Fluid Phase Equilibria 409 (2016) 354–360

Contents lists available at ScienceDirect

Fluid Phase Equilibria

journal homepage: www.elsevier.com/locate/fluid

A molecular structure based model for predicting optimal salinity of anionic surfactants

Marzieh Mavaddat, Siavash Riahi*

Institute of Petroleum Engineering, Faculty of Chemical Engineering, College of Engineering, University of Tehran, Tehran, Iran

ARTICLE INFO

Article history:
Received 13 July 2015
Received in revised form
17 September 2015
Accepted 5 October 2015
Available online 19 October 2015

Keywords:
Optimal salinity
Quantitative structure property relationship
Microemulsion flooding
Surfactant selection
Molecular description

ABSTRACT

Optimal salinity (Sopt) is an important property that influences the relative phase volume and solubilization parameters in microemulsion flooding. Selecting an appropriate surfactant for chemical flooding with high Sopt requires characterizing lots of different chemicals and a sophisticated interpretation procedure. The main objective of this paper is to apply quantitative structure property relationship (QSPR) technique to model the Sopt of 20 different surfactant molecules. The surfactant molecules are sorted into two different groups according to their experimental conditions. Geometrical optimization of surfactants was performed at RM1 level. Then, many structural and quantum chemical descriptors were calculated using different computer software programs. Using variable selection of the genetic algorithm (GA-MLR), two descriptors were introduced as independent variables. The squared correlation coefficient (R^2) and standard deviation (s) calculated for the selected model were 0.940 and 0.643 for molecular group-A; and 0.984 and 0.445 for molecular group-B, respectively. The results demonstrate high estimation accuracy and strong generalization capacity of the models. The descriptors in both models are related to polarizability and ionization potential of the molecules, thus they are conceptually related to the changes in Sopt. The performance of QSPR was further compared with some commonly used constitutional descriptors in previous literature reports. By comparing the results, one can conclude that the estimation of Sopt can be improved significantly by using QSPR compared with other correlations.

1. Introduction

With the development of oil fields, microemulsion flooding has been introduced as one of the most appealing methods in chemical enhanced oil recovery [1]. Microemulsion is a thermodynamically stable dispersion of surfactant, oil and water (brine). It has been used in various fields due to its very low interfacial tension (IFT) and good solubilization capacity [2]. The phase behavior of the surfactant-oil-brine system is one of the key properties characterizing a microemulsion flooding [3,4]. In a microemulsion system, phases are changed from Winsor type I to Winsor type II through Winsor type III by salinity variation at a particular temperature and pressure [5]. Ultralow IFT is obtained when a microemulsion middle phase is created between the oil and water phases. Optimal salinity (Sopt) is the salinity that indicates a minimum in IFT between oil and water phases [2,6]. Estimation of Sopt is of great importance in designing economical microemulsion flooding [7]. Lots of efforts, which are focused on either academic research or specific reservoir applications, have been made to identify the microemulsion phase behavior. These studies represent that the phase behavior mostly depends on the surfactant type [2,8–16].

A wide variety of chemicals including those designed for challenging conditions like high temperature and high salinities are now available for the purpose of surfactant flooding [8,13,17–19]. Many different chemicals should be characterized to identify the best surfactant formulation for a case study [5]. Therefore, a tool for a good pre-selection of surfactants based on the reservoir conditions, which makes the formulation design faster, is highly desirable.

Estimating optimal salinity of surfactants has been studied for several decades. A number of studies were conducted to introduce ethylene oxide number (EON), propylene oxide number (PON) and alkane carbon number (ACN) as surfactant descriptors [13,17,19,20]. Later, the hydrophilic-lipophilic deviation (HLD) approach was introduced to characterize nonionic and ionic surfactant molecules [21,22]. Besides, Solairaj proposed another correlation for surfactant optimal salinity prediction, based on the HLD approach [23].

* Corresponding author.
E-mail address: riahi@ut.ac.ir (S. Riahi).

http://dx.doi.org/10.1016/j.fluid.2015.10.010
0378-3812/

This study added hydrophilic carbon number as another surfactant descriptor. All these attempts have some constrains in common: (I) The method can be applied only if the surfactant molecules contain the functional group described in the models. (II) The aforementioned descriptors may not be a good representation of a chemical structure. (III) There is no evidence of possible interactions between different groups which are present in the molecules. (IV) No validation technique was used for the given correlations. Therefore, a scientific approach is needed to structurally investigate the surfactants and find a correlation for optimal salinity. For this purpose, a quantitative structure property relationship (QSPR) technique has already been introduced.

