Extraction and characterization of crude oil acids for increased accuracy in computational microemulsion predictions

I C C Chai¹, N A A M Fauzi¹, A Khalid², C H Voon² and V Y Hon¹
¹Hydrocarbon Recovery Technology R&D, GR&T, PD&T, PETRONAS, 43000 Selangor, Malaysia.
²Lab Operations, GR&T, PD&T, PETRONAS, 43000 Selangor, Malaysia.
Corresponding author email: ivy_chai@petronas.com

Abstract. Microemulsion modelling is a computational tool developed to predict the optimal salinity of a crude-surfactant-brine ternary phase system. Whilst optimal salinity is associated with the desirable Windsor Type-III microemulsion, other types of microemulsion can be determined from the model given the surfactant type, the brine concentration, and the components of the crude oil. The main challenge is to identify the surface active acid compounds of the crude oil that is responsible for the formation of emulsion in a ternary system. Therefore, it is important to characterize the full range of acidic species from crude oil because their chemical structures are crucial inputs to accurately predict the type of microemulsion formed. In this work, crude oil acids were extracted using two different methods – liquid-liquid extraction and acid-ion exchange resin (acid-IER), and were characterized using gas chromatography-mass spectrometry (GC-MS) and tandem liquid-chromatography tandem mass spectrometry (LC-MS/MS), respectively. For liquid-liquid extraction with derivatization, GC-MS results showed that low molecular weight acid (C4) can be identified. With the IER-extracted acids, the LC-MS/MS separated out low to medium molecular weight acids (C5-C18) with 1 cyclic acid (naphthenic acid) identified. Limitations of both extraction methods and characterization of crude oil acids using different instrumentation are discussed.

1. Introduction
In a ternary system of surfactant/brine/hydrocarbon mixture, phase behaviour studies are conducted to identify the optimal salinity by varying the salinity from low to high (i.e. salinity scan) using a series of test tubes or glass pipettes. At optimal salinity, Type III microemulsion is formed and where the volume of surfactant adsorbed to both oil and brine are equal, is highly desirable because this region is where low interfacial tensions and high oil recoveries in chemical enhanced oil recovery (CEOR) can be obtained [1,2]. Laboratory phase behaviour screening are relatively simple in selecting favourable surfactant formulations at optimal salinity regions but is often laborious especially when it involves a wide range of surfactant choices. To reduce the huge experimental test matrix in shortlisting promising surfactants for a CEOR design, computational chemistry method incorporating molecular modelling i.e. methods of moments (MoM) and dissipative particle dynamics (DPD) is employed. Zhang et al. (2010) [3] defines DPD as a coarse-grained technique to describe the primarily hydrodynamic behaviour of complex fluids. The advantage of coarse graining is that fewer particles have to be considered than there are atoms in the system, which translates into higher simulation speed [4]. This technique has been applied extensively in the oil and gas industry [3,5,6] including modelling of microemulsions for ternary systems [7]. In DPD, the challenge in microemulsion modelling is to obtain the structural information of the brine, surfactant and crude oil components especially its acids; the latter is of often of high complexity in its elucidation and requires thorough analytical and instrumentation investigation.

Crude oil acids, which are surface active, are largely responsible for the formation of microemulsion in a ternary system. They are primarily acyclic (carboxylic acids), and cyclic (monocyclic and polycyclic...
carboxylic acids), where the latter are primarily known as naphthenic acids (NA), having the general chemical formula \( C_nH_{2n+2}O_2 \), where \( n \) indicates the carbon number and \( Z \) specifies a homologous series. Mass spectrometry techniques are commonly used to identify, profile, and characterize NAs in crude oil. Several authors [8,9] have described the extraction of NA from crude oil using mass spectrometry techniques in combination with other analytical methods – and are able to profile and categorize the NAs based on functional groups, molecular weight, and degree of saturation. However, the exact molecular structures of the NAs were not elucidated.

