from a capillary needle and a counter electrode resulting in a large charged droplet produced by 'pneumatic nebulization'; i.e. the forcing of the analyte solution through a needle. The potential applied is sufficiently high to disperse the emerging solution into a very fine spray of charged droplets all at the same polarity. During the flight from the point of generation towards the ion sampling orifice, volatile and neutral solvent molecules evaporate from the droplets. This process is stimulated by a counter current gas (e.g. nitrogen); which shrinks the droplet size and increases the electric charge density at the droplet's surface. Eventually, at the Rayleigh limit, Coulombic repulsion overcomes the droplet's surface tension and ions begin to leave the droplet through what is known as a “Taylor cone”. This 'Coulombic explosion'

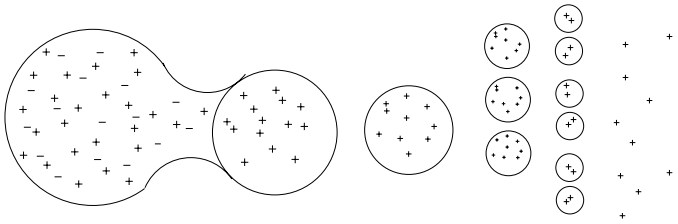


Figure 25. A Taylor cone

forms a series of smaller, lower charged droplets. The process of shrinking followed by explosion is repeated until individually charged 'naked' analyte ions are formed. The charges are statistically distributed amongst the analyte's available charge sites, leading to the possible formation of multiply charged ions. Increasing the rate of solvent evaporation, by introducing a drying gas flow counter current to the sprayed ions, increases the extent of production of multiple charged species. Decreasing the capillary diameter and lowering the analyte solution flow rate i.e. in nanospray ionization, will create ions with higher $^{m}/_{z}$ ratios (i.e. it is a softer ionization technique) than those produced by 'conventional' ESI and are of much more use in the field of bioanalysis. The ionized droplet can be detected as either a positive or negative ion (depending on the chemistry of the solvent and its modifiers). Only volatile modifiers and buffers can be used in this configuration. Complex mixtures can be separated using a HPLC. The sensitivity is dependent on the target compound, but usually small amounts (e.g. in the nanomole to femtomole range) can be detected. The number of charges retained by an

analyte can depend on such factors as the composition and pH of the electrosprayed solvent as well as the chemical nature of the sample. For small molecules (< 2000 Daltons) ESI typically generates singly or doubly charged ions, while for large molecules (> 2000 Daltons) the ESI process typically gives rise to a series of multiply charged species. Because mass spectrometers measure the mass-to-charge (m/z) ratio, the resultant ESI mass spectrum contains multiple peaks corresponding to the different charged states.

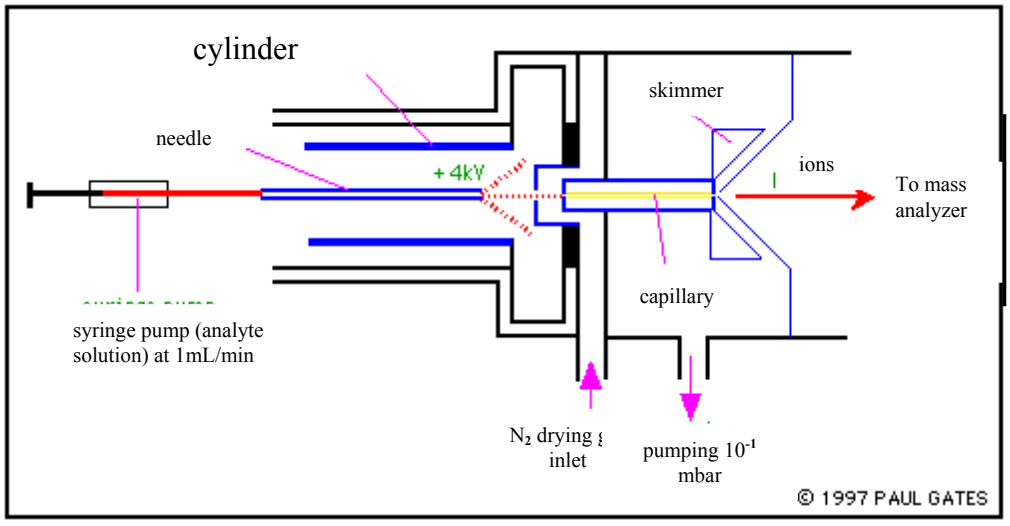


Figure 26. A schematic diagram of the typical layout of an electrospray source.

Different manufacturers use variations on this layout, and many manuals and research papers refer to parts of the source with different names. This layout is based on the Analytical design used with the Bruker BioApex FT-ICR mass spectrometer.

ESI characteristics in summary

1. Soft ionization method provides molecular weight information.
2. Suitable for analyzing large bio- or synthetic polymers.
3. Sensitivity depends strongly upon the analyte.
4. Suitable for analyzing polar and even ionic compounds (e.g. metal complexes).
5. Less fragmentation.
6. Enables LC / MS and CE (Capillary Electrophoresis) coupling.

Atmospheric Pressure Chemical Ionization (APcI)

Atmospheric pressure chemical ionization is a relative of ESI. The ion source is similar to the ESI ion source. In addition to the electrohydrodynamic spraying process, a corona-discharge needle at the end of the metal capillary creates plasma. In this plasma proton transfer reactions and to a small amount fragmentation can occur. Depending on the solvents, only quasi-molecular ions like $[M+H]^+$, $[M+Na]^+$ and M^+ (In the case of aromatics), and/or fragments can be produced. Multiply charged molecules $[M+nH]^{n+}$, as in ESI, are typically not observed.

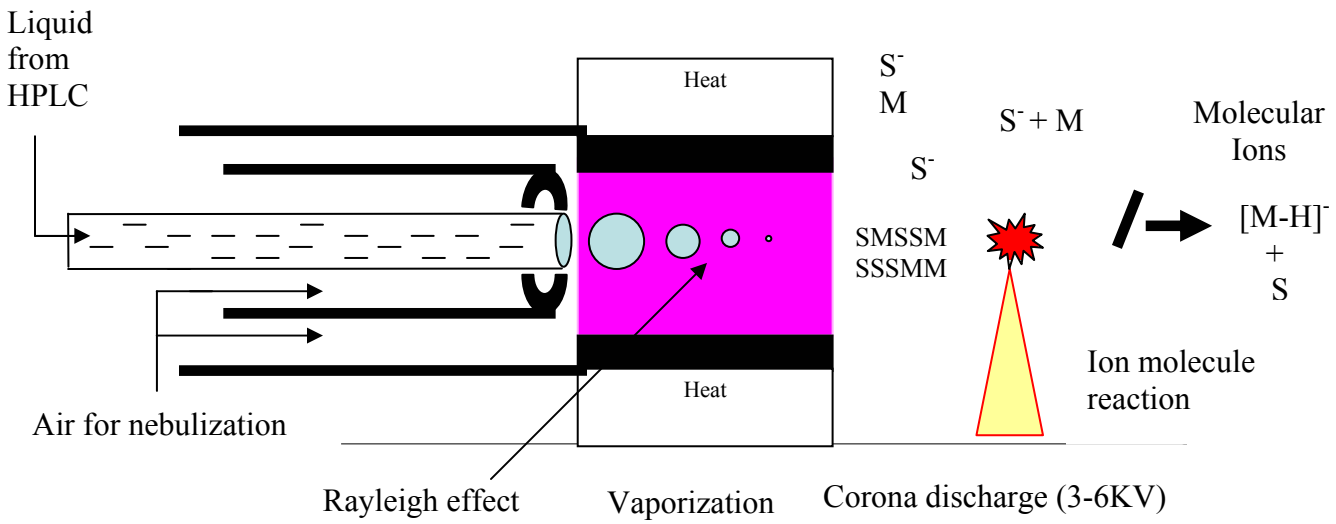


Figure 27. Working of APCI in negative ion mode

In Atmospheric Pressure Chemical Ionization the target compound is delivered in a liquid stream by an HPLC pump, which enters the heated region of an APcI probe. This facilitates rapid desolvation/vaporization of the droplets. Vaporized sample molecules are carried through an ion-molecule reaction region at atmospheric pressure. The ionization occurs through a corona discharge, creating reagent ions from the solvent vapor. Chemical ionization of sample molecules is very efficient at atmospheric pressure due to the high collision frequency. Proton transfer (protonation MH^+ reactions) occurs in the positive mode, and either electron transfer or proton transfer (proton loss, $[M-H]^-$) in the negative mode. The moderating influence of the solvent clusters on the reagent ions, and of the high gas pressure, reduces fragmentation during ionization and results in primarily molecular ions. APCI is widely used in the pharmaceutical industry to analyze relatively

non-polar, semi volatile samples of less than 1200 Daltons and it is an especially good ionization source for liquid chromatography. Complex mixtures can be separated using HPLC. The sensitivity is dependent on the target compound but usually small quantities of the target compound can be detected. Solvent adducts or radical cations can be observed.

The following conditions are required for APCI to work:

1) the analyte must be volatile and thermally stable
2) the mobile phase must be suitable for gas phase acid-base reactions
3) for working in positive mode, the proton affinity of the analyte must be higher than the proton affinity of the eluent (i.e. the analyte can take a proton from the protonated solvent).
4) for working in negative mode, the gas phase acidity of the analyte must be lower than the gas phase acidity of the eluent (i.e. the analyte can give a proton to the deprotonated solvent)

Choosing Between Ionization Techniques

General rules can be given to choose between ESI, APCI, or EI techniques.

- ESI is preferred for compounds, which are ionic or very polar or thermo labile, or with masses higher than 1000
- APCI is preferred for compounds, which are not very polar.

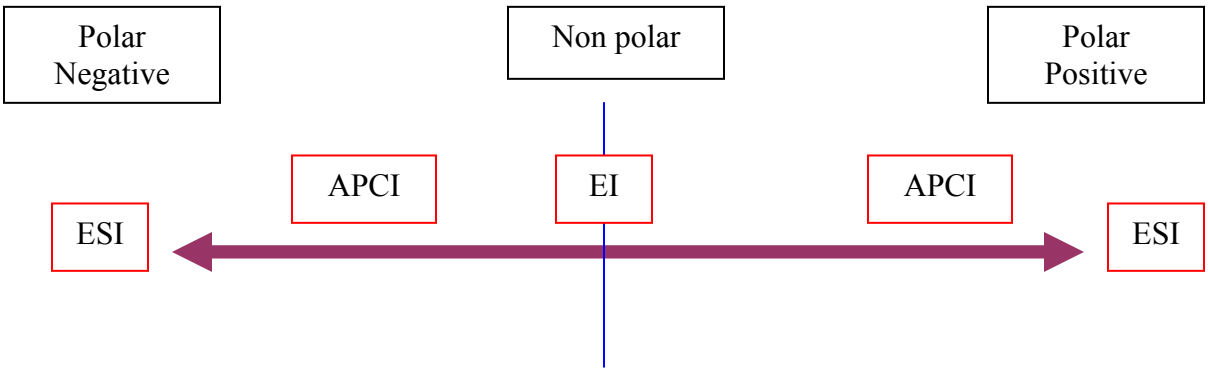


Figure 28. Choosing between the MS techniques

Practical considerations are also important; changing from ESI to APCI cannot be done automatically during the separation. The instrument itself has an influence. Depending on the supplier, some instruments might show better performance in APCI, while other ones will perform better in ESI. So, when it comes to practice, and when the choice is not obvious, the best approach for determining the most suitable ionization method is still to try to inject in various conditions, or to refer to existing work.

APCI characteristics in summary

1. Provides molecular weight information.
2. Sensitivity depends strongly upon the analyte.
3. Suitable for analyzing less polar compounds compared to ESI.
4. Increased fragmentation compared to ESI.
5. Enables coupling MS and LC with flow rate up to 2 ml/min.

Laser Ionization (LIMS)

A laser pulse ionizes some of the sample constituents. There are a number of LIMS techniques, e.g. RIMS or MALDI, of which MALDI is more commonly used.

Resonance Ionization (RIMS)

One or more laser beams (photons) are tuned in resonance to transitions of a gas-phase atom or molecule to promote it in a stepwise manner above its ionization potential to create an ion. Since each element has a unique energy level structure, RIMS provides a selective ionization method.

Matrix Assisted Laser Desorption Ionisation (MALDI)

MALDI is a LIMS method of vaporizing and ionizing large biological molecules, such as proteins or DNA fragments. The biological molecules are embedded in a chemical matrix such as nicotinic acid. A laser light energy ablates the matrix, which carries some of the

large molecules into the gas phase in an ionized form so they can be extracted into a mass spectrometer. MALDI allows determining the molecular weight of up to 500 kDa, routinely 5 to 100 kDa molecules (polymers, biomolecules, complexes, enzymes) are analyzed.

MALDI characteristics in summary

1. Soft ionization method provides molecular weight information.
2. Suitable for analyzing very large bio- or synthetic polymers.
3. Sensitivity depends strongly upon the analyte.
4. Suitable for analyzing polar and even ionic compounds (e.g. metal complexes).
5. Less fragmentation.
6. Pulsed ionization technique, in contrast to EI, CI, FAB, ESI, and APCI.

Collision Induced Dissociation (CID)

MS employs CID to fragment a precursor ion. In this the first quadrupole mass analyzer generates a spectrum of ions. The ion(s) of interest are then selectively sent into the second quadrupole also called as CID chamber containing an inert gas such as argon which is allowed to collide with the selected ion(s). This induces fragmentation and the third quadrupole selects from the product ions that is/are to be sent to the detector. In a

single quadrupole the collision induced dissociation occurs between the sample cone and the extractor cone.

3.1.2.3 Mass Analyzers

Immediately following ionization, gas phase ions enter a region of the mass spectrometer known as the mass analyzer. The mass analyzer is used to separate ions within a selected range of mass-to-charge (m/z) ratios. The analyzer is an important part of the instrument because of the role it plays in the instrument's accuracy and mass range. Ions are typically separated by magnetic fields, electric fields, or by measuring the time it takes an ion to travel a fixed distance. Some of the common mass analyzers include double focus magnetic mass filter quadrupole ion trap and time of flight.

Double Focusing Magnetic Mass Analyzer

In magnetic analysis, the ions are accelerated (using an electric field) and are passed into a magnetic field. A charged particle traveling at high speed through a magnetic field will experience a force, and travel in a circular motion with a radius depending upon the m/z and speed of the ion. A magnetic analyzer separates ions according to their radii of curvature, and therefore only ions of a given m/z will be able to reach a point detector at

any given magnetic field. A primary limitation of typical magnetic analyzers is their relatively low resolution.

Quadrupole Mass Filter

Quadrupoles are four precisely parallel rods with a direct current (DC) voltage and a superimposed radio-frequency (RF) potential. And by scanning a pre-selected radio-frequency field one effectively scans a mass range. Quadrupole mass analyzers in conjunction with electron ionization sources are the most common mass spectrometers in existence today. Quadrupoles have three primary advantages. First, they are tolerant of relatively poor vacuums (~5 x 10^{-5} torr), which make them well suited to electrospray ionization since the ions are produced under atmospheric pressure conditions. Secondly, quadrupoles are now capable of routinely analyzing up to m/z of 3000, which is useful because electrospray ionization of biomolecules commonly produces a charge distribution below m/z 3000. Finally, the relatively low cost of quadrupole mass spectrometers makes them attractive as electrospray analyzers.

Quadrupole Ion Trap (Ion Storage)

The physics behind both the quadrupole mass analyzers and quadrupole ion trap is very similar. In an ion trap, the ions are trapped in a radio frequency quadrupole field. The ions are then ejected and detected as the radio frequency field is scanned. It is also possible to isolate one ion species by ejecting all others from the trap. The isolated ions can subsequently be fragmented by collision induction and detected to generate a fragmentation spectrum.

Time of Flight

A time-of-flight (TOF) analyzer is one of the simplest mass analyzing devices and is commonly used with MALDI. Time-of-flight analysis is based on accelerating a set of ions to a detector with the same amount of energy. Because the ions have the same energy, yet a different mass, the ions reach the detector at different times. The smaller ions reach the detector first because of their greater velocity and the larger ions take longer, thus the analyzer is called time-of-flight and the mass is determined at the ions' time of arrival. The arrival time of an ion at the detector is dependent upon the mass, charge, and kinetic energy of the ion. Since kinetic energy (KE) is equal to $1/2mv^2$ or velocity $v = (2KE/m)^{1/2}$, ions will travel a given distance, *d*, within a time, *t*, where *t* is dependent upon their m/z.

3.1.2.4 Ion Detector

Once the ion passes through the mass analyzer it is then detected by the ion detector, the final element of the mass spectrometer. The detector allows a mass spectrometer to generate a signal current from incident ions by generating secondary electrons, which are further amplified. Alternatively, some detectors operate by inducing a current generated by a moving charge. Among the detectors described, the electron multiplier and scintillation counter are the most commonly used and convert the kinetic energy of incident ions into a cascade of secondary electrons. Some of the common detectors include Faraday cup, an electron multiplier and a photomultiplier conversion dynode.

3.2 Structural Characterization of Lignin

Lignin is a heterogeneous, highly cross-linked, extremely complex three-dimensional macromolecule that is particularly resistant to biological degradation. Lignification is the process by which units are linked together via radical coupling reactions. The main "end-wise" reaction couples a new monomer (usually a monolignol and usually at its β position) to the growing polymer, giving rise to structures **A** and **B (figure 30)** (both of which are β-linked). Coupling between preformed lignin oligomers results in units linked 5–5 **D** and 5–O–4 **E (figure 30)**. The coupling of two monolignols is a minor event, with resinol (β–β) units **C** or cinnamyl alcohol end groups **M** as the outcome **(figure 29).**[1]

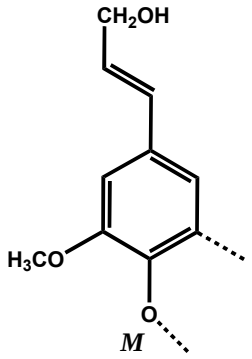


Figure 29. Cinnamyl alcohol end group containing lignin unit

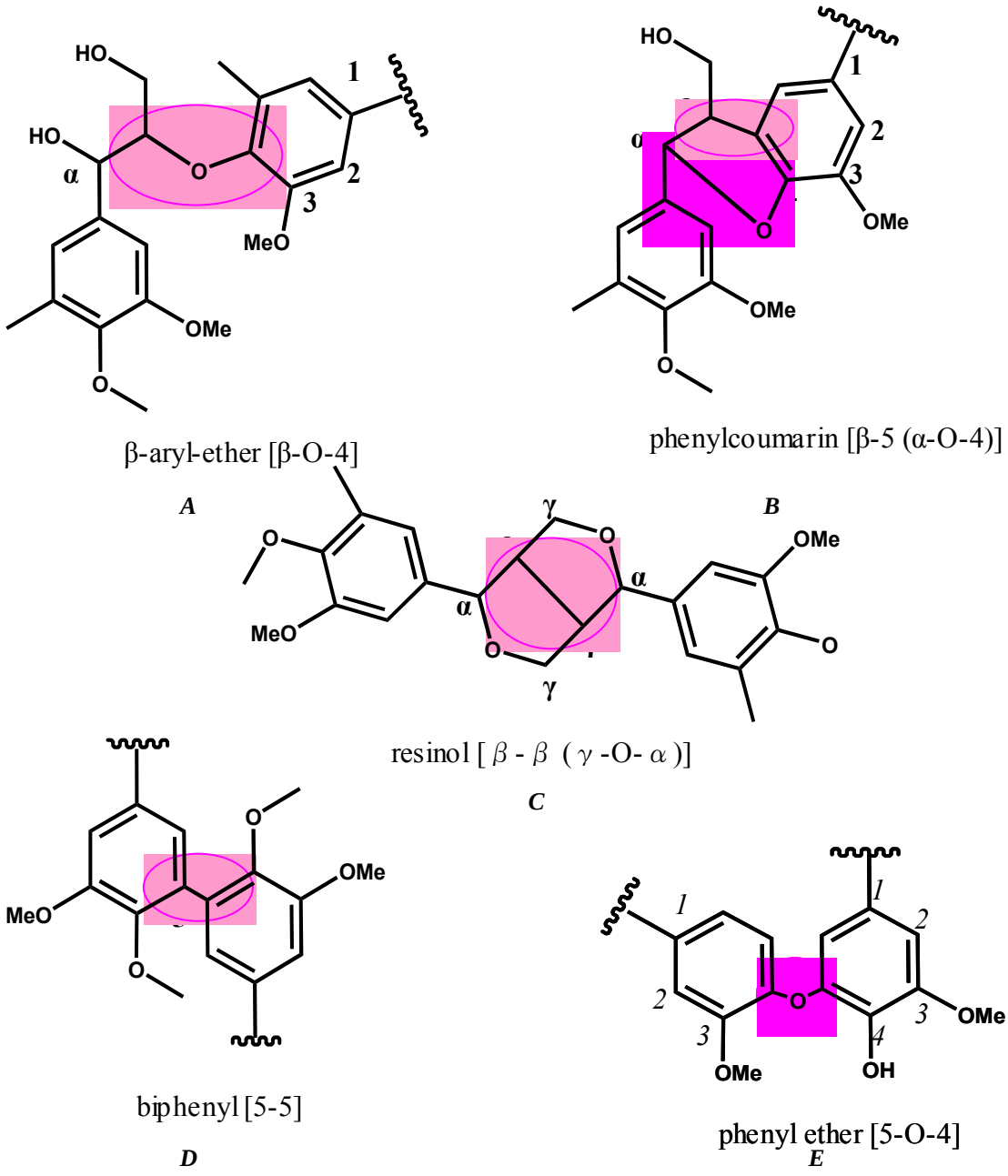
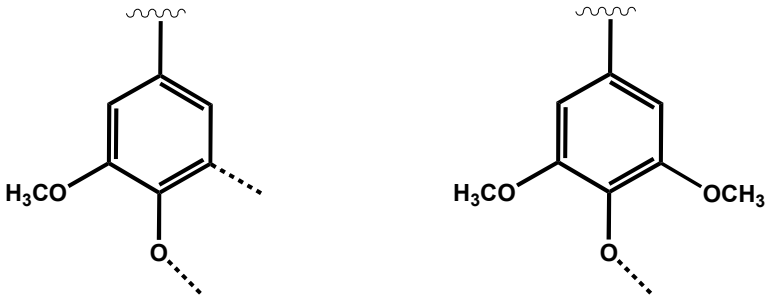


Figure 30. Most common type of linkages found in lignins

The lignins from woody gymnosperms are mainly coniferyl alcohol derived together with small amounts of *p*-coumaryl alcohol, whereas those from woody angiosperms mostly contain coniferyl and sinapyl alcohols as well as lower levels of *p*-coumaryl alcohol.

