Available online at www.sciencedirect.com
SCIENCE DIRECT®

Polymer 44 (2003) 5671–5679

polymer

www.elsevier.com/locate/polymer

Correlation of binary interaction energies for polymer repeat unit pairs via connectivity indices

S. Zhu, D.R. Paul*

Department of Chemical Engineering and Texas Materials Institute, The University of Texas at Austin, Austin, TX 78712, USA

Received 4 April 2003; received in revised form 27 June 2003; accepted 2 July 2003

Abstract

Departures from the geometric mean assumption for the interaction energy between unlike polymer repeat unit pairs *i/j* computed from experimental data, were correlated with the molecular connectivity indices for the repeat units. The effects of various factors on the correlation results were investigated. A better correlation is obtained when terms involving the solubility parameters for the repeat units are included. This empirical scheme provides a viable way to predict interaction energies, and, thus, blend phase behavior and interfacial parameters. The limitations of this scheme and possible ways to improve the correlation are discussed.

Keywords: Binary interaction energies; Connectivity indices

1. Introduction

Thermodynamic interactions determine many aspects of the structure and properties of polymer blends and block copolymers. As a result, accurately quantifying these interactions is important to the rational design of the multi-component blend and copolymer systems and has received considerable attention in the past [1,2]. Many experimental techniques are described in the literature for determining the energetic interaction between unlike repeat units including inverse gas chromatography, calorimetry, melting point depression, sorption, cross-sectional microscopy and small-angle neutron scattering (SANS), etc [1]. For about a decade, our laboratory has successfully employed the concepts of the critical molecular weight, copolymer composition mapping and the combination of these two methods to obtain interaction energies within the framework of the binary interaction model [1,3–5]; the latter resolves the interaction energy down to the individual monomer segment and considers the pair-wise interactions that may occur intermolecularly and intramolecularly. However, all these experimental approaches are costly and time-consuming. Therefore, it would be desirable to develop some systematic methodologies, even semi-empirical ones, to accurately predict blend phase behavior and the responsible individual thermodynamic interaction energies. A common method to estimate the binary interaction energies is the simple solubility parameter approach derived from regular solution theory. However, this method is only useful for very limited cases where the intermolecular forces between unlike repeat units are dispersive in nature [6]. The main objective of this paper is to explore the possibilities for globally correlating binary interaction energy densities with relevant structural parameters of the repeat units leading to a semi-empirical predictive scheme. Our ultimate goal is to understand the polymer–polymer interactions at a deeper level from a systematic perspective. We limit our consideration to systems that do not involve obvious hydrogen bonding; other methodologies have already been developed for these cases [2]. This paper explores the possibility of correlating experimentally determined interaction energies in terms of molecular connectivity indices.

2. Background and theory

The phase behavior of polymer blends is governed by the Gibbs free energy of mixing per unit volume (Δg_{mix}) which, for a blend of monodisperse homopolymers A and B, can be

---

* Corresponding author. Tel.: +1-512-471-5392; fax: +1-512-471-0542.
E-mail address: drp@che.utexas.edu (D.R. Paul).

0032-3861/03/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0032-3861(03)00617-7

simply modeled by the Flory–Huggins (FH) theory [7,8]

$$\Delta g_{\text{mix}} = B\phi_{\text{A}}\phi_{\text{B}} + RT\left[\frac{\rho_{\text{A}}\phi_{\text{A}}\ln\phi_{\text{A}}}{M_{\text{A}}} + \frac{\rho_{\text{B}}\phi_{\text{B}}\ln\phi_{\text{B}}}{M_{\text{B}}}\right] \quad (1)$$

where R is the universal gas constant, T the absolute temperature, ϕ_i, ρ_i and M_i, the volume fraction, density, and molecular weight of component i, respectively, and B is the net binary interaction energy density. Stone and Sanchez recently gave new insights about why the FH model is so enduring even though far more sophisticated theories are rapidly evolving [9]. The interaction energy density B is related to the Flory–Huggins interaction parameter χ by

$$B = \frac{\chi RT}{V_{\text{ref}}} \quad (2)$$

where V_{ref} is an arbitrary reference volume. For present purposes, introduction of this arbitrary reference volume causes considerable problems that are simply avoided by use of the actual interaction energy B.

The binary interaction model has proven quite successful for describing the phase behavior for blends containing copolymers [1], at least when highly specific interactions such as hydrogen bonding are not present. The binary interaction model accounts quite well for the multiple intramolecular and intermolecular interactions that contribute to the net interaction energy density B in Eq. (1) [3–5]. The most general form of this model for mixing two copolymers A and B is given by

$$B = \sum_{i>j}^{\text{Inter}} B_{ij}\phi_i\phi_j - \sum_{i>j}^{\text{Intra}} B_{ij}\phi_i\phi_j \quad (3)$$

where ϕ_i is the volume fraction of monomer i in the copolymer A, ϕ_j, the volume fraction of monomer j in copolymer B, and B_{ij} is the binary interaction energy densities between repeat units i and j. The B_{ij} values can be evaluated from the miscibility behavior of polymer blends using an appropriate thermodynamic theory, such as the FH theory, combined with the binary interaction model.