Our paper describes the different extraction methods and instrumentation to identify and obtain structural information of acids present in a Malaysian petroleum crude oil that becomes crucial input for molecular parameterization to predict optimal salinity. We discuss the strengths and limitations of each extraction method and instrumentation when analysing the acid samples using gas chromatography mass spectrometer (GC-MS) and liquid chromatography tandem mass spectrometer (LC-MS/MS).

2. Materials and methods
The crude oil acids extraction and characterization scheme using two different mass spectrometry methods i.e. GC-MS and LC-MS/MS are as shown in figure 1. The crude oil, having an acid number of 0.16 mg KOH/g, API = 41°, a viscosity of 3.47 mm²/s at 40°C, was obtained from a Malaysian oilfield. Extraction solvents hexane, ethanol (Merck), diethyl ether (VWR International, Ltd) were of reagent grade and used without further purification. Salts potassium hydroxide (Analar) and sodium sulfate (Merck) were used as is. For GC-MS and LC-MS/MS, solvents were all of HPLC quality. All standards used for GC-MS (purity 99.5%) were purchased from Sigma-Aldrich.

![Figure 1. Sample extraction and instrumentation scheme for extraction and characterization of crude oil acids.](image)

2.1. Liquid-liquid extraction and derivatization
Liquid-liquid extraction method using alkaline solution is described in Sun et al. (2008) [10]. Acid standards commonly found in crude and organic oils [11,12] - pentanoic acid (C5), nonanoic acid (C9), hexadecanoic (C16) acid, and octadecanoic acid (C18) were spiked into crude oil before liquid-liquid extraction process. These standards were used to enlarge the active components in the crude [10] and to validate the capability of the extraction and derivatization methods to recover low to medium MW acids [13]. 250 g of crude oil was dissolved in 250 g n-hexane and the solution extracted three times with 250 mL portions of 1 wt percent KOH in 70 v/v % ethanol. The combined alkaline solution was extracted with hexane to remove trace of dissolved or admixed oil, followed by concentrating then acidifying by HCl to pH 3. Finally, the acidic
mixtures were extracted by diethyl ether, dried with anhydrous sodium sulfate. The diethyl ether was removed by evaporation under reduced pressure, and acidic fraction was recovered into a bottle for derivatization by methylation. Derivatization by converting acids into esters was performed using BF$_3$-methanol complex as methylation reagent at 80°C. The derivatized acid samples were injected directly onto a non-polar column for GC-MS analysis.

2.2. Acid-IER extraction
Acids were extracted from crude oil by the use of ion-exchange resin as described by Mediaas et al. (2003) [14]. QAE Sephadex A-25 is used as a solid-phase material for the ion exchange of the crude oil acids. The capacity of the ion-exchange material is 2.5 mmol acid equivalents per gram of ion-exchange material, calculated from the TAN of the oil. A buffer solution (1 M Na$_2$CO$_3$/NaHCO$_3$) is added to the ion-exchange material and allowed to run slowly through a filter (GF/C, Whatman glass microfiber filter, Sigma-Aldrich), at a ratio of 75 mL of the buffer solution per gram of ion-exchange material. The ion-exchange material is then cleaned with distilled water until pH is 7. The water is removed by vacuum filtration, and methanol (MeOH) is added, at a ratio of 25 mL of MeOH per gram of ion-exchange material. The ion-exchange material is then transferred to the crude oil and left to stir overnight. The oil is filtered, and the ion-exchange material is washed with toluene several times. Toluene aliquots of 5 mL per gram of ion-exchange material are used. The ion-exchange material is then washed with a mixture of toluene and MeOH, 2:1 (v/v), until the filtrate is colorless. The filtration procedure is repeated two times using a filter paper having a smaller pore diameter (GF/C the first time and by GF/F for subsequent filtrations). The acids from the crude oil are recovered from the ion-exchange material by adding 1 M formic acid (3.5 mL per gram of ion-exchange material) and a mixture of toluene and MeOH, 1:1 (v/v; 50 mL per gram of ion-exchange material). The mixture is left to stir for 3 h before the ion-exchange mass is removed by filtration, first through a GF/C filter, and then through a GF/F filter. The ion-exchange mass is washed with 2:1 toluene/methanol (v/v; portions of 10 mL per gram of ion-exchange material). To ensure that all the petroleum acids are recovered, the residue from the filtrations is collected and the acid recovery procedure is repeated using formic acid (0.5 mL per gram of ion-exchange material) and 1:1 toluene/MeOH (v/v; 50 mL per gram of ion-exchange material). The mixture is stirred and filtered in the same way as described above. The solvents in the filtrates are removed by a rotary evaporator, and the acid extract is re-dissolved in MeOH (v/v) for LC-MS/MS analysis.

