Malaysian Journal of Science **20**: 113 - 120 (2001)

Glycolipids chemical informatics

Rauzah Hashim and **Sri Ganesh Janarthanan**

Chemistry Departments, University Malaya, 50603 Kuala Lumpur

ABSTRACT Information on new chemicals and biologically active compounds increase tremendously through the Internet database. Our current project is to build a comprehensive database on Malaysian Natural Product that has vast potentials in medicinal and commercials values. Building a database means development of all the necessary entities to define suitable properties for the Natural Product database. We have used and object-oriented database engine for this information system. From this system we shall also develop a glycolipid chemical information system. Glycolipid forms a class of carbohydrate liquid crystals, commonly found in nature and living cells. Many can be also be synthesized chemically and enzymatically. Their liquid crystalline behavior is mostly unknown; hence their potential as specialty fine chemicals is under exploited. The study of liquid crystal properties of glycolipid is essential both fundamentally and more importantly for their applications.

ABSTRAK Informasi mengenai bahan kimia dan sebatian yang aktif secara biologikal telah meingkat melalui Internet. Sehingga kini projek kita telah memberi tumpuan dalam membina satu pangkalan data komprehensif mengenai Produk Semulajadi Malaysia yang mempunyai potensi luas dalam bidang perubatan dan nilai-nilai komersial. Pembinaan satu pangkalan data terutamanya untuk Produk Semulajadi memerlukan ciri-ciri dan struktur kimia yang sesuai. Oleh yang demikian, kami telah menggunakan satu penggerak (engin) pangkalan data berasaskan objek untuk melaksanakan dalam pembinaan suatu sistem informasi. Daripada sistem ini juga kami akan membina satu pangkalan data untuk Glaikolipid. Glaikolipid sebenarnya membentuk satu kelas yang dipanggil sebagai cecair hablur karbohidrat, yang biasanya ditemui secara semulajadi dan sel-sel hidup. Kebanyakkanya boleh juga disintesiskan secara tindak balas kimia dan enzim. Kelakuan sifat cecair hablurnya kebanyakkannya masih tidak diketahui; maka potensinya sebagai bahan kimia halus istimewa masih dalam proses kajian. Kajian mengenai ciri-ciri cecair hablur bagi glaikolipid adalah penting bagi kedua-dua asas dan paling penting untuk aplikasinya.

(Glycolipids, natural product, informatics systems, liquid crystal, surfactant, colloid)

INTRODUCTION

Development of an Informatics system for Malaysian (hence tropical) natural products integrated with complete physical and chemical information is a major task to be undertaken [1] since the natural diversity is so vast and almost limitless. For plant natural products (for example) there are at least 15000 species have been identified [2]. Out of these only a small sample genus and species have been studied despite numerous research activities focused on the study of natural products in Malaysia and the surrounding countries. Up to the present, the Informatics system comprises of 900 plant families (include of the folklore information); over 2000 bibliographic information; 33,000 chemical entries. Out of these only several hundreds chemical properties have been completely recorded. Obviously the development of the informatics system is a continuous effort. Up to now the decision of which record to store is quite random, for the purpose of sampling the diversity. A long term planning for its development is underway. Currently also the development of this database is extended to include chemical systems such as glycolipids with certain target properties and applications.

Glycolipids are collectively part of a larger family of substances known as glycoconjugates, produced when carbohydrates interact with biomolecules e.g. with lipid group to form glycolipids or with protein to form glycoproteins [3]. Glycolipids are amphiphilic molecules containing hydrophilic (the sugar) components and hydrophobic (i.e. lipid) parts. Glycolipids exist in nature, but can as well be synthesized either chemically or enzymatically [4]. Glycolipids are commonly found in cell membrane. They play an important role in many biological functions such as in cell fusion, transport mechanism and also in intercellular recognition processes [5]. Natural glycolipids can be classified as glycosphingolipids, glycoglycerolipids, and glycosyl phospho-polyprenols. To date there are over 3500 glycolipids compounds (both natural and synthetics) that have been identified and characterized physically and chemically [6].