As mentioned earlier, solubility parameters can be useful for estimating binary interaction energies in some cases. In general, for random mixing, the binary interaction energy density for the i/j pair, B_{ij}, can be written as [6,10,11]

$$B_{ij} = C_{ii} - 2C_{ij} + C_{jj} \quad (4)$$

where C_{ii} and C_{jj} are the cohesive energy densities for homopolymers i and j, respectively, and C_{ij} refers to interactions between the unlike i/j pair. The cross-interaction term, C_{ij}, can be calculated by invoking the geometric mean assumption introduced by Berthelot and London [12,13]. This assumption states that

$$C_{ij} = \delta_i\delta_j \quad (5)$$

where $\delta_i = \sqrt{C_{ii}}$ and $\delta_j = \sqrt{C_{jj}}$ represent the solubility parameters of pure component i and pure component j, respectively. By inserting Eq. (5) into Eq. (4), the following well-known relation for estimating B_{ij} from pure component properties is obtained

$$B_{ij} = (\delta_i - \delta_j)^2 \quad (6)$$

However, this is only a rough approximation since the geometric mean assumption is strictly applicable only for non-polar systems. Furthermore, it is impossible to predict negative interactions by only using Eq. (6). The geometric mean rule can be generalized by defining an empirical term that accounts for the deviation from this assumption [6], i.e. $C_{ij} = (1 - k_{ij})\delta_i\delta_j$, where k_{ij} is a small value compared to unity which is characteristic of the unlike i/j pair interaction. This leads to the following [5]:

$$B_{ij} = (\delta_i - \delta_j)^2 + 2k_{ij}\delta_i\delta_j \quad (7)$$

Interestingly, Eq. (7), unlike Eq. (6), does not preclude negative values of B_{ij}. Values of k_{ij} can be calculated from Eq. (7) by removing the contribution of the solubility parameters approach from experimentally determined B_{ij}; in principle, it may be easier to correlate k_{ij} values with molecular structure than B_{ij}. A critical question is what structural parameters should be used in such a correlation. Based on successes in other structure–property correlations, we propose that the so-called connectivity indices might be the appropriate structural parameters for this purpose. Connectivity indices stems from graph theory and provide a method of molecular structure quantitation in which weighted counts of substructure fragments are incorporated into numerical indices. Structural features such as size, branching, unsaturation, heteroatom content and cyclicity are encoded [14]. Due to the fact that the set of connectivity indices conveys pertinent structural information about the molecules, connectivity indices have been widely used in the study of quantitative structure–property relationships (QSPR) developed for the physical and chemical properties of small molecules and polymers [14–16]. For example, they have proven very successful to predict the chemical reactivity and biological activity of molecules in medicinal, drug and agricultural chemistry [15,16]. Recently, they have also been successfully extended to homopolymers to predict volumetric, mechanical, optical, electrical and magnetic properties, etc., of homopolymers [14]. Therefore, it is reasonably expected that connectivity indices may also be applied to polymer blend systems to produce promising QSPR results. Since details about connectivity indices can be found elsewhere [14–16], only a brief review is provided here to facilitate discussions in later sections. Each connectivity index can be computed from valence bond structures of the molecule or repeat unit of interest. In the literature, most of the homopolymer properties are correlated using four indices, the zero order connectivity indices, ${}^0\chi$ and ${}^0\chi^{\text{v}}$, and the first order connectivity indices, ${}^1\chi$ and

${}^{1}\chi^{\mathrm{v}}$, which are defined as follows [14]

$$ {}^{0}\chi \equiv \sum \frac{1}{\sqrt{\alpha}}, \qquad {}^{0}\chi^{\mathrm{v}} \equiv \sum \frac{1}{\sqrt{\alpha^{\mathrm{v}}}}, \qquad {}^{1}\chi \equiv \sum \frac{1}{\sqrt{\beta}}, \qquad {}^{1}\chi^{\mathrm{v}} \equiv \sum \frac{1}{\sqrt{\beta^{\mathrm{v}}}} \tag{8} $$

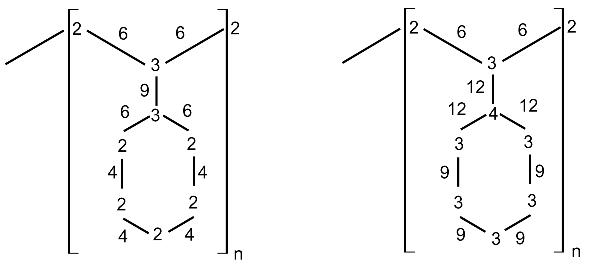
where α is equal to the number of non-hydrogen atoms to which a given non-hydrogen atom is bonded in the molecule and

$$ \alpha^{\mathrm{v}} \equiv \frac{Z^{\mathrm{v}} - N_{\mathrm{H}}}{Z - Z^{\mathrm{v}} - 1} $$

where Z is the atomic number, Z^{v}, the number of valence electrons of the atom, and N_{H} is the number of hydrogen atoms bonded to the atom. The two atomic indices α and α^{v} describe the electronic environment and the bonding configuration for each non-hydrogen atom in the molecule. The two bond indices β and β^{v} are defined as products of the atomic indices, i.e. $\beta_{ij} \equiv \alpha_i \alpha_j$, $\beta_{ij}^{\mathrm{v}} \equiv \alpha_i^{\mathrm{v}} \alpha_j^{\mathrm{v}}$, for the ith and jth adjacent atoms. In Eq. (8), χ terms are the sums of the corresponding $\alpha(\alpha^{\mathrm{v}})$ or $\beta(\beta^{\mathrm{v}})$ terms over all the non-hydrogen atoms in the repeat unit. An example of the calculations of these χ terms for a styrene repeat unit is shown in Fig. 1 [14].

3. Results and discussion

The approach used here is to calculate the deviation from the geometric mean rule for an i/j pair, k_{ij}, from the experimental B_{ij} value using Eq. (7), and then correlate these k_{ij} values with the structural parameters chosen. To predict a value of B_{ij} for an unknown pair, one would calculate k_{ij} from this correlation, and then compute B_{ij} from Eq. (7) using values of δ_i and δ_j. There have been some very limited reports in the literature on the correlation of k_{ij} for small molecule mixtures [17] and polymer blends [18]. Funk et al. found some systematic variation of k_{ij}, determined from the phase equilibrium data, with the degree of branching of the saturated component for mixtures of aromatic and saturated hydrocarbons [17]. However, as they pointed out, their correlation is restricted to a narrow class of mixtures. In general, the attempts to globally correlate k_{ij} with structural parameters have met with little success in the past as pointed out by Barton and by Prausnitz et al. [6,19–21]. This failure may be due to not choosing the most appropriate structural parameters. In this study, we propose that k_{ij} may correlate with connectivity indices since these parameters encompass a large amount of structural information as indicated earlier. A significant advantage of the connectivity indices is that they are easily calculated and available for all the repeat units of interest.



