Available online at www.sciencedirect.com
ScienceDirect
Polymer 48 (2007) 6115–6124
polymer
www.elsevier.com/locate/polymer

Glass transition temperature prediction of polymers through the mass-per-flexible-bond principle

J. Schut, D. Bolikal, I.J. Khan, A. Pesnell, A. Rege, R. Rojas, L. Sheihet, N.S. Murthy, J. Kohn*

New Jersey Center for Biomaterials, Rutgers, The State University of New Jersey, 145 Bevier Road, Piscataway, NJ 08854, United States

Received 25 March 2007; received in revised form 10 July 2007; accepted 13 July 2007
Available online 28 July 2007

Abstract

A semi-empirical method based on the mass-per-flexible-bond (*M/f*) principle was used to quantitatively explain the large range of glass transition temperatures (T_g) observed in a library of 132 L-tyrosine derived homo, co- and terpolymers containing different functional groups. Polymer class specific behavior was observed in T_g vs. *M/f* plots, and explained in terms of different densities, steric hindrances and intermolecular interactions of chemically distinct polymers. The method was found to be useful in the prediction of polymer T_g. The predictive accuracy was found to range from 6.4 to 3.7 K, depending on polymer class. This level of accuracy compares favorably with (more complicated) methods used in the literature. The proposed method can also be used for structure prediction of polymers to match a target T_g value, by keeping the thermal behavior of a terpolymer constant while independently choosing its chemistry. Both applications of the method are likely to have broad applications in polymer and (bio)material science.

Keywords: Glass transition temperature; Flexible bond; Mass-per-flexible-bond

1. Introduction

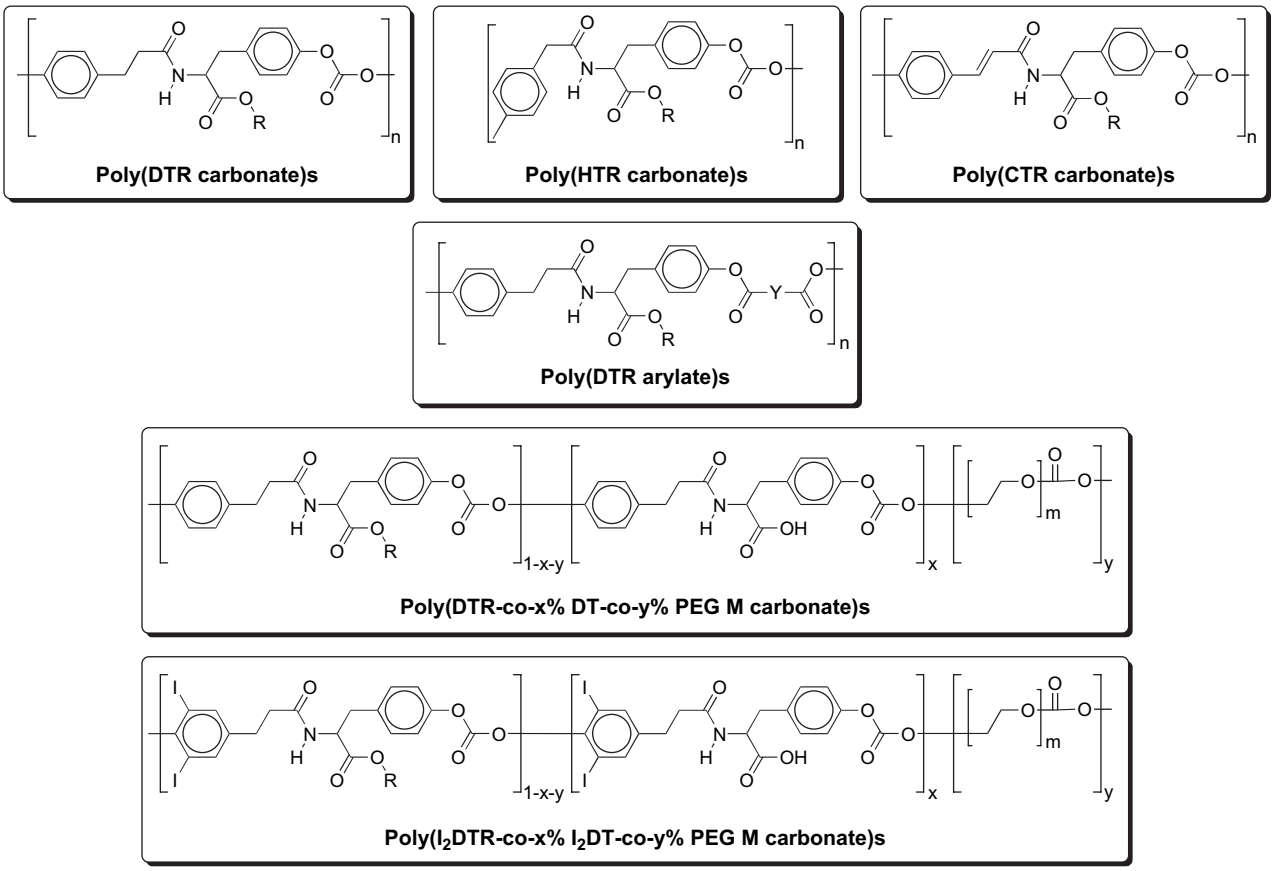
With the advent of parallel synthesis robots, combinatorial methodology and high-throughput screening techniques, it is possible to explore rapidly a large reaction and composition space in a combinatorial fashion [1,2]. It is now possible to synthesize a large number of new polymer structures in a minimal amount of time, which will require analysis and characterization. In order to reduce the number of structures that would appear as noise in a pool of newly synthesized materials, limits for relevant polymer properties should be specified a priori to exclude synthesis of those materials that fall outside these limits. The ability to predict relevant polymer properties is therefore of paramount importance. To this end, our group previously reported on protein adsorption experiments as a rapid screening method for the biological response of polymer surfaces in a biodegradable library [3], and a surrogate modeling approach to correlate polymer structure with biological response data [4–8]. From these results it was evident that the glass transition temperature had a strong correlation with protein adsorption. Additionally, the glass transition temperature is a good indicator for determining whether a polymeric biomaterial can be used as a hard or soft tissue device.