Determining the chemical structures and properties are important and novel for new compounds. However in the information age, this vast amount of gathered data could be analyzed critically to shed new knowledge and insight, which may be useful in the process of application development. Therefore it is imperative to store information about known glycolipids in a suitable data structure for the intended application.

The Glycolipids Informatics systems is a subset of the Natural Products Informatics and using the same object oriented database engine SciDex (see figure 1).

The biological application of glycolipids is well known and well studied by biologists and biochemists [7] More importantly glycolipids also form a special class of liquid crystal mesophases which display both the character of amphiphilic and thermotropic liquid crystals i.e. they are amphotropic [8]. The high diversity of carbohydrate mesogens has been described where hundreds of different compounds can be formed for example by connecting a dodecyl chain with a monosaccharide by a simple permutation and combination of different diastereomers (glucosides, mannosides, tallosides,...), constitutional isomers (pyranosides [6membered rings], furanosides [5membered rings], acyclic compounds), regioisomers (2-alkyl, 3-alkyl, 4,6-dialkyliden) etc [6]. Therefore glycolipids have the potential of giving rise to a variety of interesting applications [3, 4, 5] both biologically and technologically. For this reason the informatics system we developed focuses on two particular fields of application properties, namely liquid crystals and surfactants.

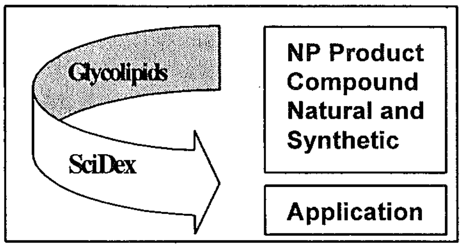


Figure 1. Glycolipid (as one of natural product compounds (NP)) exists in SciDex (database engine).

2. Component of Glycolipids Informatics.

Glycolipids Informatics comprises of several components similar to the Natural Products Information Systems (figure 2), namely:

(a) **Chemical names, formulae and structures** list down glycolipids from both synthetic and natural sources. Both IUPAC and common names are given.

(b) **References/bibliography** lists reported publications on glycolipids in journal, patent, dissertation, conferences, reviews and other resources.

(c) **Chemical information** includes properties such as optical rotation, chemical spectra (H and C-13NMR, IR, UV etc), crystal structure, melting and boiling point, etc.

In addition properties important to applications are also stored and these are:

(d) **Liquid Crystal properties:** such as example magnetic susceptibility, refractive index, dielectric constant, viscosity, diffusion...etc

(e) **Surfactant properties** which are properties of surfactant mixtures.

Figure 3 illustrates the assimilation of glycolipids data into the natural product conceptual diagram. The data mining relates to the statistical techniques used to discover prepositional and associations from data to classification on different accounts.

3. Physical Properties of Glycolipids

The physical property of a substance is also a function of its state (temperature, pressure, volume and amount) [9]. As matter appears in different phases (of which the more common ones are, solid, liquid and gas), properties that depend on the amount of substances in the sample are called *extensive properties*, whereas *intensive properties* are independent of the amount of substance in the sample. Physical properties of a substance describe how the substance obeys natural physical laws due to external perturbations for example mechanical such as gravity and surfaces (i.e. frictional force), thermal, electrical and magnetic forces.

This work attempts to gather physico-chemical properties that have been measured for glycolipid systems. These are broadly categorized into chemical nomenclatures, structures, spectroscopy, thermal, flow, electrical, optical, elastic and magnetic. (See figure 4).

The gathered information both at macroscopic and microscopic levels of glycolipids provides insight and understanding to their liquid crystalline and surfactant nature. It is wise to discuss further on each of these properties. Table 1 describes a general definition of the physical properties liquid crystal and their attributes. These attributes are recorded in the database. Figure 5 illustrate an example output from the glycolipid informatics.

