Chemometrics and Intelligent Laboratory Systems 144 (2015) 11–16

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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

Materials design and control synthesis of the layered double hydroxide with the desired basal spacing

Pan Xiong [a], Xiaobo Ji [a], Xin Zhao [a], Wei Lv [b], Taiang Liu [b], Wencong Lu [a,1]

[a] Department of Chemistry, College of Sciences, Shanghai University, Shanghai 200444, China
[b] College of Materials Science and Engineering, Shanghai University, Shanghai 200072, China

ARTICLE INFO

Article history:
Received 22 December 2014
Received in revised form 11 March 2015
Accepted 14 March 2015
Available online 21 March 2015

Keywords:
QSPR
Support vector regression
Layered double hydroxides
Materials genome initiative

ABSTRACT

Efficient and effective prediction of the basal spacing is of great importance to materials design of layered double hydroxides (LDHs). In this work, the QSPR model was constructed to predict the basal spacing of LDHs from 7.5 to 8.0 Å by using the support vector regression (SVR) algorithm. The genetic algorithm (GA)–support vector regression (SVR) method was used to filter the main molecular descriptors in modeling. The QSPR model available was tested by an external test set consisting of 8 compounds. As a case study of controllable synthesis based on the QSPR model, the new LDH of Mg–Al–CO_3 system with the desired basal spacing 7.6 Å, which was screened out from a list of LDH dataset consisting of 30 different kinds of samples, was verified by our experiment with the relative error equal to 0.93%. The method outlined here can be served as a new computational template for the materials design and control synthesis of the LDH with the desired basal spacing based on QSPR model for the first time.

1. Introduction

Layered double hydroxides (LDHs), also known as hydrotalcite-like compounds (HTLCs) or anionic clays, offer potential applications in a diverse range of areas including catalysts [1–3] and catalyst supports [4], storage and triggered release of functional anions [5], carriers for drugs [6,7], optical materials [8,9], supercapacitors [10,11], and electrode modifiers [12]. The crystal structures of LDHs consist of positively charged mixed metal hydroxide sheets, alternating with interlayers containing interlayer exchangeable anions and water molecules. The general formula of LDHs can be represented by $[(M^{2+}_{1-x}M^{3+}_{x}(OH)_2)]^{x+}(A^{n-}_{x/n})^{x-}\cdot mH_2O$, where M^{2+} and M^{3+} are divalent and trivalent metallic cations, such as Mg^{2+}, Ni^{2+}, Mn^{2+}, or Zn^{2+} and Al^{3+}, Cr^{3+}, or Fe^{3+}, and so forth. A^{n-} is an interlayer exchangeable inorganic anion such as CO_3^{2-}, NO_3^-, SO_4^{2-}, Cl^- or various organic anions, and the value of x is equal to the molar ratio $M^{3+}/(M^{2+}+M^{3+})$; m is the number of water molecules located in the interlayer region together with anions [13].

Studies have shown that many applications of LDHs, such as drug sustained release, wastewater treatment and adsorbents, are closely related to its basal spacing [14]. The basal spacing is defined as

$$d_{spacing} = d_{layer} + d_{inter}$$

where d_{layer} represents the thickness of the hydrotalcite brucite-like LDH sheet, and the d_{inter} includes the length of intercalated species and absorptive water in the interlayer (Fig. 1).

The basal spacing of the layered hydroxide cobalt hydroxide nanocones (NCs) can be adjusted by intercalating different amounts of dodecyl sulfate (DS) [15], which in turn impacts the exfoliating property of resultant products into unilamellar nanosheets. It is also important to rationally design the basal spacing of layered transition metal hydroxides for high-performance pseudocapacitors [16], where the basal spacing and capacity are extremely related. Therefore, accurate measurement of the basal spacing and control synthesis of the LDH with the desired basal spacing is a very rewarding job.

One of the big challenges is how we can obtain a LDH with desired basal spacing. One way is called trial and error method, in which batches of sample are empirically prepared and analyzed for a desired compound, and if not desirable, preparation of a new set of samples is needed to try again. Although the experimental value of the basal spacing can be determined by using the powder X-ray diffraction (PXRD), it is not a smart way to try to synthesize all different types of LDHs with different basal spacing because the experiments are very costly and time consuming. Another way is the computational simulation for basal spacing of LDHs even before experiments. Up to now, a number of force-field based simulations and quantum chemical calculations have been reported for predicting the basal spacing of LDHs [17–20]. However, it is almost impossible to simulate a lot of LDHs since the simulation work based on the first principle for every compound is also time-consuming.

Except for experimental measurement and computational simulation based on the first principle, QSPR (Quantitative Structure Property

E-mail address: wclu@shu.edu.cn (W. Lu).
[1] Author information: Address: 311 Room, HA building, 99 Shangda Road, Shanghai 200444, China. Tel./fax: +86 21 6613 3513.

http://dx.doi.org/10.1016/j.chemolab.2015.03.005
0169-7439/

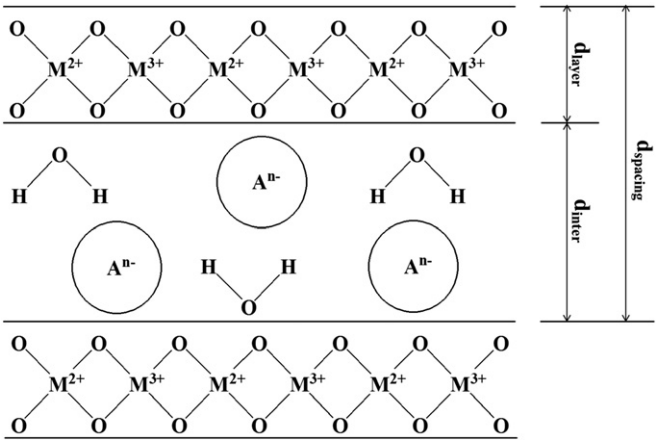

Fig. 1. The illustration of the structure of LDHs.