(a) (b)

$$ {}^{0}\chi = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} = 5.3973 $$

$$ {}^{0}\chi^{\mathrm{v}} = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{4}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{3}} = 4.6712 $$

$$ {}^{1}\chi = \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{9}} + \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{4}} + \frac{1}{\sqrt{4}} + \frac{1}{\sqrt{4}} + \frac{1}{\sqrt{4}} = 3.9663 $$

$$ {}^{1}\chi^{\mathrm{v}} = \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{6}} + \frac{1}{\sqrt{12}} + \frac{1}{\sqrt{12}} + \frac{1}{\sqrt{12}} + \frac{1}{\sqrt{9}} + \frac{1}{\sqrt{9}} + \frac{1}{\sqrt{9}} + \frac{1}{\sqrt{9}} = 3.0159 $$

(c)

Fig. 1. Calculation of the zero order and first order connectivity indices for polymer chains, using polystyrene (PS) as an example. (a) hydrogen-suppressed graph with α values at the vertices and β values along the edges; (b) hydrogen-suppressed graph with α^{v} values at the vertices and β^{v} values along the edges; (c) summation of the reciprocal square roots of the α, α^{v}, β and β^{v} values to calculate ${}^{0}\chi$, ${}^{0}\chi^{\mathrm{v}}$, ${}^{1}\chi$ and ${}^{1}\chi^{\mathrm{v}}$ by Eq. (8).

3.1. Calculation of experimental k_{ij} and connectivity indices for repeat units

In order to calculate k_{ij} from B_{ij}, or vice visa, from Eq. (7), an appropriate approach to determine the solubility parameters for the repeat units i and j is required. Numerous group contribution methods for calculating solubility parameters were investigated and compared with each other. The method proposed by Van Krevelen and Hoftyzer was found to provide a more consistent correlation and to be more useful since it is the most comprehensive tabulation of molar cohesion parameters [22]. Most of the experimentally determined B_{ij} values [1,23] used here were determined in our laboratory by using critical molecular weight, copolymer composition mapping and copolymer/critical molecular weight methods. For accuracy and consistency purposes, only those B_{ij} values obtained by these three approaches are used in the correlation. Table 1 tabulates the k_{ij} values computed from the experimental B_{ij} database mentioned (53 B_{ij} data) and the Van Krevelen–Hoftyzer δ_i and δ_j values shown. The connectivity indices for the repeat units i and j calculated by using Eq. (8) are also shown in Table 1.

3.2. Correlation strategy and results

In the literature, the intensive connectivity index ξ defined by [14]

$$ \xi = \chi/N \tag{9} $$

where N is the number of non-hydrogen atoms in the repeat unit is also widely used to correlate the intensive properties of homopolymers. In this work, it was found that the extensive connectivity indices χ lead to better correlation of

Table 1
Connectivity indices, interaction information for the repeat unit pairs used in correlation