The most frequent inter-unit linkage is the β–O–4 or 8-O-4 (β-aryl ether) linkage **A**. In lignin, 8-O-4′ aryl alkyl-ether linkages are most abundant and account for approximately half of all interunit linkages. It is also the one most easily cleaved chemically. The other linkages are β–5, β–β, 5–5 and 5–O–4, which are all more resistant to chemical degradation **(figure 30)**. [2]

The relative abundance of the different linkages depends largely on the relative contribution of a particular monomer to the polymerization process. For example, lignins composed mainly of **G** units, such as conifer lignins, contain more resistant (β–5, 5–5, and 5–O–4, **see figure 30**) linkages than lignins incorporating **S** units because of the availability of the C5 position for coupling **(figure 31)**.



G ***S***

Figure 31. General guaiacyl unit (G) and general syringyl unit (S) found in lignins

There are numerous methods that have been described in literature for chemical degradation of lignin followed by its structural characterization. A variety of chemical degradation methods have been developed over the years and new methods continue to emerge. No single chemical degradation method of analysis yields complete information on the content and composition of lignins in plant samples because some types of bonds are very difficult to break. Being ostensibly three-dimensional structural biopolymers, they cannot readily be depolymerized into smaller units.

For analyzing lignin macromolecules, studies with mass spectrometric techniques such as, matrix-assisted laser desorption ionization (MALDI), electronspray ionization (ESI), atmospheric pressure chemical ionization (APCI) and ToF-SIMS analysis have been reported. The polymerization mechanism is not fully understood and the structural complexity and morphological heterogeneity of the lignin polymer makes it difficult to deduce its structure, as it exists *in vivo*. Basic building blocks of the lignin polymer have been used to deduce the structure of lignin.[3-5] Complete structural characterization has not been achieved because the entire lignin cannot be degraded to monomeric fragments by chemical degradation methods, the tree-to-tree variation in the syringyl/guaiacyl ratio and the presence of other polyphenols, like tannins, complicates the process.

ESI-MS has been successfully applied to the structural analysis of lignin. Evtuguin *et al.* have used tandem mass spectrometry for structural elucidation of lignin oligomers and the structural information has been deduced based on the fragmentation patterns provided by experiments on the dimeric model compounds.[6] The data clearly shows a significant abundance of β-O-4 linked lignin monomers, syringyl/guaiacyl units. They have also identified structures with β-β, β-5 and γ-O-α linkages. The highest degree of polymerization that was assigned was a tetramer with highest molecular mass of 839 **(figure 32)**. Lignin macromolecule fragmentation occurs predominantly through the disruption of the ether linkages. Banoub *et al.* have used combined techniques such as atmospheric pressure chemical ionization mass spectrometry (APCI-MS), tandem mass

spectrometry (MS/MS) and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOFMS)[7]. APCI-MS both in positive and negative ion modes showed good results, but failed to identify higher polymeric lignin fragment masses. The highest degree of polymerization that was identified was a trimer with highest molecular mass being 507 **(figure 33a & b)**. MS/MS gave similar results as APCI-MS. MALDI-TOFMS in the positive ion mode detected peaks with highest molecular weight 860 tentatively assigned as a pentamer of lignin fragment. The linkages that were identified were α-O-4, and β-5 connecting guaiacyl moieties. Saito *et al.* have shown applicability of ToF-SIMS to the structural analysis of lignin by detecting the basic building blocks of lignin **(figure 34)** and dimeric models for depolymerized fragments of lignin **(figure 35)**.[8,9] Basic building units, the monomers, guaiacyl and syringyl units have been identified with highest molecular weight being 181. Camarero *et al.* have used pyrolysis-GC/MS for identification of monomeric lignin units, *p*-hydroxyphenyl, guaiacyl and syringyl. The highest molecular weight of the monomers identified was 194. Erik *et al.* successfully put in use pyrolysis-LC-MS for identification of lignin oligomers. The types of linkages identified in lignin were β-O-4, β-5, α-O-4 and 5-5 with considerable abundance of *p*-hydroxyphenyl, guaiacyl and syringyl units. The highest degree of polymerization that was identified was a trimer with highest molecular being 418.

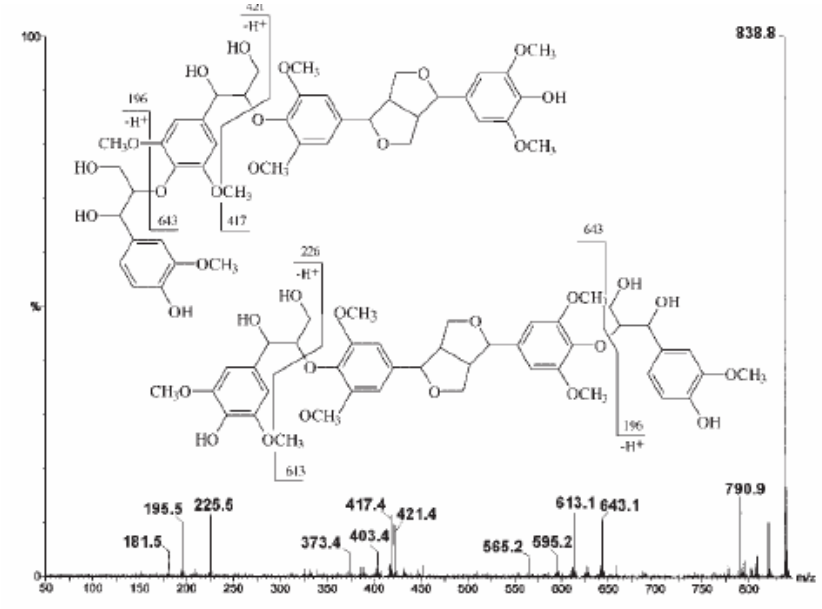


Figure 32. MS/MS spectra of lignin oligomers with m/z 839 and correspondent inferred structures

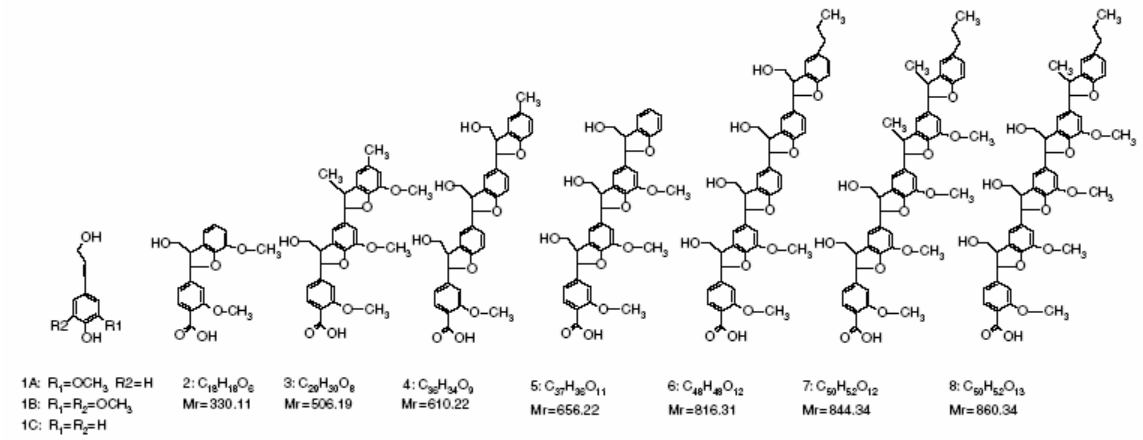


Figure 33a. Chemical structures of the various wheat straw lignin polymeric fragments.

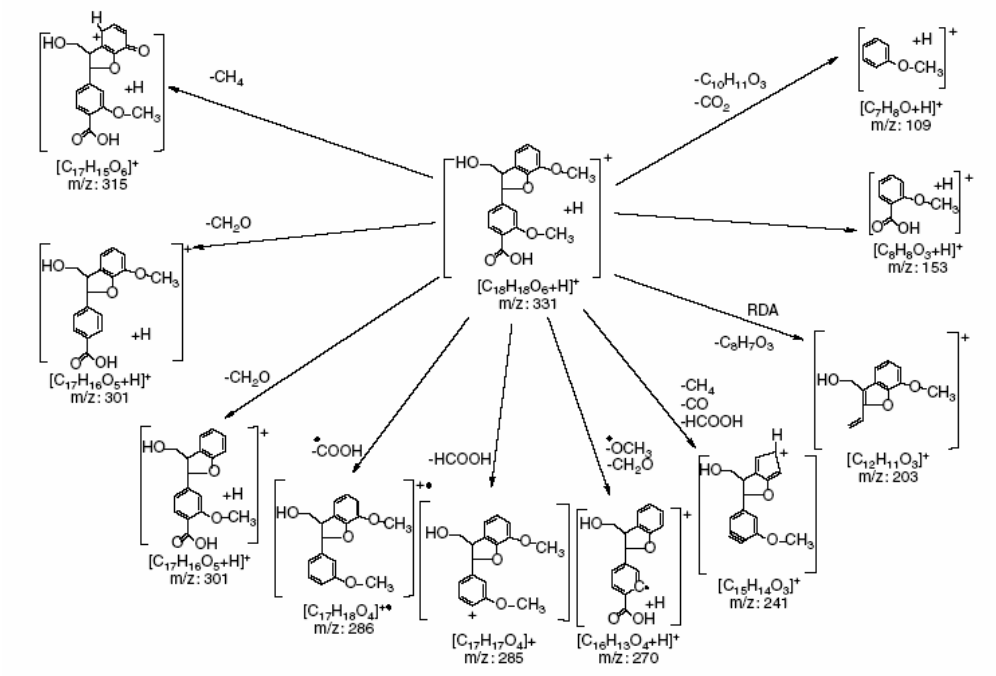


Figure 33b. Proposed fragment routes of the product ion tandem mass spectrum of the protonated molecular ion $[M + H]^+$ at m/z 331.

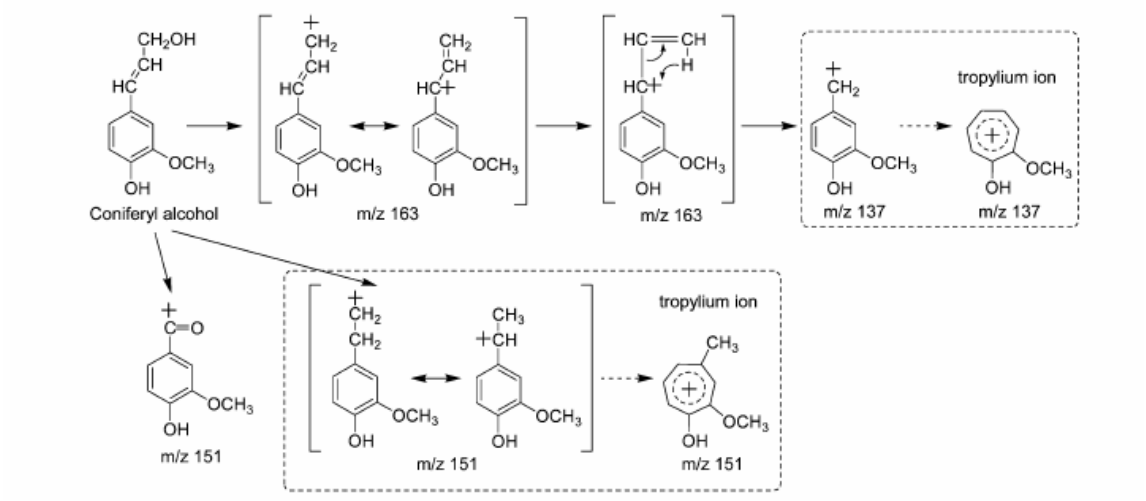


Figure 34. Proposed secondary ion formation in the positive ToF-SIMS spectrum of coniferyl alcohol.

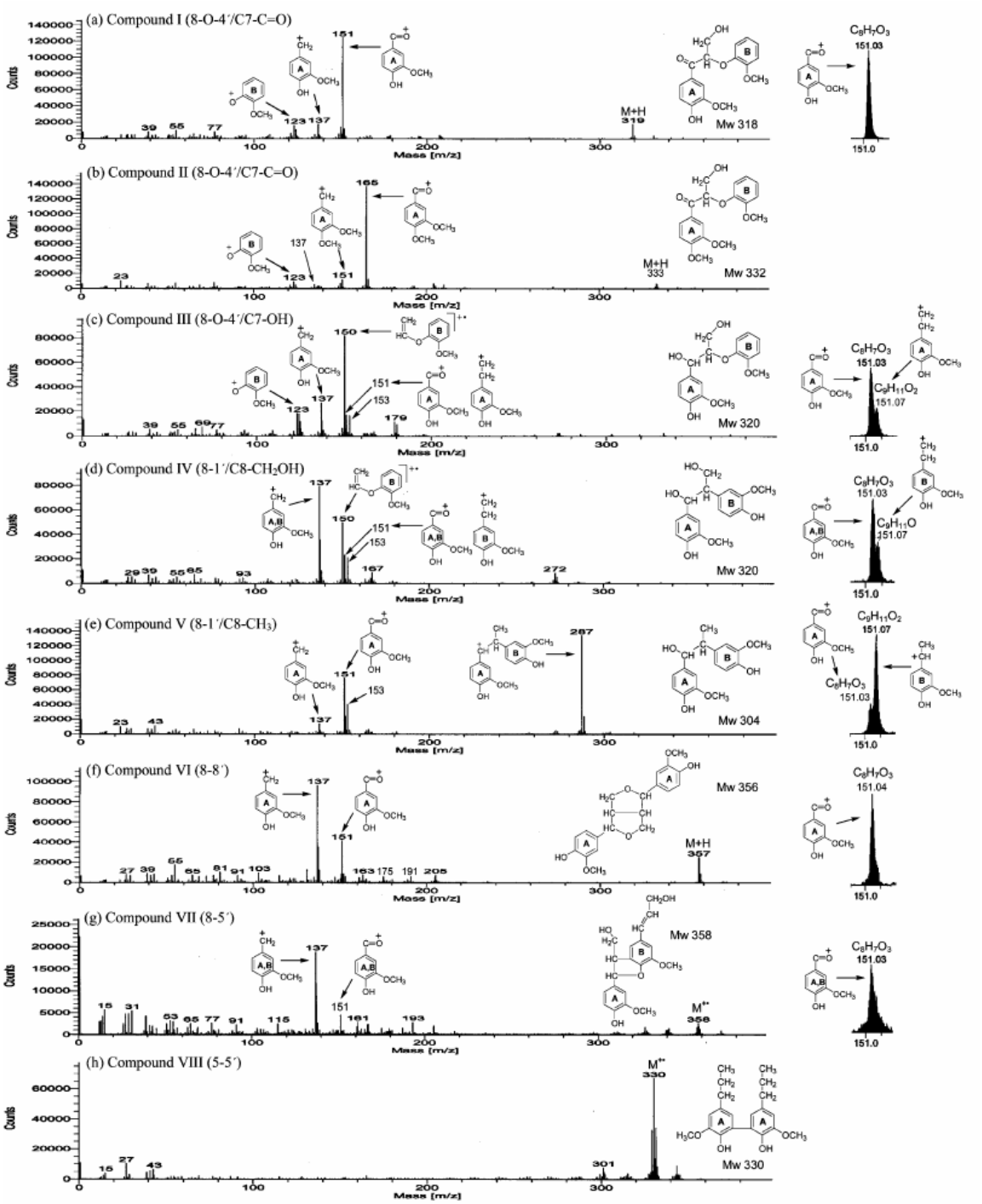


Figure 35a. Positive ToF-SIMS spectra of lignin dimer models.

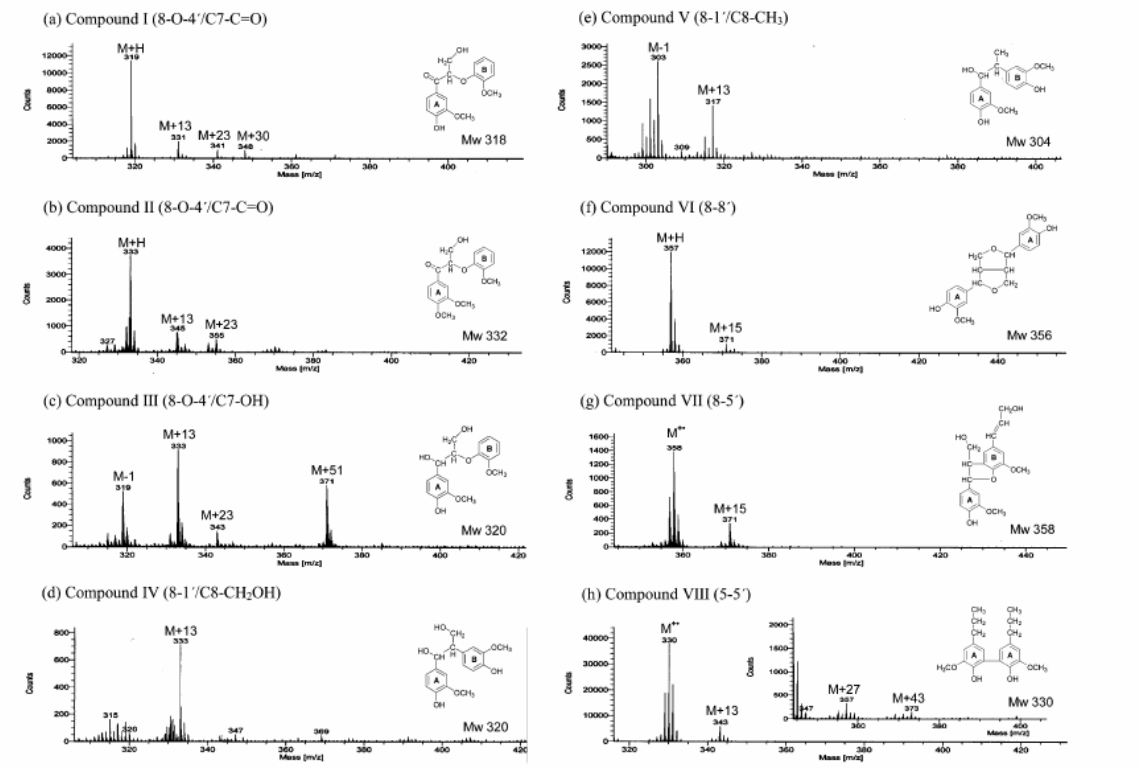


Figure 35b: Adduct ions observed in positive ToF-SIMS spectra of lignin dimer models.

