Chemometrics and Intelligent Laboratory Systems 139 (2014) 121–131

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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

Prediction of elongation at break for linear polymers

Damián Palomba [a,b], Gustavo E. Vazquez [b,c], Mónica F. Díaz [a,b,*]

[a] Planta Piloto de Ingeniería Química (PLAPIQUI) CONICET-UNS, La Carrindanga km. 7, Bahía Blanca 8000, Argentina
[b] Laboratory for Research and Development in Scientific Computing (LIDeCC), DCIC, UNS, Avenida Alem 1253, Bahía Blanca 8000, Argentina
[c] Universidad Católica del Uruguay, Facultad de Ingeniería y Tecnologías,Av. 8 de Octubre 2801, Montevideo 11600, Uruguay

ARTICLE INFO

Article history:
Received 15 July 2013
Received in revised form 13 August 2014
Accepted 20 September 2014
Available online 26 September 2014

Keywords:
Quantitative structure–property relationships
Molecular modeling
Elongation at break
Mechanical properties
Polymers

ABSTRACT

In this paper we present results on prediction of *elongation at break* (target property) for a group of 77 amorphous polymers of high molecular weight. Novel descriptors are proposed in order to better represent structural features related to the target property. These proposed descriptors along with the classic ones, were calculated for the set of polymers. The final descriptors of the predictive model were obtained by using a combination of variable selection method and domain knowledge. The model consisted of three descriptors: *Cross-head Speed* (*CHS*), *Number Average Molecular Weight/Main Chain Surface Area ratio* (Mn/SA_{MC}), and *Normalized Main Chain Mass* (nM_{MC}). By means of a multi-layer perceptron (MLP) neural network a good prediction model ($R^2 = 0.88$ and $MAE = 1.89$) was achieved, which was internally and externally validated. The model shows the advantages of using well-known parameters in the field of polymers and of capturing the structural characteristics of the main and side chains. Thus, more intelligent tools are developed for the design of new materials with a specific application profile.

1. Introduction

Any engineering activity is dependent on a careful and intelligent selection of materials, including the extraction or preparation of raw products, the design of manufacturing and consumer equipment, the operation and maintenance of a plant, to name but a few. A selection among materials must often be made in order to satisfy requirements of performance and/or cost. In approaching a design problem, the engineer will consider first the desired properties of a specific material. The material performance requirements can be divided into five broad categories, namely functional requirements, processability requirements, cost, reliability, and resistance to service conditions [1]. Although the typical approach in the design of new materials has been empirical (formulation, assembling, synthesis, processing and testing), at present there has been much progress in the knowledge of relationships between the molecular structure of a material and its properties [2]. These advances led to improve the ability to predict the material properties prior to synthesis, which in turn is translated into tremendous savings in time and cost. Nevertheless, it is not easy to achieve these predictions as the variables involved are very complex from a quantitative and quality point of view. Subsequently, the design and synthesis of new materials with specific and novel properties have resulted in one of the most dynamic fields of the modern science [3].

In the materials science, the plastics or polymers are everywhere and their use has been increased almost 20-fold in the last 30 years [3]. Polymers have been modified so as to improve their utility and consequently synthetic polymers were developed. Plastics, fibers, elastomers, adhesives, and coatings have come on the scene as a result of a continual search for man-made substances that can either perform better or be produced at lower cost than natural materials. As a consequence, there was an extraordinary growth in the macromolecule field [4]. In particular, thermoplastics are an interesting set of polymers that become liquid when heated and return to the solid state when cooled. This cycle of melting and freezing can be repeated, so that the plastic can be reshaped by heating it. They are useful for a wide variety of applications, including consumer goods, machine parts, medical equipment, packaging and storage materials. They can be classified as amorphous or semicrystalline plastics, according to their molecular arrangement [5]. Even though amorphous polymers are hard and rigid below the glass transition temperature (T_g), they become soft, flexible and can be shaped above the T_g. Thus, mechanical properties exhibit profound changes in the temperature range where this transition occurs. Semicrystalline polymers have melting points that are above their glass transition temperature. The degree of crystallinity and the morphology of the crystalline phase have an important effect on mechanical properties. Semicrystalline plastics become less rigid above their glass transition temperature yet do not flow until the temperature is above the crystalline melting point. Therefore, when it is attempting to predict the mechanical properties of polymers, it is reasonable to assume that the amorphous polymers behave very differently from semicrystalline ones.

* Corresponding author at: C.C. 717, 8000 Bahía Blanca, Argentina. Tel.: +54 291 4861700x255; fax: +54 291 4861600.
E-mail address: mdiaz@plapiqui.edu.ar (M.F. Díaz).

http://dx.doi.org/10.1016/j.chemolab.2014.09.009
0169-7439/

There are numerous mechanical properties that define the profile of applicability of a polymer and, among them, the ability to resist breaking under tensile stress is one of the most important and widely measured of material properties used in structural applications [6]. For polymers, the tensile test provides vital information related to the ductility and strength through the modulus of elasticity, tensile strength at break and elongation at break, among others. The latest is a measure of material ductility. The *elongation at break* value for brittle materials can be vanishingly small — typically is assumed to be zero. While rigid plastics, especially fiber reinforced ones, often exhibit values fewer than 5%, elastomers and some particularly soft thermoplastics tend to have values above 100%. Materials with higher elongation than 100% have a better capacity to handle an excessive load without failure. Furthermore, not only the rate (cross-head speed), affects the final value of *elongation at break*, but also the ambient temperature does [5]. Values of *elongation at break* reported at specific temperatures and cross-head speeds are typical for these test conditions.