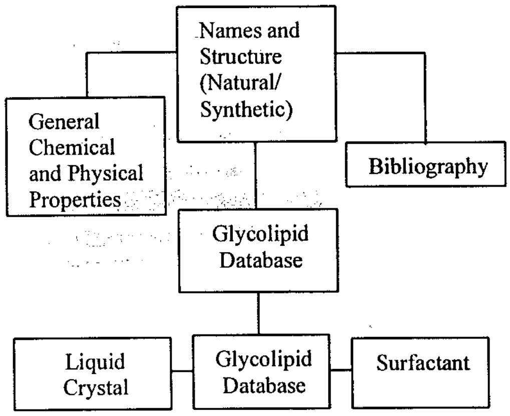


Figure 2. Components of Glycolipid

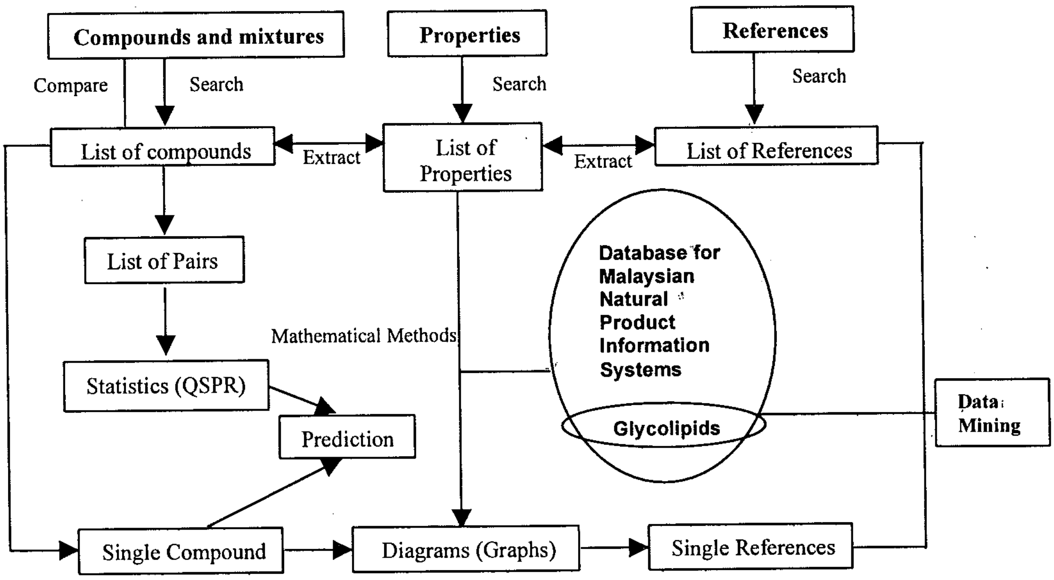


Figure 3. Glycolipids in the Malaysian Natural Product Database Conceptual Diagram