Table 5: Literature review of lignin characterization using different mass spectroscopy techniques

	Technique	Degree of polymerization	Linkages Identified	Type of monomers
1)	ESI-MS –ve	Tetramer	β-O-4, β-β, β-5, γ-O-α	S and G
2)	ESI-MS –ve	Trimer	β-O-4, β-β, β-5, 4-O-5	S, G and H
3)	APCI –ve/+ve	Trimer	α-O-4, β-5	G
4)	MALDI-TOFMS +ve	Pentamer	α-O-4, β-5	G
5)	ToF-SIMS +ve	Monomer		G, S
6)	Py-GC-MS	Monomer		G, S and H
7)	Py-LC-MS	Trimer	β-O-4, α-O-4, β-5, 5-5	G, S and H

S = syringyl, G = guaiacyl and H = *p*-hydroxyphenyl

3.3 Results

Reversed-phase HPLC-MS of the degraded products of lignin

The typical total ion chromatogram (TIC) of the depolymerized lignin products showed at least 10 different analyzable peaks (figure 37). Most of these peaks were characterized based on the fragmentation pattern in the MS. The base line in the entire chromatogram has moved upwards. The chromatogram with UV diode array detection at 279nm is comparable to that with the total ion chromatogram with mass detection.

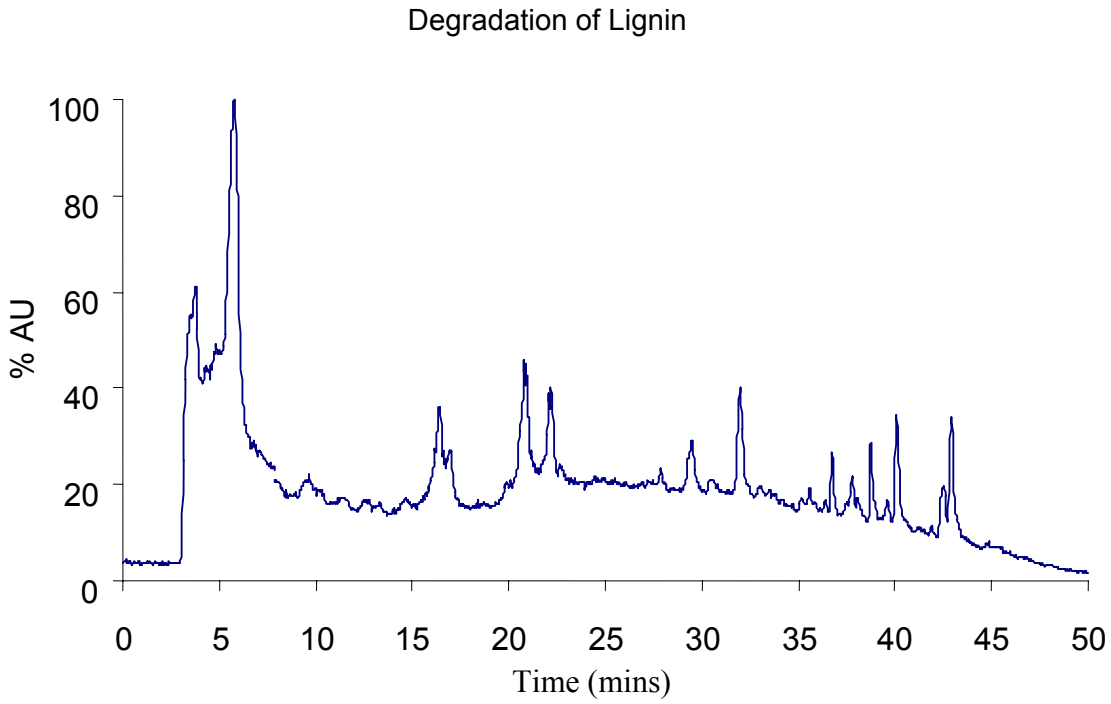


Figure 36. Chromatogram with MS detection of the degraded lignin showing numerous peaks

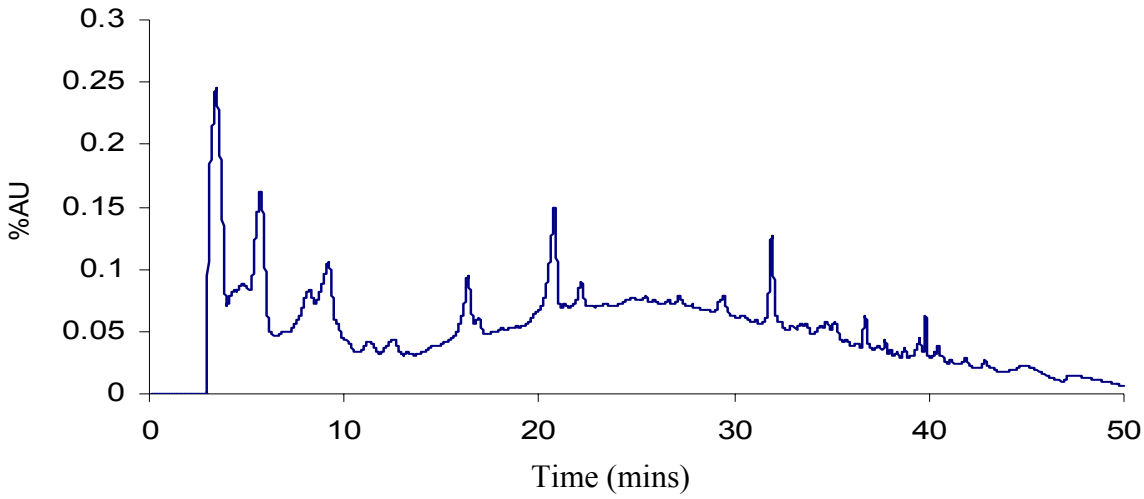


Figure 37. Chromatogram with UV- diode array detection at 279nm of the degraded lignin showing numerous peaks

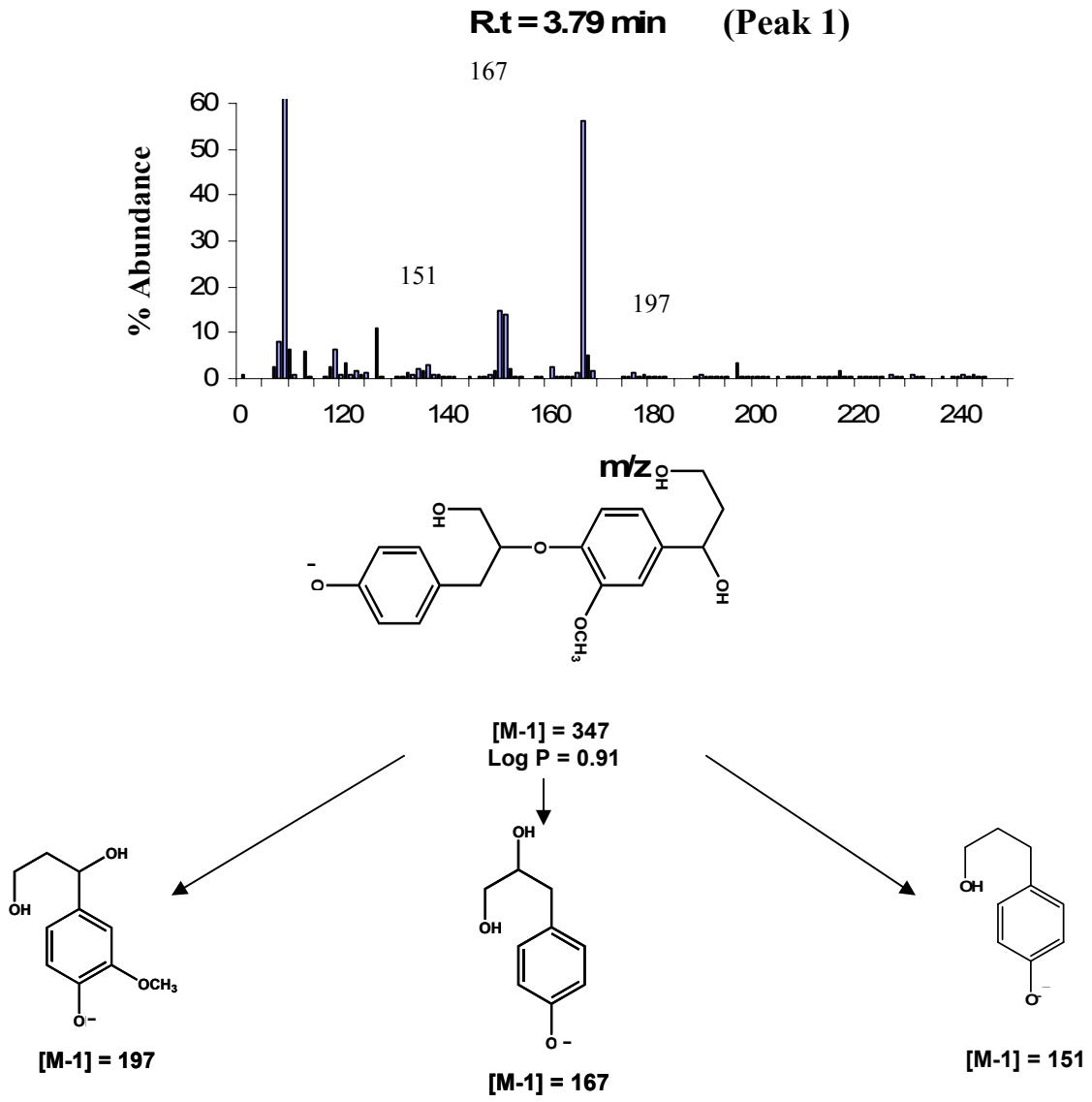


Figure 38. Mass fragmentation pattern at retention time 3.79 min gave a unique dimeric structure consisting of coumaryl and coniferyl alcohol monomers

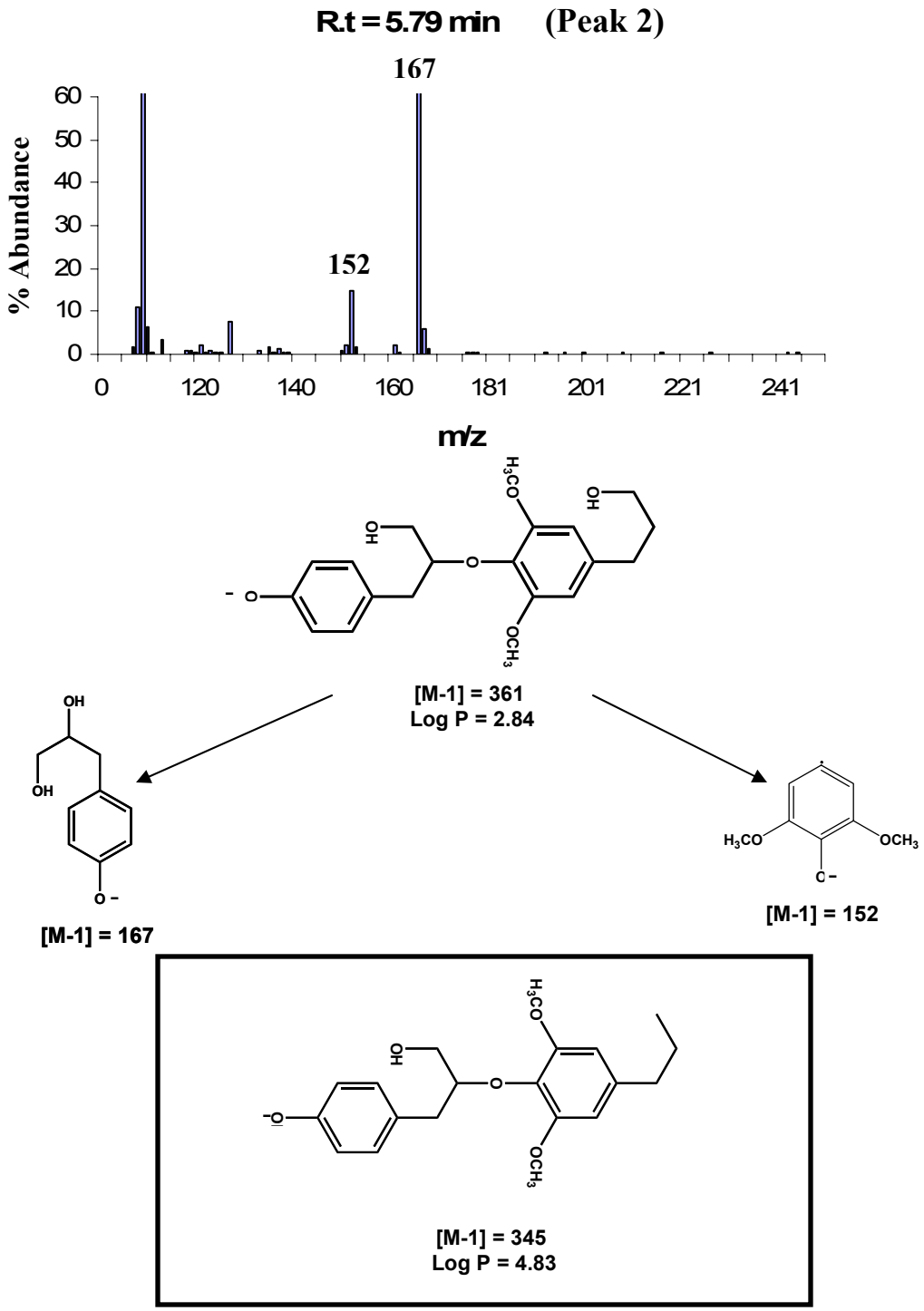


Figure 39. Mass fragmentation pattern at retention time 5.79 min gave rise to two possible structures differing in their log P values

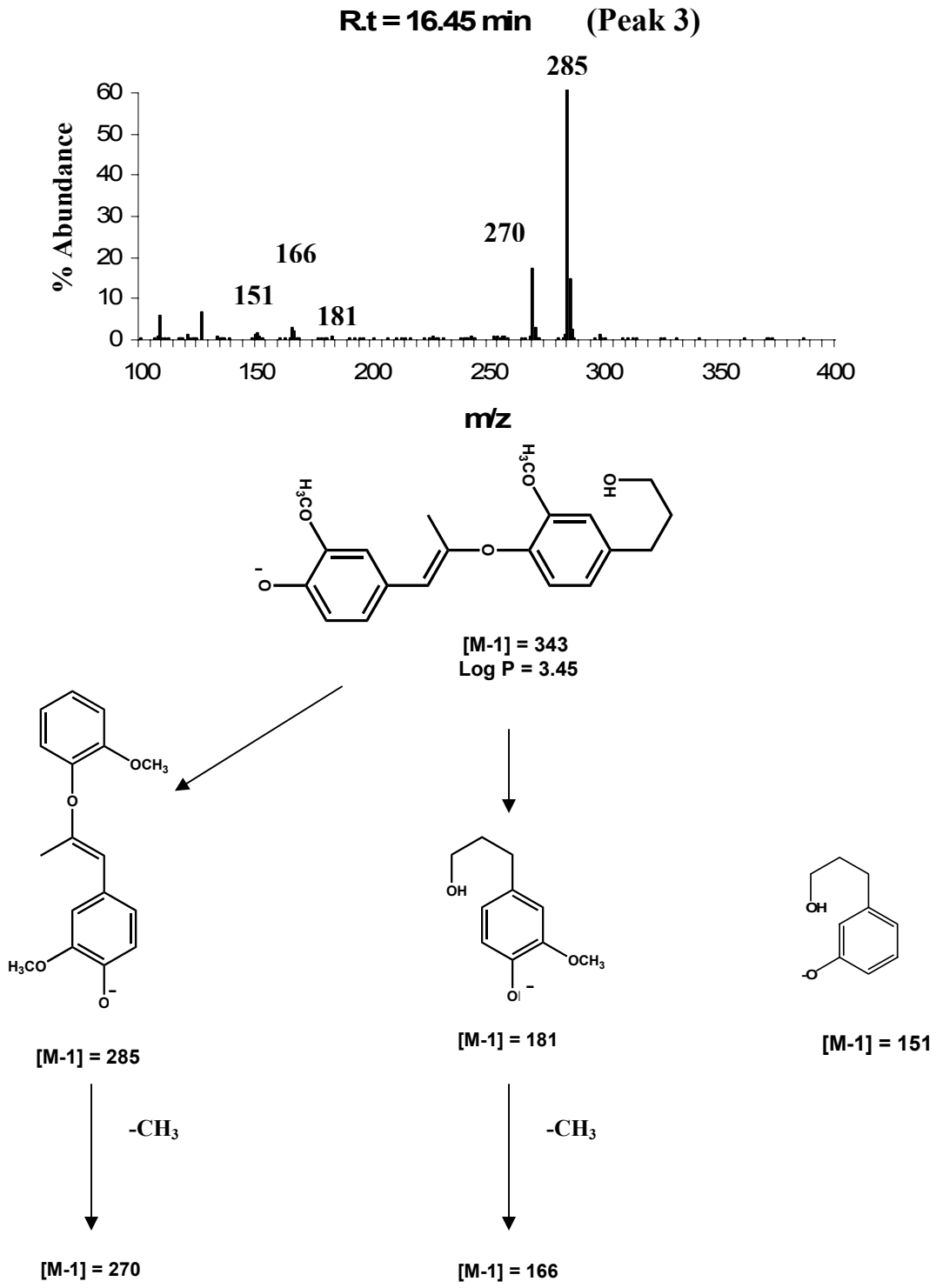


Figure 40. Mass fragmentation pattern of a dimeric coumaryl structure at retention time 16.49 min

