Minerals Engineering 39 (2012) 140–148

Contents lists available at SciVerse ScienceDirect

Minerals Engineering

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

QSAR analysis of selectivity in flotation of chalcopyrite from pyrite for xanthate derivatives: Xanthogen formates and thionocarbamates

Fan Yang, Wei Sun *, Yuehua Hu

School of Resources and Bioengineering, Central South University, Changsha 410083, Hunan, China

ARTICLE INFO

Article history:
Received 6 March 2012
Accepted 1 June 2012
Available online 4 October 2012

Keywords:
Sulfide ores
Flotation collectors
Modeling

ABSTRACT

A quantitative structure–activity relationship (QSAR) study of selectivity in flotation of chalcopyrite from pyrite for xanthates and xanthate derivatives was performed. The genetic function approximation (GFA) algorithm was applied to select the descriptors and to generate the correlation models. Three combinations of topological, structural, physicochemical, spatial and electronic descriptors were applied. Six QSAR models with acceptable R^2 correlation coefficients and Q^2 cross-validation correlation coefficients ($R^2 > 0.75$, $Q^2 > 0.6$) were developed. The models show that shadow area fraction *YZ*, LUMO density of carbonyl oxygen atom and Hirshfeld Fukui indices (+) of thiocarbonyl sulfur atom are the three most significant variables for the selectivity. The descriptor shadow area fraction *YZ* revealed that the planar structure is important for the selectivity. The descriptors LUMO density of carbonyl oxygen atom and Hirshfeld Fukui indices (+) of thiocarbonyl sulfur atom demonstrated LUMO distributions over carbonyl oxygen atom should be large for favorable selectivity. This study concluded that selectivity of xanthates and xanthate derivatives is determined by the strength of the dative bonding between copper atom on the chalcopyrite surface and carbonyl oxygen atom of these collectors and the stability of the six-member ring formed as a result.

1. Introduction

Flotation separation chalcopyrite from pyrite has always been a problem as it lacks effectively selective collectors. Traditionally, xanthates are used as collector though a large amount of lime is required to obtain a highly alkaline condition to achieve the separation (Liu, 2004). Addition of lime, however, may depress chalcopyrite and other coexisting precious metal minerals and cause the erosion of the equipments (Li and Sun, 2000). To overcome this problem, several series of xanthate derivatives such as xanthogen formats and thionocarbamates were invented. They work well in the natural circumstances of copper sulfide ores, i.e. at pH of about five sometimes, with high selectivity against pyrite (Crozier, 1978; Aplan and Chander, 1987; Harris and Fishback, 1954; Seryakova et al., 1975; Woods and Hope, 1999). Especially, in 1980s, based on N-alkyl thionocarbamate N-alkoxycarbonyl thionocarbamates were invented (Fu and Wang, 1987), the selectivity against pyrite in neutral circumstance of which is much better than the former ones.

All along, most of researches concerning xanthates and its derivatives have been focused on their reaction mechanism with sulfide minerals (Fairthorne et al., 1997, 1998; Nagaraj and Brinen, 2001; Hope et al., 2005, 2007; Liu et al., 2006), while few tried to reveal the relation between their structures and selectivity. Ackerman et al. (2000) evaluated the influence of different substituent structures on the flotation of copper sulfide minerals and pyrite and demonstrated a relation between them does exist. Liu et al. (2008) found the relation of efficiencies of N-substituent thionocarbamates with the properties such as energies and compositions of frontier molecular orbital and atomic charges and explained why IBECTC and IBACTC have excellent collecting power for copper sulfide minerals and selectivity against iron sulfide minerals on the basis of the order of electron-donating ability and feedback-electron-accepting ability. In these work, the relations obtained are mostly qualitative and without statistical validations. This study was to systematically investigate the relation between molecular structures and selectivity using quantitative structure–activity relationship (QSAR) modeling technique.

A quantitative structure activity relationship (QSAR) is a method that correlates an activity of a set of compounds quantitatively to chemicals descriptors (structure or property) of those compounds, which are generally obtained from experiments and quantum chemical calculations. The exact form of the relationship between structure and activity can be determined using a variety of statistical methods and computed molecular descriptors. QSAR has the objective of prediction but maintaining a relationship to mechanistic interpretation. Recent years, quantitative structure–activity

---

* Corresponding author. Tel.: +86 371 88830482.
E-mail address: sunmenghu@mail.csu.edu.cn (W. Sun).

0892-6875/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.mineng.2012.06.001

relationship (QSAR) modeling has been applied extensively in pharmaceutical chemistry and predictive toxicology (Yangali-Quintanilla et al., 2010; Wang et al., 2011; Mandal and Roy, 2009; Iyer et al., 2007; Moorthy et al., 2011). There are also several applications in flotation. Natarajan and Nirdosh (2008) selected 10N-arylhydroxamic acids from a virtual database of 3800 compounds with the method of molecular similarity clustering and synthesized and tested them as collectors for sphalerite. Hu et al., (2012) carried out a QSAR research on the quaternary ammonium salt collectors for bauxite reverse flotation and established a robust model.

The purpose of this work was to establish QSAR models of the selectivity in flotation of chalcopyrite from pyrite for xanthates and xanthate derivatives, including xanthogen formates and thionocarbamates, and to give the models mechanistic interpretations to reveal the favorable structural characteristics that enhance their selectivity.