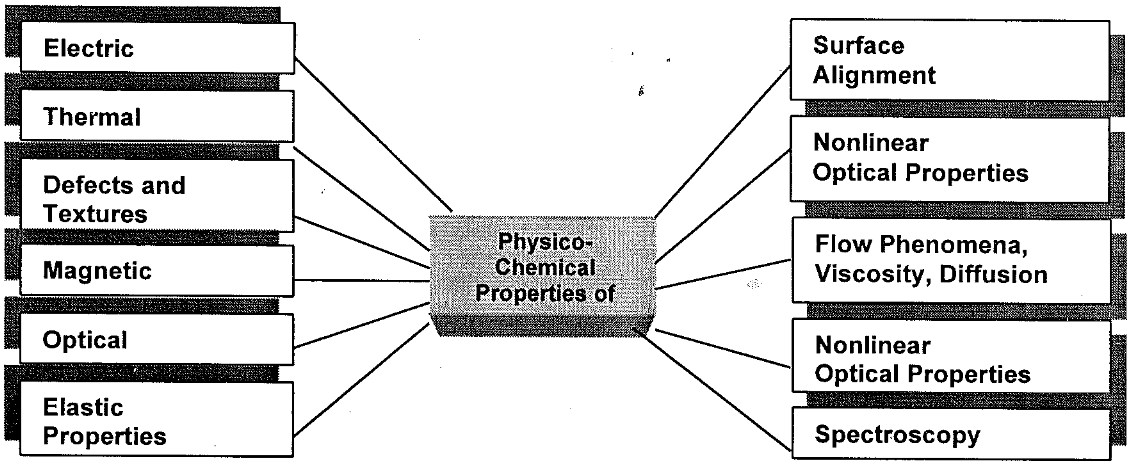


Figure 4. Classification of physical properties according to the nature of external perturbation.

Physical Properties of Surfactant and Colloids

The ideal surfactant would be infinitely soluble in water at all temperature, has a very good wetting and detergency at low concentrations and does not corrode or otherwise affect the parts being cleaned. These characteristics are all controlled by the physical parameters of the surfactant molecule. For example, the linear alcohol nonionic surfactants can be characterized by a number of physical parameters including the hydrophilic-lipophilic balance (HLB), ethoxylation Level (EO), critical micelle concentration (CMC) and cloud point. These properties are well documented in the literature and briefly these are:

i. Ethoxylation Level (EO)

Ethylene oxide (EO) may be added to improve the detergency property of linear alkyl surfactant [10]. Ethoxylation level (EO level) measures the amount of ethylene oxide added as a ratio of the amount of EO and alcohol groups (in mole per mole). The ethylene oxide chain is the water-soluble portion of the surfactant molecule. Therefore a high EO-level results in a higher water-solubility [11].

ii. Chain Length (CL)

Refers to the number of carbon atoms contained in the longest direct line in the surfactant molecule.

iii Cloud Point

Refers to the temperature where an anionic surfactant becomes insoluble in the solution. They are usually determined at a 1% (w/w) concentration.

iv. Critical Micelle Concentration (CMC)

A micelle is a colloidal aggregate of surfactant molecules. The critical micelle concentration (CMC) of a surfactant is the concentration at which micelles begin to form. At this point, several properties of the solution change abruptly, including surface tension and solubilization. The surface tension decreases with increasing surfactant concentration until the CMC is reached, at which point it levels off. Solubilization remains fairly constant with increasing concentration of surfactant until the CMC is reached. Then it increases rapidly. Detergency also increases with increase in concentration of surfactant but not as much as solubilization.

v. Hydrophilic-lipophilic balance value (HLB)

This is a concept for choosing emulsifiers. Hydrophilic refers to the portion of the surfactant that is soluble in the aqueous phase, while liphophilic refers to the oil soluble portion of the surfactant. The value of HLB ranges from 1-20. Low HLB emulsifiers are soluble in the oil while high HLB emulsifiers are soluble in water. Bancroft's rule tells us that the type of emulsion (i.e. oil in water or water in oil) is dictated by the emulsifier ad that the emulsifier should be soluble in the continuous phase, (Bancroft, 1913). Low

HLB emulsifier's are soluble in oil and give rise to water in oil emulsions. The value of HLB is calculated by:

$$HLB = 20(1-(S/A))$$

Where S is the saponification number of the ester and A is the acid number of the resulting acid. The saponification number is the mass in mg of potassium hydroxide required to saponify one gram of the acid plus ester. The acid number is the number in mg of potassium hydroxide required to neutralize one gram of acid plus ester