Binary pair (*i*/*j*)	B_{ij}[a] (cal/cm^3)	Temperature (°C)	δ_i[b] (cal/cm^3)$^{0.5}$	δ_j[b] (cal/cm^3)$^{0.5}$	k_{ij}	$^0\chi_i$	$^0\chi_i^v$	$^1\chi_i$	$^1\chi_i^v$	$^0\chi_j$	$^0\chi_j^v$	$^1\chi_j$	$^1\chi_j^v$
S/MMA	0.22	120	9.34	9.32	0.00126	5.3973	4.6712	3.9663	3.0159	5.4916	4.5236	3.1885	2.2736
S/AN	7.02	120	9.34	12.60	−0.0153	5.3973	4.6712	3.9663	3.0159	2.9916	2.2317	1.9319	1.3288
S/EA	0.86	120	9.34	9.58	0.00448	5.3973	4.6712	3.9663	3.0159	5.2760	4.3081	3.3425	2.5092
S/*n*BA	0.28	120	9.34	9.16	0.00146	5.3973	4.6712	3.9663	3.0159	6.6902	5.7233	4.3425	3.5092
S/*n*BMA	0.21	120	9.34	8.76	−0.000772	5.3973	4.6712	3.9663	3.0159	7.6129	6.6449	4.6885	3.8611
S/*t*BMA	0.75	130	9.34	8.82	0.00291	5.3973	4.6712	3.9663	3.0159	7.9916	7.0236	4.3349	3.5695
S/CHMA	−0.03	130	9.34	8.92	−0.00122	5.3973	4.6712	3.9663	3.0159	8.6044	7.6365	5.7061	4.9176
S/PhMA	0.2	130	9.34	9.67	0.000517	5.3973	4.6712	3.9663	3.0159	8.6045	6.9104	5.7061	3.9802
S/EMA	−0.036	30	9.34	9.12	−0.000490	5.3973	4.6712	3.9663	3.0159	6.1987	5.9916	3.6052	3.3005
S/MA	10.7	170	9.34	11.63	0.0250	5.3973	4.6712	3.9663	3.0159	5.0165	3.3794	3.3045	2.0605
S/*n*PMA	−0.031	30	9.34	8.98	−0.000970	5.3973	4.6712	3.9663	3.0159	6.9058	5.9378	4.1885	3.3611
S/PC	0.37	180	9.34	8.34	−0.00404	5.3973	4.6712	3.9663	3.0159	13.4579	10.8435	9.0537	6.1634
S/PPO	−0.38	140	9.34	10.33	−0.00715	5.3973	4.6712	3.9663	3.0159	6.4307	5.5629	4.2709	3.0629
S/VC	0.27	100	9.34	9.67	0.000897	5.3973	4.6712	3.9663	3.0159	2.2845	2.4183	1.3938	1.4711
S/TMPC	−0.14	180	9.34	8.83	−0.00242	5.3973	4.6712	3.9663	3.0159	16.9389	14.5341	10.6964	7.8301
α-Me S/AN	7.96	150	8.97	12.60	−0.0229	6.3200	5.5939	4.3123	3.3678	2.9916	2.2317	1.9319	1.3288
α -MeS/MMA	0.12	150	8.97	9.32	0	6.3200	5.5939	4.3123	3.3678	5.4916	4.5236	3.1885	2.2736
α -MeS/PPO	0.06	25	8.97	10.33	−0.00961	6.3200	5.5939	4.3123	3.3678	6.4307	5.5629	4.2709	3.0629
α -Me S/TMPC	0.18	180	8.97	8.83	0.00101	6.3200	5.5939	4.3123	3.3678	16.9389	14.5341	10.6964	7.8301
α -Me S/VC	0.37	130	8.97	9.67	−0.000660	6.3200	5.5939	4.3123	3.3678	2.2845	2.4183	1.3938	1.4711
MMA/AN	4.51	120	9.32	12.60	−0.0265	5.4916	4.5236	3.1885	2.2736	2.9916	2.2317	1.9319	1.3288
MMA/methyl acrylate	0.2	120	9.32	9.73	0.000165	5.4916	4.5236	3.1885	2.2736	4.5689	3.6010	2.8425	1.9217
MMA/EA	0.11	120	9.32	9.58	0.000236	5.4916	4.5236	3.1885	2.2736	5.2760	4.3081	3.3425	2.5092
MMA/*n*BA	0.45	120	9.32	9.16	0.00249	5.4916	4.5236	3.1885	2.2736	6.6902	5.7233	4.3425	3.5092
MMA/*t*BMA	1.35	130	9.32	8.82	0.00668	5.4916	4.5236	3.1885	2.2736	7.9916	7.0236	4.3349	3.5695
MMA/CHMA	0.73	130	9.32	8.92	0.00344	5.4916	4.5236	3.1885	2.2736	8.6044	7.6365	5.7061	4.9176
MMA/PhMA	0.23	130	9.32	9.67	0.000612	5.4916	4.5236	3.1885	2.2736	8.6045	6.9104	5.7061	3.9802
MMA/MA	7.18	170	9.32	11.63	0.00843	5.4916	4.5236	3.1885	2.2736	5.0165	3.3794	3.3045	2.0605
TMPC/AN	5.67	180	8.83	12.60	−0.0382	16.9389	14.5341	10.6964	7.8301	2.9916	2.2317	1.9319	1.3288
TMPC/MA	10	170	8.83	11.63	0.0104	16.9389	14.5341	10.6964	7.8301	5.0165	3.3794	3.3045	2.0605
TMPC/PC	−0.3	140	8.83	8.34	−0.00365	16.9389	14.5341	10.6964	7.8301	13.4579	10.8435	9.0537	6.1634
PPO/AN	10	140	10.33	12.60	0.0187	6.4307	5.5629	4.2709	3.0629	2.9916	2.2317	1.9319	1.3288
AN/MA	−0.35	120	12.60	11.63	−0.00436	2.9916	2.2317	1.9319	1.3288	5.0165	3.3794	3.3045	2.0605
AN/EA	5.01	120	12.60	9.58	−0.0169	2.9916	2.2317	1.9319	1.3288	5.2760	4.3081	3.3425	2.5092
AN/*n*BA	6.63	120	12.60	9.16	−0.0223	2.9916	2.2317	1.9319	1.3288	6.6902	5.7233	4.3425	3.5092
AN/*t*BMA	6.42	130	12.60	8.82	−0.0353	2.9916	2.2317	1.9319	1.3288	7.9916	7.0236	4.3349	3.5695
AN/CHMA	6.66	130	12.60	8.92	−0.0304	2.9916	2.2317	1.9319	1.3288	8.6044	7.6365	5.7061	4.9176
AN/PhMA	4.47	130	12.60	9.67	−0.0169	2.9916	2.2317	1.9319	1.3288	8.6045	6.9104	5.7061	3.9802
AN/*n*PMA	5.85	30	12.60	8.98	−0.0321	2.9916	2.2317	1.9319	1.3288	6.9058	5.9378	4.1885	3.3611
AN/VC	4.24	130	12.60	9.67	−0.0178	2.9916	2.2317	1.9319	1.3288	2.2845	2.4183	1.3938	1.4711
PC/MA	8.3	180	8.34	11.63	−0.0130	13.4579	10.8435	9.0537	6.1634	5.0165	3.3794	3.3045	2.0605
PC/MMA	0.074	150	8.34	9.32	−0.00567	13.4579	10.8435	9.0537	6.1634	2.9916	2.2317	1.9319	1.3288
MMA/*n*PA	0.40	120	9.32	9.17	0.00208	5.4916	4.5236	3.1885	2.2736	5.9831	5.0152	3.8425	3.0092
MMA/*n*HA	0.82	120	9.32	8.75	0.00305	5.4916	4.5236	3.1885	2.2736	8.1044	7.1365	5.3425	4.5092
MMA/*n*DA	0.98	120	9.32	8.48	0.00175	5.4916	4.5236	3.1885	2.2736	10.9329	9.9649	7.3425	6.5092

Table 1 (*continued*)

Binary pair (*i*/*j*)	B_{ij}[a] (cal/cm^3)	Temperature (°C)	δ_i[b] (cal/cm^3)$^{0.5}$	δ_j[b] (cal/cm^3)$^{0.5}$	k_{ij}	$^0\chi_i$	$^0\chi_i^{\mathrm{v}}$	$^1\chi_i$	$^1\chi_i^{\mathrm{v}}$	$^0\chi_j$	$^0\chi_j^{\mathrm{v}}$	$^1\chi_j$	$^1\chi_j^{\mathrm{v}}$
AN/methyl acrylate	4.2	120	12.60	9.73	−0.0167	2.9916	2.2317	1.9319	1.3288	4.5689	3.601	2.8425	1.9217
AN/*n* PA	5.71	120	12.60	9.17	−0.0237	2.9916	2.2317	1.9319	1.3288	5.9831	5.0152	3.8425	3.0092
AN/*n* HA	7.68	120	12.60	8.75	−0.0322	2.9916	2.2317	1.9319	1.3288	8.1044	7.1365	5.3425	4.5092
AN/*n* DA	8.47	120	12.60	8.48	−0.0396	2.9916	2.2317	1.9319	1.3288	10.9329	9.9649	7.3425	6.5092
S/methyl acrylate	1.12	120	9.34	9.73	0.00520	5.3973	4.6712	3.9663	3.0159	4.5689	3.601	2.8425	1.9217
S/*n* PA	0.49	120	9.34	9.17	0.00268	5.3973	4.6712	3.9663	3.0159	5.9831	5.0152	3.8425	3.0092
S/*n* HA	0.42	120	9.34	8.75	0.000455	5.3973	4.6712	3.9663	3.0159	8.1044	7.1365	5.3425	4.5092
S/*n* DA	0.49	120	9.34	8.48	−0.00156	5.3973	4.6712	3.9663	3.0159	10.9329	9.9649	7.3425	6.5092

[a] See Refs. [1,23].