In the following sections we will present experimental T_g values for a large number of polymers in our library comprising a collection of 132 L-tyrosine derived homo-, co-, and terpolymers. Classification of these polymers into distinct classes and empirical formulae for predicting the T_g of any polymer within the library based on mass-per-flexible-bond-principle are discussed. Such a practical tool for predicting the glass transition temperature of homo-, co- and terpolymers will be valuable as a pre-synthesis polymer selection tool. This is one of the necessary steps towards accelerating the discovery of practically applicable polymeric biomaterials.

The general structures of the polymers used in this study are shown in Fig. 1. These poly(ester amide)s, termed

* Corresponding author. Tel.: +1 732 445 0488; fax: +1 732 445 5006.
E-mail address: joachim@rutchem.rutgers.edu (J. Kohn).

0032-3861/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.polymer.2007.07.048



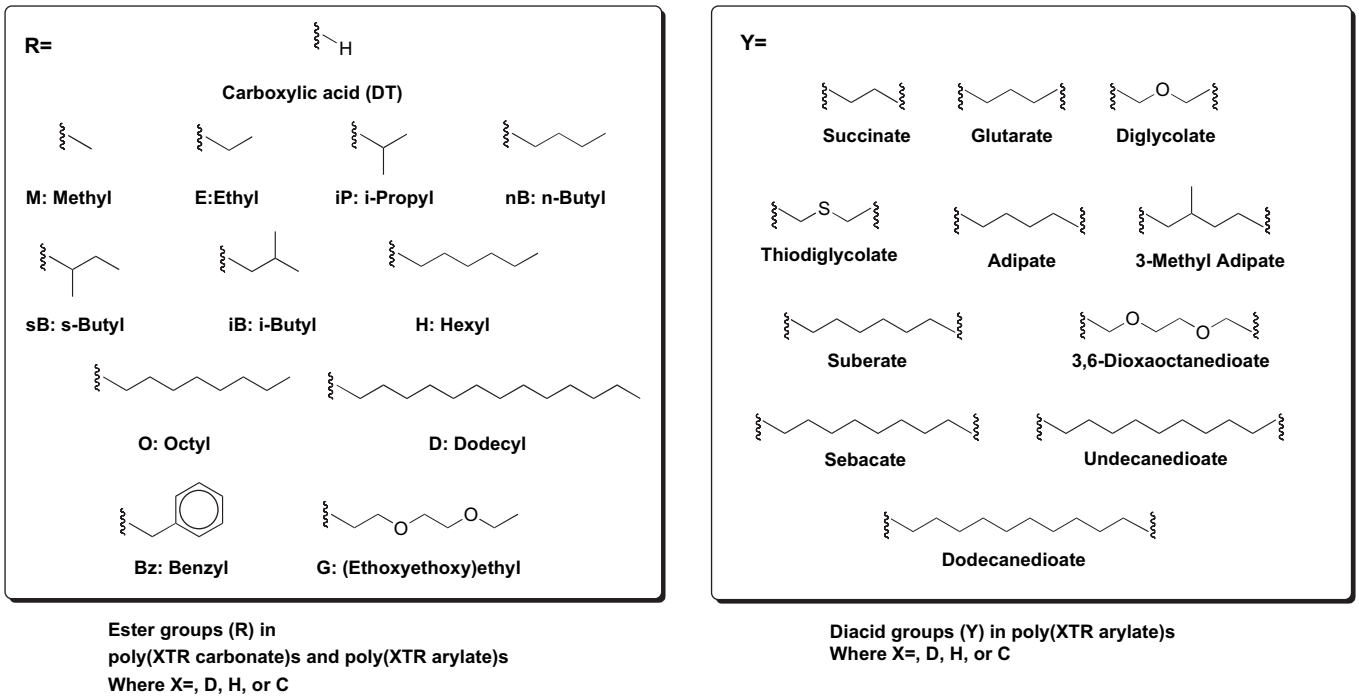


Fig. 1. The library of L-tyrosine derived polymers studied in the present work.

polyarylates, and poly(carbonate amide)s, termed polycarbonates, contain a wide variety of functional groups and heteroatom containing side chains, carboxylic acid groups, alkyl and heteroatom containing main chain functionalities, poly (ethylene glycol) (PEG) comonomers, and covalently bound iodine atoms. This wide variety of polymer structures serves as an excellent test bed for the robustness and general applicability of the method. Although this polymer library has been designed to target biomedical applications (PEG reduces protein adsorption and cell adhesion, and increases water uptake and cell motility [9]; iodine imparts radio-opacity [10]; and the carboxylic acid (DT) groups allow tuning of the polymer's degradation rate [11,12]), it should be stressed that the method presented in this work is universally applicable to any polymer with any number of comonomers, with no limitation to their size or structure, provided that for each polymer class a few initial T_g values are known.