2.3. GC-MS analysis (non-polar column)
The sample was injected into a Perkin Elmer Clarus 500 GC gas chromatography equipped with a 30 m HP-1 fused-silica capillary column, with 100% dimethyl polysiloxane as stationary phase, with 0.25 mm inner diameter and 0.25 μm film thickness with a temperature limit of 350°C. The mass spectrometer is Perkin Elmer Clarus 560S MS operated in electron impact ionization mode (70 eV) at 1.2 scans/s, within the m/z 25-350 mass range using helium as carrier gas. The oven temperature was programmed to begin at 40°C, then ramped at 7°C/min to temperature of 200°C, maintained for 0 min. Then the temperature was ramped at 10°C/min to a final temperature of 300°C and held for 5 min. Identification of derivatives was supported by comparison with mass spectra obtained from mass spectral library.

2.4. LC-MS/MS analysis
Group-structural analysis was performed by mass-spectrometric fragmentation to quasimolecular ions by a soft ionization technique in the negative ion mode. Liquid chromatography coupled with electrospray ionization mass spectrometry (LC–MS/MS) spectra were recorded on a hybrid triple quadrupole linear/ion trap mass spectrometer (AB SCIEX QTRAP®) equipped with electrospray ionization source. The spectra encompassed a molecular series of deprotonated molecular ions of the acid [M–H]$^-$ recorded in the negative ion mode in Enhanced-MS scan, with scan range m/z 50 to 1700. 1 scan/s. A gradient mobile phase composed of eluent A [methanol+0.1 μM ammonium acetate] and eluent B [methanol+0.1 μM ammonium
acetate] starting with composition of [A:B = 90:10 v/v], then increasing eluent B linearly to 95% for 35 minutes and hold for 5 minutes. Eluent B was then lowered linearly to 10% and hold for 5 minutes. The injected sample volume was 50 μL. Spectra were analyzed by the Multiquant Software v2.1 (AB SCIEX). The LC-column used were a Kinetex C18 column (4.6 mm ID X 250 mm, 5 μm particle size, 100A pore size, by Phenomenex (Torrance, CA)). All samples were filtered through a 45 μm membrane PTFE filter (Millipore) before analysis.

3. Results and discussion

3.1. Liquid-liquid extracts and analysis using GC-MS

The acid standards that were spiked prior to liquid-liquid extraction and methylation were all recovered in their ester form. Figure 2 shows that acid standards at different retention time (RT) – pentanoate at 4.24 min, nonanoate at 12.39 min, hexadecanoate at 24.29 min, and octadecanoate at 26.62 min. The recovery of the standards showed that the liquid-liquid extraction followed by derivatization is suitable for acid determination in crude oil.

Figure 2. Chromatogram from GCMS (non-polar column) showing recovery of acid standards from crude oil by liquid-liquid extraction with derivatization.

In a separate sample as shown in figure 3 where standards are excluded, three distinct peaks in the chromatogram at RT= 1.45, 1.52 and 1.88 min are identified as our target compounds: methyl ester. The peaks that eluded out at RT=2.7–3.0 min are identified as heptane impurities and heptane solvent used during the derivatization process.