Table 1. Physical Properties of Liquid and their attributes

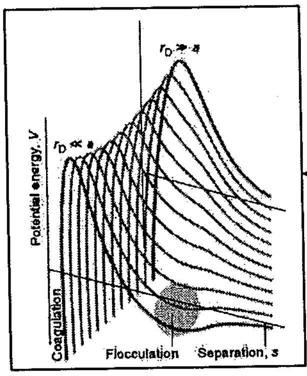
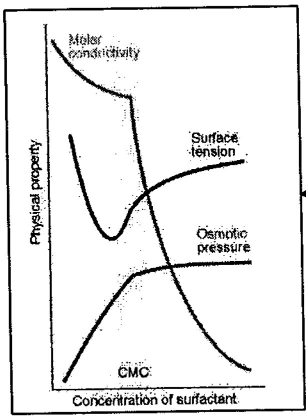
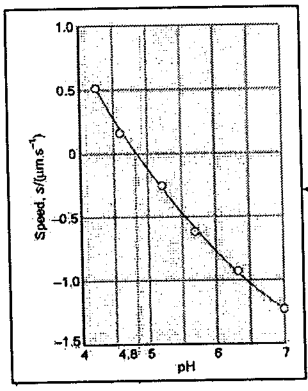
Physical Properties (L.C)	Attributes	Subfield (suitability)
Magnetic Properties Magnetic field introduced to macroscopic anisotropy	Magnetic susceptibility, χ Diamagnetic, $\kappa^{dia}//$ Paramagnetic, $\bar{\kappa}^{para}$ Strength of the magnetic field, **M** Magnetic Induction, **B**	Field Method Solvent Temperature
Optical Properties Response to high frequency electromagnetic radiation	Wavelength of selective reflection, **£** Refractive index, n_i Optical absorption, **A** Optical activity, α Polarizibility, **D** or **L** Induced polarization **P**	Value Wavelength Solvent Concentration
Electric Properties Measurement of the response of a charge free system to an applied electric field	Applied electric field, **E** Relative permitivity, ε_r Dielectric constant, ε Electric dipole moment, μ Electric conductivity, σ	Field Method Solvent Temperature
Elastic Properties Applied static stress to the strain or deformation produced in the material Intermolecular interaction	Stress tensor, $\sigma_{\alpha\beta}$ Strain tensor, $e_{\alpha\beta}$ Torsional elasticity, $\gamma\delta$ Director order parameter $\langle a \rangle$ Elastic constant, σ_I	Method Solvent Temperature Pressure
Flow Phenomena and Viscosity Related to hydrodynamics of L.C.	Shear Viscosity Coefficient $\eta_1, \eta_2, \eta_3, \eta_{12}$ Rotational Viscosity coefficient, $\gamma 1$	Method Solvent Temperature Pressure
Surface Alignment Phenomenon of orientation of liquid crystal by surfaces	Interfacial region of thickness, ξ_1 Anchoring energy, γ_s	Method Solvent Temperature Pressure
Diffusion Spreading of one species to another as a molecular mass transport by thermally activated particles motions	Diffusion sensor or diffusivity, **D**	Method Solvent Temperature Pressure

Table 2. Components of Glycolipids Informatics

Physical properties	Surfactant A (810-60)	Surfactant B (91-6)	Surfactant C (25-12)
CL	C_8-C_{10}	C_9-C_{11}	C_{12}-C_{15}
EO	5	6	12
HLB	12	12.5	14.4
CMC (% vol)	0.007	0.02	0.0009

In the example tabulated above, optimum surfactant properties were listed. These experimental values were chosen to see the affect they have on precision cleaning. The surfactants listed were chosen to cover a reasonable range in each of the physical properties to be tested, as opposed to changing one property at a time [11].

Table 3. Examples of surfactant properties [9]

	Description
	**A.** The potential energy of interaction as a function o the separation of the centers of the two particles and it variation with the ratio of the particle size to th thickness a of the electrical double layer r_D. The region labeled coagulation and flocculation show the dips i the potential energy curves where these processe occur.
	**B.** The typical variation of some physical propertie (e.g. molar conductivity, surface tension, osmoti pressure of an aqueous solution of sodiu dodecylsulfate close to the critical micelle concentratio (CMC).
	**C.** The plot of drift speed against pH by which th isoelectric point of a macromolecule can be determined it corresponds to the pH at which the drift speed in th presence of an electric field is zero.