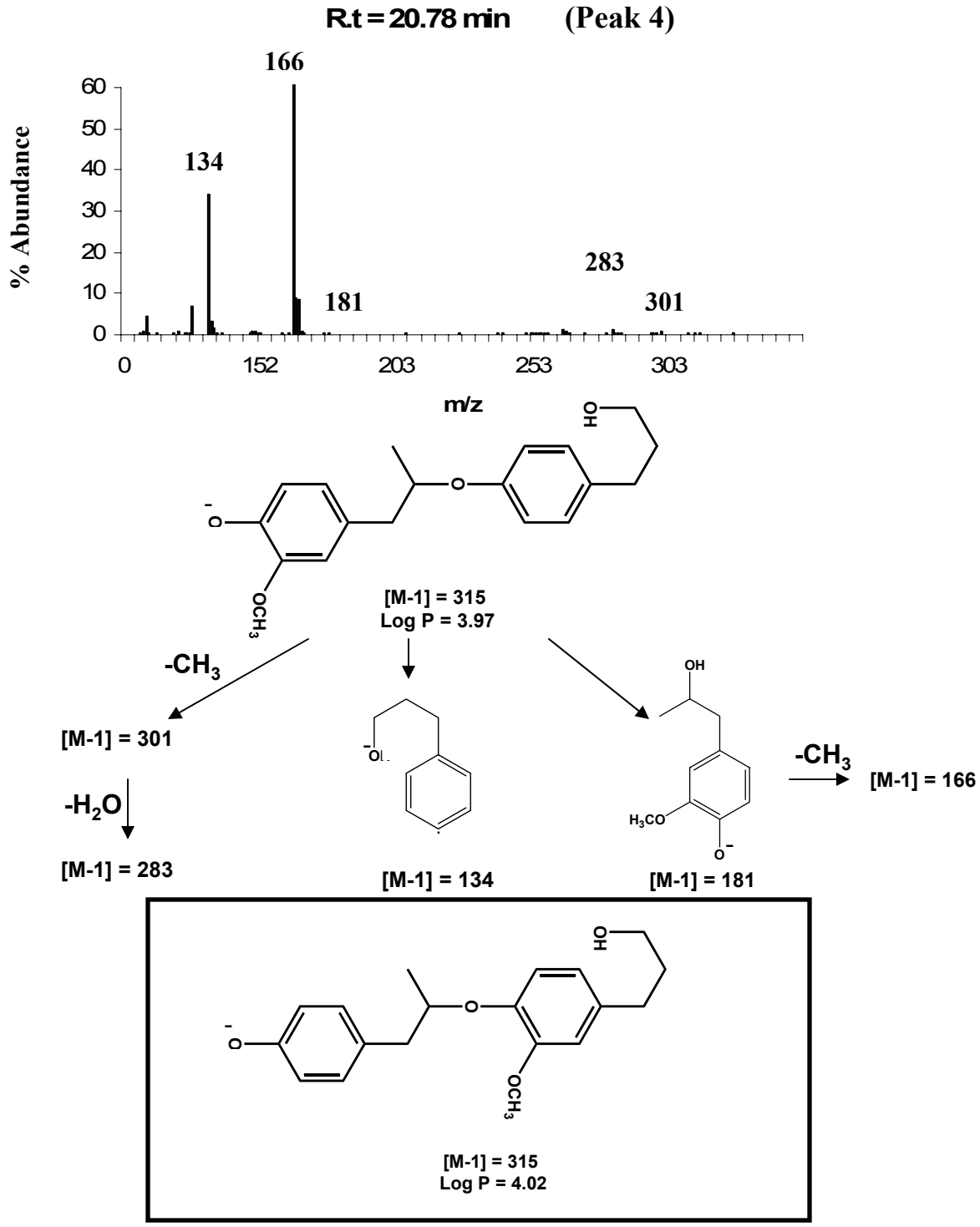


Figure 41. Mass fragmentation pattern of two possible dimeric structures at retention time 20.78 min

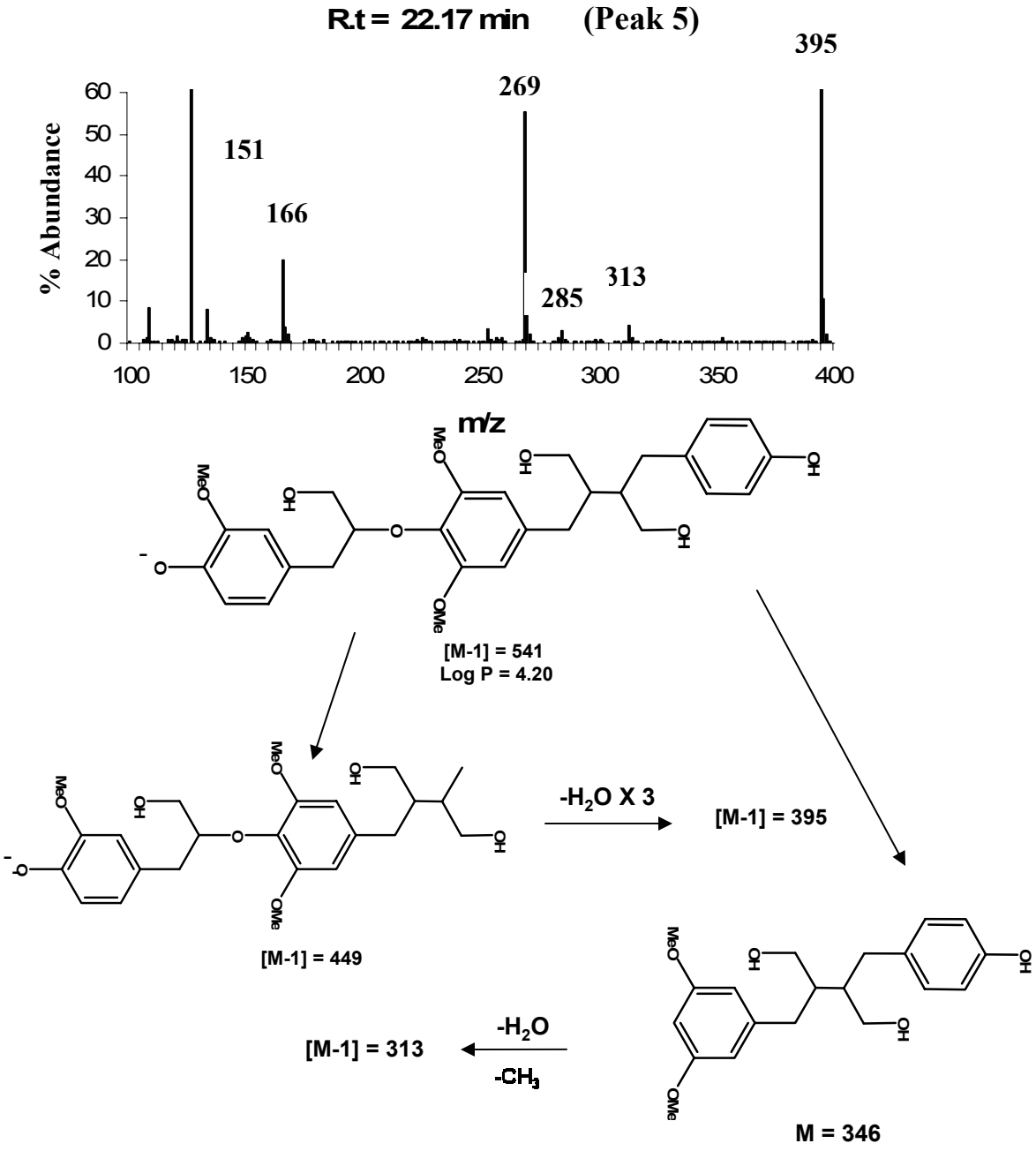


Figure 42. Mass fragmentation pattern of a trimer having β-O-4 and β-β linkage at retention time 22.7 min

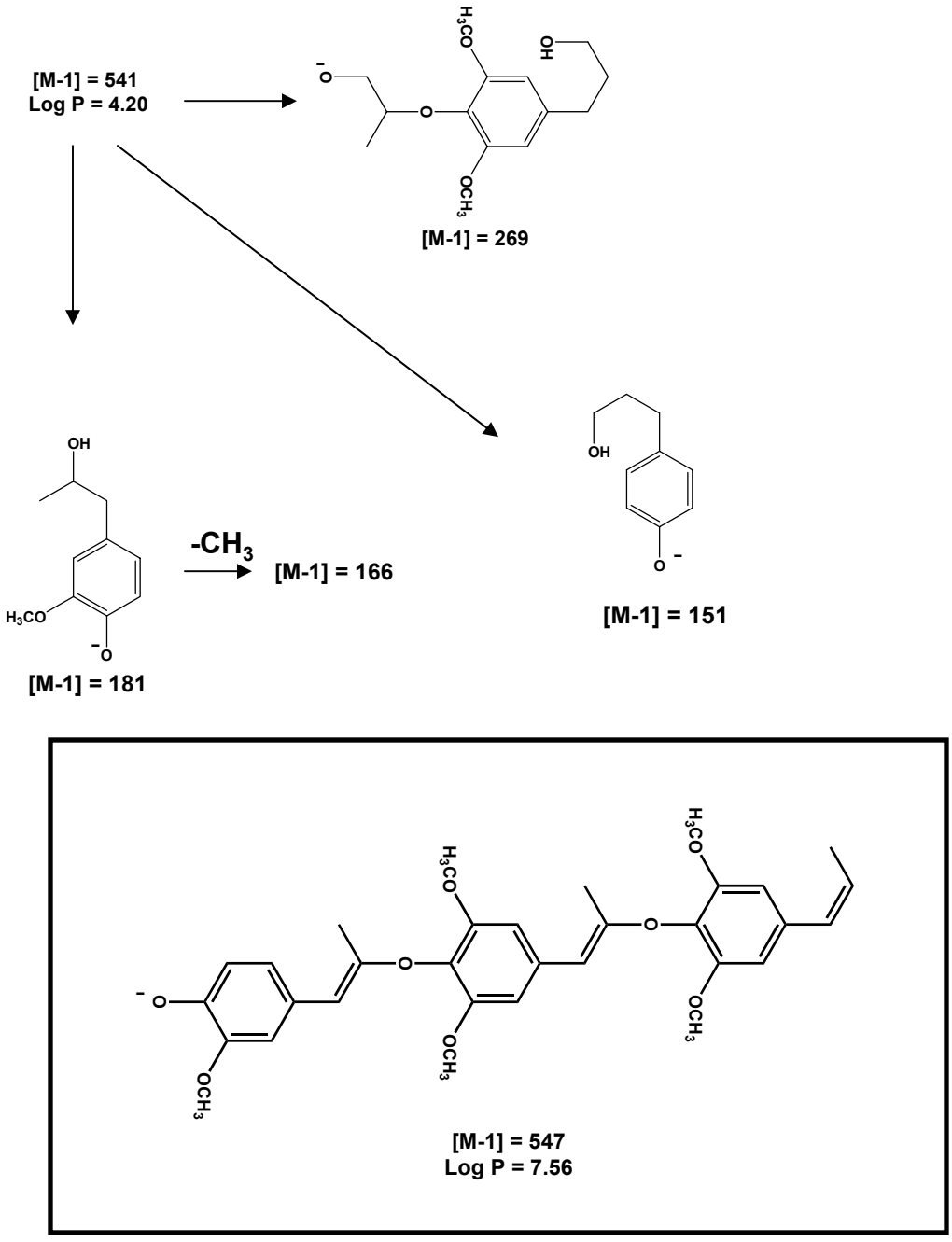


Figure 42. Mass fragmentation pattern of a trimer having β-O-4 and β-β linkage (continued) and another possible trimer with β-O-4 linkage at retention time 22.7 min

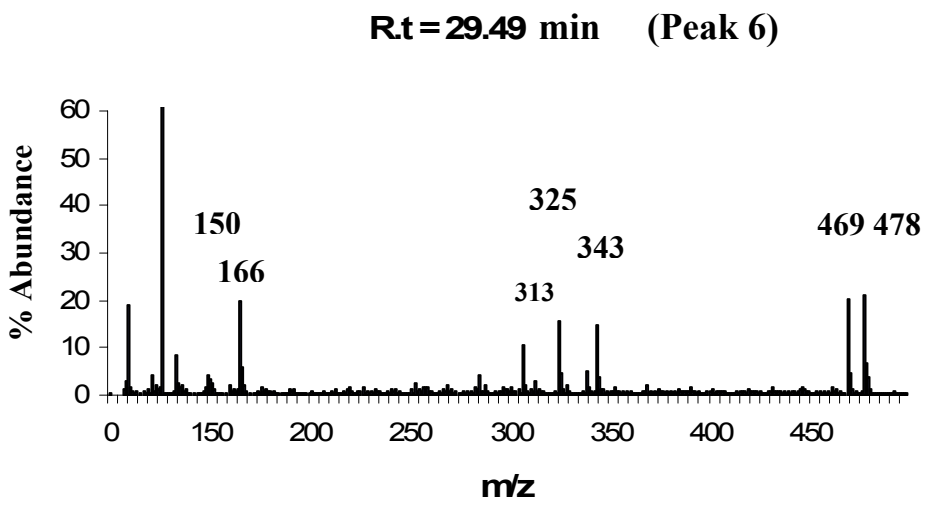


Figure 43. Mass fragmentation pattern at retention time 29.49 min

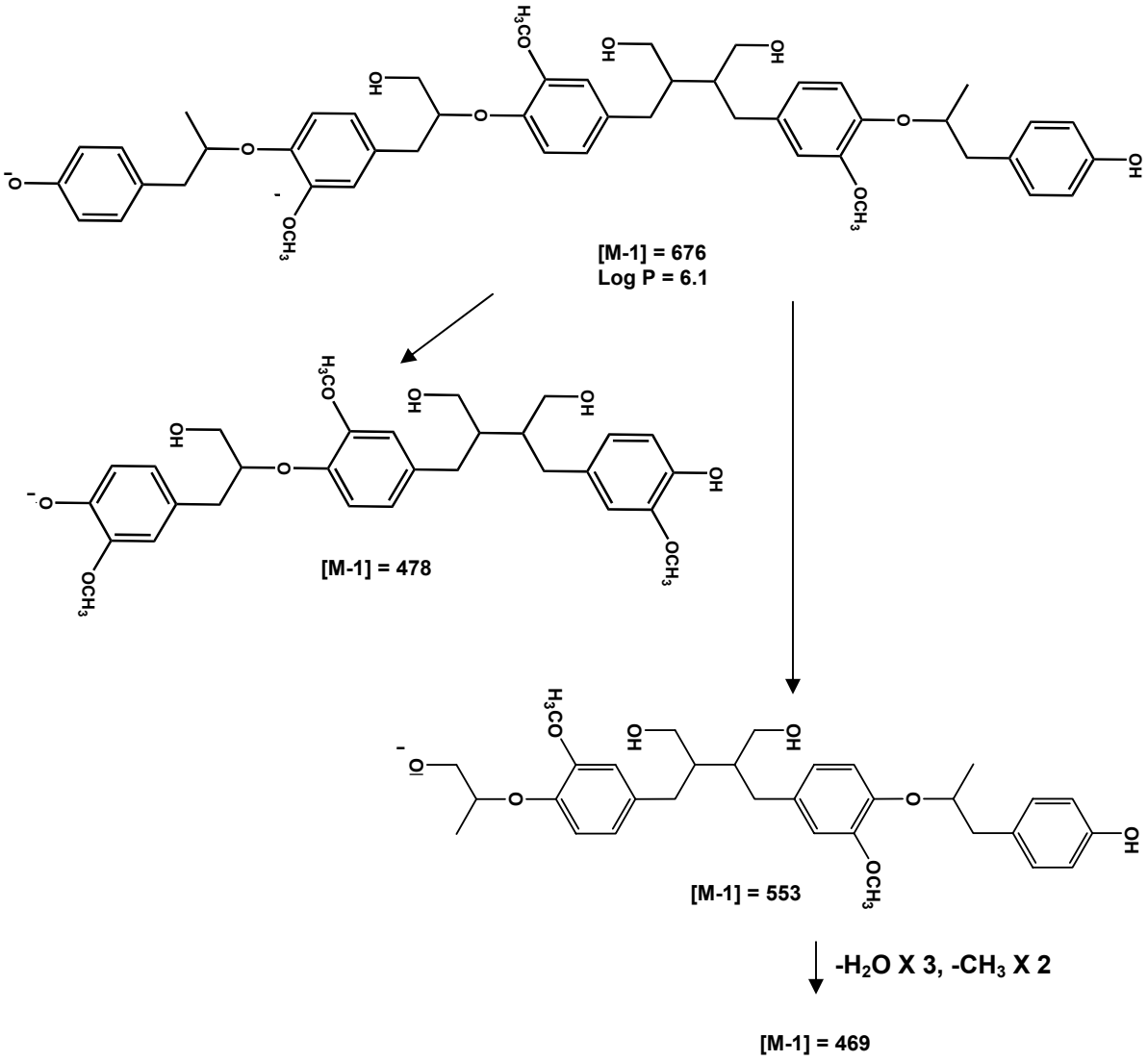


Figure 43. Mass fragmentation pattern of a unique tetrameric structure having β-O-4 and β-β linkage at retention time 29.49 min (continued)

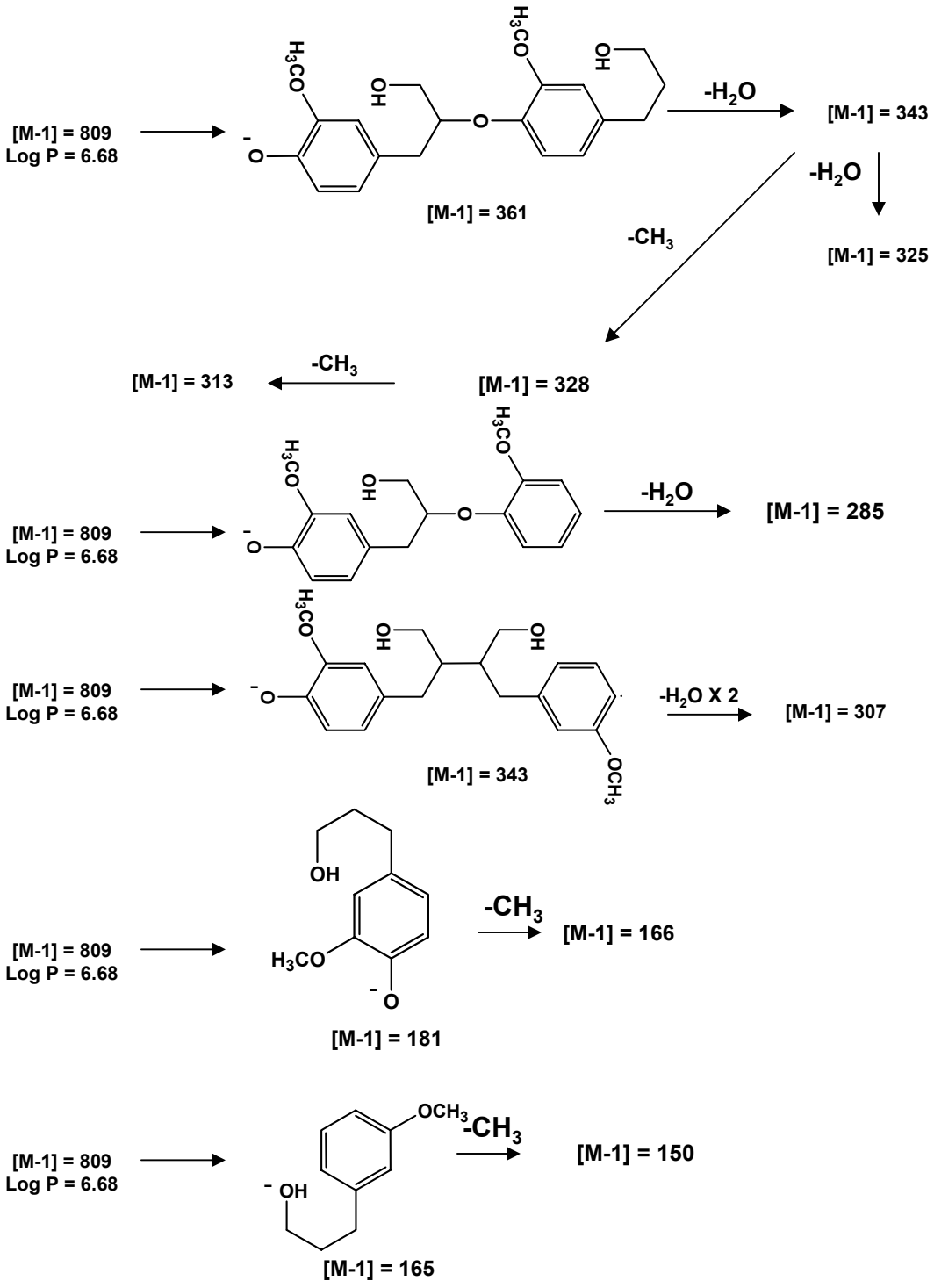


Figure 43. Mass fragmentation pattern of a unique tetrameric structure having β-O-4 and β-β linkage at retention time 29.49 min (continued)