The QSPR technique describes a mathematical relationship between structural attributes and a macroscopic property in a set of chemicals [24]. In spite of its wide use in many chemical and biological fields of study; very few researchers have focused on surfactant optimal salinity predictions. Barnes et al. first proposed a semi-empirical model for predicting optimal salinity of commercial internal olefin sulfonate surfactants [25]. They introduced a two-variable model with a square correlation coefficient (R^2) of 0.950. The selected descriptors were limited to thermodynamic descriptors and they did not mention any model validation check in their study. Later, Moreau et al. proposed a QSPR model for mixtures of industrial surfactants [26]. In their study, mixing rules were used as a tool to calculate descriptors for each surfactant mixture. They represented a model for each of their surfactant groups with correlation coefficients of more than 0.9. The models had 7 and 9 variables, which made them hard to describe. Besides, their variables were not well introduced and could not be repeated or calculated for other molecular structures. Moreover, validation was not completely studied in their work. To summarize, earlier works on characterizing Sopt using the QSPR method have the following limitations in common: (I) The relationship between the selected descriptors and dependent variables was not logically explained. (II) The validation parameters were not fully examined.

In this article, a new model has been developed to characterize Sopt in microemulsion systems. The modeling has been carried out with more than 240 individual descriptors for 20 different surfactant molecules. The selected descriptors of each surfactant were correlated to provide a simple reference tool for describing the surfactants optimal salinity. Besides the applicability domain of the results, several validation techniques were examined. Finally, the correlation between Sopt and some constitutional descriptors was checked.

2. Materials and methods

The Quantitative Structure Property Relationships (QSPR) method is a computational tool which deals with the correlation between any measured property of a molecule and its structural features. The method usually includes the following steps [27]: (I) Acquiring a data set consisting of a number of molecules along with one of their uniform properties. (II) Extracting structural properties of each molecule in terms of different descriptors. (III) Revealing the best models by various optional algorithms such as stepwise or genetic algorithm. (IV) Validating the model using several validation techniques.

2.1. Data set

We initially extracted 270 different components from phase behavior studies in the literature, each of them might be at different experimental condition, i.e. different temperature, oil phase composition and surfactant concentration [2,6,13,14,17,21,23,28,29]. In order to limit the changing variables and focus the molecular structure descriptors, it is needed to set a data at the same experimental condition. This approach, led to selection of 20 different surfactant molecules to be exactly at the same condition. They were all 2 wt% surfactant solutions in the aqueous phase, n-Decane as the oil phase and sodium chloride for the salinity scan, and temperature is 60 °C. It's worth mentioning that this condition is known to be representative of surfactant solution for use in chemical EOR process at reservoir condition [5,15,17,30–34]. The detailed microemulsion phase behavior tests have been described before [35].

The selected components were sorted into two groups, each having 10 molecules. Group-A contains 0.2 wt% Na_2CO_3 as alkaline and Group-B contains no alkaline. Molecular modeling computations were done separately on each group. The molecular structures and the experimental optimal salinity values for each molecular group are presented in Tables S1 and S2 (supplementary materials).

2.2. Computational details

Chemical structures of the selected surfactants were created using HyperChem [36] and the molecular structures were optimized to their minimum molecular energy at RM1 level. In order to calculate the descriptors, all of the surfactants were introduced to a proper descriptor-calculating software [37]. In addition, a collection of quantum chemical descriptors such as the energy of the highest occupied molecular orbital (HOMO), molecular dipole moments and local charges, were calculated by Gaussian 98 [38]. All of the Gaussian calculations were carried out by employing density functional theory (DFT) at RHF/6-31G level. Finally, 1590 descriptors were calculated for each molecule.

2.3. Data process and modeling

At first, the constant and pseudo constant (more than 95% constant) descriptors were removed. To decrease the redundancy of descriptors, a filtration process was performed to examine the correlations among the descriptors. Among each detected inter-correlated descriptors (i.e. $R^2 > 0.95$), the one with the highest correlation with Sopt was retained and the other was removed. The above process resulted in 241 descriptors for each molecule in group A and 282 descriptors in group B.

Furthermore, a GA-MLR procedure was performed to select the best models with one, two and three variables. The best models are presented in Table 1 along with their appropriate statistical parameters. To check the co-linearity between the variables of each model, R^2 was calculated for any two descriptors. The result from Table 2 shows that there is no significant correlation between the descriptors in each molecular group. This means each variable encodes different aspects of the molecular structure and three of them can be kept.

3. Results and discussion

3.1. Model selection

Comparing the statistical parameters in Table 1 for the one variable models, it can be observed that for both molecular groups, the models with only one variable have a proper statistical fit and good prediction ability. The quality of the models or the anomalies was detected by investigating the scatter plot of experimental versus predicted Sopt in Fig. 1. Although the plot analysis confirms that the models with one variable can be used as predictive models, models with two and three variables have also been discussed to obtain a descriptive model for the surfactant optimal salinity. Models with more than three variables were not checked because