1. Introduction

Generally, biomarkers are naturally occurring, ubiquitous and stable complexes that are objectively measured and evaluated as an indicator of a certain state. It is used in many scientific fields; medicine, cell biology, exposure assessment, astrobiology, geology and petroleum.

Biomarkers "biological markers" in medicine are complex compounds that can be used as an indicator of a particular disease state or some other physiological state of an organism. Biomarkers have been defined also as cellular, biochemical or molecular alterations that are measurable in biological media such as human tissues, cells or fluids (Hulka et al., 1990). Broader definitions include biological characteristics that can be objectively measured and evaluated as an indicator of normal or abnormal biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention (Naylor, 2003).

Biomarkers can also reflect the entire spectrum of disease from the earliest manifestations to the terminal stages. Characterization of the healthy and diseased cells, when identifying a specific biomarker as an indicator of cancer, is a needed research strategy for validating biomarkers. For the nervous system, as an example, there is a wide range of techniques used to gain information about the brain in both the healthy and diseased state. These may involve measurements directly on biological media (e.g. blood or cerebrospinal fluid) or measurements such as brain imaging which do not involve direct sampling of biological media but measure changes in the composition or function of the nervous system (Mayeux, 2004). Moreover, the pharmaceutical industry is beginning to rely on biomarkers information and their importance to the future of drug discovery and development.

In practice, biomarkers science includes tools and technologies that can aid in understanding the prediction, cause, diagnosis, progression, regression, and/or the outcome of disease treatment which provide a dynamic and powerful approach to understand the spectrum markers and offer the means for homogeneous classification of a disease and risk factors which can extend the base information about the underlying pathogenesis of disease.

2. Petroleum biomarkers

Due to the variety of geological conditions and ages under which oil was formed, every crude oil exhibits a unique biomarker fingerprint. Crude oils compositions vary widely depending on the oil sources, the thermal regime during oil generation, the geological migration and the reservoir conditions. Crude oils can have large differences in:
1- distribution patterns of the n- alkanes, iso-alkanes and cyclic- alkanes as well as the unresolved complex mixture (UCM) profiles
2- relative ratios of isoprenoid to normal alkanes
3- distribution patterns and concentrations of alkylated polynuclear aromatic hydrocarbons (PAHs) homologues. Most of these constituents undergo changes in their chemical structure by time as an effect of several factors among which are the biodegradation and weathering. Relative to other hydrocarbon groups in oil, there are some compounds that are more degradation-resistant in the environment as for example; Pristane, phytane, steranes, triterpanes and porphyrins. These undegradable compounds are known as Biomarkers.

Trebs (1934) was the first one to develop the biomarkers concept, with his pioneering work on the identification of porphyrins in crude oils suggesting that these porphyrins are generated from chlorophyll of plants. Blumer et al. (1963) and Blumer & Thomas (1965) isolated pristane from recent marine sediments and concluded that it was derived from the phytol side chain of chlorophyll. Later, other workers reported the present of various classes of degradation-resistant organic compounds and recognized their biomarker implementations.

Petroleum biomarkers can thus be defined as complex organic compounds derived from formerly living organisms found in oil (Mobarakabad et al., 2011). They show little or no changes in their structure from the parent organic molecules and this distinguishes biomarkers from other compounds (Maioli et al., 2011). Various biomarkers formed under different geological conditions and ages can occur in different carbon ranges exhibiting different biomarker fingerprints.

From the identification point of view, biomarkers are the most important hydrocarbon groups in petroleum because they can be used for chemical fingerprinting which provides unique clues to the identity of source rocks from which petroleum samples are derived, the biological source organisms which generated the organic matter, the environmental conditions that prevailed in the water column and sediment at the time, the degree of microbial biodegradation and the thermal history (maturity) of both the rock and the oil. The information from biomarker analysis can be used also to determine the migration pathways from a source rock to the reservoir for the correlation of oils in terms of oil-to-oil and oil-to-source rock and the source potential. Also chemical analysis of biomarkers generates information of great importance to environmental forensic investigations in terms of determining the source of spilled oil, differentiating and correlating oils and monitoring the degradation process and weathering state of oils under a wide variety of conditions.

Terpanes and steranes are highly resistant to biodegradation but few studies have shown that they can be degraded to certain degree under severe weathering conditions i.e, extensive microbial degradation (Chosson, 1991).

2.1 Biomarkers analysis

The commercial availability of Gas Chromatography- Mass Spectroscopy (GC-MS) and associated data systems in the mid-1970s led to use the biomarkers for a wide variety purposes. The complex structure of biomarkers and the possible presence in low concentrations make a pressing need for more sensitive and precise analysis. The development of analytical methodologies and the combination between these methods are of great importance to separate, monitor and detect the absolute concentrations and structure of petroleum biomarkers. The use of "hybrid" or "hyphenated" techniques, which are a combination of different separate techniques, increases the analytical power of the
used methods. GC-MS can be considered as the most popular method used in the characterization of major biomarker groups. GC provides the significant advantage of the separation of different structures of biomarkers while MS can accurately detect and identify these structures. The concept and the development of these instrumentations will be briefly mentioned.

2.1.1 Separation by chromatographic techniques

Chromatography is the separation of a mixture of compounds into their individual components primarily according to their volatilities. There are numerous chromatographic techniques but gas chromatography (GC) is the most important one. It has a number of advantages over other separation techniques. It can identify (qualitate) and measure the amount (quantitate) various sample components.

