CHAPTER-1

Introduction

1.1 Introduction

Natural products are important source of biologically active molecules; more than 50% drug molecules are either natural isolates or synthetic but inspired from the natural product. The humans have used plant products since ancient times to treat the diseases. It has revealed that humans were using plants to cure diseases back at least 60,000 years [1]. The traditional knowledge can be utilized to overcome several medical problems of present generations. Particularly, plants have served as the foundation for highly developed traditional medical systems; the earliest known accounts, from approximately 2600 BCE, describe the uses of about a thousand plant-derived substances in Mesopotamia. These include the oils of the following: *Commiphora species* (myrrh), *Glycyrrhiza glabra* (licorice), *Cedrus species* (cedar), *Cupressus sempevirens* (cypress), and *Papaver somniferum* (poppy juice). These are still used today to treat a variety of illnesses, from inflammation and parasitic infections to coughs and colds. Recording more than 700 medications, the majority with a plant basis. Similarly, the Indian Ayurveda system was documented prior to 1000 BCE (Charka; Sushruta and Samhitas, containing 341 and 516 medicines, respectively). Plant-based systems continue to play an essential role in healthcare, and their use by different cultures has been extensively documented. Ayurveda, Unani, Chinese and Siddha are some well documented sources of traditional medicine systems. Apart from the plants, microbes have been explored for the discovery of various antibiotics. The modern natural product chemistry started with the isolation of morphine from opium (1804). Many similar developments led the discovery of various important drug molecules like atropine, quinine, codeine, reserpine and taxol. The importance of natural products in modern drug discovery has been extensively reviewed and it was found that out of the new drugs introduced between 1981 and 2019, 33% of the 1394 small molecule approved drugs were either natural products or derived from natural products and another 35% were created around a pharmacophore from a natural product [2]. Several approaches have been introducing to explore the natural products for the drug discovery. NPs offer characteristic features that makes them drug like molecules in comparison with conventional synthetic molecules. The chemical diversity of occurring molecules falls within the criteria which is suitable for modern drug discovery as follows. More scaffold diversity and and structural complexity, higher numbers of H-bond acceptors and donors, lower calculated octanol–water partition coefficients (cLogP values, indicating higher hydrophilicity) and greater molecular rigidity compared with synthetic compounds [3]. The emerging fields such as genomics, bioinformatics, mass spectrometry, proteomics, transcriptomic, metabolomics, and gene expression, and network pharmacology has extensive applications in the discovery and development of natural products [4-6].

1.2 Natural products as a potential source of drug

The World Health Organization (WHO) estimated in 1985 that approximately 65% of the population of the world predominately relied on plant-derived traditional medicines for their primary health care. Only about 6% of the roughly 300,000 species (some estimates put that number as high as 500,000 species) of higher plants have been thoroughly studied pharmacologically, and 15% have been studied phytochemicals, despite the extensive study of terrestrial flora [2, 7]. Some of natural product use as drugs including vinblastine **(1.1)** and vincristine **(1.2)** from *vinca rosea*, podophyllotoxin **(1.3)** from *podophyllum peltatum Linn*, cocaine **(1.4)** from *Erythoxylon coca*, codeine **(1.5)** from *opium poppies*, digitoxin **(1.6)** from *Digitalis purpurea L*, quinine **(1.7)** from *Cinchona officinalis* and pilocarpine **(1.8)** from *Pilocarpus Microphyllus*. Apart from plants, the microbes are the excellent source of the antibiotics like penicillin G and other semi-synthetically developed beta-lactam antibiotics. These discoveries revolutionised microbial natural products-based drug discoveries that led the discovery of streptomycin **(1.10)**, chloramphenicol **(1.11)**, chlortetracycline **(1.12)**, erythromycin **(1.13)** [2, 8]. Not only the antibiotics microbes have been well explored to discover anticancer molecules.