Figure 3. Chromatogram from GC-MS (non-polar column) showing possible crude oil acids from liquid-liquid extraction with derivatization.
The mass spectrum and NIST library matches (listed partially below the spectrum) as shown in figure 4 at RT=1.45 min points to 2-hydroxy methyl cyclopropane carboxylic acid, methyl ester; boiling point (BP) = 134.6°C as the likeliest match. Although the compound has the highest reverse fit by NIST library search, the BP at a short RT=1.45 min does not match the behaviour of the compound eluded through a non-polar column. The BP of a compound are often used as verification method for compounds eluded using non-polar column [15,16] and therefore caution should be exercised when selection is conservatively based on high reverse fit of the compound. For RT=1.52 min, the mass spectrum did not show any probable match to target compound i.e. methyl-ester. In figure 5, the highest reverse fit was the compound methylene chloride did not fit the target compound. Methylene chloride has a BP of 39.6°C. At RT=1.52 min, the BP of components eluding out at this retention time should range from 50-70°C.

Figure 4. MS spectrum of compound eluded at RT=1.45 min with NIST library search and matches given.

Figure 5. MS spectrum of compound eluded at RT=1.52 min and NIST library search and matches given.

Figure 6 shows the MS spectrum and NIST library search at RT=1.88 min with the highest reverse fit corresponds to that of boronic acid ethyl-, dimethyl ester. Although an ester compound, the boronic acid moiety may be an impurity from the derivatization process using BF$_3$-methanol complex as methylation reagent. Another probable ester eluded at this retention time is butyric acid, 2-hydroxy, 3-methyl, methyl ester (C$_6$H$_{12}$O$_3$) but the high BP=165°C obtained from literature could rule out this compound as one of the
target ester-derived acid when using non-polar column. Heptane impurities (i.e. cycloheptane) from heptane solvent used in the derivatization process was identified at RT=2.7–3.0 min. Table 1 summarizes the compounds eluded and identified from mass spectrum derivatized crude oil extracts.

Table 1. Summary of compounds from liquid-liquid extracts and GC-MS analysis using non-polar column.

| Retention Time (min) | Compound identified                                      |
|----------------------|----------------------------------------------------------|
| 1.46                 | no possible matches to acid compound                     |
| 1.52                 | no possible matches to acid compound                     |
| 1.88                 | Ester compound, probable acid (C4)                       |
| 2.93                 | Heptane impurities                                       |
| 2.5-3.0              | Solvent heptane                                          |
3.2 Acid-IER extracts and analysis using LC-MS/MS

The acid samples from IER extraction was first separated with liquid chromatogram and molecular ion mass spectra from MRM mode and fragment ion mass spectra from induced fragmentation. Identification of probable acids were referenced using online MS library, Massbank.eu. The acidic components’ molecular structure and composition were interpretation by determining the parent/molecular ion (M) and fragment/daughter ions: –COOH and –OH represented as [M − 45] and [M − 18] (to account for alcohols, phenols), respectively, in a negative ion mode. Table 2 lists the molecular ions that have fragment ions that confers carboxylic or phenolic compounds were analysed by hand and cross-referenced the ion peaks intensities to ensure the closest match to compounds suggested in the library. Figure 7 shows the nine acids identified, conferring the carbon numbers of between C5 to C18. The compound 3,4,5-trihydroxy-1-cyclohexene-1-carboxylic acid (shikimic acid) (C7H10O5) was the only cycloaliphatic NA identified (Z=1). Two molecular ions could not be characterized due to poor match to the library based on relative peak intensities.