2. Polymer flexibility and T_g

Factors that affect the chain mobility, and thus also influence the T_g, are mainly the molecular weight (chain ends being more mobile than the rest of the chain), chemical structure (stiffness of the chain and interactions with neighboring chains), diluents or plasticizers (which increase the free volume), and cross-links and crystallization (both of which limit the chain mobility) [13]. Here we limit our discussion to perhaps the most important factor, the chemical structure. An empirical parameter that quantifies the influence of structure on chain flexibility, proposed by Di Marzio [14] and further developed by Schneider and Di Marzio [15–17], is the mass-per-flexible-bond of a polymer repeat unit (abbreviated as *M/f*). This is the mass of a monomer divided by the number of flexible bonds in it. *M/f* is a simplified representation of the polymer's conformational entropy, and as such there is a linear correlation between this value and the polymer's T_g, as expressed in Eq. (1) [16],

$$T_g = A\left(\frac{M}{f}\right)_p + C \tag{1}$$

where T_g is the glass transition temperature of polymer (in K), M is the molecular weight of the polymer repeat unit (in g/mol), f is the "number" of flexible bonds in the repeat unit, and $(M/f)_P$ is the mass-per-flexible-bond of the (average) repeat unit. The coefficients A and C are polymer class specific constants that originate from the different densities, steric hindrances and intermolecular interactions of chemically distinct polymers, the factors that collectively affect the cooperative segmental mobility and hence determine the T_g.

It should be emphasized that f is an *effective* number of "flexible" bonds contributing to rotation and conformational changes of the repeat unit from known steric effects and unknown barriers due to inter/intramolecular interactions [16]. As such, *M/f* does not depend on the thermodynamic state of the system; rather it signifies the ability to promote energetically stimulated rotations to release stresses [18]. Although f depends on the temperature, as the polymer is cooled below T_g, f freezes to a value appropriate to that temperature and pressure, and remains constant [14]. Thus, the *M/f* in Eq. (1) is not dependent on temperature.

The *M/f* principle is based on the configurational entropy of the polymer. This equilibrium theory is successful in predicting the value of the T_g but not the kinetics. The alternate theory based on the free volume theory emphasizes the kinetic aspect of the glass transition. There is no convincing evidence that proves the superiority of one over the other, and a complete understanding of the glass transition is not yet available [19]. With no pretension of developing a complete theory for T_g, our goal has been to search for an empirical relationship between the observed T_g and the known structural parameter. For example, in a previous publication of this laboratory, the T_gs of polymers in a 16-member polymer library were correlated with their structure using a flexibility index [20]. This simpler approach was limited to polymers containing aliphatic chains and was not adequate for T_g prediction. The *M/f* parameter turns out to be more robust and useful in predicting the T_g, and will be used in this paper.

Other commonly used methods for T_g prediction are *ab initio* quantum mechanical calculations [21,22], Monte Carlo [23,24], and molecular dynamics simulations [25–29] and semi-empirical or empirical methods based on group contribution methods, often using the QSPR approach through neural network computation [30–44]. While some of these methods yield good results by predicting glass transition temperatures with errors as good as 3–10 K, most predictive accuracies are on the order of 20–100 K. Furthermore, all these methods suffer from the disadvantage that they require a substantial amount of experimental data to construct the models, such as well defined force fields, bond rotation potentials or group contribution parameters, and they are labor intensive and not readily used by non-experts.

The *M/f* principle that will be used in this report can be viewed as a refinement of the widely used group contribution approach for T_g prediction [41]. The *M/f* approach has been used to explain polymer T_g [16,18] and melting points [18], as well as trends in the T_gs of binary, miscible homopolymer blends [14,15,17]. We will here extend these ideas to co- or terpolymers, and develop this principle into a simple and practical, semi-empirical T_g prediction tool for use in homo-, co- and terpolymers. We will first present rules for the assignment of bond flexibility, provide formulae for the calculation of the *M/f* values of co- and terpolymers (and beyond), and describe the actual T_g prediction method as well as the method used to evaluate the accuracy of the predicted T_g values. We then show that the method may also be used in the reverse manner: as a structure prediction tool to match terpolymers to a target T_g. The method would be validated by applying it to a polymer library of 132 L-tyrosine derived homo-, co-, and terpolymers, of which the experimental T_g values have been experimentally determined.

3. Experimental methods

3.1. Polymer synthesis

The polymers were synthesized via solution polycondensation reactions according to previously published procedures or modifications thereof [45–50]. Polymers were repeatedly precipitated in a suitable non-solvent and vacuum dried to constant weight.

3.2. Glass transition temperature measurements

A large number of the experimental T_g values used in this work were obtained from a previously published polymer library [51]. The T_g values of newly synthesized polymers were measured using a TA Instrument DSC 2920 calibrated with an indium standard. Samples (5–15 mg) were equilibrated against ambient temperature and humidity prior to measurement. Measurements were made under a nitrogen atmosphere at 10 °C/min. Samples were heated in the first run to at least 50 °C above the T_g, annealed for 1 min, cooled to −55 °C and heated in a second scan to approximately 100 °C above T_g. The glass transition temperature was determined from the second heating scan by the half C_p extrapolated tangent method.