![Fig. 1. Schematic diagram of gas chromatograph.](image)

The GC column is the heart of the system; the structure of the stationary phase and the packed material greatly influence the separation of the compounds and affect the time of separation (retention time). Two types, packed and capillary columns, have been used. The advantages in capillary columns over packed columns are in obtaining practically improved resolution in order to give fine structured chromatographic fingerprints. The column is placed in an oven where the temperature can be controlled very accurately over a wide range of temperatures. As compounds come off the column, they enter a detector for identification. Figure 2 represent the carbon number range distribution of common hydrocarbons in crude oil and petroleum products.

2.1.2 Identification by spectroscopic techniques

There are different types of detectors that can be employed depending on the compounds to be analyzed. Mass spectrometry (MS) has very common use in analytical laboratories that study a great variety of compounds and provides a satisfactory tool for obtaining specific fingerprints for classes and homologous series of compounds resolved by gas chromatography. The technique has both qualitative and quantitative uses include identifying and determining the structure of a compound by observing its fragments.

The mass spectrometer has long been recognized as the most powerful detector for gas chromatography. Typical MS instruments consist of three modules; an ion source: which can
convert the separated constituent into ions, a mass analyzer: which sorts and separates ions by their masses by applying electromagnetic fields and a detector: which calculate the abundances of each ion present by a quantitative method to generate signals. The size of the signals corresponds to the amount the compound present in the sample.

Characterization of some major biomarker groups is largely achieved using the following MS fragment ions:

- alkyl-cyclohexanes: $m/z$ 83
- methyl-alkyl-cyclohexanes: $m/z$ 97
- isoalkanes and isoprenoids: m/z 113, 127, 183
- sesquiterpanes: m/z 123
- adamantanes: m/z 135, 136, 149, 163, 177, and 191
- diamantanes: m/z 187, 188, 201, 215 and 229
- tri-, tetra-, penta-cyclic terpanes: m/z 191
- 25-norhopanes: m/z 177
- 28,30-bisnorhopanes: m/z 163, 191
- steranes: m/z 217, 218
- 5α(H)-steranes: m/z 149, 217, 218
- 5β(H)-steranes: m/z 151, 217, 218
- X diasteranes: m/z 217, 218, 259
- methyl-steranes: m/z 217, 218, 231, 232
- monoaromatic steranes: m/z 253
- triaromatic steranes: m/z 231

The m/z 191 fragment is often the base peak of mass spectra of biomarkers. In general, GC/MS chromatograms of terpanes (m/z 191) are characterized by the terpane distribution in a wide range from C19 to C35 with C29 αβ- and C30 αβ-pentacyclic hopanes and C23 and C24 tricyclic terpanes being often the most abundant. As for steranes (at m/z 217 and 218), the dominance of C27, C28 and C29 homologues. Figure 4, as a representative model, shows

[Images of GC-MS chromatograms]

API: American Petroleum Institute. The larger the API, the greater the amount of the light components the oil contains and with decreasing the API, the amounts of medium and heavy weight components increase.

Fig. 4. GC-MS chromatograms at m/z 191 for light to heavy crude oils.
GC-MS chromatograms at m/z 191 for light (API > 35), medium (API: 25–35), and heavy (API < 25) crude oils (Wang et al., 2006).

3. Diagnostic ratios of biomarkers

Biomarker diagnostic parameters have been long established and are widely used by geochemists for oil correlation; determination of organic input and precursors, depositional environment, assessment of thermal maturity and evaluation of in-reservoir oil biodegradation (Peters et al., 2005). Diagnostic ratios (DRs) can either be calculated from quantitative (i.e., compound concentrations) or semi-quantitative data (i.e., peak areas or heights). Also, many diagnostic ratios currently used in oil spill and environmental studies. Oil–oil correlations are based on the concept that the composition of biomarkers in spill samples does not differ from those of the candidate source oils. Most biomarkers in spill samples and source oils, in particular those homologous series of biomarkers with similar structure, show little or no changes in their diagnostic ratios. An important benefit of comparing diagnostic ratios of spilled oil and suspected source oils is that concentration effects are minimized. In addition, the use of ratios (rather than absolute values) tends to induce a self-normalizing effect on the data because the variations due to the fluctuation of instrument operating conditions day-to-day, operator, and matrix effects are minimized. Therefore, comparison of diagnostic ratios reflects more directly differences of the target biomarker distribution between samples.

<table>
<thead>
<tr>
<th>Biomarker classes</th>
<th>Diagnostic ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acyclic Isoprenoids</td>
<td>pristane/phytane</td>
</tr>
<tr>
<td></td>
<td>pristane/n-C17</td>
</tr>
<tr>
<td></td>
<td>phytane/n-C18</td>
</tr>
<tr>
<td>Terpanes</td>
<td>C21/C23 tricyclic terpane</td>
</tr>
<tr>
<td></td>
<td>C23/C24 tricyclic terpane</td>
</tr>
<tr>
<td></td>
<td>C23 tricyclic terpane/C30αβ hopane</td>
</tr>
<tr>
<td></td>
<td>C24 tricyclic terpane/C30αβ hopane</td>
</tr>
<tr>
<td></td>
<td>C24 tertracyclic/C26 tricyclic (S)/C26 tricyclic (R) terpane</td>
</tr>
<tr>
<td></td>
<td>C2718α, 21β-trisnorhopane/C27 17α, 21β-trisnorhopane</td>
</tr>
<tr>
<td></td>
<td>C28 bisnorhopane/C30αβ hopane</td>
</tr>
<tr>
<td></td>
<td>C29αβ-25-norhopane/C30αβ hopane</td>
</tr>
<tr>
<td></td>
<td>C29αβ-30-norhopane/C30αβ hopane</td>
</tr>
<tr>
<td></td>
<td>oleanane/C30 αβ hopane</td>
</tr>
<tr>
<td></td>
<td>moretane(C30 βα hopane)/C30αβ hopane</td>
</tr>
<tr>
<td></td>
<td>gammacerane/C30αβ hopane</td>
</tr>
<tr>
<td></td>
<td>tricyclic terpanes (C19-C26)/C30 αβ hopane</td>
</tr>
<tr>
<td></td>
<td>C31 homohopane (22S)/C31 homohopane (22R)</td>
</tr>
<tr>
<td></td>
<td>C32 bishomohopane (22S)/C32 bishomohopane (22R)</td>
</tr>
<tr>
<td></td>
<td>C33 trishomohopane (22S)/C33 trishomohopane (22R)</td>
</tr>
<tr>
<td></td>
<td>Relative homohopane distribution</td>
</tr>
<tr>
<td></td>
<td>(C31-C35)/C30 αβ hopane</td>
</tr>
<tr>
<td></td>
<td>homohopane index</td>
</tr>
</tbody>
</table>
Biomarkers