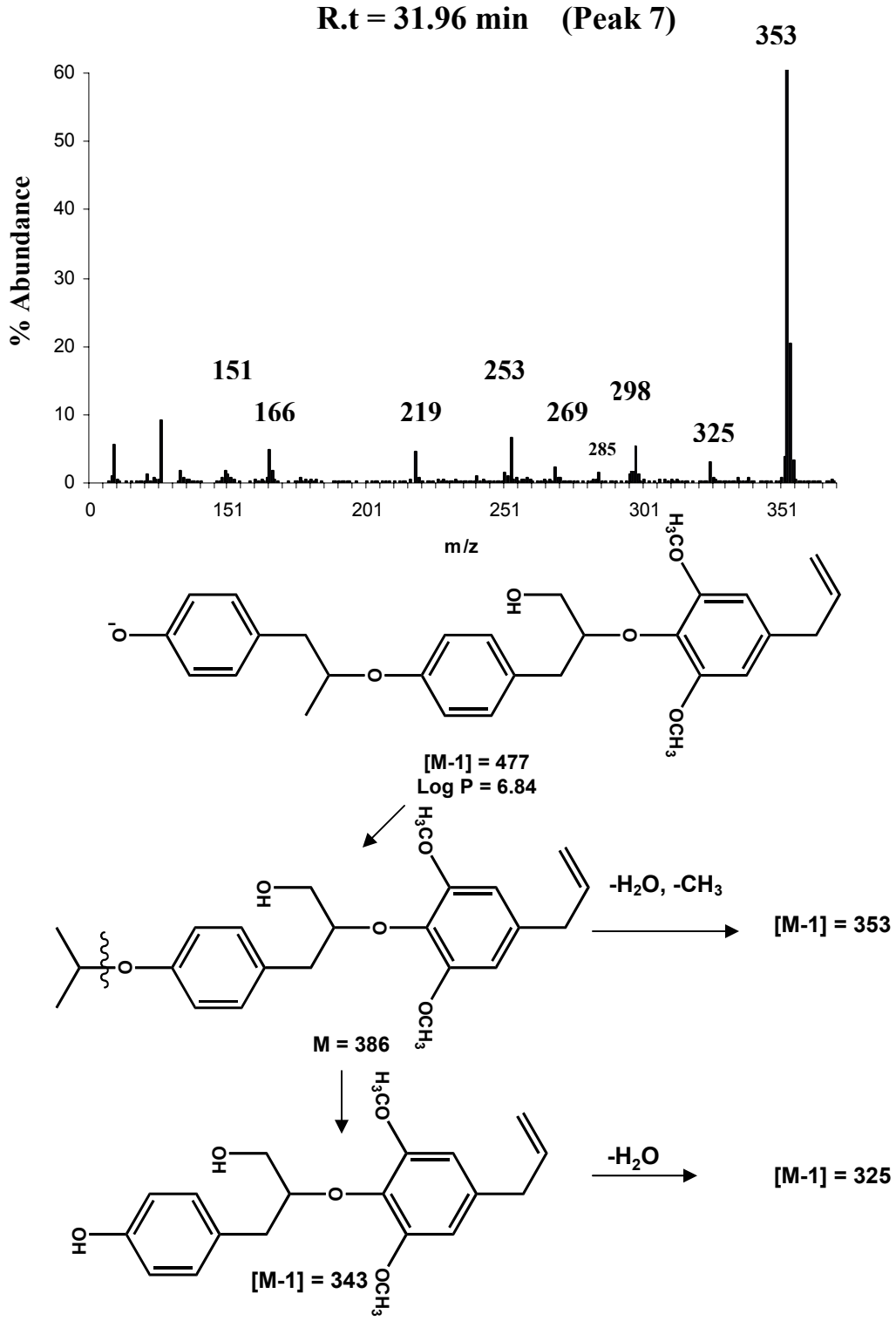


Figure 44. Mass fragmentation pattern of a trimer at retention time 31.96 min (continued)

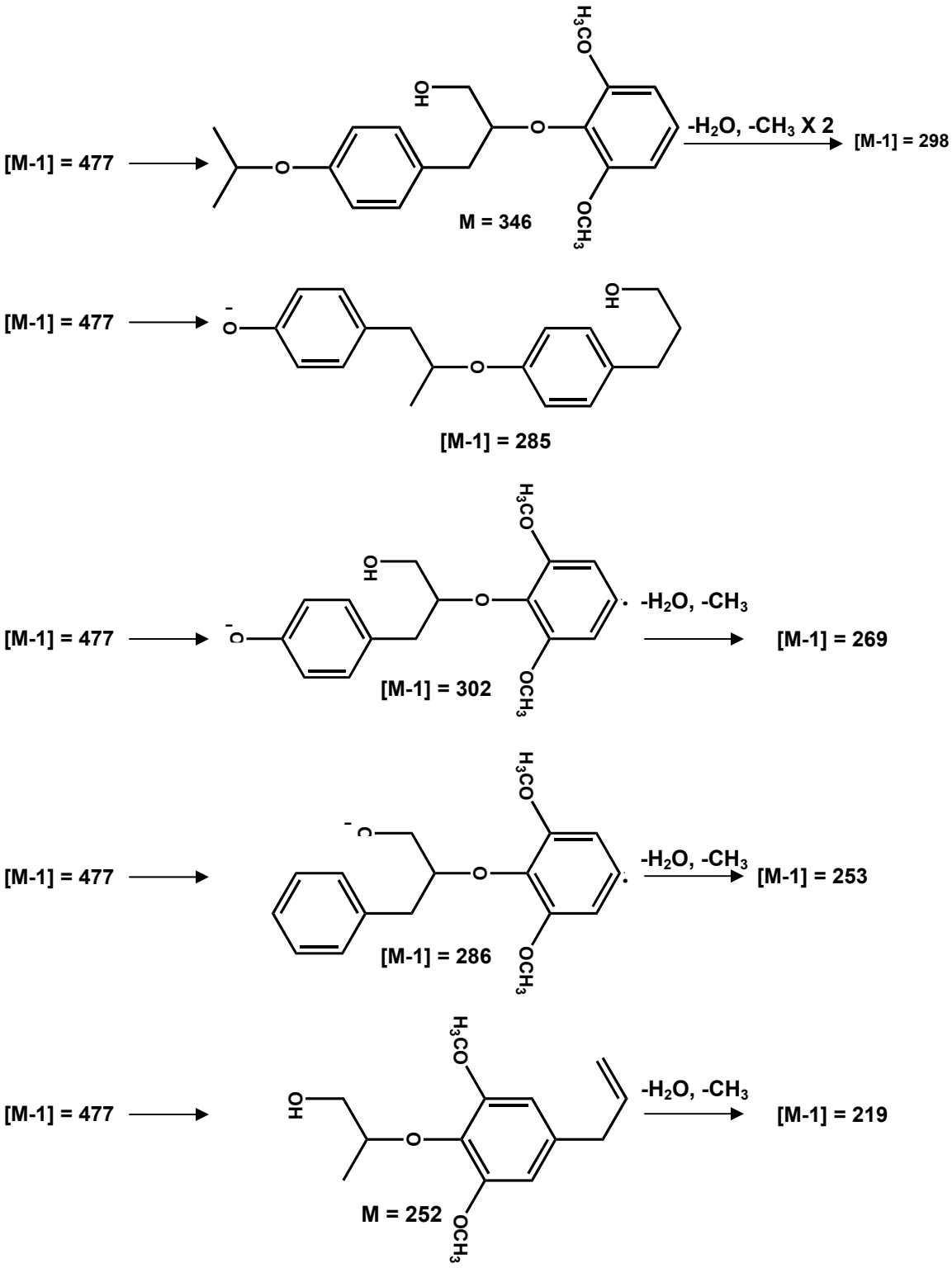


Figure 44. Mass fragmentation pattern of a trimer at retention time 31.96 min (continued)

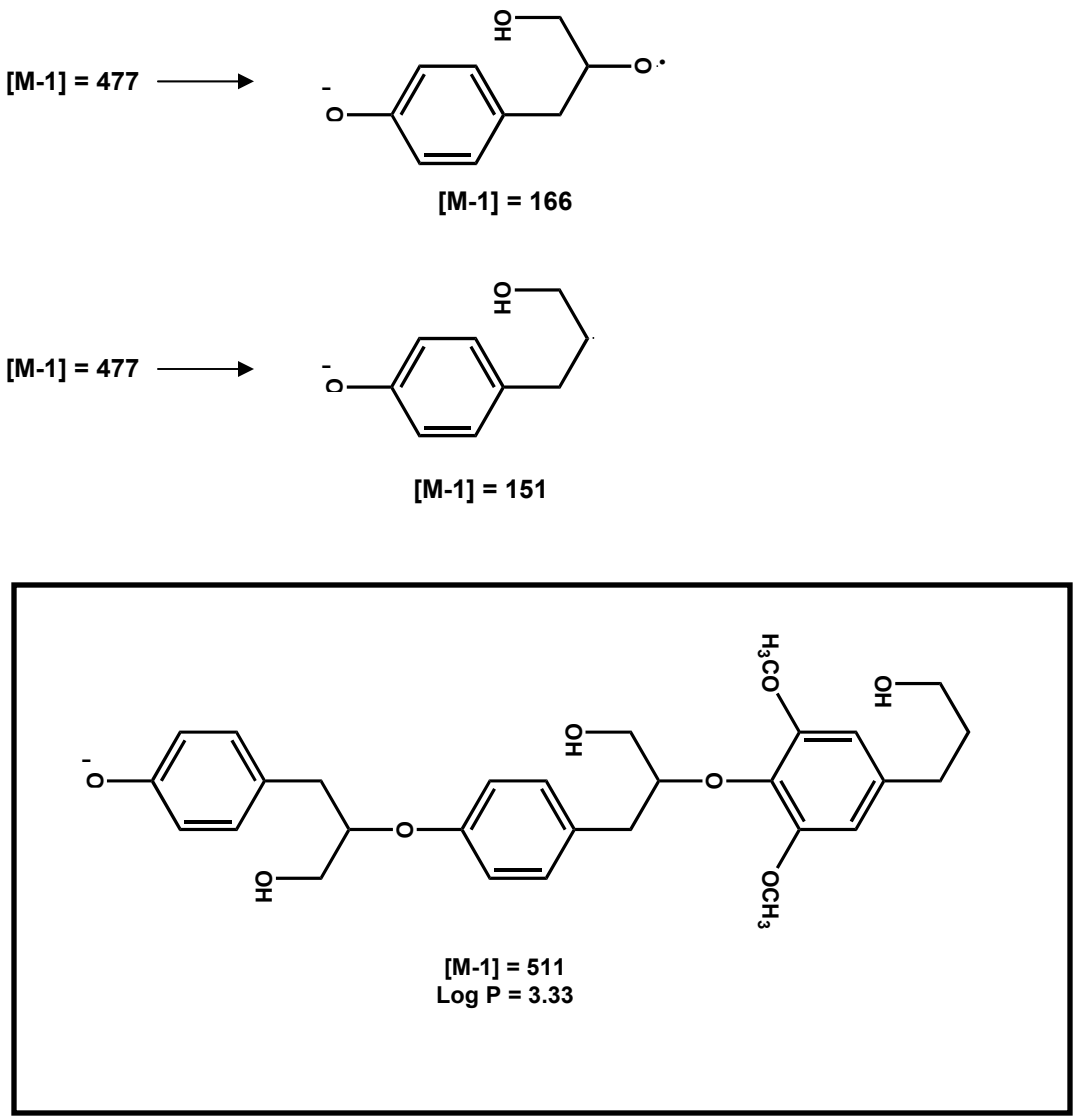


Figure 44. Mass fragmentation pattern of a trimer (continued) and another possible structure at retention time 31.96 min differing significantly in its log P value

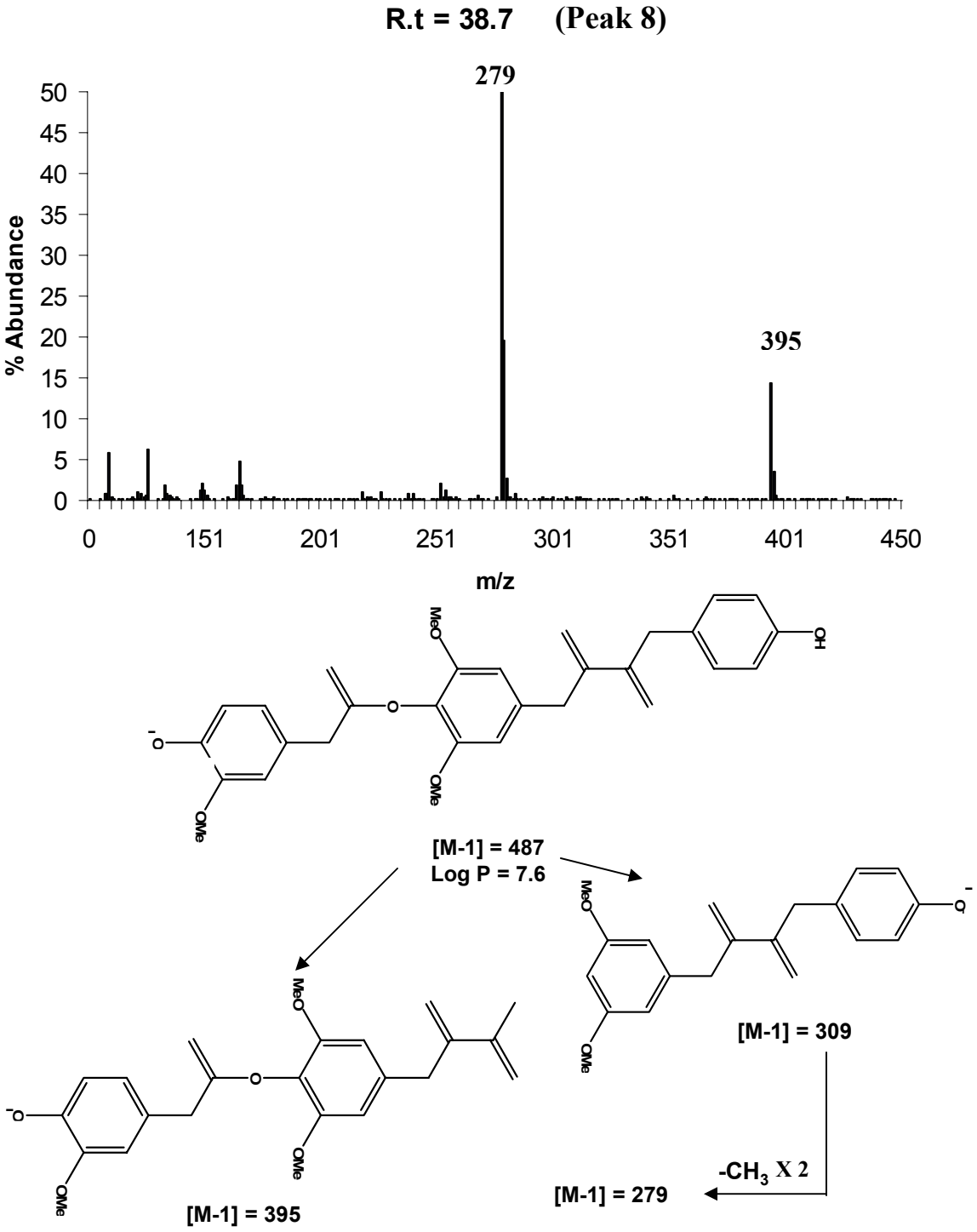


Figure 45. Mass fragmentation pattern of a β-O-4 and β-β linked trimer at retention time 38.7

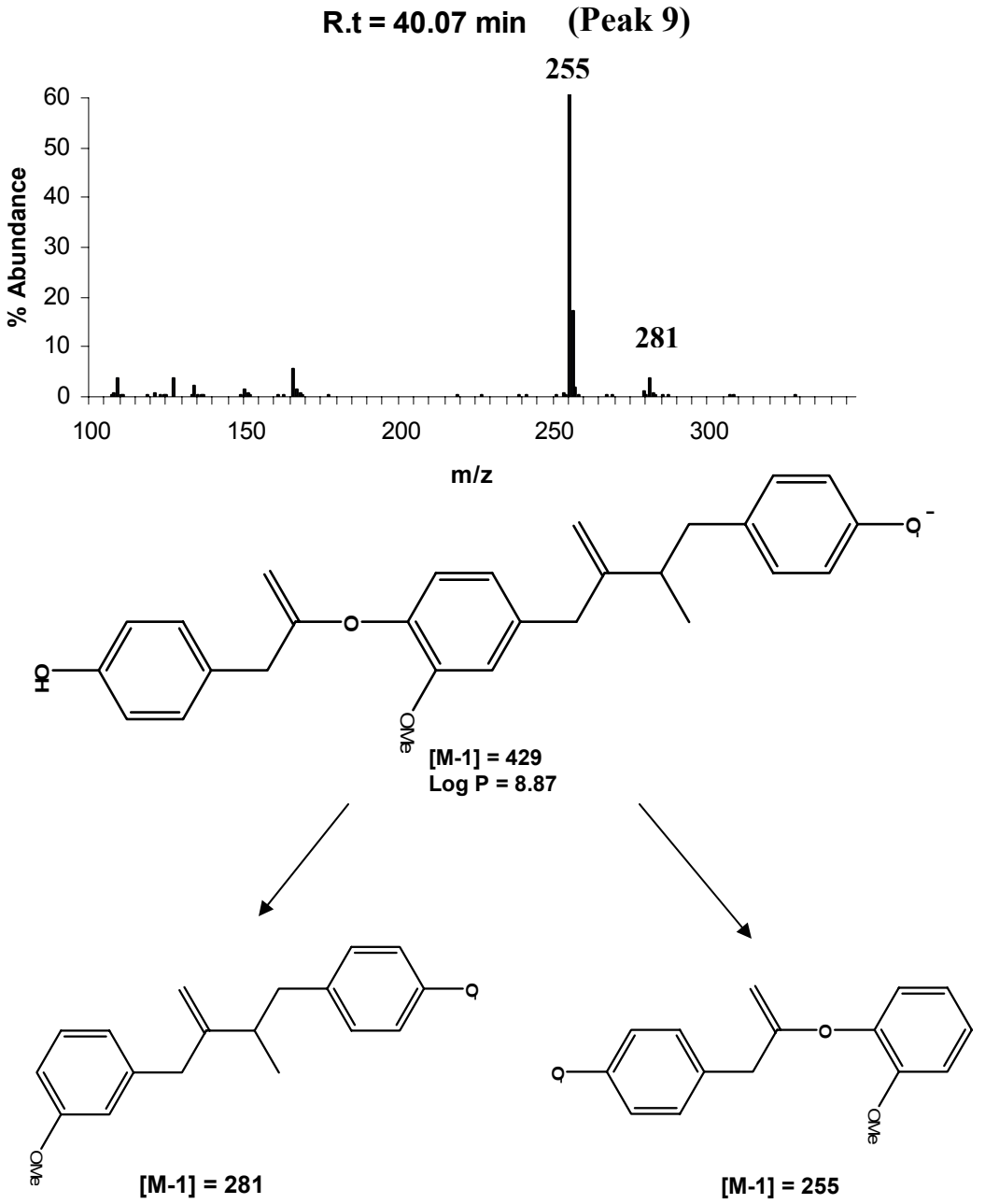


Figure 46. Mass fragmentation pattern of a β-O-4 and β-β linked trimer at retention time 40.07 min