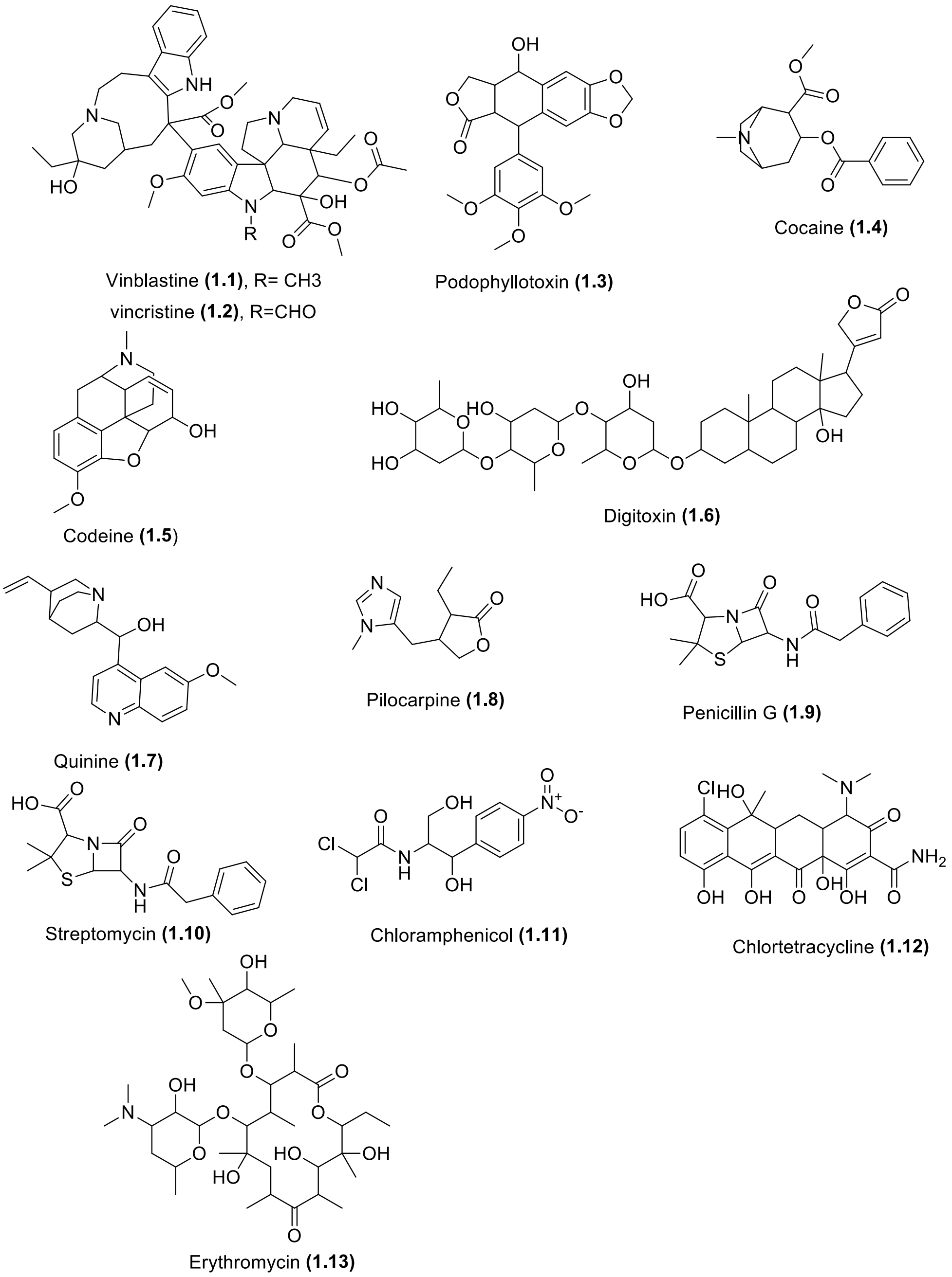


Figure 1. 1 Chemical structures of plant and microbial derived natural products used drug.

1.3 Challenges and approaches associated with natural products-based drug discovery

The plant secondary metabolites obtained from the plants, microbes, and marine sources are often biologically active and they can be considered as environmentally friendly in contrast synthetic molecules. Major underline challenges with natural product-based drug discovery have been reviewed [9, 10]. Building up and maintaining a high-quality natural product library requires a skill. Natural products are often biosynthesized in small quantities and present as mixtures in extracts, which require labour-intensive and time-consuming purification procedures. Obtaining further quantities for preclinical development requires large-scale reacquisition or fermentation that would have a substantial impact on the development timeline. Rediscovery of known compounds is a major problem when screening natural product libraries. This is caused by lack of efficient dereplication methodologies. Natural products are often structurally complex. Modification of complex natural products using organic chemistry is frequently challenging.