4. Analysis

4.1. Assignment of bond flexibility

We will use Schneider's approach [16,18] with several modifications and additions to assign bond flexibility to the polymers in our library. We note again that the parameter f is an index of bond flexibility, not a numerical count of the number of flexible bonds, and includes factors such as rotation number. A covalent bond is termed flexible if rotation around it causes a conformational change in the molecule [14]. We assumed that the following four bonds are not flexible ($f = 0$) due to the small size of the hydrogen atoms: a C–OH and a C–CH_3 bond. The amide bond is not flexible ($f = 0$) due to its partial double bond character. A *p*-phenyl ring is assigned to a flexibility of 1.5 because a rotation around the principle axis of this linear and almost rotationally symmetric unit does not change the overall molecular shape as dramatically as most other bond rotations. When studying a series of poly(methacrylate)s, Schneider invoked cooperative crankshaft like motions in aliphatic chains to reduce the effective number of flexible bonds in alkyl side chains longer than six carbon atoms and to account for what has been referred to by others as side chain crystallization. This justification was not germane to our work because the T_g prediction of our alkyl side chain containing polymers was either accurate enough without it or the data scatter was too large to decide on the effectiveness of this approach. Moreover, side chain crystallization has not been observed in any of our polymers. An illustration of the bond flexibility assignment using these rules is shown in Fig. 2 for poly(DTE carbonate) (E stands for ethyl). The flexibility of an individual bond is indicated by the number next to it (f'), leading to a number of flexible bonds of 12 ($f = 12$) for this repeat unit. With these basic rules in place it is now possible to objectively and consistently assign bond flexibility to any polymer repeat unit in- or outside of the current library.

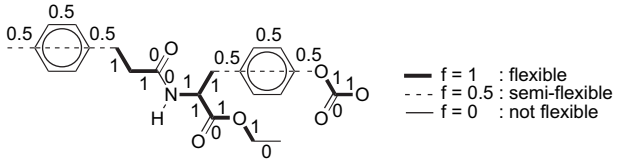


Fig. 2. An example of bond flexibility bond assignment: the repeat unit of poly (DTE carbonate) has 12 flexible bonds ($f = 12$) (the flexibility of each individual bond (f') is indicated by the number next to it).

4.2. Calculation of the mass-per-flexible-bond of polymers

The mass-per-flexible-bond (M/f) of a homo-, co, or terpolymer is calculated as the weighed average of the M/fs of each constituent monomeric unit, where the polymer composition is expressed as the mass fraction of its comonomers. This is shown in Eq. (2):

$$\left(\frac{M}{f}\right)_p = x_1\left(\frac{M_1}{f_1}\right) + x_2\left(\frac{M_2}{f_2}\right) + x_3\left(\frac{M_3}{f_3}\right) + \ldots + x_w\left(\frac{M_w}{f_w}\right) = \sum_{i=1}^{w} x_i\left(\frac{M_i}{f_i}\right) \quad (2)$$

where M_i is the molecular weight of the polymer repeat unit i (in g/mol), f_i is the number of flexible bonds in this repeat unit i, x_i is the mass fraction of the repeat unit, and w is the total number of comonomers in polymer. The conversion of the polymer composition from mole to mass fraction is given by Eq. (3):

$$x_i = \frac{m_i M_i}{\sum_{i=1}^{w}(m_i M_i)} \quad (3)$$

where x_i is the mass fraction of the repeat unit i, m_i is the mole fraction of the repeat unit i, M_i is the molecular weight of the repeat unit i (in g/mol), and w is the total number of comonomers in polymer. This averaging scheme was shown to be sufficiently accurate by Di Marzio [14]. Because the effect of chain ends is not taken into account, the M/f approach is strictly valid only at infinite molecular weights. This does not diminish the value of this approach since T_g is expected to reach a plateau at large molecular weights [52,53]. Because the polyarylates (see Fig. 1) are strictly alternating copolymers between an XTR diol and a diacid, they are treated as homopolymers in this approach: the XTR and ester unit are counted as one single repeat unit. The structures of all other co- and terpolymers are random and composition dependent, and as such Eq. (2) applies to them.

5. Results and discussion

5.1. Flexibility assignment and M/f calculation

Tables 1a and b show the molecular weight, number of flexible bonds, *M/f* values and the experimental T_gs of several selected repeat units and polymers (the structures of polymers 1–5 are shown in Fig. 3). It can be seen from Table 1a that the DTE and CTE carbonate repeat units have almost identical molecular weights (383.4 and 381.4 g/mol, respectively) while their mass-per-flexible-bond values (31.95 and 40.15, respectively) differ by 26%. This is because the carbon–carbon double bond present in poly(CTE carbonate) reduces the number of flexible bonds (f) in this repeat unit from 12 to 11 compared to poly(DTE carbonate). Electronic conjugation between the phenyl ring, the carbon–carbon double bond and the amide group stiffens the entire subunit and reduces the flexibility even more. Based on the best fit of the *M/f* value of this polymer with the rest of the library, the number of flexible bonds in CTE carbonate was found to be $f = 9.5$. This is the minimum possible value for f in this repeat unit, which signifies the occurrence of total electronic conjugation.