[b] Calculated by using the group contribution method proposed by Van Krevelen and Hoftyzer [22].

the experimental k_{ij} values. This could be due to the fact that χ terms include the contribution from the number of non-hydrogen atoms in the repeat unit, i.e. the N term in Eq. (9), and thus, the structural information about the molecular size for the repeat units is more completely considered in the correlation. Molecular size for the repeat units is an important factor affecting the k_{ij} values as reported from the literature [6]. For typical molecules, polarizability is roughly proportional to molecular size, a parameter important to the intermolecular forces between unlike repeat unit pair. Therefore, only the correlation results obtained by using extensive connectivity indices χ are presented in the following discussions.

Various combinations of the connectivity indices were tested to obtain a correlation, and the following general mathematical form proved most promising:

$$\begin{aligned}k_{ij} = {} & \sum_{n=1}^{N_1}\sum_{k=0}^{1}[{}^kA_n|{}^k\chi_i - {}^k\chi_j|^n + {}^kA_n^{\mathrm{v}}|{}^k\chi_i^{\mathrm{v}} - {}^k\chi_j^{\mathrm{v}}|^n] \\ & + \sum_{n=1}^{N_2}\sum_{k=0}^{1}[{}^kB_n|{}^k\chi_i + {}^k\chi_j|^n + {}^kB_n^{\mathrm{v}}|{}^k\chi_i^{\mathrm{v}} + {}^k\chi_j^{\mathrm{v}}|^n] \\ & + \sum_{n=1}^{N_3}\sum_{k=0}^{1}[{}^kC_n|\ln({}^k\chi_i/{}^k\chi_j)|^n + {}^kC_n^{\mathrm{v}}|\ln({}^k\chi_i^{\mathrm{v}}/{}^k\chi_j^{\mathrm{v}})|^n] \\ & + \sum_{n=1}^{N_4}[|D_n\ln(\delta_i/\delta_j)|^n] + \text{Intercept Term} \qquad (10)\end{aligned}$$

where A, B, C and D are regression coefficients, χ_i and χ_i, the connectivity indices defined by Eq. (8) for i and j repeat units, δ_i and δ_j, the solubility parameters for repeat unit i and j, respectively, n, an integer exponent, and N_1, N_2, N_3 and N_4 are the maximum values of n used and were restricted to be no greater than 3 for simplicity. In Eq. (10), the square brackets represent that the expressions inside the brackets are optional terms; N_1, N_2, N_3 and N_4 may be different from each other. Absolute values are used in the correlation equation due to the fact that k_{ij} must be equal to k_{ji}. It should be emphasized there are no cross terms of connectivity indices expressions in Eq. (10). For each set of regression coefficients A, B, C and D values in Eq. (10), the correlation constructs an objective function defined as the sum of the orthogonal distances between experimentally determined k_{ij} and those defined by Eq. (10); and, then determines the set of A, B, C and D that minimizes this objective function.

Many cases of Eq. (10) were tested to see which offered the better correlation; some of them are shown graphically in what follows. Fig. 2 shows a plot of the experimentally determined k_{ij} (Table 1) versus the predicted k_{ij} from Eq. (10). The solid line represents perfect correlation where the experimentally determined k_{ij} is equal to the predicted k_{ij}. Eight connectivity indices expressions, not including the δ terms, were used in the correlation in Fig. 2. This is a very poor correlation since the corresponding correlation

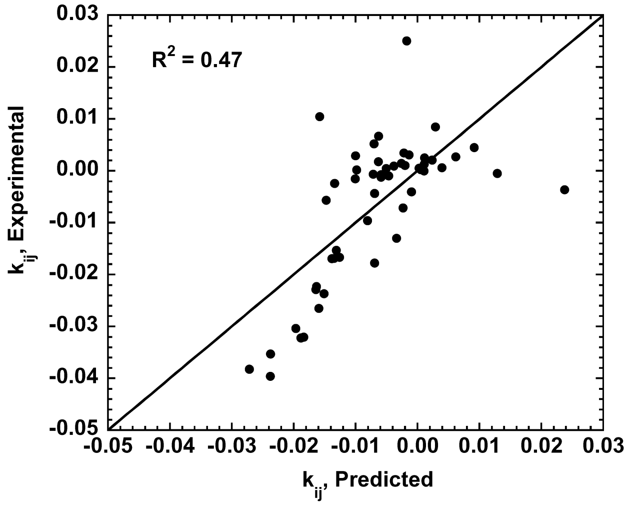


Fig. 2. Predicted k_{ij} as a function of experimental k_{ij} for case 1 in Table 2.

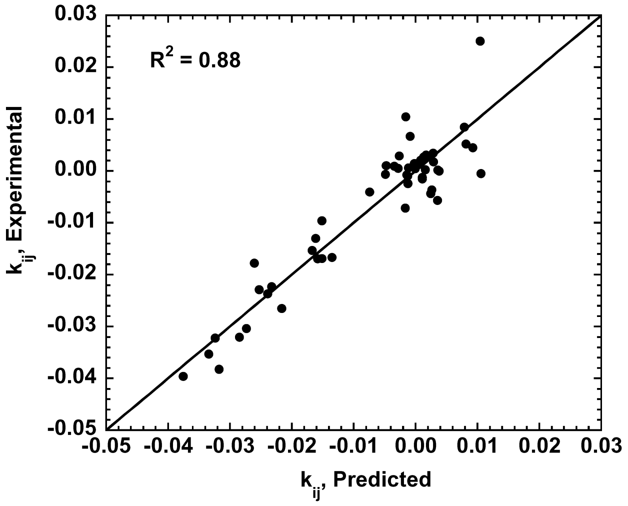


Fig. 3. Predicted k_{ij} as a function of experimental k_{ij} for case 6 in Table 2.

coefficient R^2 is only 0.47. As can be seen in Fig. 2, most of the data points scatter far away from the 45° diagonal line.