Biomarker classes

Diagnostic ratios

Steranes

Relative distribution of regular C27-C28-C29 steranes
C27 αββ/C29 αββ steranes (at m/z 218)
C27 αββ/(C27 αββ + C28 αββ + C29αββ) (at m/z 218)
C28 αββ/C29 αββ steranes (at m/z 218)
C28 αββ/(C27 αββ + C28 αββ + C29 αββ) (at m/z 218)
C29 αββ/(C27 αββ + C28 αββ + C29 αββ) (at m/z 218)
C30 sterane index: C30/(C27 to C30) steranes selected
diasteranes/regular steranes
Regular C27-C28-C29 steranes/C30αβ-hopanes

Monoaromatic steranes

C27-C28-C29 monoaromatic steranes (MA) distribution.

Triaaromatic steranes

C20 TA/(C20 TA + C21 TA)
C26 TA (20S)/sum of C26 TA (20S) through C28 TA (20R)
C27 TA (20R)/C28 TA (20R)
C28 TA (20R)/C28 TA (20S)
C26 TA (20S)/[C26 TA (20S) +C28 TA (20S)]
C28 TA (20S)/[C26 TA (20S) +C28 TA (20S)]

Table 1. Examples of some diagnostic ratios of biomarkers frequently used for the environmental forensic studies.

4. Examples of parameters used in fingerprinting

4.1 Normal -alkanes characteristics

The distribution of n-alkanes in crude oils can be used to indicate the organic matter source (Duan and Ma, 2001). For example, the increase in the n-C15 to n-C20 suggests marine organic matters with contribution to the biomass from algae and plankton (Peters and Moldowan, 1993). Oil samples characterized by uniformity in n-alkanes distribution patterns suggest that they are related and have undergone similar histories with no signs of biodegradation (Ficken et al. 2000 and Duan and Ma, 2001).

4.2 Carbon preference index (CPI)

Carbon preference index, obtained from the distribution of n-alkanes, is the ratio obtained by dividing the sum of the odd carbon-numbered alkanes to the sum of the even carbon-numbered alkanes. CPI is affected by both source and maturity of crude oils (Tissot and Welte, 1984). CPI of petroleum oils ranging about 1.00 generally shows no even or odd carbon preference indicates mature samples. Also, it can be used in source identification; petroleum origin contaminants characteristically have CPI values close to one (Maioli et al., 2011).

4.3 Degree of waxiness

The degree of waxiness can be expressed by the ΣC21-C31/ΣC15-C20 ratios. The oils characterized by high abundance of n-C15to n-C20 n-alkanes in the saturate fractions
reflecting low waxy (Moldowan et al., 1994). Generally, the degree of waxiness < 1 reveals low waxy nature and suggests marine organic sources (Peters and Moldowan, 1993) mainly of higher plants deposited under reducing condition.

5. Examples of parameters used in biomarker fingerprinting

5.1 Pristane/phytane ratio

Both pristine (2,6,10,14- tetramethyl pentadecane) and phytane (2,6,10,14- tetramethyl hexadecane) are derived from the phytol side chain of chlorophyll, either under reducing conditions (phytane) or oxidizing conditions (pristane). Also both pristine and phytane became dominant saturated hydrocarbon components of highly weathered crude oils until they are degraded (Moustafa et al., 2004).

The pristane/phytane (Pr/Ph) ratio is one of the most commonly used correlation parameters which have been used as an indicator of depositional environment (Peters et al., 2005). It is believed to be sensitive to diagenetic conditions; Pr/Ph ratios substantially below unity could be taken as an indicator of petroleum origin and/or highly reducing depositional environments. Very high Pr/Ph ratios (more than 3) are associated with terrestrial sediments. Pr/Ph ratios ranging between 1 and 3 reflect oxidizing depositional environments (Hunt, 1996).

According to Lijmbach (1975) low Pr/Ph values (<2) indicate aquatic depositional environments including marine, fresh and brackish water (reducing conditions), intermediate values (2–4) indicate fluviomarine and coastal swamp environments, whereas high values (up to 10) are related to peat swamp depositional environments (oxidizing conditions).