Table 2. Carboxylic acids identified from molecular ions (M) and fragment ions obtained from LC-MS/MS mass spectrum.

| No. | Molecular ions (M) (m/z) | Fragment ions (m/z) | Probable acid match based on peak intensities | Molecular formula |
|-----|------------------------|--------------------|---------------------------------------------|-------------------|
| 1   | 111.1                  | 66.9               | 2-furancarboxylic acid                      | C5H4O3           |
| 2   | 173.1                  | 137 155.2          | shikimic acid (3,4,5-trihydroxy-1-cyclohexene-1-carboxylic acid) | C7H10O5          |
| 3   | 157.1                  | 113.1 139.1       | nonanoic acid                              | C9H18O2          |
| 4   | 185.2                  | 141.1 167.1       | undecanoic acid                            | C11H22O2         |
| 5   | 199.1                  | 155.2 181.1       | dodecanoic acid                            | C12H24O2         |
| 6   | 203.1                  | 159.1 -           | (2S)-2-amino-3-(1H-indol-3-yl) propanoic acid | C11H12N2O2       |
| 7   | 213.2                  | 169.1 195.5       | 3-phenoxybenzoic acid                      | C13H10O3         |
| 8   | 236.2                  | 192.1 218.2       | no possible match in library               |                  |
| 9   | 241.3                  | 197.1 223.3       | 2-(3-phenoxyphenyl)propanoic acid          | C15H14O3         |
| 10  | 277.3                  | 233.3 259.1       | linolenic acid (all-cis-6,9,12-octadecatrienoic acid) | C18H30O2         |
| 11  | 433.3                  | 389.3 415.6       | no possible match in library               |                  |

Electrospray ionization is a ‘soft ionization’ technique, well-suited to the characterization of NA mixture although complete resolution of the mixture is not straightforward as GC-MS when the extraction matrix is of petroleum crude oil source. For crude oil matrixes, only partial characterization is generally attainable because the current mass spectrometric techniques cannot differentiate between isomers in the same NA group nor can it be separated [21]. Nevertheless, adopting the technique of difference in ions’ mass-to-charge ratio (m/z) between molecular ions and fragment ions, and evaluating the corresponding relative intensities of each fragmentation peaks of the mass spectrum, are common best practices in elucidating isomers and structure. Several molecular identifications, metabolic pathway determination and postulating
of molecular structure purposes using tandem mass-spectrometry were described [22,23] where fragment ions generated from the parent ions help to predict the fragmentation pattern of the molecule and are useful in confirmation of the target analytes.

![Chemical structures](image1.png)

Figure 7. Identity and molecular structure of acids from acid-IER extract using LC-MS/MS.

In this study, GC-MS and LC-MS/MS both yielded different carboxylic acid(s) for the same crude oil. Depending on the extraction method and whether derivatization is used, methyl ester formation of most common form of derivatization of carboxylic acid but limited to low MW (i.e. C4) acids due to elucidation limitations from a non-polar column (i.e. boiling point dependent) and presence of derivatization artefacts. The derivation procedure using reagent BF₃-methanol complex was not optimized for petroleum crude oil acids (but for fatty acids of palm oil); other derivatization methods can be considered for NA isomers such as using the reagent N-(3-dimethylaminopropyl)-N-ethylcarbodiimide (EDC) as described by Woudneh et al. (2013) [24]. From this investigation, LC-MS/MS remains a favourable choice in elucidating crude oil acid structural information as inputs for the modelling of microemulsion using DPD.

A clear advantage of LC-MS/MS is that derivatization is not necessary for although experience in postulating the relationship between molecular ions and fragment ions is important, coupled with a reference to a comprehensive mass spectral libraries [13]. The amount of acids from the crude oil extract remains a crucial input to the simulation which was not explored in this research. Quantitation of acids found from the LC-MS/MS from the analysis was not performed in this study because the response factor for each acid and their corresponding standard have to be obtained and determined.