There is a large difference between *M/f* values of the DTD dodecanedioate and I_2DTE carbonate repeat units: 20.97 vs. 60.49, or 288%. This is of course because the mass of DTD dodecanedioate is made up by flexible, aliphatic groups in side and main chains ($f = 33$), while the mass of I_2DTE is made up by the two heavy iodine atoms that additionally hinder the rotation vis-à-vis the carbonate bond through their size, reducing f from 12 to 10.5 compared to DTE carbonate. The value of 10.5 was established after comparing various polycarbonates having different configurations of iodine substituted phenyl rings, a discussion that is outside of the scope of the current publication. The highest *M/f* of all monomers used in this study is found in I_2DT carbonate (71.43), due to the effect of the 3,5-diiodo substitution just discussed, and due to the lack of flexible side chains (it has a carboxylic acid group) or flexible main chain units.

The PEG_{2000} carbonate repeat unit has the lowest *M/f* of all monomers used in this study (14.90), and this is a result of the low molecular weight of its backbone and the lack of any bulky substituents that make it maximally flexible. Table 1b shows that the *M/f* values of the homopolymers 1, 2, and 3 are the same as their repeat unit, while the *M/f* values of polymers 4 and 5 lie in-between that of their constituent monomers, as dictated by Eq. (2). The effect of iodination is evident when comparing polymers 4 and 5: the *M/f* value increases from 30.49 to 58.52, and the experimental T_g increases from 344 to 384 K.

Table 1a
Calculated and experimental properties of selected repeat units (for a description of the abbreviations see Fig. 1)

Repeat unit	M[a] (g/mol)	f[b]	M/f[c]
DTE carbonate	383.4	12	31.95
DT carbonate	355.4	10	35.54
CTE carbonate	381.4	9.5	40.15
DTD dodecanedioate	692.0	33	20.97
I_2DTE carbonate	635.2	10.5	60.49
I_2DT carbonate	607.1	8.5	71.43
PEG_{1000} carbonate	1073.2	71[d]	15.12
PEG_{2000} carbonate	2086.3	140[e]	14.90

[a] The molecular weight of the repeat unit.
[b] The number of flexible bonds in a repeat unit.
[c] The mass-per-flexible-bond of the repeat unit.
[d] $3 \times 23 = 69$ flexible bonds in PEG_{1000}, add 2 for carbonate = 71.
[e] $3 \times 46 = 138$ flexible bonds in PEG_{2000}, add 2 for carbonate = 140.

Table 1b
Calculated and experimental properties of polymers from the library discussed in this work (for a description of the abbreviations see Fig. 1)

Polymer composition (mol%)	Polymer composition[a] (mass%)	$(M/f)_P$[b]	T_g[c] (K)
Poly(DTE carbonate)	100	31.95	369
Poly(CTE carbonate)	100	40.15	426
Poly(DTD dodecanedioate)	100	20.97	283
Poly(DTE-*co*-10%DT-*co*-4%PEG_{1000} carbonate)	80.8/8.7/10.5	30.49[d]	344
Poly(I_2DTE-*co*-10%I_2DT-*co*-4%PEG_{1000} carbonate)	84.1/9.3/6.6	58.52[e]	384

[a] Calculated from Eq. (3).
[b] The average mass-per-flexible-bond of the polymer (from Eq. (2)).
[c] The experimental T_g.

5.2. Glass transition temperatures vs. mass-per-flexible-bond plot

On the basis of the arguments presented above, we plotted the experimental T_g values against the computed *M/f* values of all 132 homo-, co- and terpolymers to verify the general applicability of the method (Fig. 4). Some of the poly(DTE-*co*-DT-*co*-PEG carbonate)s are known to exhibit nanoscale phase separation [54], and this manifests itself through the occurrence of a very small 2nd T_g above or below the main T_g. In these instances, the main T_g value was used. This is justified by the absence of any abnormal errors in T_g prediction or otherwise unusual behavior in this subset of polymers. A linear relationship between *M/f* and T_g confirms the results of Schneider [15] and proves the validity of our approach as a basis for a T_g prediction method.

It can be seen in Fig. 4 that the library can naturally be divided into three main classes (I, II, and III). This polymer class specific behavior has been observed and explained previously [16] and is a result of the different densities, steric hindrances and intermolecular interactions in different polymer classes. In this specific library, the three classes are the following: (I) the XTR carbonate homopolymers and all the poly(DTR arylate)s; (II) all poly(DTE-*co*-*y*%DT-*co*-*z*%PEG_M carbonate)s; and (III) all poly(I_2DTE-*co*-*y*%I_2DT-*co*-*z*%PEG_M carbonate)s, the linear fit coefficients of which are listed in Table 2.