5.2 Isopreniods/n-alkanes

Waples (1985) stated that by increasing maturity, n-alkanes are generated faster than isoprenoids in contrast to biodegradation. Accordingly, isopreniods/n-alkanes (Pr/n-C17 and Ph/n-C18) ratios provide valuable information on biodegradation, maturation and diagenetic conditions. The early effect of microbial degradation can be monitored by the ratios of biodegradable to the less degradable compounds. Isoprenoid hydrocarbons are generally more resistant to biodegradation than normal alkanes. Thus, the ratio of the pristane to its neighboring n-alkane C17 is provided as a rough indication to the relative state of biodegradation. This ratio decreases as weathering proceeds.

5.3 Steranes (m/z 217) distribution

The distribution of steranes is best studied on GC/MS by monitoring the ion m/z=217 which is a characteristic fragment in the sterane series. It is agreed that the relative amounts of C27-C29 steranes can be used to give indication of source differences (Lijmbach, 1975). For example, predominance of C28, C29 and C30 steranes indicate an origin of the oils derived mainly from mixed terrestrial and marine organic sources, while oils show slightly low abundance of C28 and C29 and relatively higher concentrations of C27 steranes indicate more input of marine organic source.
5.4 Triterpanes (m/z 191) distribution

Together with steranes, triterpanes belong to the most important petroleum hydrocarbons that retain the characteristic structure of the original biological compounds. Tricyclic, tetracyclics hopanes and other compounds contribute to the terpane fingerprint mass chromatogram (m/z=191) are commonly used to relate oils and source rocks (Hunt, 1996). Mass fragmentogram at m/z=191 can be used to detect triterpanes in the saturate hydrocarbon fraction.

5.4.1 Tricyclic terpanes

Aquino et al. (1983) indicated that tricyclic terpanes are normally associated with marine source. In addition it has been used as a qualitative indicator of maturity (Van Grass, 1990). In high mature oils, the tricyclic terpanes is dominated more than in low mature oils (Hunt, 1996).

5.4.2 Homohopanes

The homohopanes (C31 to C34) are believed to be derived from bacteriopolyhopanol of prokaryotic cell membrane. C35 homohopane may be related to extensive bacterial activity in the depositional environment (Ourisson et al., 1984). Homohopane index can be used as an indicator of the associated organic matter type, as it can also be used to evaluate the oxic/anoxic conditions of source during and immediately after deposition of the source sediments (Peters and Moldowan, 1991). Low C35 homohopanes is an indicator of highly reducing marine conditions during deposition whereas high C35 homohopane concentrations are generally observed in oxidizing water conditions during deposition, consistent with the oxic conditions (Peters and Moldowan, 1991).

5.4.3 Gammacerane

Gammacerane, originally thought to be as hypersalinity indicator (Sinninghe-Damste et al., 1995), is associated with both marine and lacustrine environments of increasing salinity (Waples and Machihara, 1991; and Peters and Moldowan, 1993).

Fig. 5. Gammacerane chemical structure.
5.4.4 Ts/Tm
The ratio of Ts (trisnorneohopane) to Tm (trisnorhopane) more than (0.5) was found to increase as the portion of shale in calcareous facies increases (Hunt, 1996). Van Grass (1990) stated that Ts/Tm ratios begin to decrease quite late during maturation but Waples and Machihara (1991) reported that Ts/Tm ratio does not appear to be appropriate for quantitative estimation of maturity.

5.4.5 C29/C30 hopanes ratios
C29/C30 hopanes ratios are generally high (>1) in oils generated from organic rich carbonates and evaporates (Connan et al., 1986).

5.4.6 Steranes/17α (H)-hopanes ratio
The regular steranes /17α(H)-hopanes ratio reflects input of eukaryotic (mainly algae and higher plants) versus prokaryotic (bacteria) organisms to the source rock. The sterane/hopane ratio is relatively high in marine organic matter with values generally approaching unity or even higher. In contrast, low steranes and sterane/hopane ratios are more indicative of terrigenous and/or microbially reworked organic matter (Suzuki et al., 1996).

5.4.7 Bisnorhopanes
It is believed that sediments containing large amounts of bisnorhopane were deposited under anoxic conditions (Mello et al., 1988). Bisnorhopanes are types of pentacyclic triterpanes present in significant concentrations in oil. Bisnorhopanes are observed in Guatemalan evaporites (Connan et al., 1986) and frequency reported in other biogenic siliceous rocks of the circum-Pacific region (Katz and Elrood, 1983).

5.5 Metalloporphyrins
Porphyrins are the tetrapyrole compounds; the porphyrin nucleus consists of four pyrrole rings joined by four methine bridges giving a cyclic tetrapyrole structure. The majority of these compounds are thought to originate from various chloropigments produced by phototrophic organisms of the geological past (Yui et al., 2007). Metalloporphyrins has become a valuable tool in the determination of the origin and maturity of the organic matter (Doukkali et al., 2002; Chikaraishi et al., 2005 and Ohkouchi et al., 2006). The porphyrin structure consists of a porphyrin nucleus with various groups of side chains occupying some or all of its peripheral positions.

Metalloporphyrins were extracted from asphaltene and maltene fractions using adsorption column chromatography (Faramawy et al., 2010). Porphyrins occur as etioporphyrin (EtiO), Benzo-etiO, deoxophyloerythroetioporphyrin (DPEP), Benzo-DPEP and tetrahydrobenzo-DPEP (THBD). The distribution of different types of metalloporphyrins is useful for interpreting transformation of kerogen into bitumen, depositional environments and maturation levels of deposited organic matters.
6. Developments in GC-MS instrumentation

The low biomarker concentrations in oils (often in the range of several parts per million) in the presence of a highly complex petroleum hydrocarbon matrix especially weathered oils, the variety of chemical classes present in oils and the possible co-elutions in conventional chromatographic separations make the identification of biomarkers a more difficult task.