Table 2 summarizes the results of more selected correlations. It includes both the cases where the δ terms are not included and those where the δ terms are employed in the correlation equation. As can be clearly seen from Table 2, the correlation coefficient R^2 can be increased by increasing the number of connectivity indices expressions used in the correlation. For example, when the δ terms are not included, R^2 is 0.81 when the number of terms is 20 and increases to 0.99 when the number of terms is 36. Of course, it should be expected that the quality of the correlation or R^2 will strongly depend upon the number of the predictor terms; in principle, a perfect correlation is possible if the number of parameters is equal to the number of experimental data points. The improved correlations obtained without the δ terms are mostly because more predictor terms are employed.

On the other hand, as seen in Table 2, inclusion of solubility parameter terms greatly improves the correlation. For example, with only nine predictor terms, the correlation coefficient R^2 can reach as large as 0.88 if the δ terms are used (Fig. 3); whereas, 22 connectivity indices expressions are required to obtain about the same quality of correlation if the δ terms are not used. With 26 predictor terms in Eq. (10), R^2 can be as large as 0.98 if the δ terms are used in the correlation, which suggests that the global correlation is quite good (Fig. 4). However, the improvement due to adding more terms has a point of diminishing returns. It is quite possible that some other mathematical function than those employed in Eq. (10) would lead to a more efficient correlation in terms of the number of adjustable parameters needed; however, we have not fully explored this option since the main objective of this study is to develop an empirical predictive correlation instead of to obtain the minimum number of parameters.

The effects of including solubility parameter terms on the correlation results may be explained by revisiting the factors that contribute to the deviation from the geometric mean rule from the molecular perspective. As pointed out in the literature [6], mismatch of the molecular characteristics for the repeat units is generally the key to the k_{ij} value. It should be noted that there are still some structural features for the repeat units that are missing in the connectivity indices since they do not completely specify all information about the molecular geometry or polarity of the repeat units which may be important to k_{ij}. Recently [21], we showed some evidence that polarity, polarizability, the first ionization potential, etc, have appreciable effects on k_{ij} values. Therefore, it is highly expected that the correlation may be greatly improved by incorporating terms involving such molecular characteristics in the correlation equation. Unfortunately, quantitative information about the dipole moment, polarizability, or first ionization potential of polymer repeat units are quite limited which as a practical matter precludes including such terms in the correlation. However, solubility parameters can be easily calculated by

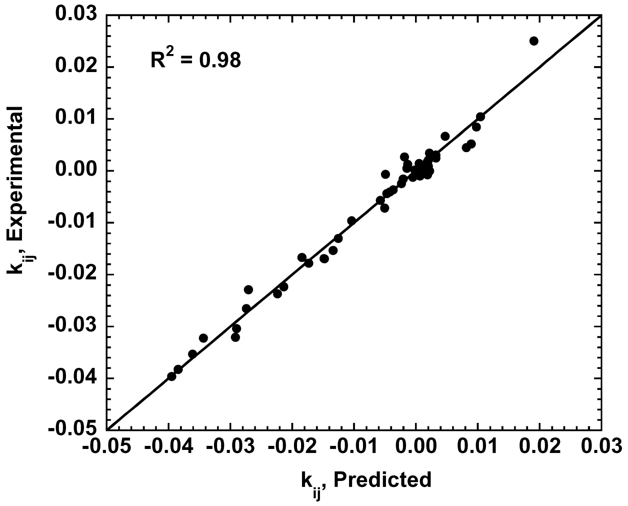


Fig. 4. Predicted k_{ij} as a function of experimental k_{ij} for case 10 in Table 2.

Table 2
Summary of selected correlation results by using Eq. (10)