The division of our library into three classes can be rationalized as follows: (1) PEG and iodine containing co- and terpolymers are estimated to have a lower and a higher density, respectively, than pure XTR polymers that only contain

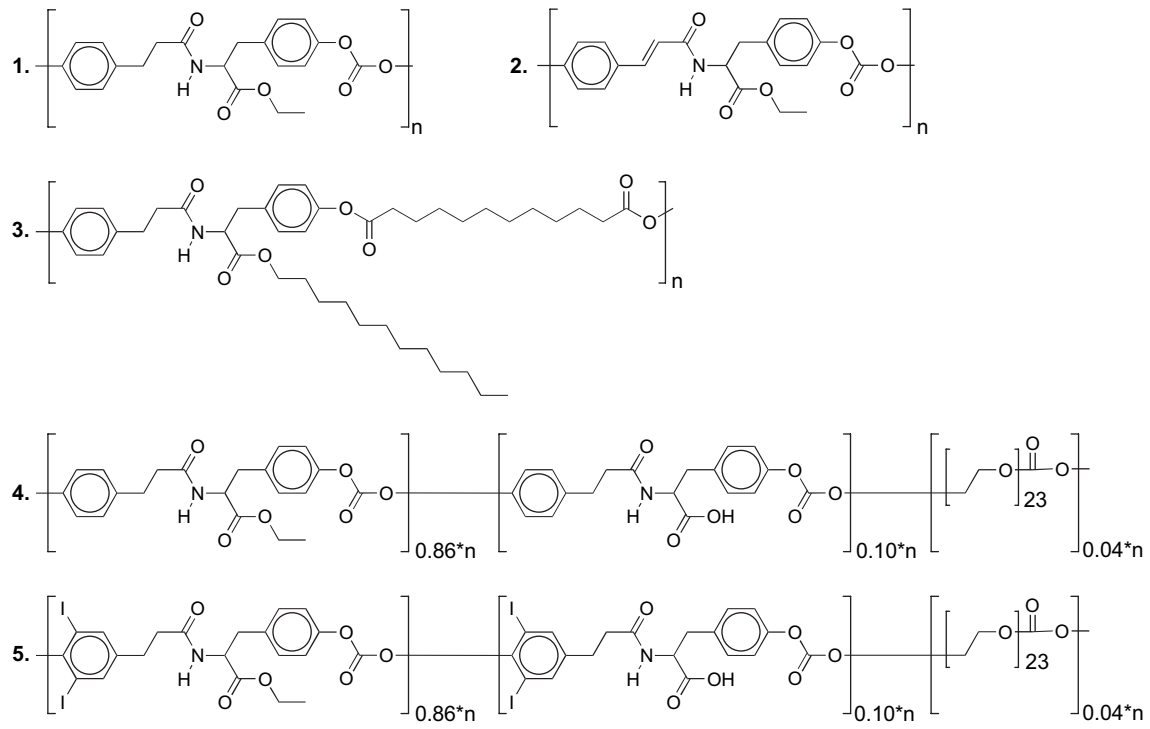


Fig. 3. Five selected polymer structures from the library discussed in this work. (1) Poly(DTE carbonate); (2) poly(CTE carbonate); (3) poly(DTD dodecanedioate); (4) poly(DTE-*co*-10%DT-*co*-4%PEG_{1000} carbonate); (5) poly(I_2DTE-*co*-10%I_2DT-*co*-4%PEG_{1000} carbonate). The numbers refer to the arrows in Fig. 4.

aliphatic and aromatic groups, and the mass-per-flexible-bond approach does not account for this density difference resulting in a shift of the line towards high *M/f* values, as well as a change in slope; (2) iodine increases the molecular weight of monomers through its high atomic mass, while simultaneously decreasing the number of flexible bonds through steric hindrance, thereby increasing its *M/f*, again shifting the line towards high *M/f* values; and most importantly, (3) the interchain hydrogen bonding in the three classes is of a different magnitude.

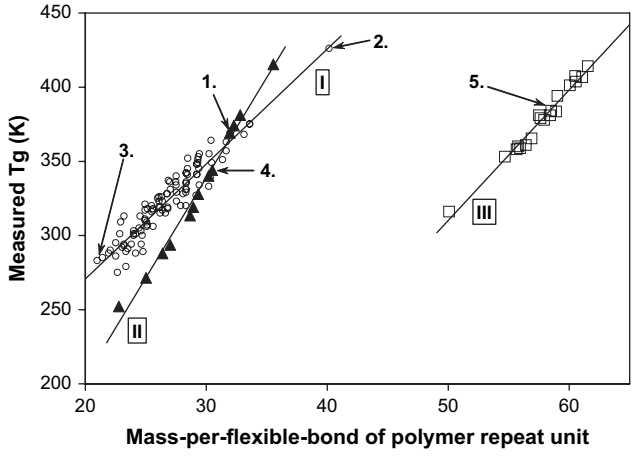


Fig. 4. The experimental glass transition temperature plotted against the calculated mass-per-flexible-bond for all 132 polymers from the library discussed in this work, showing class specific behavior. (I) Polycarbonates and polyarylates (open circles); (II) poly(DTE-*co*-DT-*co*-PEG carbonate)s (filled triangles); (III) poly(I_2DTE-*co*-I_2DT-*co*-PEG carbonate)s (open squares). The bold lines are linear fits (see Table 2 for fit coefficients). The numbers 1–5 refer to the polymer structures in Fig. 3 and Table 1.