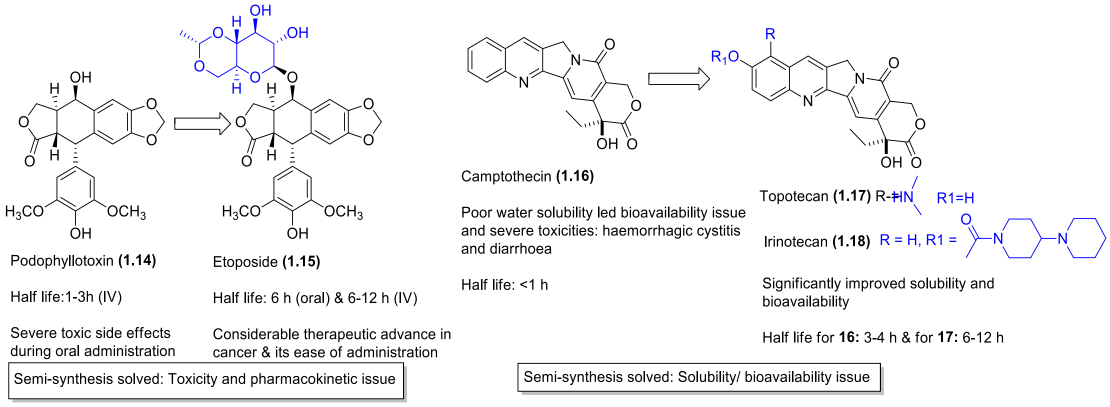
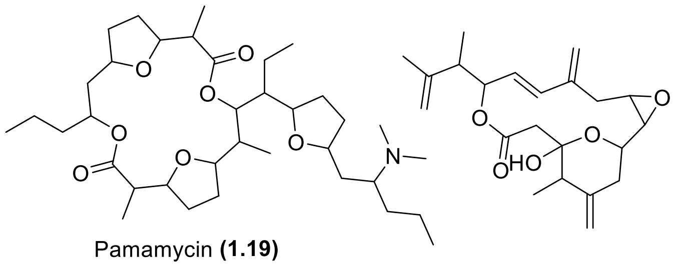


Figure 1. 2 Semi-synthetically developed drug candidates

Additional critical issues in their therapeutic uses, including supply issues, pharmacokinetic properties, drug resistance, and toxicity. These natural molecules are often considered favorable candidates for drug development due to their drug-likeness. Semi-synthetic modification has been adopted to quickly access large number of drugs like molecules around the bioactive scaffold [11, 12]. These approached led the several successful drug molecules such as topotecan, azithromycin, and etoposide.

Total synthesis also offers an alternative approach to synthesize natural products like molecules. A lot of natural products are only obtainable in insufficient amounts from their original sources. Their consistent supply and reliable access is essential for converting their biological activity into useful applications [13, 14]. Therefore, one of the most difficult tasks facing synthetic chemists today is the development of effective synthetic routes for the scalable synthesis of complex natural products. Natural product total synthesis, is a rapidly developing field. The diversity-oriented synthesis (DOS) also an effective approach in which a suitable intermediate can be utilized to divert the synthetic route to new series of natural product like molecules. This is alternative approach to rapidly generate the SAR and further identification of the lead [13, 14].



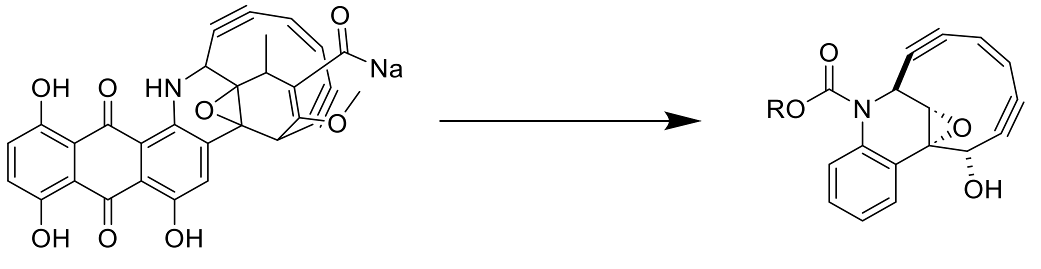
Pamamycin **(1.19)**

Amphidinolide p **(1.20)**

Figure 1. 3 Example of total synthesis

Predication of drug like property is also offering a way for the selection of natural products for the drug discovery. The natural products (NPs) pharmacophores have encouraged scientists to use the drug discovery process, which includes computational evaluation of safety, pharmacokinetics, and pharmacodynamics parameters [11].