Regression coefficient terms	Predictor terms	Values of regression coefficients									
		Without δ terms					With δ terms				
		Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10
${}^0A_1$	$\lvert {}^0\chi_i - {}^0\chi_j \rvert$	−0.0173	−0.0704	−0.1130	0.0474	−0.2932	−0.0193	−0.0169	−0.0142	−0.0147	−0.0184
${}^0A_1^v$	$\lvert {}^0\chi_i^v - {}^0\chi_j^v \rvert$	0.0101	0.0672	0.1211	0.0032	0.1617	0.0231	0.0200	0.0188	0.0181	0.0221
${}^1A_1$	$\lvert {}^1\chi_i - {}^1\chi_j \rvert$	0.0321	0.0716	0.0804	−0.0080	0.2866	0.0097	0.0092	−0.0022	0.0032	−0.0033
${}^1A_1^v$	$\lvert {}^1\chi_i^v - {}^1\chi_j^v \rvert$	−0.0334	−0.0642	−0.1064	−0.0257	−0.1040	−0.0130	−0.0127	−0.0026	−0.0057	0.0050
${}^0A_2$	$\lvert {}^0\chi_i - {}^0\chi_j \rvert^2$	–	–	–	−0.0056	0.0287	–	–	–	0.0001	0.0024
${}^0A_2^v$	$\lvert {}^0\chi_i^v - {}^0\chi_j^v \rvert^2$	–	–	–	0.0052	−0.0177	–	–	–	0.0002	−0.0024
${}^1A_2$	$\lvert {}^1\chi_i - {}^1\chi_j \rvert^2$	–	–	–	0.0043	−0.0418	–	–	–	−0.0011	−0.0005
${}^1A_2^v$	$\lvert {}^1\chi_i^v - {}^1\chi_j^v \rvert^2$	–	–	–	−0.0065	0.0247	–	–	–	0.0004	−0.0026
${}^0A_3$	$\lvert {}^0\chi_i - {}^0\chi_j \rvert^3$	–	–	–	–	−0.0008	–	–	–	–	−0.0003
${}^0A_3^v$	$\lvert {}^0\chi_i^v - {}^0\chi_j^v \rvert^3$	–	–	–	–	0.0003	–	–	–	–	0.0003
${}^1A_3$	$\lvert {}^1\chi_i - {}^1\chi_j \rvert^3$	–	–	–	–	0.0016	–	–	–	–	0.0002
${}^1A_3^v$	$\lvert {}^1\chi_i - {}^1\chi_j \rvert^3$	–	–	–	–	−0.0007	–	–	–	–	0.0002
${}^0B_1$	$\lvert {}^0\chi_i - {}^0\chi_j \rvert$	0.0201	0.0193	−0.1008	−0.0295	−0.0124	0.0096	0.0023	−0.0598	−0.0704	−0.1058
${}^0B_1^v$	$\lvert {}^0\chi_i^v - {}^0\chi_j^v \rvert$	−0.0185	−0.0237	0.0382	0.0113	−0.0901	−0.0144	−0.0147	0.0074	0.0268	0.0469
${}^1B_1$	$\lvert {}^1\chi_i - {}^1\chi_j \rvert$	−0.0265	−0.0184	0.1464	0.0353	0.0761	−0.0046	0.0013	0.0291	0.0406	0.0678
${}^0B_1^v$	$\lvert {}^1\chi_i^v - {}^1\chi_j^v \rvert$	0.0334	0.0275	−0.0622	−0.0186	0.0903	0.0107	0.0221	0.0573	0.0293	0.0287
${}^0B_2$	$\lvert {}^0\chi_i - {}^0\chi_j \rvert^2$	–	–	0.0084	0.0014	−0.0039	–	0.0004	0.0048	0.0056	0.0095
${}^0B_2^v$	$\lvert {}^0\chi_i^v - {}^0\chi_j^v \rvert^2$	–	–	−0.0053	−0.0012	0.0098	–	−0.0001	−0.0021	−0.0039	−0.0069
${}^1B_2$	$\lvert {}^1\chi_i - {}^1\chi_j \rvert^2$	–	–	−0.0139	−0.0026	0.0014	–	−0.0006	−0.0031	−0.0043	−0.0092
${}^1B_2^v$	$\lvert {}^1\chi_i^v - {}^1\chi_j^v \rvert^2$	–	–	0.0077	0.0027	−0.0166	–	−0.0006	−0.0065	−0.0021	−0.0006
${}^0B_3$		–	–	−0.0002	–	0.0003	–	–	−0.0001	−0.0001	−0.0002
${}^0B_3^v$	$\lvert {}^0\chi_i^v - {}^0\chi_j^v \rvert^3$	–	–	0.0001	–	−0.0004	–	–	0.0000	0.0001	0.0002
${}^1B_3$	$\lvert {}^1\chi_i - {}^1\chi_j \rvert^3$	–	–	0.0004	–	−0.0006	–	–	0.0001	0.0001	0.0004
${}^1B_3^v$	$\lvert {}^1\chi_i^v - {}^1\chi_j^v \rvert^3$	–	–	−0.0001	–	0.0010	–	–	0.0003	0.0001	−0.0001
${}^0C_1$	$\lvert \ln({}^0\chi_i/{}^0\chi_j) \rvert$	–	0.3307	0.5621	−0.3336	1.5692	–	–	–	–	–
${}^0C_1^v$	$\lvert \ln({}^0\chi_i^v/{}^0\chi_j^v) \rvert$	–	−0.2899	−0.5160	0.0553	−0.6832	–	–	–	–	–
${}^1C_1$	$(\lvert \ln({}^1\chi_i/{}^1\chi_j) \rvert$	–	−0.2178	−0.2765	−0.0238	−1.1361	–	–	–	–	–
${}^1C_1^v$	$\lvert \ln({}^1\chi_i^v/{}^1\chi_j^v) \rvert$	–	0.1458	0.2718	0.1053	0.3604	–	–	–	–	–
${}^0C_2$	$\lvert \ln({}^0\chi_i/{}^0\chi_j) \rvert^2$	–	–	–	0.2554	−0.7099	–	–	–	–	–
${}^0C_2^v$	$\lvert \ln({}^0\chi_i^v/{}^0\chi_j^v) \rvert^2$	–	–	–	−0.2133	0.3025	–	–	–	–	–
${}^1C_2$	$(\lvert \ln({}^1\chi_i/{}^1\chi_j) \rvert^2$	–	–	–	−0.0191	0.4469	–	–	–	–	–
${}^1C_2^v$	$\lvert \ln({}^1\chi_i^v/{}^1\chi_j^v) \rvert^2$	–	–	–	0.0480	−0.1896	–	–	–	–	–
${}^0C_3$	$\lvert \ln({}^0\chi_i/{}^0\chi_j) \rvert^3$	–	–	–	–	0.2211	–	–	–	–	–
${}^0C_3^v$	$\lvert \ln({}^0\chi_i^v/{}^0\chi_j^v) \rvert^3$	–	–	–	–	−0.0949	–	–	–	–	–
${}^1C_3$	$(\lvert \ln({}^1\chi_i/{}^1\chi_j) \rvert^3$	–	–	–	–	−0.0888	–	–	–	–	–
${}^1C_3^v$	$\lvert \ln({}^1\chi_i^v/{}^1\chi_j^v) \rvert^3$	–	–	–	–	0.0246	–	–	–	–	–
D_1	$\lvert \ln(\delta_i/\delta_j) \rvert$	–	–	–	–	–	–	0.0718	0.0427	0.0616	0.0609
D_2	$\lvert \ln(\delta_i/\delta_j) \rvert^2$	–	–	–	–	–	−0.3357	−0.5068	−0.4359	−0.4973	−0.4880
Intercept term	–	−0.0342	−0.0027	−0.0131	0.0560	−0.0257	0.0048	−0.0106	0.0342	0.0322	0.0360

(*continued on next page*)

Table 2 (*continued*)

Regression coefficient terms	Predictor terms	Values of regression coefficients									
		Without δ terms					With δ terms				
		Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10
Number of χ and δ expressions	–	8	12	20	24	36	9	14	18	22	26
Correlation coefficient R^2	–	0.47	0.65	0.81	0.94	0.99	0.88	0.90	0.95	0.96	0.98

various group contribution methods such as the one used earlier [22].