The hydrogen bonding interactions between the mobile PEG ether groups and the DT carboxylic acid groups in class II (the poly(DTE-*co*-DT-*co*-PEG carbonate)s) are expected to be strong, while in class III (the same polymers but iodinated) these interactions will most probably be much weaker due to steric hindrance and the hydrophobic nature of the bulky iodine substituents. The total absence of both PEG and free carboxylic acid groups in the polymers of class I decreases the hydrogen bonding interactions in these polymers to even lower levels. This explanation is supported by the changes in the slopes of the fitted lines among the three classes (Table 2): I < III <<II (7.7505, 8.8031 and 13.512, respectively), which correspond to an increase in intra- and interchain interactions [18]. Furthermore, it can be concluded that the hydrophobic iodine substituents reduce the effect of the hydrophilic PEG blocks, as seen by the lower slope of class III that is closer to that of class I than II although a density effect cannot be excluded.

From Fig. 4 and Table 2 it can be seen that the data scatter in class I is considerably larger than that in classes II and III, as reflected by order of their increasing correlation coefficients: R^2(I)<< R^2(III) < R^2(II) (0.9050, 0.9989, and 0.9775, respectively). There are a number of explanations for this data scatter: (1) the materials in classes II and III have been synthesized relatively recently compared to those in class I, leading to less operator dependent variations present in that sub library, e.g. experimental T_g values; (2) because the polyarylates are essentially alternating copolymers, they are

Table 2
Linear fit parameters obtained from Eq. (1) for the three polymer classes shown in Fig. 5

Polymer class	Polymers	Linear fit parameters for Eq. (1)		
		A	C	R^2
I	All polycarbonates and polyarylates	7.7505	116.17	0.9050
II	Poly(DTE-*co*-DT-*co*-PEG carbonate)s	13.512	−66.374	0.9889
III	Poly(I_2DTE-*co*-I_2DT-*co*-PEG carbonate)s8.8031		−130.02	0.9775

extremely sensitive to stoichiometry and their molecular weights are therefore difficult to control, which might manifest itself in T_g variations; and (3) some of the polyarylates have been shown to display liquid crystalline like behavior [18,55–57], which will influence the T_g values through intermolecular hydrogen bonding, and these interactions are not included in the mass-per-flexible-bond approach.

5.3. T_g prediction and evaluation of predictive accuracy

Despite the scatter in the available data in this polymer library, the correlations of T_g with *M/f* for these homo, co- and terpolymers are very good, suggesting that the proposed method can be used as a T_g prediction tool. A linear fit of the plot of the polymer's T_g (in K) vs. *M/f* value can be used to obtain the constants A and C of Eq. (1). These values can then be used for the prediction of the T_g of a polymer by calculating its *M/f* value. However, in the present case all the T_g values in the library had already been measured and were available at the moment of the model construction. We therefore chose an a posteriori approach to evaluate the predictive accuracy of the proposed method.

An algorithm was written in Matlab (MathWorks, Natick, MA; version 6.1.0.450, release 12.1) that selected, from a data set of n polymers, a construction set of k polymers, leaving $n-k$ polymers as the test set. The data consisted of the *M/f* and T_g (in K) values of the polymers as the X and Y values, respectively. A least squares linear fit was performed on the T_g vs. *M/f* plot of the construction set polymers, giving the fit coefficients A and C for this set. These coefficients were then used to calculate the glass transition temperatures of each polymer in the test set via Eq. (1). For each polymer in this test set we calculated ΔT, the absolute difference between the predicted and the experimentally measured T_g values. These individual ΔT values were averaged over the whole test set to give ΔT^*, the average, absolute error in T_g prediction for that test set. Note that this value ΔT^* results from only one of many possible combinations of choosing k out of n polymers to form this specific construction and test set. Thus, for reasons of completeness and objectivity we therefore generated all possible combinations of k out of n polymers that can make up that test set size, and averaged all the ΔT^* values of those combinations to give ΔT_{av}, the average error in T_g prediction for that specific test set (of size $n-k$). In order to investigate how many T_g values need to be known to accurately predict new T_g values, the size of the construction set was now varied from $k=3$ to $k=n-1$ with a step size of 1, leading to a progressively smaller test set in each case. The ΔT_{av} values were calculated for all construction set sizes k, and the trend of ΔT_{av} vs. k was investigated to establish the minimum construction set size (k) necessary to accurately predict new T_g values.

5.4. Accuracy of the prediction

In order to objectively evaluate the predictive accuracy of this method, an a posteriori data selection and analysis procedure, described in Section 3, was employed. Fig. 5a–d shows a selection of four representative plots of the predicted vs. measured glass transition temperature in the class of poly(DTE-*co*-*y*%DT-*co*-*z*%PEG_M carbonate)s (13 data points) for four different sizes of the construction set k: 3, 4, 5, and 7. It should be noted that each plot represents *one* random selection from all possible combinations P of choosing k out of n values for that construction set of size k. Increasing the construction set size obviously reduces the number of predicted values in the test set from 10 to 6. It is clear from these plots that increasing k from 3 to 7 gives more accurate prediction results for the test set, as seen from the decreasing data point deviation of the predicted values from the line $y=x$ that represents a perfect prediction. This fact is reflected in the average, absolute prediction error (ΔT_{av}) associated with each plot that decreased from 6.6 to 5.3 K upon increasing k from 3 to 7. For small polymer classes ($n \leq 15$) all possible combinations P were selected and used to calculate the corresponding ΔT_{av} values. However, for large classes ($n > 15$) P becomes very large; the number of possible unordered combinations P to choose k polymers from a set of n values is given by $P = n!/((k!(n-k)!)$, written as ${}^nC_k$. For example, using $n=100$ and taking half of the values for the construction set (i.e. ${}^{100}C_{50}$) results in $P \sim 10^{29}$ different combinations of k out of n polymers. A computation of this magnitude is beyond the memory capacity of a regular personal computer, and in these cases we therefore reduced the number of data points by limiting the number of construction sets that were investigated. For large data sets the size of the construction set was therefore varied from $k=10$ to $k=n-10$ with a step size of 10. Additionally, instead of using all possible combinations P per construction set, we randomly selected only 2000 combinations from which ΔT_{av} was calculated.