2. Methods

2.1. Data set

In this study the selectivity of 16 sample molecules including xanthates, xanthogen formates and thionocarbamates was collected from Fu et al. (1989). The values are reported in the form of I_{cu}, which was defined and calculated in accordance with Eq. (1):

$$I_{Cu} = \frac{(100 - \%\ \text{Pyrite recovered})}{(100 - \%\ \text{Copper recovered})^2} \tag{1}$$

The names and structures of the 16 collectors were listed in Table 1 with their I_{cu} values.

Generally, the data set is divided into a training set to develop models and a test set to externally validate the constructed QSAR models. As there are no such large data sets available, the set of the 16 molecules were wholly used as the training set and as the test set flotation results of a new set of 6 collectors were obtained. The names and structures of the 6 collectors were listed in Table 2.

The individual-mineral flotation tests were carried out in the SFG micro flotation cell (40 ml). Chalcopyrite and pyrite used in this test were obtained from Donggua Hill copper mine, Anhui, China. They were crushed with hammer wrapped up with cloth to avoid contaminating the minerals before hand picking the pure minerals, which, afterwards, were milled by porcelain ball mill respectively. And the part with size of $-150\sim+200$ mesh of each was used in flotation tests. Measured by X-ray Diffraction and chemical elemental analysis, the purity of chalcopyrite is up to 85.32% containing trace amount of silica and fluorite while the purity of pyrite is as high as 93.56% and no other impurities are tested. First, 2 g minerals were placed in the cell and added 30 ml distilled water. Then, the mixture was agitated for 1 min. Collectors and pine oil were added successively at the interval of 3 min and every time the mixture was blended for 3 min. Last, the froth was collected for 3 min. All tests were run at collector concentration of 1E−4 mol/l and frother concentration of 35 mg/l at natural pH. The flow sheet of flotation is listed in Fig. 1.

Obviously, the two data sets were obtained in two different experimental conditions. And their absolute values are not comparable. So the external validation cannot be performed quantitatively in the general way. Actually, selectivity is different from many other physical and chemical properties. Flotation results that selectivity is obtained from depend largely on the artificial operations and they vary from person to person. Hence, even the values of the predicted selectivity with externally validated models are meaningless for the compounds investigated. The models of selectivity only make sense qualitatively. Therefore, it is acceptable to externally validate the models qualitatively with a data set from a different source for the training set. And in this case, it was performed through comparing the order of selectivity of the test set predicted with QSAR models constructed with the one obtained from flotation tests above.

2.2. Quantum chemical calculations

All molecules were built using the Material Studio 4.0 software. The geometry optimizations and single point calculations were both carried out with the Dmol3 module in Material Studio with Generalized Gradient Approximation (GGA) functional PBE (Perdew et al., 1996) and the basis set DND. Conformers were evaluated to make sure the optimized structures were the global-minimum structures. Xanthates have no stable conformers and the local-minimum structure is global-minimum. For xanthogen formates and thionocarbamates, three most stable conformers (Tobón et al., 2009) were all optimized and evaluated with the single point energies to identify the global-minimum conformations. Finally, vibrational analysis was performed to confirm the structures obtained.

In this study, QSAR models were developed using topological, structural, physicochemical, spatial and electronic descriptors. These descriptors were divided into three categories using different combinations of descriptors. The first category was a combination of topological and structural descriptors while the second category included structural, physicochemical, spatial and electronic descriptors. All these parameters were clubbed together in the third category. All descriptors were calculated using Material Studio 4.0 software and are listed categorically in Table 3. The meanings of all descriptors are available in Material Studio 4.0 Tutorials (Accelrys, 2005).

2.3. QSAR analysis

The correlation coefficients for all pair of descriptor variables used in the models were evaluated to identify highly correlated descriptors in order to detect redundancy in the data set. Hence, some highly correlated ($r \geqslant 0.9$) and constant descriptors were removed from the further consideration. Furthermore, the models with the descriptors that had cross-correlation coefficient of more than 0.6 were also removed from consideration.

Typically, a ratio of five or more measured values for every descriptor should be sought in order to prevent overfitting (Walker et al., 2003). Hence, only two and three-parameter models were considered in this study. For the development of equations, GFA regression analysis was utilized with Material studio 4.0 software.

The genetic function approximation (GFA) algorithm is a useful technique for searching in a large parameter space when the data set is small. This method provides multiple models that are created by evolving random initial models using different descriptors. Models are improved by performing a crossover operation to recombine terms providing better scoring models. The GFA algorithm approach has a number of important advantages over other techniques such as, it builds multiple models rather than a single model, it automatically selects which features are to be used in the models, and it can build models using either a linear relation or higher order polynomial, splines, and Gaussians. (Ponnurengam et al., 2006).

The application of GFA to QSAR/QSPR studies has been successfully used in a wide number of QSPR/QSAR research (Rogers and Hopfinger, 1994; Shi et al., 1999; Hou et al., 2009; Couling et al., 2006; Fan et al., 2001). This method combines Friedman's multivariable adaptive regression splines MARS (Friedman, 1991) and Holland's genetic algorithm (Holland, 1975). This method also used the Friedman's lack-of-fit (LOF) along with the common R-squared, cross validated R-squared, and *F*-value to evaluate the significance