The development of more reliable, highly selective, fast and sensitive separation and identification tools for biomarker analysis purposes can be considered as one of the most important research points in this field for a meaningful biomarker analysis.
The use of comprehensive two-dimensional gas chromatography (GC × GC) coupled to time-of-flight mass spectrometry (TOFMS) was found to be a powerful tool for overcoming some problems and limitations since it (i) separates substances using two interconnected capillary columns containing different stationary phases and (ii) uses the fast data acquisition of time-of-flight analyzer as a robust registry for GC × GC (Aguiar et al., 2001).

In their work, Aguiar et al. (2001) used this technique to overcome the co-elution between tri- and pentacyclic terpanes separated by extracted ion chromatograms (EIC) for ions of mass-to-charge ratio (m/z) 191. The biomarker analysis by GC × GC–TOFMS was much better than in previous works using one-dimensional GC. Co-elutions between tri- and pentacyclic terpanes were clearly resolved in the second column. Noteworthy separation between the C30 hopane and C30 dimethylated homohopane was achieved and overlap of hopanes with steranes in the m/z 217 was eliminated. Besides hopanes, dimethylated tri- and tetracyclic terpanes were identified. These findings indicate the superiority of GC × GC–TOFMS as a technique for separation and identification of biomarkers in oils due to its high sensitivity, specificity and capability to elucidate compounds structure with high spectral resolution.

Comprehensive two-dimensional gas chromatography (GC×GC) has also been used to separate and identify alkylated aromatics (naphthalenes, biphenyls, fluorenes, phenanthrenes and chrysenes), sulfur-containing aromatics (dibenzothiophenes, benzonaphthothiophenes), steranes, triterpanes, and triaromatic steranes. These biomarkers were separated into easily recognizable bands in the GC×GC chromatogram. Methods used to identify the bands included peak matching with chemical standards and comparison with GC/MS extracted ion chromatograms (Frysinger and Gaines, 2001). By designing mass spectrometers that can determine m/z values accurately to four decimal places, it is possible to distinguish different formulas having the same nominal mass. Since a given nominal mass may correspond to several molecular formulas, lists of such possibilities are especially useful when evaluating the spectrum of an unknown compound.

GC/MS/MS is an operation based on the covariant scan of electrostatic magnetic fields on the trisector double focusing mass spectrometer providing more accurate data. The quadrupole is a common mass separator gives a sufficient sensitivity and selectivity however, high resolution mass spectrometry (HRMS) is also used due to its ability to provide quantitative data for compounds present in complex mixtures for biomarkers analysis. Triple quadrupole GC/MS offers a viable alternation for the rapid, routine analysis providing excellent precision, sensitivity, selectivity, and dynamic range (Thermo-application note 10261).

Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) benefits from ultra-high mass resolving power (greater than one million), high mass accuracy (less than 1 ppm) and rapid analysis which make it an attractive alternative for the analysis of different and wide range of petroleum products (Klein et al., 2003).

It should be noted, however, that there is no single fingerprinting technique that can fully and readily meet the objectives of biomarkers investigation and quantitatively allocate hydrocarbons to their respective sources, particularly for complex hydrocarbon mixtures or extensively weathered and degraded oil residues. Combined and integrated multiple tools are often necessary under such situations.
7. Data analysis by computerized techniques

Data analysis is an important part of chemical fingerprinting and a broad collection of statistical techniques has been used for evaluation of data.

After separation and identification of biomarkers, principal component analysis PCA, a mathematical procedure, can be used for analyses of chromatograms using a fast and objective procedure with more comprehensive data usage compared to other fingerprinting methods. The discriminative power of PCA can be enhanced by deselecting the most uncertain variables or scaling them according to their uncertainty.

For example, preprocessing of GC-MS chromatograms followed by principal component analysis (PCA) of oil spill samples collected from the coastal environment in the weeks after the Baltic Carrier oil spill and from the tank of the Baltic Carrier (source oil) was carried out (Christensen et al., 2005). The preprocessing consists of baseline removal by derivatization, normalization, and alignment using correlation optimized warping. The method was applied to chromatograms of \( m/z \) 217 (tricyclic and tetracyclic steranes) of oil spill samples and source oils. The four principal components were interpreted as follows: boiling point range (PC1), clay content (PC2), carbon number distribution of sterols in the source rock (PC3), and thermal maturity of the oil (PC4). The method allows for analyses of chromatograms using a fast and objective procedure and with more comprehensive data usage compared to other fingerprinting methods.

8. Previous studies in Egypt

8.1 Egyptian Western Desert

The Western Desert covers about 700,000 square kilometers (equivalent in size to Texas) and accounts for about two-thirds of Egypt's land area. This immense desert to the west of the Nile spans the area from the Mediterranean Sea south to the Sudanese border. The chemical fingerprinting of oils in this area is a great interesting research area for many proposes as for example identifying the sources of petroleum oil or complex environmental pollutants. The original sources of complex mixtures can often be identified by the relative abundance of some major individual compounds (e.g. n-alkanes) forming a chemical pattern by ratios of specific constituents or by identifying source-specific compounds or markers (e.g. triterpanes) in the environmental sample being investigated (Peters et al., 2005). These parameters depend mostly on the preburial environments of the living organisms, the depositional environments of the organic matter and the diagenetic processes in the source rocks.