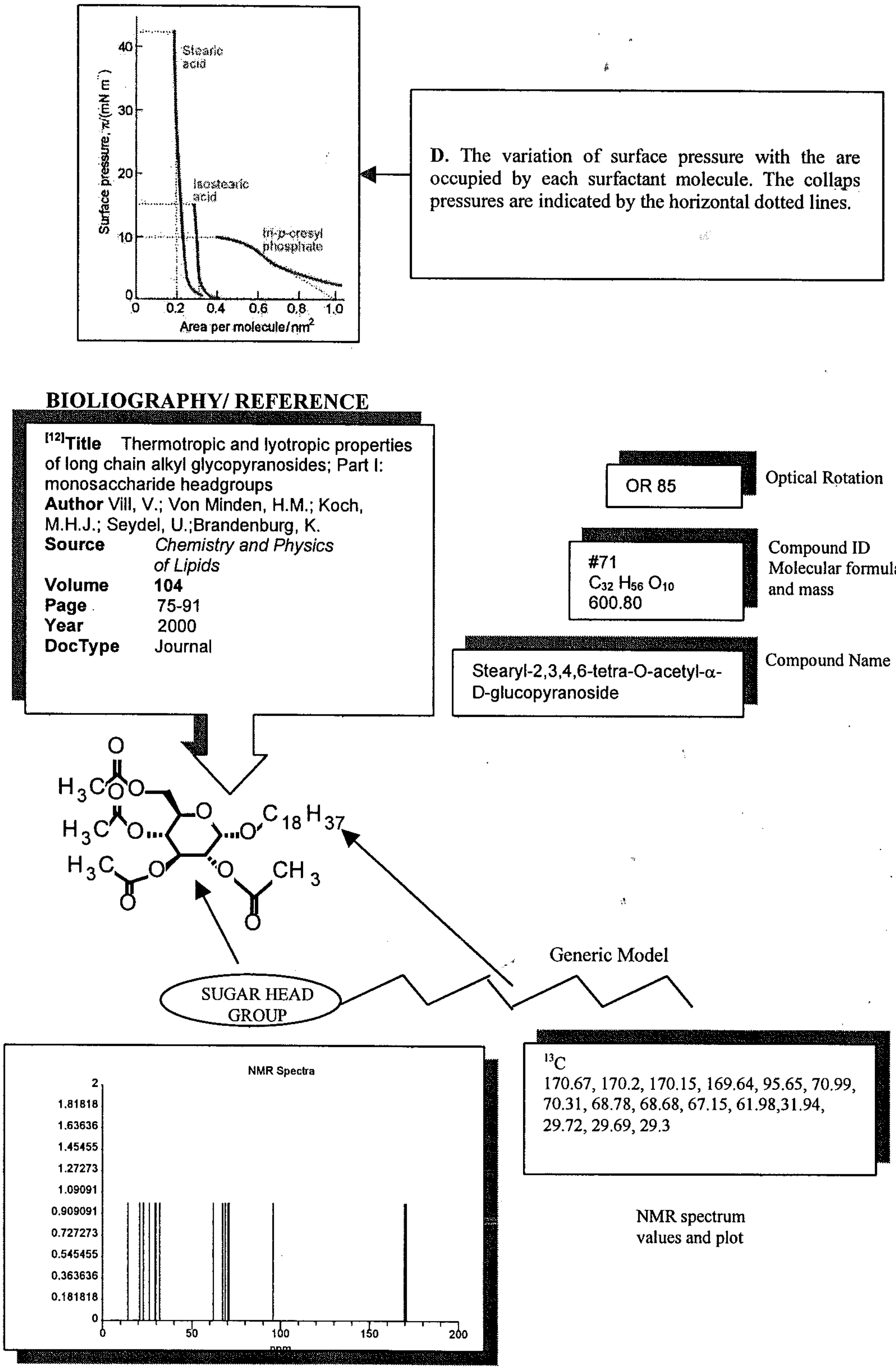


Figure 5. An example output from glycolipids informatics

CONCLUSION

Cheminformatics is not a newly developed area of scientific endeavor. It covers both: the modeling and simulation aspects; either using atomistic (or generic models) to represent the chemical system or gathered and stored information of the chemical system obtained from experimental measurements. Both methods allow new knowledge to be derived about the chemical systems. In modeling and simulation average behaviors are predicted (or calculated) from mathematically formulated equations and laws for example the Newton laws of motion. In the later, insight about the systems is obtained by manipulating voluminous data using statistical analysis and chemometric techniques.

In present the work we focus on the gathering of experimental of glycolipids. The often complex and contradicting requirements of industrial applications demand mixtures of chemical to be used. In the case of simple TN-LCD, no less 20 different chemicals systems to be used for a particular LCD materials. Hence it is not too difficult to realize the importance of an informatics system to the development of any smart materials. The storage, manipulation, presentation and application of chemical information lie at the heart of contemporary scientific research. All chemical, fine chemical, biochemical, agrochemical, and pharmaceutical industry are increasingly reliant upon electronic data management and application both 'in house' and for interfacing with external information sources.

REFERENCES

1. Sri Ganesh Janarthanan; Rauzah Hashim, Development of an Integrated Chemical Information System of Natural and Chemical Resources of Malaysia, The Proceedings of Asia Pacific Advanced Network Conference 2001, Penang, August 20-23rd, 2001,.