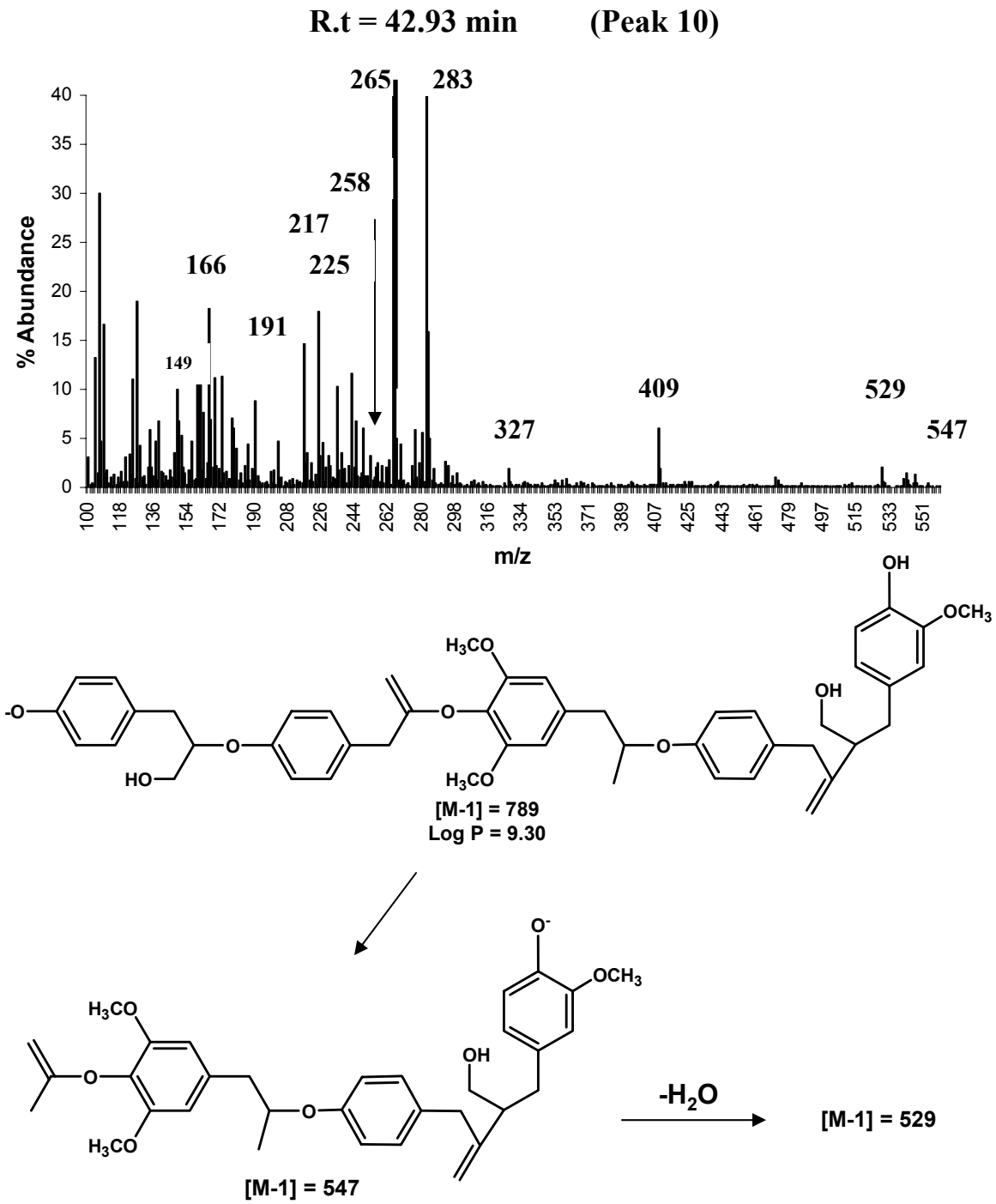


Figure 47. Mass fragmentation pattern of a unique pentamer having β-O-4 and β-β linkages at retention time 42.93 min

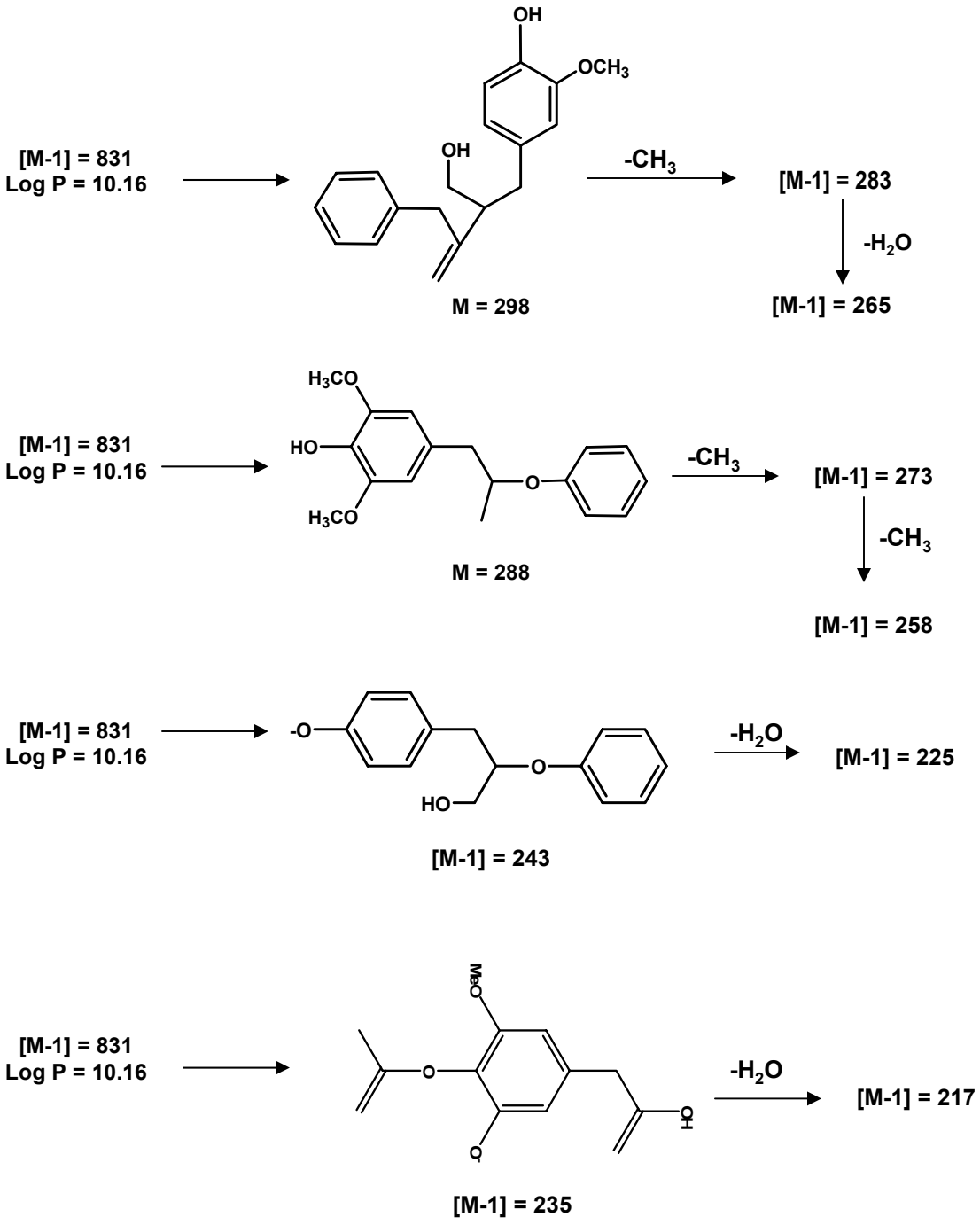


Figure 47. Mass fragmentation pattern of a unique pentamer having β-O-4 and β-β linkages at retention time 42.93 min (continued)

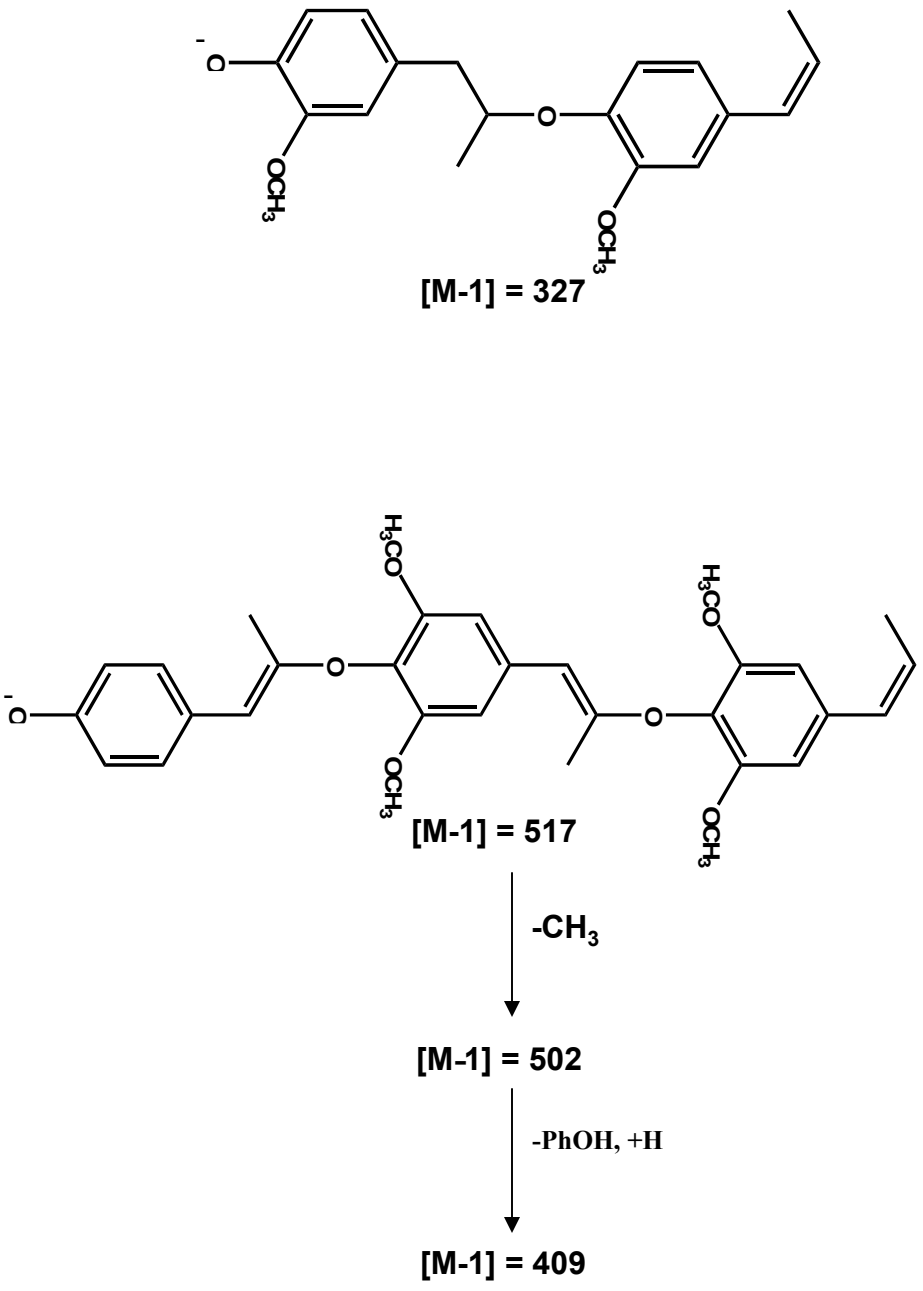


Figure 48. Structures (M-1 327 & 409) arise due to other than the parent molecule (M-1 831)

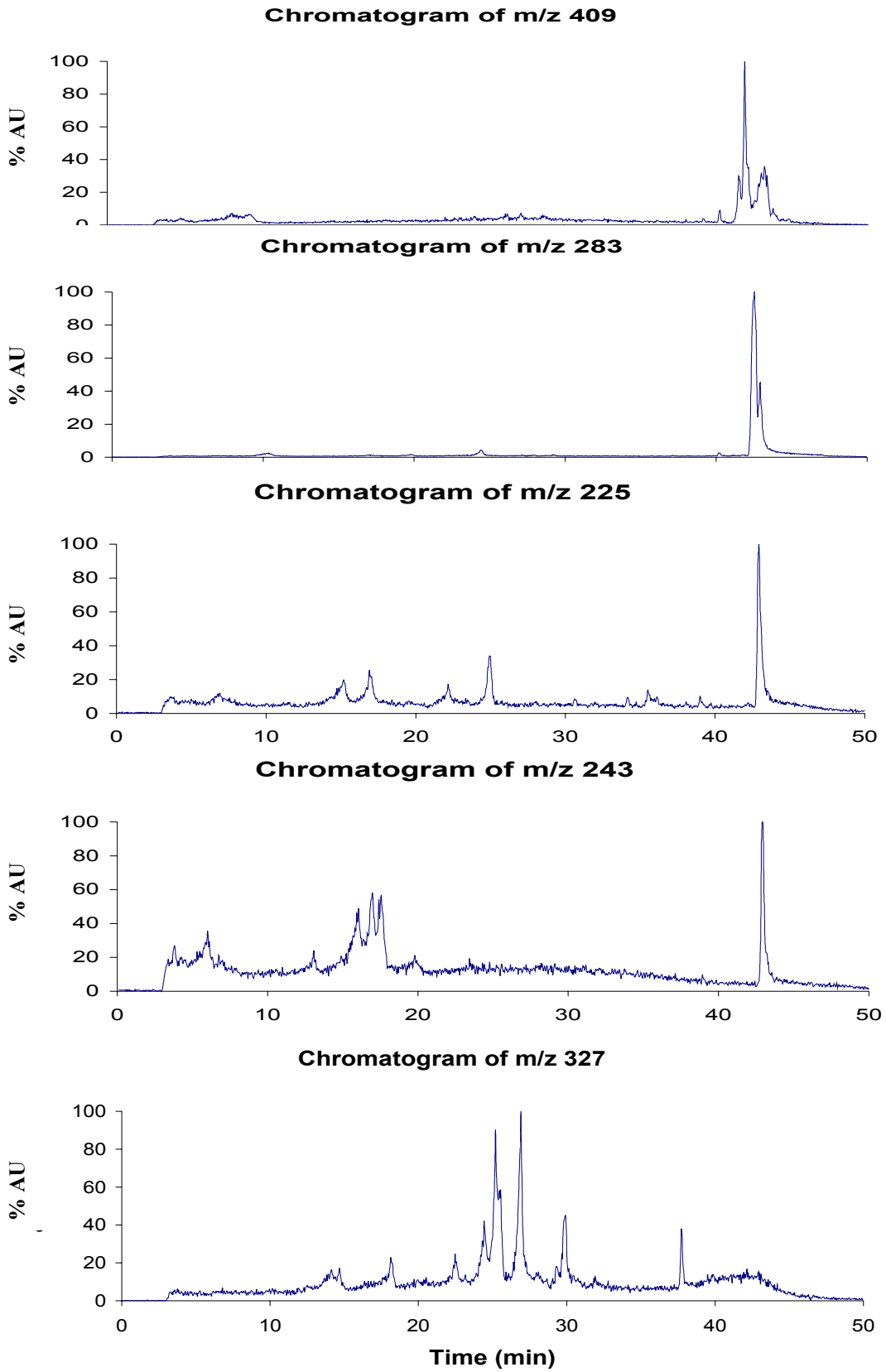


Figure 49. Chromatogram for different mass peaks from the same spectra

3.4 Disussion

Acid catalyzed degradation method has been developed for the structural identification, based on selective cleavage of the aryl alkyl-ether bonds that are the most abundant linkages in lignin. Lignins can be characterized through the analysis of their low molecular weight acidolysis products because these products can be related with confidence to specific lignin sub-structure. Acidolysis can be exploited to determine occurrence of β-O-4 and β-5 structures and to determine other characteristic substructures in lignin such as β-β.

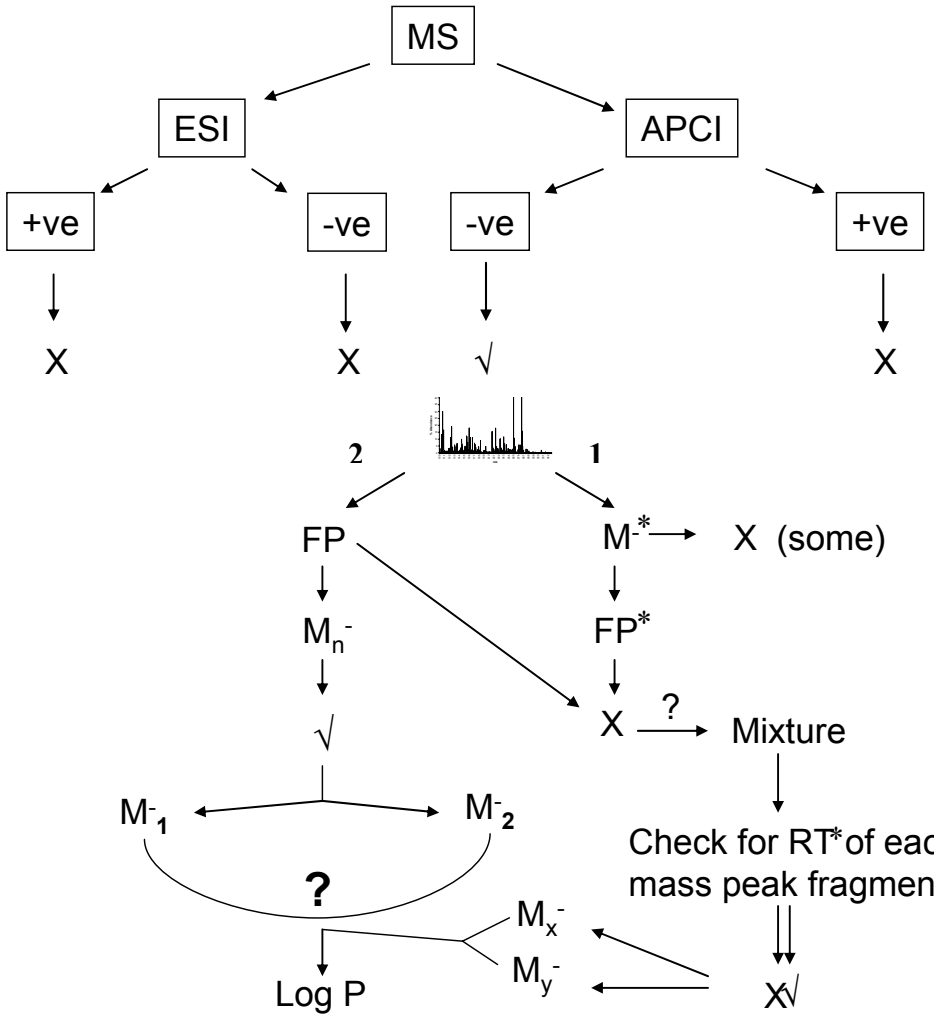


Figure 50: Schematic representation of protocol for assignments of the structures to the mass peaks

*FP: fragmentation pattern; M^-: molecular ion peak; RT: retention time

To identify the monomeric constituents present in our underivatized lignin, the polymer was partially depolymerized under acidic conditions. Resolution was achieved only with methanolic mobile phases under reverse-phase conditions, while ionization of the analytes appeared to work only in APCI negative mode. The acidolysis products were recorded in negative mode of APCI in the *m/z* 100-1100 region, connected to an LC. ESI-MS both in positive and negative mode with different solvent system was tried with little or no success. APCI in positive mode did not yield high sensitivity spectra. APCI was found to be working well in the negative mode but that was only after changing the solvent system. The earlier method for the HPLC was developed using acetonitrile:water (8:2) mixture which did not provide high sensitivity spectra due to lack of formation of ionizable groups on the molecule. Methanol:water (1:1) which was consequently developed seemed to work well in the negative mode of APCI with the presence of easily ionized phenolic groups. Comparable chromatographic resolution was obtained in the chromatogram obtained with UV-diode array detection at 279nm and total ion chromatogram (TIC) with mass detection **(figure 36, 37)**. The reverse-phase chromatogram of depolymerized lignin indicated the presence of at least 10 major peaks in addition to numerous minor peaks and residual polymer hump.

The negative ion APCI-MS spectra of all of the major peaks showed significant similarity of the mass peaks in the region 100-200 *m*/*z* suggesting a common monomeric backbone.

Intense peaks were observed in the region of 100-400. Parts of spectra are shown in **figure 38-47.** The base line is moved up as the polymer is not fully degraded and residual polymer is still present throughout. Each peak on the chromatogram was carefully analyzed from the mass fragmentation obtained. The protocol followed for the assignments of the structures was according to the scheme shown in **figure 50.** We started assigning the parent molecular ion in the mass spectra and then tried to fit the fragments. In many cases we could not find a parent molecular ion peak and for those the parent molecular ion peak was assigned the fragments could not be assigned. The fragments that could not be assigned might possibly arise due to the fact that each peak is a bad mixture. To assure that this is what was happening, we used a simple method to demonstrate this, based on the shape of the co-eluting peaks that has been discussed on page 122. But in this case the daughter ions were coming from the same parent molecule so we used a different route, we assigned parent molecular ion based on the fragmentation pattern. This seemed to work well but there were more than one possible structure that could be assigned. We assigned all possible structures and further narrowed it down using the hydrophobicity factor. We calculated the log P values for all the structures and chose the one which best fit the elution time pattern on the chromatogram. Each peak being a mixture factor was also explored using the same method based on fragmentation pattern and log P values to assign new structures in the same spectra. Final structures assigned were reached after a careful review of all the possible structures that

would show the mass fragmentation that was recorded. In most cases there were structures that showed many possible structures of the parent compound but when looked at the mass fragmentation, in most cases only one corroborated well with the fragmentation pattern. Thus indicating that the only possibility. Structural assignment was started with the most common lignin structure with β-O-4 linkage and then if no structure was possible with β-O-4, β-β was used.