In their work, Roushdy et al. (2011) utilize biomarkers characteristics together with bulk geochemical parameters to identify and characterize the crude oils and to assess the respective depositional environments and maturation. Variation of crude oil-gravities in the Western Desert reflects different stages of oil migration and accumulation as well as different oil source rocks in the same and different ages (Zein El Din et al., 1990). The authors attempt to assess the correlation between the crude oil samples and the potential source rocks to confirm the indigenous sources for the petroleum generation of some oilfields of the North Western Desert. This target was made throughout the study in detail of the analytical results for three crude oil samples collected from three oilfields in the North Western Desert oilfields (Meleiha, Misaada and Qarun) as well as three extract samples (Baharia, Kharita and Khtataba) from formations ranging in age from Upper Cretaceous to Middle Jurassic.
The specific geochemical parameters have been assessed by the aid of gas chromatography and gas chromatographic-mass spectrometric analyses of the saturated fractions. The degree of the correlation between crude oils and the extracted samples was determined by studying the correlation scores for both oils and extracts. Eight correlation parameters have been studied for this purpose includes: saturates%, saturates/aromatics ratio, $C_{\text{max}}$, $C_{21}+C_{22}/C_{28}+C_{29}$, CPI, pristane/phytane, pristane/$n$-$C_{17}$ and pristane+$n$-$C_{17}$/phytane+$n$-$C_{18}$. An overall correlation score was obtained for each oil and extract by summing up the contribution from each parameter. The GC/FID chromatogram of the Meleiha crude oil sample is characterized by a monotonically decreasing homologous series of heavy normal alkanes ($n$-$C_{25}$ to $n$-$C_{30}$) and display odd carbon preference at $n$-$C_{15}$ which reflects mature oils originated mainly from non-marine origin mainly terrestrial organic matters deposited under slightly oxidizing environment and slightly mixed with inputs from marine source (Hunt, 1996). The mode of distribution of $n$-paraffins in the crude oils of Misaada and Qarun oilfields show that the maximum abundance is at $n$- $C_{15}$ to $n$-$C_{25}$ reflecting marine origin.

The steranes distribution of crude oils was studied. Meleiha crude oil was found to be characterized by low predominance of $C_{27}$ steranes and slightly high abundance of $C_{28}$ and $C_{29}$ indicating that the Meleiha oil is believed to be generated from both marine shales and carbonates enriched in marine algae with more contribution from terrestrial organic sources deposited under saline conditions. It reveals that the Meleiha oil is derived mainly from terrestrial organic sources.
The high concentrations of C27 - C29 diasteranes in case of Meleiha oil indicate input of marine organic source with more contribution from terrestrial organics (Waples and Machihara, 1992). The high diasteranes concentration compared to regular steranes suggest a clay rich source rock because the clay is required to catalyze the steroids transformation to diasteranes (Peters and Moldowan, 1991). Misada and Qarun oils are characterized by slightly lower predominance of C27 steranes and higher abundance of C28 and C29 steranes indicating inputs from marine organic sources (Waples and Machihara, 1992). Moreover, the distributions of regular steranes C29, C27 and C28 on the ternary diagram reveal also more contribution from marine organic sources. The diasteranes concentrations compared to regular steranes is low, suggesting a clay rich source rock.

The crude oil samples of Qarun and Misaada oilfields have Pr/n-C17 and ph/n-C18 ratios 0.28, 0.47 and 0.1, 0.1, respectively reflecting mostly mature and originated mainly from marine organic sources deposited under reducing environment. The crude oil of Meleiha oilfield has Pr/n-C17 and ph/n-C18 of 0.40 and 0.28 indicating mixed organic sources.

Terpanes biomarkers distributions derived from the m/z 191 mass chromatograms show that the C21-C25 tricyclic terpanes of Meliha oil appear to be the largest components which may support that the oil of Meleiha oilfield is more mature and sourced mainly from marine carbonate source rocks. At the same time, the C23, C24 and C25 tricyclic terpanes are generally of lower values compared with C22 indicating that the oil has some inputs from terrestrial organic materials (Hunt, 1996). The unusual low amounts of C30 extended hopanes seem to be associated with mixed organic sources (Moldowan et al., 1985). This phenomenon can be displayed by the low ratio of C29/C30 extended hopanes.

On the other hand, The C30 hopanes are the largest components in the series C27-C34 in oil samples from Misaada and Qarun oilfields. This indicates that the organic materials in these oils were originated mainly from saline and hypersaline environments (Peters and Moldowan, 1993). The extended hopanes are available as paleo environmental indicator (Waples and Machihara, 1992). The unusual large amounts of C30 hopanes seem to be associated with marine sources (Moldowan et al., 1985).

Bisnorhopanes are types of pentacyclic triterpanes present in significant concentrations in oil. Bisnorhopanes are observed in Guatemalan evaporites and frequency reported in other biogenic siliceous rocks of the circum-Pacific region (Connan et al., 1986). It is believed that sediments containing large amounts of bisnorhopane were deposited under anoxic conditions (Mello et al., 1988). The crude oils of Misaada oil field have relatively higher amounts of C28 bisnorhopane indicating more anoxic environment than Qarun and Meleiha oils.

Carbon preference index (CPI) values of the studied crude oils are close to unity, ranging from (0.94 to1.04) indicating mature crude oils.

Pr/Ph ratio of the oil sample from Meleiha oilfield is 3.0 indicating oxidizing depositional environment of the crude oil while the crude oils from Qarun and Misaada oil fields have Pr/Ph ratios of 0.63 and 2.00 respectively reflecting that these crude oils were deposited under transitional (reducing– oxidizing) environments. These results indicate good correlation between crude oils from Qarun and Misaada oilfields with slight correlation to crude oil from Meleiha oilfield.