2. http://www.arbec.com.my/biotech. htm.
3. Allen HJ and Kisailus EC. Glycoconjugates: composition, structure and function. Marcel Dekker Inc USA; 1992.
4. Zhang Z, Fukunaga K, Sugimura Y, Nkao K., and Shimizu T. Synthesis of glycolipids: dialkyl-N-[N-(4-lactonamidobutyl) succinamoyl]-L-glutamates. Carbohydrates Research 1996; **290**: 225-232.
5. Lochmit G, Geyer R, Heinz E, Rietschel ET, Zähringer U and Müthing J. Glycolipids and Glycosphingolipids: Chemical Biology and Biomedicine, Plant Glycolipids, in Glycoscience: Chemistry and Chemical Biology. Edited by Fraser-Reid B, Tatsuta K and Thiem J. Springer-Verlag, Berlin; 2001; **111**: 2183-2252.
6. Vill V. and Hashim, R; Carbohydrate Liquid Crystals: structure property relationship of thermotropic and lyotropic glycolipids. Submitted for publication to Curr Opin in Colloid and Interface Science 2002.
7. Goodby JW. Liquid crystals and life. Liq Cryst 1998; **24(1)**: 25-38.
8. Baron M (Ed). Definition of basic terms relating to low molecular mass and polymer liquid crystal, IUPAC recommendation 2001. International Union of Pure and Appl Chem 2001; **73**: 849-895.
9. Atkins P.W., Physical Chemistry, 6th edition, Oxford University Press;
10. http//www.mastey.com/avoid1.asp
11. Kenneth R. Monroe, Elizaberth A. Hill, K. David Carter, Surfactants Parameters effects on Cleaning Efficiency, Surfactant Cleaning Technology Program, Research Triangle Program.
12. Vill, V., Von Minden, H.M., Koch, M.H.J., Seydel, U., Brandenburg,K, Thermotropic and lyotropic properties of long chain alkyl glycopyranosides; Part I:. monosaccharide head groups, Chemistry and Physics of Lipid, 2000, **104**: 75-91.