The improvement in the quality of the correlation when terms involving solubility parameters are included implies that Eq. (7) removes only part of their contribution to B_{ij}. No doubt, adding terms with greater information content about the polarity of the repeat units would improve the correlation, but as a practical matter such information is rarely available.

3.3. Effect of temperature

It should be noted that the attempts described here to correlate interaction energies to repeat unit structures have completely ignored the fact that B_{ij} is temperature dependent [1]. Indeed the B_{ij} values employed in this correlation were determined at various temperatures; however, most values were determined in the temperature range from 120 to 170 °C (Table 1). One possible way to solve this problem would be to remove the temperature dependence of B_{ij} by using equation-of-state theories, such as the Sanchez–Lacombe lattice fluid theory, before calculating k_{ij}. The interaction energy, B_{ij}, in the Flory–Huggins framework can be translated into a bare interaction energy, ΔP_{ij}^*, in the Sanchez–Lacombe lattice fluid framework and vice versa [24–27]. One might then replace Eq. (7) with the following as the definition of k_{ij}

$$\Delta P_{ij}^* = (\delta_i - \delta_j)^2 + 2k_{ij}\delta_i\delta_j \qquad (11)$$

Unfortunately, the equation-of-state characteristic parameters for many polymers of interest here are not available [9], which precludes resolving the problem of temperature dependence in this way at the current time.

4. Conclusions

Deviations from the geometric mean assumption for *i*/*j* repeat unit pairs, expressed in terms of k_{ij} defined by Eq. (7), were calculated from the experimentally determined binary interaction energy densities, B_{ij}, reported in the literature, see Table 1. Attempts were made to correlate these k_{ij} values with the connectivity indices for repeat units *i* and *j*. The effects of various factors on the correlation results were examined in detail. The quality of the correlation improved as the number of connectivity indices terms was used (Table 2). However, correlation was significantly improved by incorporating terms involving the solubility parameters of repeat units *i* and *j*. It was suggested that even better correlations might be achieved if terms describing issues of polarity and polarizability were included; however, the lack of information of this type for a broad range of polymer repeat units practically precludes use of this possible improvement. The methodology presented here should provide a useful estimate of k_{ij} and then B_{ij} for repeat unit pairs when no experimental information is currently available. This scheme employs only information that is widely available once the repeat unit structures are known, i.e. the connectivity indices and the solubility parameters; note that this correlation demands that δ_i be calculated by the Van Krevelen–Hoftyzer scheme. The limitations of this method were discussed.

Acknowledgements

This research was funded by National Science Foundation grant number DMR 97-26484 administered by the Division of Materials Research—Polymers Program.

References

[1] Merfeld GD, Paul DR. In: Paul DR, Bucknall CB, editors. Polymer blends: formulation and performance, vols. 1 and 2. New York: Wiley; 2000. p. 55.
[2] Coleman MM, Graf JF, Painter PC. Specific interactions and the miscibility of polymer blends. Lancaster, PA: Technomic Publishing; 1991.
[3] Kambour RP, Bendler JT, Bopp RC. Macromolecules 1983;16:753.
[4] Ten Brinke G, Karasz FE, MacKnight WJ. Macromolecules 1983;16: 1827.
[5] Paul DR, Barlow JW. Polymer 1984;25:487.
[6] Prausnitz JM, Lichtenthaler RN, Gomez de Azevedo E. Molecular thermodynamics of fluid-phase equilibria, 3rd ed. 1999.
[7] Flory PJ. J Chem Phys 1942;10:51.
[8] Huggins ML. J Chem Phys 1941;9:440.
[9] Stone MT, Sanchez IC. In: Paul DR, Bucknall CB, editors. Polymer blends: formulation and performance, vols. 1 and 2. New York: Wiley; 2000. p. 15.
[10] Hildebrand JH, Prausnitz JM, Scott RL. Regular and related solutions. The solubility of gases, liquids, and solids; 1970.
[11] Scatchard G. Chem Rev 1931;8:321.
[12] Berthelot D. Compt Rend 1898;126:1703.
[13] London F. Trans Faraday Soc 1937;33:8.
[14] Bicerano J. Prediction of polymer properties, 3rd ed. New York: Marcel Dekker; 2002.
[15] Kier LB, Hall LH. Molecular connectivity in chemistry and drug research. New York: Academic Press; 1976.
[16] Kier LB, Hall LH. Molecular connectivity in structure–activity analysis. New York: Wiley; 1986.
[17] Funk EW, Prausnitz JM. Ind Engng Chem 1970;62:8.
[18] Graessley WW, Krishnamoorti R, Reichart GC, Balsara NP, Fetters LJ, Lohse DJ. Macromolecules 1995;28:1260.
[19] Barton AFM. Chem Rev 1975;75:731.
[20] Barton AFM. CRC handbook of solubility parameters and other cohesion parameters. 1983.
[21] Zhu S, Paul DR. Macromolecules 2002;35:8227.
[22] Van Krevelen DW. Properties of polymers: their estimation and correlation with chemical structure, 2nd ed. New York: Elsevier; 1976.
[23] Zhu S. PhD Dissertation. The University of Texas at Austin; 2003.
[24] Sanchez IC, Lacombe RH. J Phys Chem 1976;80:2352.
[25] Sanchez IC, Lacombe RH. J Polym Sci, Polym Lett Ed 1977;15:71.
[26] Sanchez IC, Lacombe RH. Macromolecules 1978;11:1145.
[27] Callaghan TA. PhD Dissertation. The University of Texas at Austin; 1992.