Oil: source correlation reflect a good correlation between the extract samples of Kharita and Khatatba source rocks and crude oils from Meleiha and Qarun oilfields. The extract of
Bahariya source rock shows slight correlation with Meleiha oil and differ from the other oil samples. These evidences indicate that Kharita and Khtataba source rocks seem to act as sources and reservoirs for oil generation in the Qarun and Misaada oilfields while the oil generation of Meleiha oilfield seems to be migrated from Bahariya source rocks.

8.2 Suez Gulf

The Gulf of Suez occupies the northwestern arm of the Red Sea between Africa proper (west) and the Sinai Peninsula (east) of Egypt. The length of the gulf, from its mouth at the Strait of Jubal to its head at the city of Suez, is 195 miles (314 km) and it varies in width from 12 to 20 miles (19 to 32 km). Because the importance occurrence of crude oil in the Gulf of Suez, the biological markers was analyzed to evaluate the geochemical relationships between the oils recovered from some oil fields within the Gulf of Suez to assess and investigate oil characterization, maturation, source depositional environments and oil families.

Roushdy et al. (2010) evaluate the geochemical relationships between the oils recovered from some oil fields within the Gulf of Suez. This target was achieved through analytical results of GC and GC-MS analysis for seven crude oil samples collected from seven oilfields namely: Ras Badran, Belayim marine, Belayim Land, Rahmi, West Bakr, Esh El Mellaha and Geisum distributed within the Gulf of Suez. These samples are representative for the producing horizon zones (Belayim, Rudies and Nuhkul formations.) of Upper- Lower Miocene age characterized by limestone facies with depths ranging from 2250 to 8286 ft. Geochemical parameters based upon acyclic isoprenoids, steranes and terpanes coupled with bulk geochemical parameters indicated whether the crude oils are of marine, terrestrial or mixed marine-terrestrial origin.

Biomarkers analyses of crude oils from the Gulf of Suez suggest that oils are more mature and derived mainly from mixed organic sources from terrestrial and marine inputs contribution to the biomass from algae and plankton in different saline environments.

A few discrepancies that appear between the results obtained by using the different parameters can be related to the alteration caused by the number of processes (physical, chemical and/or biological) affecting part of the source related biomarkers pattern of the oil after generation and/or primary migration from the source rock.

In another study, two genetic families based on biomarker analyses of oils were isolated from the Gulf of Suez, Egypt. Oils from Ras Fanar and East-Zeit wells have high gammacerance, low diasterances and high C33/C34 hopanes, consistent with an origin from the Brown Limestone. Oils from the Gama and Amal-9 wells have low gammacerance, high diasterances and oleanane indices > 20 %, indicating an angiosperm-rich Tertiary siliciclastic source rock, probably the Lower Miocene Rudeis Formation (Peters et al., 2005).

(Younes et al., 2004) evaluated the depositional environments and maturation assessments of source rocks from the central Gulf of Suez, Egypt utilizing the biomarker distributions in nine crude oils derived from a synrift tectonic sequence of the central Gulf of Suez province. No obvious variations were observed amongst the studied crude oils, suggesting that these oils are all of the same genetic type. These oils features, a predominance of oleanane, reaching 24 %, and a relatively low gammacerane concentration of 10%, suggested that these oils were derived from a terrigenous organofacies source rock with a significant angiosperm higher land plants input deposited within the marginally mature syn-rift shale of Lower Miocene Nukhul, Rudeis and Kareem formations of mixed kerogen types II-III. Maturity
parameters based on various sterane isomerisation distributions and polycyclic aromatic compounds indicate a low thermal maturation level for the generated hydrocarbons within the syn-rift lithostratigraphic succession. These similarities in geologic occurrences and biomarker characteristics suggest the possibility that the hydrocarbon expulsion could have been initiated from deeply buried Miocene source rocks and trapped within the syn-rift structures throughout the extensional faults of the central Gulf of Suez province.

(Barakat et al., 2000) studied the aliphatic and aromatic fractions of a beach tar sample from the Mediterranean coast of Sidi Kreir, 37 Km west of the city of Alexandria by GC and GC/MS techniques. A complete analysis was carried out to investigate chemical composition changes, fate of weathered oil residue and possible source identification. The distribution of sterane, hopane, mono-and triaromatic steroids, \( C_2 \) and \( C_3 \) phenanthrenes and dibenzothiophenes and chrysenes, however, had remained unaltered by weathering. The beach tar possessed geochemical features consistent with a marine carbonate or evaporite source depositional environment under normal saline reducing conditions.

9. Current work

Applications of biomarkers in oil spill source identification

Although oil is the dominant energy source, oil spill occurs worldwide causing a severe global environmental problems (Abostate et al., 2011). Egypt is suffering from oil pollution owing to the increasing petroleum activities in the last decades. Environmental protection is currently an important subject of increasing public and research concern and as a result, special efforts have already been done so as to develop oil spill detection and fingerprinting. Therefore, to unambiguously characterize, identify, categorize, and quantify all sources of hydrocarbons entering the environment is very important for environmental damage assessment, evaluation of the relative risks to the ecosystem posed by each spill and selecting appropriate spill response and taking effective cleanup measures. Biomarkers are the most important hydrocarbon groups for chemical fingerprinting which play a very important role in source identification in environmental forensic investigations of oil spills. It was a useful analogy to explain this type of forensic analyses for spilled oil. However, it was recognized then, and remains true today, that the analyses of spilled oils do not have the statistical discriminating power of the human fingerprint in the sense that each human has an individual fingerprint. Analyses of spilled oils and potential sources are usually undertaken by increasingly sophisticated chemical analyses until either all but one potential source oil remains that cannot be distinguished from the spilled oil, or all potential sources have been eliminated and the spill is then a "mystery". The presumption for success using fingerprinting is that a complete collection of possible sources has been secured for the matching analyses. The term "passive tagging" has been used in place of fingerprinting in the past to describe the chemical analyses of oils. The term derives from the process of using the chemicals naturally present in the oil as "tags". The "passive" part of the term was used because there were proposals and some experiments conducted in the late 1960s and early 1970s to introduce "active tags" into various oil cargos to allow for identifying the oils if they were spilled (Adlard, 1972). Various chemicals were proposed as active tags, but the obvious international administrative and logistical effort needed to keep track of such "active tags" prevented operational use of active tagging systems.