The assignment of the structures based on the MS spectrum shows lignin monomers containing different linkages and different substituents on the aromatic ring and the aliphatic side chains. The influence of the substituents on the elution order of the phenolic compounds is clearly demonstrated. The introduction of hydroxyl groups dramatically decreases the retention time as is shown e.g. for compounds 1-4 compared to compounds 6-9. Introduction of methyl group on the aliphatic side chain also results in stronger interaction with the stationary phase as is observed for compounds 6-9. To support our conclusion that on the reverse phase column more hydrophobic compounds will elute at the end of the chromatogram, we calculated the log P values of all the parent structures using HINT program. The graph clearly shows that our hypothesis was true and shows $r^2 = 0.9772$ **(figure 51)**.

Some of the peaks that have been characterized contain a mixture of compounds. To demonstrated this a simple method was employed wherein each mass peaks in the MS spectrum which were co-eluting were checked for the shape of the peak **(figure 49).** The peaks on the chromatogram were used to obtain the MS spectrum and then each of the fragments obtained were re-checked for the shape of the co-eluting peak at that retention time. If the daughter ions in the spectra have same chromatogram peak shape at that retention time as that of the parent molecule then it is an indication that they are all arising from a single parent species. In case of peak at 42.93 min, the MS peaks showed different peak shape for elution time on the chromatogram supporting the idea that the peak eluted on the chromatogram was a mixture. The mass spectrum of the mixture of peaks at 42.93 min has been shown in **figure 47**. The mass spectrum shows major peaks at 283 and 265, can be interpreted as arising from oligomers of *p*-coumaryl, coniferyl and sinapyl type structures with *β*-*O*-4 linkage. There are numerous other peaks that when checked for the peak shape at that particular retention time indicates that the daughter ion of that peak is a result of a different molecular species other than the parent molecule that was assigned strengthening the conclusion that the chromatogram is a mixture of compounds supported by the observation from **figure 49**. Based on this conclusion peak 327 and 409 were characterized as their chromatographic peak shape at retention time 42.93 min differ significantly from the assigned parent molecule **(figure 48)**. If the

fragments were of the same molecule it should have shown the same peak shape at 42.93 min, which would be more convincing.

Table 6. Log P values increase with increasing time.

Time (min)	Log P values
03.80	0.91
05.79	2.84
16.45	3.45
20.78	3.97
22.17	4.20
29.49	6.10
31.96	6.84
38.70	7.60
40.07	8.87
42.93	9.30

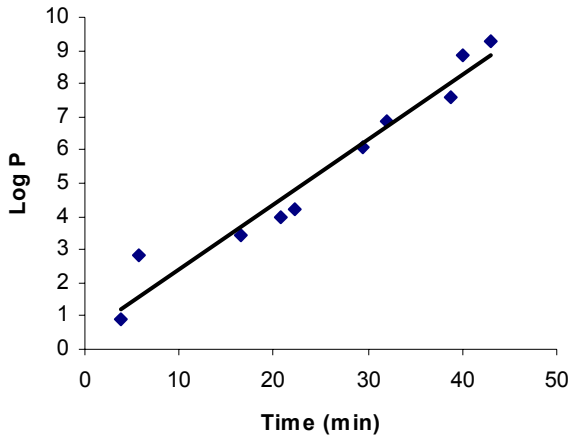


Figure 51. Plot of log P versus time gives a regression of 0.9772

3.5 Conclusion

Lignins are highly heterogeneous, complex, polydisperse polymers, which constitute the skeletal substance of all terrestrial plants. The current work identifies presence of a lignin as an impurity in the starting material, morin, through an arduous chemical and biophysical structure elucidation process. APCI-LCMS mass spectral analysis of lignin, which has been reported for the first time in here, primarily suggests the presence of a *β*-*O*-4 and *β*- *β* linked based lignin. For the first time combination of log P and LC-MS has achieved this kind of analysis. Several structural features remain to be elucidated as yet. For example, the ^{1}H and ^{13}C NMR spectra indicate the presence of hydrophobic methyls and methylenes, their location and number remain to be determined. The UV-vis study suggests the presence of underivatized phenolic groups, possibly at polymer termination points. Likewise, a number of high molecular weight peaks in the LC-MS spectra, possibly containing other intermonomeric linkages, remain unidentified. These native and non-native structures introduce significant structural complexity, yet the macromolecule that exhibits HSV-1 inhibitory activity is primarily a polymer containing an average of one sulfate group per monomer residue as can be concluded from elemental analysis. The medicinal properties of lignins remain unknown. This report constitutes the first example of a lignin derivative, a sulfated form of lignin, as an inhibitor of HSV-1 entry into cells. The IC_{50} value of 6 μg/mL determined herein compares favorably with values of 0.5-10 μg/mL measured for heparin, heparan sulfate, dextran sulfate, fucan sulfate, and other

sulfated polysaccharides. Competitive inhibition of HSV-1 entry into cells is known to be highly dependent on the sulfation level of the competitor. Further, just as with heparan sulfate, our lignin sulfate structure may possess pockets of higher charge density originating from the differential reactivity of –OH groups in parent heterogeneous lignin. Thus, the co-incidental optimal sulfation level, coupled with significant structural heterogeneity, is likely the origin for the anti-HSV-1 activity of lignin sulfate. Structurally, lignin sulfate represents a rich combination of substructures that represent many different options for interaction. The presence of multiple aromatic rings introduces hydrophobic forces, while unsulfated -OH groups make available hydrogen-bonding capability. Further, sulfation introduces the capability to form ionic interactions, while the bound water molecules, which can be released upon binding to a protein, introduce a favorable entropic factor. Thus, important enthalpy forces including hydrogen bonding, ionic, and hydrophobic, and entropic forces that govern nearly all interactions are present in lignin sulfate. Further, the structural possibilities afforded by many different types of linkages and different types of monomers suggest that lignin sulfate may possess structural richness thought to be present in heparan sulfate. Further work is needed to exploit this opportunity. Mass spectrometry forms an important tool in structural elucidation of lignins, although the heterogeneity and complexity of polymers represents a formidable challenge to overcome. Further, attempts to identify higher oligomeric lignin fragment masses appear to fail because of instability of the polyphenolic structure.

Structurally diverse lignin, which on sulfation has been shown to inhibit the entry of HSV-1 into the cells, can serve as a lead in structure based drug design.

Present knowledge about the molecular structure of lignin is based on the analysis of monomers, dimers or, at the most, tetramers of degraded isolated lignins. ^{13}C NMR spectroscopy has significant potential in providing detailed structural information for lignins, which indispensable in the quantitative determination of the amounts of different structural units in lignin. In particular, the advent of multidimensional NMR techniques has extended the prospect of lignin structural analysis considerably. Compared to the broad proton NMR signals that occur over narrow frequency range render limited quantitative information, ^{13}C NMR spectroscopy provides an elegant alternative, mainly due to its significantly larger chemical shift dispersion. However, a more detailed characterization of technical lignin side-chain structures is still needed. APCI-LC mass spectrometers are beginning to appear in environmental laboratories because of their potential for the determination of high molecular weight, polar, and reactive substances. In the present study APCI-LC-MS was useful for direct determination of degraded acidolytic congeners. APCI was substantially more sensitive with negative ion detection. One difficulty in studying lignin primary structure is that lignin chains can self-associate very strongly this property, together with the relatively harsh treatment required for lignin-derived fragment solubilization, and has made analysis difficult by conventional

means. Perhaps this explains why, until recently, there have been few attempts to analyze primary chain sequence information.

3.6 Methods

3.6.1 Chemical degradation of lignin[132-134]

A solution of 200 mg of isolated lignin in 10 mL of 0.2 M HCl in dioxane-methanol (1:1 v/v) was refluxed for 24 h, cooled, and treated with 0.2 M aqueous NaOH to neutralize the acid. The reaction mixture was worked up in a standard manner using ethyl acetate to get a mixture of lignin products.

3.6.2 LCMS

The mixture of lignin fragments from the degradation step was analyzed using LC-MS. The LC-MS system consisted of a Waters Alliance 2690 separation module and a Waters 996 photodiode array (PDA) UV detector (Waters Corp.,Milford, MA). Chromatographic separation was achieved using an analytical Discovery C18 column (Supelco, Bellefonte, PA, 4.6 _ 150 mm) and a linear binary gradient consisting of water-methanol (50:50 v/v) (solvent A) and methanol (solvent B) at a flow rate of 1 mL/min over period of 30 min. Eluent peaks were monitored at 279 nm and then analyzed by mass spectrometry. Liquid eluent was delivered to a Micromass ZMD4000 single quadrupole mass spectrometer

with ACPI ionization probe operating in negative ion mode (Waters Corp., Milford, MA). Optimized MS ionization conditions were employed; the source block temperature and the APCI probe temperature were held at 100 and 40 °C, respectively. Corona and cone voltages of 3.5 kV and 52 V were selected following optimization. The desolvation nitrogen flow was 400 L/h. Mass spectra were acquired in the mass range from 100 to 1100 Da at 400 amu/s.

3.6.3 Log P calculations

SYBYL 7.1 was used for this purpose. Structures were minimized using the following conditions:

Method – Powell

Gradient - 0.05 kcal/mol*A

Dielectric constant – 4

Maximum Iterations – 10000

Once the structures were minimized HINT program was used for calculating the log P values. Hydrogen treatment – united.

References:

(1) Corey, L. S., P. G. Infections with herpes simplex viruses (2). *N. Engl. J. Med.* **1986***, 314*, 749-757.

(2) Corey, L. S., P. G. Infections with herpes simplex viruses (1). *N. Engl. J. Med.* **1986***, 314*, 686-691.

(3) Mitchell, B. M.; Bloom, D. C.; Cohrs, R. J.; Gilden, D. H.; Kennedy, P. G. Herpes simplex virus-1 and varicella-zoster virus latency in ganglia. *J Neurovirol* **2003***, 9*, 194-204.

(4) Immergluck, L. C.; Domowicz, M. S.; Schwarts, N. B.; Herold, B. C. Viral and cellular requirements for entry of herpes simplex virus type 1 into primary neuronal cells. *J. Gen. Virol.* **1998***, 79*, 549-559.

(5) Herold, B. C.; Gerber, S. I.; Belval, B. J.; Siston, A. M.; Shulman, N. Differences in the susceptibility of herpes simplex virus types 1 and 2 to modified heparin compounds suggest serotype differences in viral entry. *J Virol* **1996***, 70*, 3461-3469.

(6) Corey L and Handsfield H H. Gential Herpes and Public Health: Affressing a Global Problem. *JAMA* **2000** (Feb); 283[6] 791-794.

(7) Cheshenko, N.; Keller, M. J.; MasCasullo, V.; Jarvis, G. A.; Cheng, H. *et al.* Candidate topical microbicides bind herpes simplex virus glycoprotein B and prevent viral entry and cell-to-cell spread. *Antimicrob Agents Chemother* **2004***, 48*, 2025-2036.

(8) Kleymann, G. New antiviral drugs that target herpesvirus helicase primase enzymes. *Herpes* **2003***, 10*, 46-52.

(9) Kleymann, G. Novel agents and strategies to treat herpes simplex virus infections. *Expert Opin Investig Drugs* **2003***, 12*, 165-183.

(10) Cai, W. H.; Gu, B.; Person, S. Role of glycoprotein B of herpes simplex virus type 1 in viral entry and cell fusion. *J Virol* **1988***, 62*, 2596-2604.

(11) Campadelli-Fiume, G.; Stirpe, D.; Boscaro, A.; Avitabile, E.; Foa-Tomasi, L. *et al.* Glycoprotein C-dependent attachment of herpes simplex virus to susceptible cells leading to productive infection. *Virology* **1990***, 178*, 213-222.

(12) Cheshenko, N.; Herold, B. C. Glycoprotein B plays a predominant role in mediating herpes simplex virus type 2 attachment and is required for entry and cell-to-cell spread. *J Gen Virol* **2002***, 83*, 2247-2255.

(13) Gerber, S. I.; Belval, B. J.; Herold, B. C. Differences in the role of glycoprotein C of HSV-1 and HSV-2 in viral binding may contribute to serotype differences in cell tropism. *Virology* **1995***, 214*, 29-39.

(14) Herold, B. C.; WuDunn, D.; Soltys, N.; Spear, P. G. Glycoprotein C of herpes simplex virus type 1 plays a principal role in the adsorption of virus to cells and in infectivity. *J Virol* **1991***, 65*, 1090-1098.

(15) Hutchinson, L.; Browne, H.; Wargent, V.; Davis-Poynter, N.; Primorac, S. A novel herpes simplex virus glycoprotein, gL, forms a complex with glycoprotein H (gH) and affects normal folding and surface expression of gH. *J Virol* **1992***, 66*, 2240-2250.

(16) Johnson, R. M.; Spear, P. G. Herpes simplex virus glycoprotein D mediates interference with herpes simplex virus infection. *J Virol* **1989***, 63*, 819-827.

(17) Pertel, P. E.; Fridberg, A.; Parish, M. L.; Spear, P. G. Cell fusion induced by herpes simplex virus glycoproteins gB, gD, and gH-gL requires a gD receptor but not necessarily heparan sulfate. *Virology* **2001***, 279*, 313-324.

(18) Roop, C.; Hutchinson, L.; Johnson, D. C. A mutant herpes simplex virus type 1 unable to express glycoprotein L cannot enter cells, and its particles lack glycoprotein H. *J Virol* **1993***, 67*, 2285-2297.

(19) Tal-Singer, R.; Peng, C.; Ponce De Leon, M.; Abrams, W. R.; Banfield, B. W. *et al.* Interaction of herpes simplex virus glycoprotein gC with mammalian cell surface molecules. *J Virol* **1995***, 69*, 4471-4483.

(20) Laquerre, S.; Argnani, R.; Anderson, D. B.; Zucchini, S.; Manservigi, R. *et al.* Heparan sulfate proteoglycan binding by herpes simplex virus type 1 glycoproteins B and C, which differ in their contributions to virus attachment, penetration, and cell-to-cell spread. *J Virol* **1998***, 72*, 6119-6130.

(21) Lycke, E.; Johansson, M.; Svennerholm, B.; Lindahl, U. Binding of herpes simplex virus to cellular heparan sulphate, an initial step in the adsorption process. *J Gen Virol* **1991***, 72*, 1131-1137.

(22) Rostand, K. S.; Esko, J. D. Microbial adherence to and invasion through proteoglycans. *Infect Immun* **1997***, 65*, 1-8.

(23) Shieh, M. T.; WuDunn, D.; Montgomery, R. I.; Esko, J. D.; Spear, P. G. Cell surface receptors for herpes simplex virus are heparan sulfate proteoglycans. *J Cell Biol* **1992***, 116*, 1273-1281.

(24) Spear, P. G. Herpes simplex virus: receptors and ligands for cell entry. *Cell. Microbiol.* **2004***, 6*.

(25) Spear, P. G.; Shieh, M. T.; Herold, B. C.; WuDunn, D.; Koshy, T. I. Heparan sulfate glycosaminoglycans as primary cell surface receptors for herpes simplex virus. *Adv Exp Med Biol* **1992***, 313*, 341-353.

(26) WuDunn, D.; Spear, P. G. Initial interaction of herpes simplex virus with cells is binding to heparan sulfate. *J Virol* **1989***, 63*, 52-58.

(27) Geraghty, R. J.; Krummenacher, C.; Cohen, G. H.; Eisenberg, R. J.; Spear, P. G. Entry of alphaherpesviruses mediated by poliovirus receptor-related protein 1 and poliovirus receptor. *Science* **1998***, 280*, 1618-1620.

(28) Montgomery, R. I. W., M. S.; Lum, B. J.; Spear, P. G. Herpes simplex virus - 1 entry into cells mediated by a novel member of the TNF/NGF receptor family. *Cell* **1996**, *87*, 427-436.
(29) Shukla, D.; Liu, J.; Blaiklock, P.; Shworak, N. W.; Bai, X. et al. A novel role for 3-O-sulfated heparan sulfate in herpes simplex virus 1 entry. *Cell* **1999**, *99*, 13-22.
(30) Spear, P. G.; Eisenberg, R. J.; Cohen, G. H. Three classes of cell surface receptors for alphaherpesvirus entry. *Virology* **2000**, *275*, 1-8.
(31) Warner, M. S. G., R. J.; Martinez, W. M.; Montgomery, R. I.; Whitbeck, J. C.; Xu, R.; Eisenberg, R. J.; Cohen, G. H.; Spear, P. G. A cell surface protein with herpesvirus entry activity (HveB) confers susceptibility to infection by mutants of herpes simplex virus type 1, herpes simplex virus type 2, and pseudorabies virus. *Virology* **1998**, *246*, 179-189.
(32) Whitbeck, J. C.; Peng, C.; Lou, H.; Xu, R.; Willis, S. H. et al. Glycoprotein D of herpes simplex virus (HSV) binds directly to HVEM, a member of the tumor necrosis factor receptor superfamily and a mediator of HSV entry. *J Virol* **1997**, *71*, 6083-6093.
(33) Williams, R. K.; Straus, S. E. Specificity and affinity of binding of herpes simplex virus type 2 glycoprotein B to glycosaminoglycans. *J Virol* **1997**, *71*, 1375-1380.
(34) Xia, G.; Chen, J.; Tiwari, V.; Ju, W.; Li, J. P. *et al.* Heparan sulfate 3-O-sulfotransferase isoform 5 generates both an antithrombin-binding site and an entry receptor for herpes simplex virus, type 1. *J Biol Chem* **2002**, *277*, 37912-37919.
(35) Xu, D.; Tiwari, V.; Xia, G.; Clement, C.; Shukla, D. *et al.* Characterization of heparan sulphate 3-O-sulphotransferase isoform 6 and its role in assisting the entry of herpes simplex virus type 1. *Biochem J* **2005**, *385*, 451-459.
(36) Liu, J.; Thorp, S. C. Cell surface heparan sulfate and its roles in assisting viral infections. *Med Res Rev* **2002**, *22*, 1-25.
(37) Shukla, D.; Spear, P.G. Herpesviruses and heparan sulfate: an intimate relationship in aid of viral entry. *J Clin Invest.* **2001**, *108*, 503-510.
(38) Bame, K. J.; Esko, J. D. Undersulfated heparan sulfate in a Chinese hamster ovary cell mutant defective in heparan sulfate N-sulfotransferase. *J Biol Chem* **1989**, *264*, 8059-8065.
(39) Bame, K. J.; Lidholt, K.; Lindahl, U.; Esko, J. D. Biosynthesis of heparan sulfate. Coordination of polymer-modification reactions in a Chinese hamster ovary cell mutant defective in N-sulfotransferase. *J Biol Chem* **1991**, *266*, 10287-10293.
(40) Herold, B. C.; Visalli, R. J.; Susmarski, N.; Brandt, C. R.; Spear, P. G. Glycoprotein C-independent binding of herpes simplex virus to cells requires cell surface heparan sulphate and glycoprotein B. *J Gen Virol* **1994**, *75*, 1211-1222.

(41) Herold, B. C.; Gerber, S. I.; Polonsky, T.; Belval, B. J.; Shaklee, P. N. *et al.* Identification of structural features of heparin required for inhibition of herpes simplex virus type 1 binding. *Virology* **1995***, 206*, 1108-1116.

(42) Liu, J.; Shriver, Z.; Pope, R. M.; Thorp, S. C.; Duncan, M. B. *et al.* Characterization of a heparan sulfate octasaccharide that binds to herpes simplex virus type 1 glycoprotein D. *J Biol Chem* **2002***, 277*, 33456-33467.

(43) Tiwari, V.; Clement, C.; Duncan, M. B.; Chen, J.; Liu, J. *et al.* A role for 3-O-sulfated heparan sulfate in cell fusion induced by herpes simplex virus type 1. *J Gen Virol* **2004***, 85*, 805-809.

(44) Patel, M.; Yanagishita, M.; Roderiquez, G.; Bou-Habib, D. C.; Oravecz, T. *et al.* Cell-surface heparan sulfate proteoglycan mediates HIV-1 infection of T-cell lines. *AIDS Res Hum Retroviruses* **1993***, 9*, 167-174.