We are interested in studying the possibility of estimating or predicting the mechanical properties of a "virtual polymer" (designed molecule) prior to synthesis. This domain has been little investigated due to its complexity. However, it would be a very useful tool in order to describe the application profile of the new polymer, thereby saving time and cost. Many researchers have adopted computational approaches to predict material behaviors [7]. In particular, the QSPR (quantitative structure–property relationship) technique relates specific parameters of molecule structure (descriptors) to the studied property by using a dataset of molecules and experimental property values. This technique began to be used to predict properties of materials in the late 80s [7]. Since then, the study of materials has been very complex as its properties depend not only on the intrinsic material properties, but also on the history of the material (how it was synthesized, processed, and prepared for testing). Therefore it is important for developing QSPR technique, to generate descriptors that take into account all these aspects; that is, physicochemical property descriptors of materials and descriptors related to the synthesis, processing, or sample preparation to develop the most predictive and useful material property models [7]. Furthermore, by addressing the problem of the synthetic polymers, the difficulty of the molecular design occurs since the depiction of polymeric structures cannot be clearly defined in contrast to the small molecules. Among other factors to consider are: the structural (e.g., chain length, tacticity, and monomer segments) and the composition ones (monomer content, blends, and additives) [8]. For these reasons, it becomes a challenge to generate a reliable associated dataset too. One of the earliest and most widely studied of polymer properties has been the T_g, and good prediction results were obtained from synthetic models [9–24]; in contrast, the mechanical properties of polymers have scarcely been explored. Seitz [25] developed semi-empirical and empirical relationships so as to estimate the mechanical properties of polymeric materials from the molecular weight, van der Waals volume, the length and number of rotational bonds in the repeat unit, besides the T_g of the polymer. Thus, he related the molecular properties of the repeating unit to the properties of the polymer. Ulmer et al. [26] reported the use of a combination of neural networks in the modeling process termed "local property experts" for predicting T_g and other physical and mechanical properties of polymeric materials. The researchers expressed special interest to the design of bisphenol-A polycarbonate (BPAPC) with improved impact resistance. An evaluation of nine BPAPC derivatives by means of the trained neural networks delivered three lead compounds. In a subsequent patent, they claimed that these materials showed improved impact resistance [27]. Eslick and Camarda [28] developed in a preliminary work QSPRs for mechanical properties (tensile strength, elongation at break and 300% modulus) of 35 polyurethane elastomers, using topological descriptors. They utilized a stochastic optimization method to find novel polymers with physical and chemical properties matching a given set of properties for electronic applications. Nevertheless, authors did not present the dataset and detailed explanations of results did not either. A similar work was carried on by Eslick et al. [29], who used computational molecular design (CMD) in cross-linked polymer networks in order to facilitate the development of improved polymethacrylate dental materials. CMD employed QSPRs and optimization techniques to design molecules possessing desired properties, among others tensile strength and modulus of elasticity. The authors used three types of graph to calculate the numerical descriptors (topological) of the polymeric structures: monomer, polymer, and full. Moreover, they computed the degree of conversion and the crosslink density as structural descriptors but any dataset was presented neither for target property nor for descriptors. Holder and Liu [30] developed a quantum mechanically based QSPR model for polymer flexural modulus from structural features of tetrameric oligomers of the polymers. A four-descriptor correlation equation with $R^2 = 0.91$ was achieved using a dataset of 25 polymers. The descriptors in the model showed that rigidity of the monomer, electrostatic interactions and branching were the most important contributors to the flexural modulus value for a particular system. As may be seen, all works had a dataset of very few molecules and most cases did not report the polymer molecular weights and other important structural features for the target property did not either. Nonetheless, these early studies formed the basis to begin the exploration of this research field.

To the best of our knowledge, this is the one of the first attempts to investigate the prediction of tensile properties for polymers by means of the QSPR technique with a reasonable number of consistent and reliable data. In this paper we present results about prediction of *elongation at break* (target property) for a group of polymers. A dataset of 77 molecules was built according to a criterion of common parameters for the tensile test. Simplified molecular models (trimers) were designed so as to depict the polymers and new descriptors were proposed. Then, a combination of variable selection method and domain knowledge was applied to choose the model descriptors. Finally, a QSPR model based on neural networks was developed in order to predict the target property and to provide reliability, interpretability and good performance.

2. Experimental section

In this section experimental aspects are explained in detail according to the usual generation process of a QSPR model: (2.1) Dataset generation, (2.2) Structure entry and optimization, (2.3) Molecular descriptors generation, (2.4) Model development and (2.5) Applicability domain. Below a scheme (Fig. 1) is presented as a guide to the reader with the aim of simplifying and summarizing the entire work.

2.1. Dataset generation

Although our original intention was to work with a dataset from the literature, as usual, we had to deal with the fact that there was none for properties derived from tensile test. Therefore, we began the task of obtaining a reliable and consistent dataset. This was built from information provided by PolyInfo [31]. The dataset polymers and their corresponding observed (experimental) and predicted (calculated) *elongation at break* (%) values are shown in Table 1. Several criteria for selection and creation of dataset were used. Next, a cleaning of dataset was applied. The criteria for selecting, cleaning and a description of dataset are presented below.

2.1.1. Criteria for data selection

The influence of average molecular weights on the behavior of the mechanical properties of polymers is well-known. Furthermore, it is known how important are cross-head speed and temperature in the tensile test on the final value of *elongation at break* of polymers [5]. For these reasons, the following polymer parameters were considered with the aim of building the dataset: number and weight average molecular weight (*Mn* and *Mw*, respectively) and polydispersity index (*PDI*); and as regards tensile test parameters, the following ones