Nothing sparks concerns about contaminates in the environment quite like a petroleum release. Unfortunately, the events of 2010 served to heighten the awareness and need to
have the capability to monitor and characterize the extent and breadth of the impact of these events. Using petroleum biomarker analysis make it possible to accurately identify the source of contaminates back to the specific origin as well as determining the absolute concentrations of priority pollutant PAHs.

Generally, gas chromatograms of two oil samples are compared by comparing the envelop shapes of the n-alkanes, the unresolved backgrounds and individual peak intensities. By means of GC/MS, a big number of compound classes of oils may be separately detected and compared. Computerized oil spill identification (COSI) may highly support analysts in GC and GC/MS results evaluation and adds a new dimension to forensic oil spill identification. It is greatly increases the possibilities for finding the sources of oil pollution.

The patterns of the biomarkers and a set of parameters based on the literature findings was chosen to be investigated for most Egyptian crude oil and stored in the database in order to construct an Egyptian computerized oil spill identification database of local crude oils. Gas-chromatograms and mass-fragmentograms are rapidly produced from raw GC- and GC/MS-data for comparing an unknown pollutant sample with any oil sample stored in the database then simultaneously a much stronger connection between a distinct oil spill and its actual source accurately established than before, as shown clearly in Figure 8. These parameters allow a more objective, provable and defensible result.

Fig. 8. Representative model of computerized oil spill identification matching.

www.intechopen.com
evaluation than the mere visual comparison of the chromatograms. In addition, these parameters may also be used for finding oils in the database, which are similar to the spill sample. The system is fast and greatly saves laboratory resources and reliable and comfortable.

10. Conclusion

Biomarkers are naturally occurring, ubiquitous and stable complexes that are objectively measured and evaluated as an indicator of a certain state. It is used in many scientific fields; medicine, cell biology, exposure assessment, geology and astrobiology.

Due to the variety of geological conditions and ages under which oil was formed, every crude oil exhibits a unique biomarker fingerprint. From the identification point of view, biomarkers are the most important hydrocarbon groups in petroleum because they can be used for chemical fingerprinting which provides unique clues to the identity of source rocks from which petroleum samples are derived and the biological source organisms which generated the organic matter, the environmental conditions that prevailed in the water column and sediment at the time, the thermal history (maturity) of both the rock and the oil, and the degree of microbial biodegradation.

GC-MS is considered the most widely used method for biomarkers detection and identification which is a true combination of its separate parts (gas chromatography, GC and mass spectrometry, MS). The mass spectrometer has long been recognized as the most powerful detector for gas chromatography due to its high sensitivity, specificity and capability to elucidate compound structure. Mass fragmentography provides a satisfactory tool for obtaining specific fingerprints for classes and homologous series of compounds resolved by gas chromatography. The development of more sensitive and selective identification tool for biomarker analysis purpose especially for crude oils containing low concentration biomarkers as weathered and light oils can be considered as one of the most important research points in this field. After separation and identification of biomarkers, principal component analysis (PCA), a mathematical procedure, can be used for analyses of chromatograms using a fast and objective procedure with more comprehensive data usage compared to other fingerprinting methods. The discriminative power of PCA was enhanced by deselecting the most uncertain variables or scaling them according to their uncertainty. Chemical analysis of biomarkers generates information of great importance to environmental forensic investigations in terms of determining the source of spilled oil. The patterns of the biomarkers and a set of parameters were used to construct an Egyptian computerized oil spill identification database. This can greatly increase the possibilities for finding the sources of oil pollution by comparing an unknown pollutant sample with any similar oil sample stored in the database. A much stronger connection between a distinct oil spill and its actual source may be established than before.

11. References


Huebschmann, H. Analysis of Molecular Fossils: Crude Oil Steroid Biomarker Characterization Using Triple Quadrupole GC-MS/MS Frank Theobald, Environmental Consulting, Cologne, Thermo Fisher Scientific, Austin, TX, USA, application note 10261.


Chromatography is a powerful separation tool that is used in all branches of science, and is often the only means of separating components from complex mixtures. The Russian botanist Mikhail Tswett coined the term chromatography in 1906. The first analytical use of chromatography was described by James and Martin in 1952, for the use of gas chromatography for the analysis of fatty acid mixtures. A wide range of chromatographic procedures makes use of differences in size, binding affinities, charge, and other properties. Many types of chromatography have been developed. These include Column chromatography, High performance liquid chromatography (HPLC), Gas chromatography, Size exclusion chromatography, Ion exchange chromatography etc. In this book contains more details about the applications of chromatography by various research findings. Each and every topics of this book have included lists of references at the end to provide students and researchers with starting points for independent chromatography explorations. I welcome comments, criticisms, and suggestions from students, faculty and researchers.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