4. Summary
With the advancement of hyphenated techniques, qualitative analysis of acids from crude oil can be achieved depending on the extraction method, whether derivatization is performed, and the spectrometric instrument
used. Recovery studies from spiked standards were able to be recovered from the GC-MS indicating the liquid-liquid extraction and methylation are suitable for characterizing acids from crude oil. GC-MS method yielded a low molecular weight acid (C4), while LC-MS/MS analysis of the acid-IER extracts showed that medium molecular weight acids (C5-C18) were identified and characterized, with one NA (Z=1) (i.e. cyclic carboxylic acid) being identified from mass spectral. LC-MS/MS will be considered in all future identification and structural determination of crude oil acids in DPD simulation. The concentration of the acids remains an essential input to the microemulsion model and will be investigated in the future.

Acknowledgement
The authors would like to thank the PETRONAS Research Fund for the financial support of this project.

References
[1] Healy R N, Reed R L and Stenmark D G 1976 Soc. Pet. Eng. J. 16 147–60
[2] Healy R N, and Reed R L I 1976 Society of Petroleum Engineers, SPE 5817. Presented at SPE Improved Oil Recovery Symposium, 22-24 March, 1976, Tulsa, Oklahoma, USA
[3] Zhang S, Sun L L, Xu J, Wu H and Wen H 2010 Energy Fuels 24 4312–26
[4] Groot R D and Rabone K L 2001 Biophys. J. 81 725–736
[5] Schulz S G, Kuhn H, Schmid G and Mund C V J 2002 Colloid Polym Sci 283 284–90
[6] Fraaije J G E M, Tandon K, Jain S, Handgraaf J W and Buijse M 2013 Langmuir 29 2136–51
[7] Rekvig L and Frenkel D 2007 J. Chem. Phys. 127
[8] Mohammed M A and Sorbie K S 2009 Colloids Surfaces A Physicochem. Eng. Asp. 349 1–18
[9] Borgund A E, Erstad K and Barth T 2007 Energy and Fuels 21 2816–26
[10] Sun J, Sun L, Liu W, Liu X, Li X and Shen Q 2008 Colloids Surfaces A Physicochem. Eng. Asp. 315 38–43
[11] Saab J, Mokbel I, Razzouk A C, Ainous N, Zydowicz N and Jose J 2005 Energy and Fuels 19 525–31
[12] Salimon J, Omar T A, Salih N 2014 The Scientific World Journal vol. 2014 p 10
[13] Halket J M, Waterman D, Przyborowska A M, Patel R K P, Fraser P D and Bramley P M 2005 J. Exp. Bot. 56 219–43
[14] Mediaas H, Grande K V, Hustad B M, Rasch A, Rueslåtten H G, Vindstad J E and Asa S 2003 SPE Int. Symp. Oilf. Scale 29-30 January, Aberdeen, UK
[15] Lai W-C and Song C 1995 Fuel 74 1436–51
[16] Ahmed M M 2010 Characterization, Modelling, Prediction and Inhibition of Naphthenate Deposits in Oilfield Thesis (Heriot-Watt University) p 251
[17] Orata F 2012 Adv. Gas Chromatogr. – Prog. Agric. Biomed. Ind. Appl. 83–156
[18] John W P S, Rughani J, Green S A and McGinnis G D 1998 J. Chromatogr. A 807 241–51
[19] Holowenko F M, MacKinnon M D and Fedorak P M 2002 Water Res. 36 2843–55
[20] Brient J A, Wessner P J D M N 1995 Naphthenic acids Encyclopedia of chemical technology, vol. 16, 4th ed. (New York, NY: Wiley) pp 1017–29
[21] Rogers V V, Liber K and Mackinnon M D 2002 Chemosphere 48 519–27
[22] Haneef J, Shaharyar M, Husain A, Rashid M, Mishra R, Parveen S, Ahmed N, Pal M and Kumar D 2013 J. Pharm. Anal. 3 341–8
[23] Li S, Lai S, Song J, Qiao C, Liu X, Zhou Y, Cai H, Cai B and Xu H 2014 J. Pharm. Biomed. Anal. 53 946–57
[24] Woudneh M B, Hamilton M C, Benskin J P, Wang G, Mceachern P and Cosgrove J R 2013 J. Chromatogr. A 1293 36–43