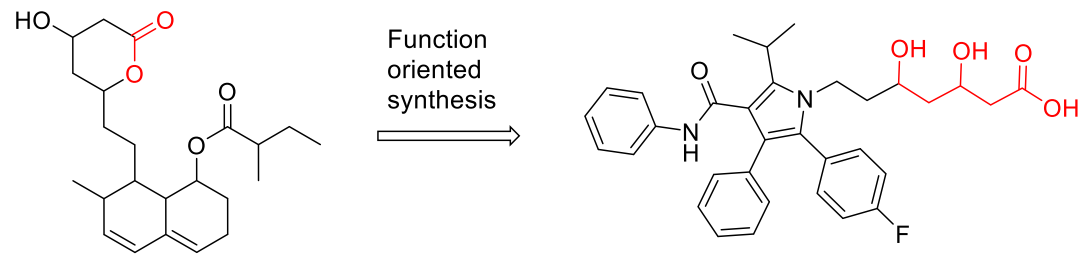
Computational tools help in the evaluation of such parameters based on the understanding of the physicochemical properties including like partition coefficient (log P), distribution coefficient at pH=7.4 (log D), topological polar surface area (TPSA), molecular weight (MW), aqueous solubility (log S), number of hydrogen bond acceptors (HBA), number of hydrogen bond donors (HBD) and number of rotatable bonds (nRot). These parameters are crucial for understanding physicochemical properties, ADME (absorption, distribution, metabolism, and elimination) attributes, and *in-silico* toxicity profile [15, 16]. Some free web resources for assisted drug-likeness analysis have emerged in recent years. For example, the Virtual Computational Chemistry Laboratory (VCCLAB) is a platform to calculate molecular properties and analysis data. The more recent Swiss ADME is a free web tool to evaluate molecular physicochemical properties [17]. Functional oriented synthesis also an excellent approach to generate simplified but effective analogues for the drug discovery. Some examples of FOS are as follows; Classical examples is discovery of Dynemicin showed significant biological profile for the drug discovery however it cannot be supplied on scale from natural sources or synthesis due to its scarcity and complexity. But as simplified analogue **1.22** showed equivalent biological activity after depriving the nonessential functionality of the dynemicin **(1.21)**.



Dynemicin (**1.21**)

Simplified analog (**1.22**)

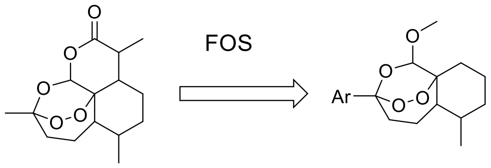
Figure 1. 4 Dynemicin and a simplified function analogue



Natural statin, Mevastatin (**1.23**) Lipitor (**1.24**)

Figure 1. 5 Design of Lipitor from the natural occurring compound

The highly complex molecules known as artemisinin, currently undergoing clinical trials for the treatment of cancer, artemisinin also exhibits synergistic effects with other antitumor agents. Additionally, intriguing new research has demonstrated that artemisinin can improve learning and memory in animals, which may have treatment implications for Alzheimer's disease. Taking advantage of FOS a new but equally potent and simplified analogues gas been developed [18, 19].



Artemisinin **(1.25)** Simplified analog **(1.26)**

IC_{50} = 9.5 nM IC_{50} = 15 nM

Figure 1. 6 Example of function-oriented synthesis

frequently encounters during the isolation and characterization of the natural products. Natural product drug discovery relies on the new chemical entity and frequent isolation of known metabolite is frustrating. Dereplication approaches have been proposed for the rapid identification of a known secondary metabolite at early stage of the discovery process e.g., extract. Dereplication helps the identification of known metabolites at the early stage such as extraction and fractionations, so that efforts can be utilized for the targeting new compounds. The term hyphenation refers to an online coupling of liquid chromatography (LC) or gas chromatography (GC) isolation technique, and one or more spectroscopic detection techniques, e.g., ultraviolet-visible (UV-Vis), infrared (IR), mass (MS), or nuclear magnetic resonance (NMR) spectroscopy. Nowadays, a number of sophisticated hyphenated techniques, e.g., GC-MS, LC-PDA, LC-MS, LC-FT-IR, LC-NMR, LC-NMR-MS, and CE-MS are available for dereplication and other purposes. This accelerates the entire progression of natural product-based drug discovery drastically. Dereplication is strategy employing spectral databases, informatics, intelligent correlation software, and hyphenated analytical apparatuses. Usually, this technique employed different steps: 1.) Chemical profiling of extract or bioactive fraction using LC-hyphenated technique (HPLC/UPLC-detection techniques; PDA, MS, HR-MS, NMR, etc), for example, HPLC coupled with PDA; give UV information of chromophore which is an important aspect of basic scaffold of compound, similarly, MS, HR-MS and MS/MS provide mass information which can be useful for the information of known and unknown compounds, 2.) Informatics or database [20-25]. These databases can be searched with minimal amounts of structural, physical, and/or biological data. Following is some examples of the databases; Dictionary of Natural Product (2.5 lac, natural products), MarinLit (approx 40,000 marine compound). and StreptomeDB (10,000 microbial isolate) etc. using similar strategy various new compound has been isolated [26].