(45) Ibrahim, J.; Griffin, P.; Coombe, D. R.; Rider, C. C.; James, W. Cell-surface heparan sulfate facilitates human immunodeficiency virus Type 1 entry into some cell lines but not primary lymphocytes. *Virus Res* **1999***, 60*, 159-169.

(46) Chen, Y.; Maguire, T.; Hileman, R. E.; Fromm, J. R.; Esko, J. D. *et al.* Dengue virus infectivity depends on envelope protein binding to target cell heparan sulfate. *Nat Med* **1997***, 3*, 866-871.

(47) Jackson, R. L.; Busch, S. J.; Cardin, A. D. Glycosaminoglycans: molecular properties, protein interactions, and role in physiological processes. *Physiol Rev* **1991***, 71*, 481-539.

(48) Fry, E. E.; Lea, S. M.; Jackson, T.; Newman, J. W.; Ellard, F. M.; Blakemore, W. E.; Abu-Ghazaleh, R.; Samuel, A.; King, A. M.; Stuart, D. I. The structure and function of a foot-and-mouth disease virus-oligosaccharide receptor complex. *EMBO J.* **1999***, 18*, 543-554.

(49) Chung, C. S.; Vasilevskaya, I. A.; Wang, S. C.; Bair, C. H.; Chang, W. Apoptosis and host restriction of vaccinia virus in RK13 cells. *Virus Res.* **1997***, 52*, 121-132.

(50) Byrnes, A. P.; Griffin, D. E. Binding of Sindbis virus to cell surface heparan sulfate. *J Virol.* **1998***, 72*, 7349-7356.

(51) Feldman, S. A.; Audet, S.; Beeler, J. A. The fusion glycoprotein of human respiratory syncytial virus facilitates virus attachment and infectivity via an interaction with cellular heparan sulfate. *J Virol.* **2000***, 74*, 6442-6447.

(52) Krusat, T.; Streckert, H. J. Heparin-dependent attachment of respiratory syncytial virus (RSV) to host cells. *Arch Virol* **1997***, 142*, 1247-1254.

(53) Goodfellow, I. G.; Sioofy, A.; Powell, R. M.; Evans, D. J. Echoviruses bind heparan sulfate at the cell surface. *J Virol.* **2001***, 75*, 4918-4921.

(54) David, G. Integral membrane heparan sulfate proteoglycans. *Faseb J* **1993***, 7*, 1023-1030.

(55) Shriver, Z.; Liu, D.; Sasisekharan, R. Emergining views of heparan sulfate glycosaminoglycan structure/activity relationships modulating dynamic biological functions. *Trends Cardiovasc. Med.* **2002**, *12*.

(56) Bernfield, M.; Kokenyesi, R.; Kato, M.; Hinkes, M. T.; Spring, J. *et al.* Biology of the syndecans: a family of transmembrane heparan sulfate proteoglycans. *Annu Rev Cell Biol* **1992**, *8*, 365-393.

(57) Bernfield, M. G., M.; Park, P. W.; Reizes, O.; Fitzgerald, M. L.; Lincecum, J.; Zako, M. Functions of cell surface heparan sulfate proteoglycans. *Annu. Rev. Biochem.* **1999**, *68*.

(58) Carey, D. J.; Evans, D. M.; Stahl, R. C.; Asundi, V. K.; Conner, K. J. *et al.* Molecular cloning and characterization of N-syndecan, a novel transmembrane heparan sulfate proteoglycan. *J Cell Biol* **1992**, *117*, 191-201.

(59) David, G.; van der Schueren, B.; Marynen, P.; Cassiman, J. J.; van den Berghe, H. Molecular cloning of amphiglycan, a novel integral membrane heparan sulfate proteoglycan expressed by epithelial and fibroblastic cells. *J Cell Biol* **1992**, *118*, 961-969.

(60) Gould, S. E.; Upholt, W. B.; Kosher, R. A. Syndecan 3: a member of the syndecan family of membrane-intercalated proteoglycans that is expressed in high amounts at the onset of chicken limb cartilage differentiation. *Proc Natl Acad Sci U S A* **1992**, *89*, 3271-3275.

(61) Esko, J. D.; Lindhal, U. Molecular diversity of heparan sulfate. *J. Clin. Invest.* **2001**, *108*, 169-173.

(62) Kjellen, L.; Lindahl, U. Proteoglycans: structures and interactions. *Annu Rev Biochem* **1991**, *60*, 443-475.

(63) Lindahl, U.; Kusche-Gullberg, M.; Kjellen, L. Regulated diversity of heparan sulfate. *J Biol Chem* **1998**, *273*, 24979-24982.

(64) Rabenstein, D. L. Heparin and heparan sulfate: structure and function. *Nat. Prod. Rep.* **2002**, *19*, 312-331.

(65) Salmivirta, M.; Lidholt, K.; Lindahl, U. Heparan sulfate: a piece of information. *Faseb J* **1996**, *10*, 1270-1279.

(66) Nyberg, K.; Ekblad, M.; Bergstrom, T.; Freeman, C.; Parish, C. R.; Ferro, V.; Trybala, E. The low molecular weight heparan sulfate-mimetic, PI-88, inhibits cell-to-cell spread of herpes simplex virus. *Antiviral Res.* **2004**, *63*, 15-24.

(67) Anderson, R. A.; Feathergill, K.; Diao, X.; Cooper, M.; Kirkpatrick, R. *et al.* Evaluation of poly(styrene-4-sulfonate) as a preventive agent for conception and sexually transmitted diseases. *J Androl* **2000**, *21*, 862-875.

(68) Carlucci, M. J.; Pujol, C. A.; Ciancia, M.; Noseda, M. D.; Matulewicz, M. C.; Damonte, E. B.; Cerezo, A. S. Antiherpetic and anticoagulant properties of carrageenans from the red seaweed Gigartina skottsbergii and their cyclized

derivatives: correlation between structure and biological activity. *Int. J. Biol. Macromol.* **1997***, 20*, 97-105.

(69) Dyer, A. P.; Banefield, B. W.; Martindale, D.; Spannier, D. M.; Tufaro, F. Dextran sulfate can act as an artificial receptor to mediate a type-specific herpes simplex virus infection via glycoprotein B. *J. Virol.* **1997***, 71*, 191-198.

(70) Feyzi, E.; Trybala, E.; Bergstrom, T.; Lindahl, U.; Spillmann, D. Structural requirement of heparan sulfate for interaction with herpes simplex virus type 1 virions and isolated glycoprotein C. *J Biol Chem* **1997***, 272*, 24850-24857.

(71) Hasui, M.; Matsuda, M.; Okutani, K.; Shigeta, S. In vitro antiviral activities of sulfated polysaccharides from a marine microalga (Cochlodinium polykrikoides) against human immunodeficiency virus and other enveloped viruses. *Int. J. Biol. Macromol.* **1995***, 17*, 293-297.

(72) Hayashi, T.; Hayashi, K.; Maeda, M.; Kojima, I. Calcium spirulan, an inhibitor of enveloped virus replication, from a blue-green alga Spirulina platensis. *J. Nat. Prod.* **1996***, 59*, 83-87.

(73) Herold, B. C.; Bourne, N.; Marcellino, D.; Kirkpatrick, R.; Strauss, D. M. et al. Poly(sodium 4-styrene sulfonate): an effective candidate topical antimicrobial for the prevention of sexually transmitted diseases. *J Infect Dis* **2000***, 181*, 770-773.

(74) Herold, B. C.; Siston, A.; Bremer, J.; Kirkpatrick, R.; Wilbanks, G.; Fugedi, P.; Peto, C.; Cooper, M. Sulfated carbohydrate compounds prevent microbial adherence by sexually transmitted disease pathogens. *Antimicrob. Agents. Chemother.* **1997***, 41*, 2776-2780.

(75) Lee, J. B.; Hayashi, K.; Hashimoto, M.; Nakano, T.; Hayashi, T. Novel antiviral fucoidan from sporophyll of Undaria pinnatifida (Mekabu). *Chem. Pharm. Bull. (Tokyo)* **2004***, 52*, 1091-1094.

(76) Mazumder, S.; Ghosal, P. K.; Pujol, C. A.; Carlucci, M. J.; Damonte, E. B.; Ray, B. Isolation, chemical investigation and antiviral activity of polysaccharides from Gracilaria corticata (Gracilariaceae, Rhodophyta). *Int. J. Biol. Macromol.* **2002***, 31*, 87-95.

(77) Ponce, N. M.; Pujol, C. A.; Damonte, E. B.; Flores, M. L.; Stortz, C. A. Fucoidans from the brown seaweed Adenocystis utricularis: extraction methods, antiviral activity and structural studies. *Carbohydr. Res.* **2003***, 338*, 153-165.

(78) Vaheri, A.; Cantell, K. The Effect Of Heparin On Herpes Simplex Virus. *Virology* **1963***, 21*, 661-662.

(79) Witvrouw, M.; De Clercq, E. Sulfated polysaccharides extracted from sea algae as potential antiviral drugs. *Gen. Pharmacol.* **1997***, 29*, 497-511.

(80) Zacharopoulos, V. R.; Phillips, D. M. Vaginal formulations of carrageenan protect mice from herpes simplex virus infection. *Clin. Diagn. Lab. Immunol.* **1997***, 4*, 465-468.

(81) Ralph, S. A. *USDA-Agricultural Research Service, Department of Forestry, University of Wisconsin, Madison, WI. Personal Communication.*

(82) Glasser, W. G. *Lignin. In: Casey, J. P. (Ed.) Pulp and Paper - Chemistry and Chemical Technology, 1980 3rd Edition*; Wiley and Sons, New York; pp. 39-111.

(83) Sarkanen, K. V.; Ludwig, C. H. *Lignins: Occurrence, Formation, Structure and Reactions*; Wiley-Interscience: New York, **1971**.

(84) Sederoff, R. R.; MacKay, J. J.; Ralph, J.; Hatfield, R. D. Unexpected variation in lignin. *Curr Opin Plant Biol* **1999***, 2*, 145-152.

(85) Oliveira, L.; Evtuguin, D. V.; Cordeiro, N.; Silvestre, A. J.; Silva, A. M. *et al.* Structural characterization of lignin from leaf sheaths of "dwarf cavendish" banana plant. *J Agric Food Chem* **2006***, 54*, 2598-2605.

(86) Susana Camarero, P. B.; Galletti, G. C.; Martı´nez, A. T. Pyrolysis-Gas Chromatography/Mass Spectrometry Analysis of Phenolic and Etherified Units in Natural and Industrial Lignins. *RAPID COMMUNICATIONS IN MASS SPECTROMETRY* **1999***, 13*, 630-636.

(87) Boerjan, W.; Ralph, J.; Baucher, M. Lignin biosynthesis. *Annu Rev Plant Biol* **2003***, 54*, 519-546.

(88) Freudenberg, K.; Neish, A. C. *Constitution and Biosynthesis of Lignin*; Springer-Verlag: New York, 1968; p 132.

(89) Hatfield, R.; Vermerris, W. Lignin formation in plants. The dilemma of linkage specificity. *Plant Physiol* **2001***, 126*, 1351-1357.

(90) Onnerud, H.; Zhang, L.; Gellerstedt, G.; Henriksson, G. Polymerization of monolignols by redox shuttle-mediated enzymatic oxidation: a new model in lignin biosynthesis I. *Plant Cell* **2002***, 14*, 1953-1962.

(91) Humphreys, J. M.; Chapple, C. Rewriting the lignin roadmap. *Curr Opin Plant Biol* **2002***, 5*, 224-229.

(92) Lewis, N. G. A 20(th) century roller coaster ride: a short account of lignification. *Curr Opin Plant Biol* **1999***, 2*, 153-162.

(93) Whetten, R. W.; MacKay, J. J.; Sederoff, R. R. Recent Advances In Understanding Lignin Biosynthesis. *Annu Rev Plant Physiol Plant Mol Biol* **1998***, 49*, 585-609.

(94) Horvath, P. J. The nutritional and ecological significance of Acer-tannins and related polyphenols.; Cornell University: Ithaca, NY., 1981.

(95) Bate-Smith, E. C. Tannins in herbaceous leguminosae. *Phytochemistry* **1973.***, 12*, 1809.

(96) Haslam, E. Natural polyphenols (vegetable tannins) as drugs: possible modes of action. *J Nat Prod* **1996***, 59*, 205-215.

(97) Reed, J. D. Nutritional toxicology of tannins and related polyphenols in forage legumes. *J Anim Sci* **1995***, 73*, 1516-1528.

(98) Hagerman, A. E. *Tannin-protein interactions. In: C. Ho, C. Y. Lee, and M. Huang (Ed.) Phenolic Compounds in Food and Their Effects on Health 1 Analysis, Occurrence, and Chemistry. p 236. ACS Symp. Series 506. American Chemical Society, Washington, DC.*, **1992**.

(99) Haslam, E. *Proanthocyanidins. In: J. B. Harborne and T. J. Mabrey (Ed.) The Flavonoids: Advances in Research.*; Chapman and Hall, London., **1982**.

(100) Haslam, E. *Plant Polyphenols-Vegetable Tannins Revisited.*; Cambridge University Press, Cambridge, U.K., **1989**.

(101) Porter, L. J.; Hrstich, L. N.; Chan, B. G. The conversion of procyanidins and prodelphinidins to cyanidin and delphinidin. *Phytochemistry* **1986***, 25*, 223-230.

(102) Ferreira, D.; Liu, X.C. Oligomeric proanthocyanidins: naturally occurring O-heterocycles. *Nat. Prod. Rep.* **2000***, 17*, 193-212.

(103) Ferreira, D.; Slade, D. Oligomeric proanthocyanidins: naturally occurring O-heterocycles. *Nat. Prod. Rep.* **2002***, 19*, 517-541.

(104) Thompson, R. S.; Jacques, D.; Haslam, E.; Tanner, R. J. N. Plant proanthocyanidins. Part I. Introduction; the isolation, structure, and distribution in nature of plant procyanidins. *J. Chem. Soc., Perkin I* **1972**, 1387-1399.

(105) Tonogai, Y. N. Analysis of Proanthocyanidins in Grape Seed Extracts, Health Foods and Grape Seed Oils. *Journal of Health Science* **2003***, 49*, 45-54.

(106) Dantuluri, M.; Desai, U. R. Capillary electrophoresis of sulfated molecules. *Recent Res. Dev. Anal. Biochem.* **2003***, 3*, 169-186.

(107) Duchemin, V.; Le Potier, I.; Troubat, C.; Ferrier, D.; Taverna, M. Analysis of intact heparin by capillary electrophoresis using short end injection configuration. *Biomed. Chromatogr.* **2002***, 16*, 127-133.

(108) Geiger, H. *In The Flavonoids. Advances in Research Since 1986; Harborne, J. B., Ed.*; Chapman and Hall: London, **1994**.

(109) Matthews, S.; Mila, I.; Scalbert, A.; Pollet, B.; Lapierre, C.; Herve´ du Penhoat, C. L. M.; Rolando, C.; Donnelly, D. M. X. Method for estimation of proanthocyanidins based on their acid depolymerization in the presence of nucleophiles. *J. Agric. Food Chem.* **1997***, 45*, 1195-1201.

110) Moe, S. T.; Ragauskas, A. J. Structural changes in kraft pulp residual lignin upon peraceticacid treatment. Norwegian University of Science and Technology, Trondheim, Norway, Institute of Paper Science and Technology, GA, USA

(111) Hassi, H. Y.; Chen, C. L.; Gratzl, J. S. *TAPPI Conference*, **1984**.

(112) Gunnarsson, G. T.; Desai, U. R. Designing small, nonsugar activators of antithrombin using hydropathic interaction analyses. *J. Med. Chem.* **2002***, 45*, 1233-1243.

(113) Gunnarsson, G. T.; Desai, U. R. Interaction of sulfated flavanoids with antithrombin: Lessons on the design of organic activators. *J. Med. Chem.* **2002***, 45*, 4460-4470.

(114) Gunnarsson, G. T.; Desai, U. R. Exploring new non-sugar sulfated molecules as activators of antithrombin. *Bioorg. Med. Chem. Lett.* **2003***, 13*, 579-583.
(115) Linhardt, R. J.; Pervin, A. Separation of negatively charged carbohydrates by capillary electrophoresis. *J. Chromatogr., A* **1996***, 720*, 323-335.
(116) Kennedy, J. A.; Jones, G. P. Analysis of proanthocyanidin cleavage products following acid-catalysis in the presence of excess phloroglucinol. *J Agric Food Chem* **2001***, 49*, 1740-1746.
(117) Downard, K. Mass spectrometry: a foundation course. *Cambridge: Royal Society of Chemistry*, **2004**.
(118) Gross, J. H. Mass spectrometry: a textbook.*Berlin; New York: Springer*, **2004**.
(119) Harrow, J. Organic spectroscopy: principles and applications. *U.K.: Alpha Science International Ltd., c2004. London: Imperial College Press; River Edge, NJ: Distributed by World Scientific Pub.,* **2004**.
(120) Ardrey, R. E. Liquid chromatography-mass spectrometry: an introduction. *West Sussex, England; Hoboken, NJ: J. Wiley,* **2003**.
(121) Schalley, C. A. Modern mass spectrometry / volume editor with contributions by Armentrout, P. B. [*et al.*]. *Berlin; New York: Springer,* **2003**.
(122) Nissen, W. M. A. Liquid Chromatography-Mass Spectrometry 2nd Ed. Revised and Expanded. *Chromatographic Science Series Vol. 79.*
(123) Davin, L. B.; Lewis, N. G. Lignin primary structures and dirigent sites. *Curr Opin Biotechnol* **2005***, 16*, 407-415.
(124) Harkin, J. M. V., 1. eds:G.W. *Lignin. In Chemistry and biochemistry of herbage. Butler and R.W. Bailey*; Academic press New York, **1973**; pp 3232-3373.
(125) Oliveira, L.; Evtuguin, D. V.; Cordeiro, N.; Silvestre, A. J.; Silva, A. M. *et al.* Structural characterization of lignin from leaf sheaths of "dwarf cavendish" banana plant. *J Agric Food Chem* **2006***, 54*, 2598-2605.
(126) Onnerud, H.; Palmblad, M.; Gellerstedt, G. Investigation of Lignin Oligomers Using Electrospray Ionisation Mass Spectrometry. *Holzforschung* **2003***, 57*, 37-43.
(127) Reale, S.; Di Tullio, A.; Spreti, N.; De Angelis, F. Mass spectrometry in the biosynthetic and structural investigation of lignins. *Mass Spectrom Rev* **2004***, 23*, 87-126.
(128) Evtuguin, D. V.; Neto, C. P.; Silva, A. M.; Domingues, P. M.; Amado, F. M. *et al.* Comprehensive study on the chemical structure of dioxane lignin from plantation Eucalyptus globulus wood. *J Agric Food Chem* **2001***, 49*, 4252-4261.
(129) Banoub, J. H.; Delmas, M. Structural elucidation of the wheat straw lignin polymer by atmospheric pressure chemical ionization tandem mass spectrometry and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *J Mass Spectrom* **2003***, 38*, 900-903.

(130) Saito, K.; Kato, T.; Takamori, H.; Kishimoto, T.; Fukushima, K. A new analysis of the depolymerized fragments of lignin polymer using ToF-SIMS. *Biomacromolecules* **2005***, 6*, 2688-2696.
(131) Saito, K.; Kato, T.; Tsuji, Y.; Fukushima, K. Identifying the characteristic secondary ions of lignin polymer using ToF-SIMS. *Biomacromolecules* **2005***, 6*, 678-683.
(132) Lundquist, K.; Kirk, T. K. Acid degradation of lignin. IV. Analysis of lignin acidolysis products by gas chromatography, using trimethylsilyl derivatives. *Acta Chem Scand* **1971***, 25*, 889-894.
(133) Lundquist, K.; Hedlund, K. Acid degradation of lignin. *Acta chemica Scandinavica.* **1967***, 21*, 1750-1754.
(134) Lundquist, K.; Lundgren, R. Acid degradation of lignin. *Acta Chemica Scandinavica.* **1972***, 26*, 2005-2023.