Relationship) models [21–24] based on data mining can be constructed to correlate the basal spacing of LDHs with their molecular descriptors derived from molecular structure alone. In order to obtain a LDH with desired basal spacing in a short time, we proposed to construct a QSPR model for screening out the compound with desired basal spacing among different kinds of LDHs. QSPR models are no doubt very helpful in materials design [25–29], which is the most important and fundamental work on the background of materials genome initiative (MGI) [30–32] for global competitiveness proposed by the National Science and Technology Council of America. The QSPR method can be used to find semi-empirical rules for predicting the properties of unknown materials.

In this work, the QSPR model was constructed to relate the atomic parameters of LDHs compounds to their basal spacings. The genetic algorithm (GA)–support vector regression (SVR) method [33–35] combined with feature selection method was used to select simultaneously the subset of features and to optimize kernel parameters. The SVR (support vector regression) method integrated with the selected descriptors was applied to develop QSPR model for predicting the basal spacing (d) of LDHs compounds. The QSPR model developed was verified by using the leaving-one-out cross-validation (LOOCV) test and an independent test set. According to the QSPR model, a new LDH of Mg–Al–CO_3 system with the desired basal spacing 7.6 Å was screened out from a list of LDH dataset consisting of 30 different kinds of samples. The desired compound was confirmed by our experiment. In this paper, the QSPR model is proven to be a simple, speedy and effective way to predict the basal spacing of LDHs compound for the first time as far as we know. In addition, the influences of molecular descriptors on the basal spacing were also discussed.

2. Methods

2.1. Computational section

2.1.1. Dataset and molecular descriptors

The dataset consisted of 36 compounds of layered double hydroxides with basal spacing between 7.5 and 8.0 Å collected from the references (Table 1) [36–46]. The full dataset was divided into two subsets: (1) 28 LDHs, used as the internal training set. (2) 8 LDHs, used for testing the predictive power of the model (external test set). We selected training set and test set by random selection. Table 1 lists the basal spacing and selected molecular descriptors of training set and test set. In Table 1, d means basal spacing of LDH; n_a and n_b are the number of atoms of the divalent metal element and the trivalent metal element respectively; z_z/r_z is the ratio of the number of valence electrons to its ionic radius for interlayer anion; and r_a/r_b is the ionic radius ratio of divalent metal elements to trivalent metal elements.

There are twenty atomic parameters (see Table 2) available for the candidates of molecular descriptors that can be used in the prediction of the basal spacing of the LDHs, and the subset of descriptors for modeling were screened by the genetic algorithm (GA)–support vector regression (SVR) method combined with feature selection method. The SVR method is employed as a regression algorithm integrated with the selected descriptors was applied to develop QSPR model for predicting the basal spacing of the LDHs. Root mean squared error (RMSE), mean relative error (MRE) and related coefficient (R) obtained in the leave-one-out cross-validation (LOOCV) method of SVR are taken as evaluation functions.

2.1.2. Implementation

The SVM software package ChemSVM including SVR has been programmed in our lab. The validation of the software has been tested in some applications in chemistry and chemical engineering. The free version of ChemSVM can also be downloaded on the website of Laboratory of Computational Chemistry in Shanghai University (http://chemdata.shu.edu.cn:8080/MyLab/Lab/download.jsp).

2.2. Experimental section

2.2.1. Control synthesis of Mg–Al–CO_3 LDH system with the desired basal spacing

Hydrotalcite was synthesized by urea hydrolysis under hydrothermal conditions according to Ogawa et al. [48] A mixture of $MgCl_2 \cdot 6H_2O$, $AlCl_3 \cdot 6H_2O$ and urea was dissolved in 80 mL of deionized water with stirring at room temperature, with x = 16/9, being $x = [Mg^{2+}]/[Al^{3+}]$. The amount of urea was fixed to be 2.5 times the total metal concentration. The homogeneous solution was transferred into a Teflon-lined autoclave and heated at 120 °C for 24 h. After cooling to room temperature, the resulting precipitate was filtered, and washed with deionized water several times, and then dried in an air oven at 80 °C overnight.

2.2.2. Characterization

X-ray diffraction (XRD) patterns were recorded on a Japan Rigaku D/Max-RB X-ray diffractometer with Cu Kα radiation (λ = 1.54178 Å). Transmission electron microscopy (TEM) images were taken on a JEOL JEM-200 CX with an accelerating voltage of 200 kV. Thermogravimetric-differential thermal analysis (TG-DTA) curves were obtained on a NETZSCH STA 409 PG/PC instrument at a heating rate of 5 °C min^{-1}. Elemental analysis of Mg and Al was carried out using Thermo-Fisher Scientific ICAP 6300 inductive coupled plasma (ICP) emission spectrometer.

3. Results and discussion

3.1. Selection of molecular descriptors

Selection of molecular descriptor is aimed at getting rid of redundant, noisy, or irrelevant descriptors for the model building tasks envisaged, in such a way that the dimensionality of input space can be reduced without loss of important information [49].

The subset of descriptors for modeling was screened by the genetic algorithm (GA)–support vector regression (SVR) method combined with feature selection method. GA has several advantages when compared to other optimization algorithms. It has the ability to move from local optima present on the response surface. It requires no knowledge or gradient information about the response surface and can be employed for a wide variety of optimization problems [49].

In general, the smaller the value of RMSE obtained the better is the set of descriptors gained. The RMSE is defined as: [50]

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(p_i - e_i)^2}{n}}, \quad (1)$$