1.4 LC-MS-based profiling to exploring the natural products

LC-MS is the most widely used system, in which HPLC (or UPLC) is coupled with MS detector, which allows the separation and detection on time dependednt manner. The separated molecules from LC column enters MS detector and get fragmented. This provides multiple informations together such as retention time, M/Z values, MS fragments, molecular formula derived from accurate or high resolution mass. Thus the mass spectra obtained contains the indispensable information about the separated molecule. Metabolite profiling focuses on the analysis of metabolites also known as chemo-profiling which may be related to a specific class of compounds. Metabolite profiling is the oldest and most established approach and is considered for the metabolomics. It is untargated metabolomics usually prefered for molecular networking. In molecular networking, metabolites present in extracts or fraction can be annotated following LC−MS/MS analysis, ultimately linking molecules into clusters based on common MS/MS fragmentation ions in individual analytes. These networks can be created using the online platform available freely to users via the University of California San Diego Global Natural Products Social Molecular Networking (GNPS) website [26, 27]. Usually in LC-MS-based dereplication, even if the molecule is known, its structure can be reconstructed by analyzing the mass spectra or it can be searched in the databases. Very less sample size is required for the LC-MS based analysis. LC-MS is very sensative nano to picogram quantity can be seen analysed. Generally, soft ionization techniques are used in MS, which produces molecular ion peaks. However, tandem mass spectrometry (MS-MS) produces collision-induced fragments of the produced molecular ion. Hyphenated techniques such as LC-UV, LC-MS have been extensively used in combination with biological screening for the discovery of natural products. Following are the some examples where this techniques have been utilized for the drug discovery and isolation of new meabolites. In recent, Network pharmacology is becoming a cutting-edge research field in current drug discovery and drug development. The targeting and network pharmacology is decorated as the new model in drug discovery. It presents strategies to search novel therapeutic predictions of natural products as drugs, where there are multiple compounds responsible for cumulative activity and synergy [35].

Table 1. 1 Novel or known compounds isolated by using LC-MS based dereplication strategies

Identified chemical classes and natural sources	Analysis	References
Alkaloids from *Cimicifuga racemosa* (black cohosh)	LC-MS	[28]
Coumarins from *Kielmeyera albopunctata* (bark)	LC-MS	[29]
Luteolin, quercetin and Kaempferol from *Punica granatum* (peel)	LC-MS	[30]
16,23-Epoxycucurbitacin derivatives from *Eleaocarpus chinensis*	LC-MS	[31]
Bioactive Flavaglines from *Aglaia perviridis*	LC-MS	[31]
Caldoramide, from the Marine Cyanobacterium *Caldora penicillata*	LC-MS	[32]
Turonicin A, a from an Australian Streptomyces sp.	LC-MS	[33]
Calothrixamides A and B from the Cyanobacterium Calothrix sp.	LC-MS	[34]

A notable feature of Ayurvedic drug is the use of herbal formulae. Understanding the mechanisms of action and combinatorial rules of herbal formula is of great significance in Ayurveda modernization and is also one of the important steps in modern drug discovery. The developing network pharmacology offers a unique chance to explore systematically not only the molecular complexity of an herbal formulae, but also the molecular relationships between an herbal formula and complex diseases. Network pharmacology is probable to be a new approach and influential tool to find the active metabolites as well as their potential targets from many herbs or herbal formulae [36]. A computational framework, based on a regression model that integrates human protein–protein interactions, disease phenotype similarities, and known gene–

phenotype associations to capture the complex relationships between phenotypes and genotypes, has been proposed to generate the network [36]. Several researchers are using network pharmacology combined with LC-MS to explore the mechanism of complex herbal formulae or poly-herbal compositions.

Current these will utilize the LC-MS as a tool to explore the phytochemicals of the selected plant material. Which may be further subjected to systematic dereplication for the identification of the known metabolites, to target the new metabolite. Identified metabolites will be subjected to network pharmacology for their mechanism of action and computational studies to support it. Extensive purification followed by detailed structure elucidation and preliminary cytotoxicity studies will be done to validate.