For reasons of brevity, we present the results of these calculations in Fig. 6 as a plot of the average, absolute prediction error (ΔT_{av}) vs. k for all three polymer classes. Note that k is expressed as the % of the available data points to better compare the behavior of the three data sets. This plot allows for the determination of the best attainable T_g prediction error in each class from the asymptotic value of ΔT_{av} with increasing k, as well as the minimum size of the construction set necessary for accurate prediction (k_{min}). It can be seen that while the value of ΔT_{av} for polymer classes I and III levels off around 30 and 50%, respectively, that of class II seems to decreases

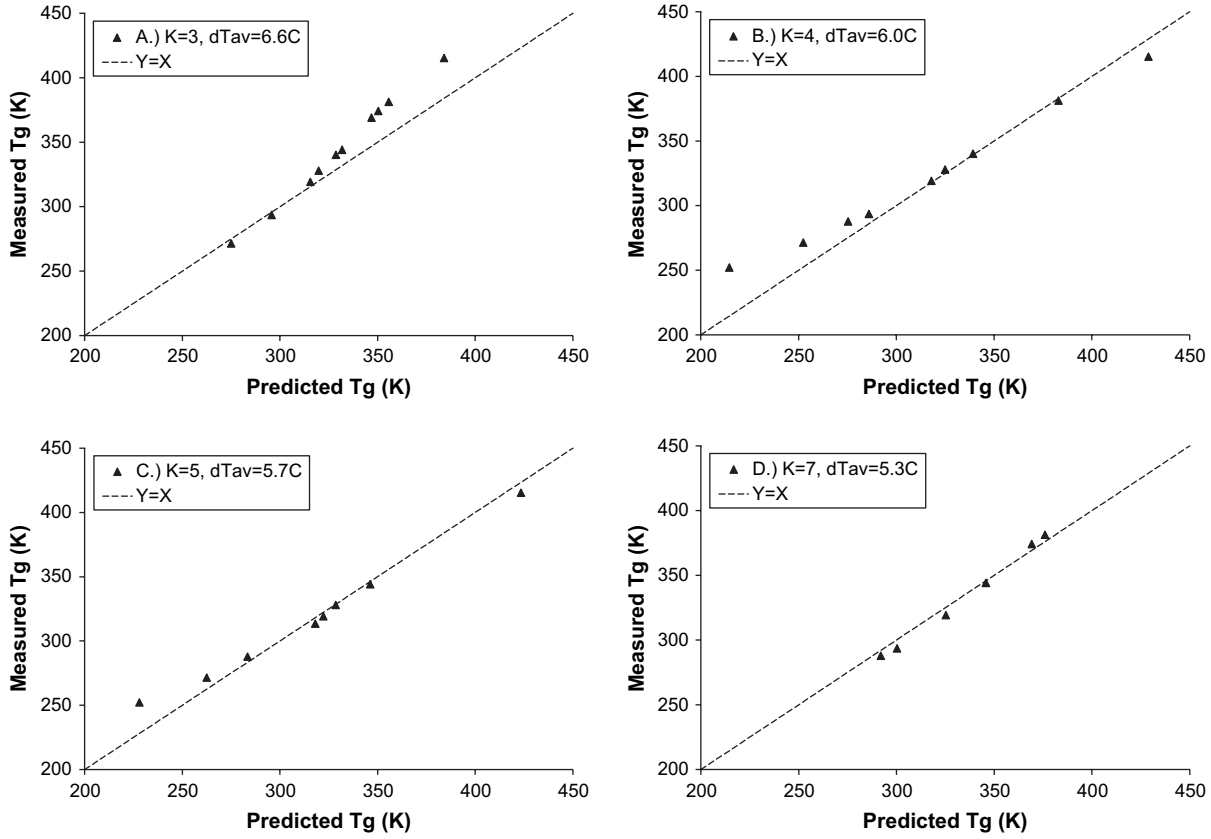


Fig. 5. The experimental T_g vs. predicted T_g for class I, the poly(DTE-*co*-DT-*co*-PEG carbonate)s ($n = 13$) for four different sizes of the construction set k. (a) $k = 3$; (b) $k = 4$; (c) $k = 5$; (d) $k = 7$. The line $Y = X$ indicates a perfect prediction. Note that each plot is one randomly selected combination out of many for choosing k out of 13 values.

continuously with k. This behavior is suspected to be due to the large data scatter in this polymer class. Table 3 lists the values of ΔT_{av} at k_{min}. It can be seen that the average error in T_g prediction is 6.4, 5.3, and 3.7 K for classes I, II, and III, respectively. These values are remarkably low considering the simplicity of the model, the large data scatter in the T_g vs. M/f plots, and the structural variety in the library. In fact, this accuracy is better than most of the results from (semi-)empirical T_g prediction methods found in the literature [30–43,50].

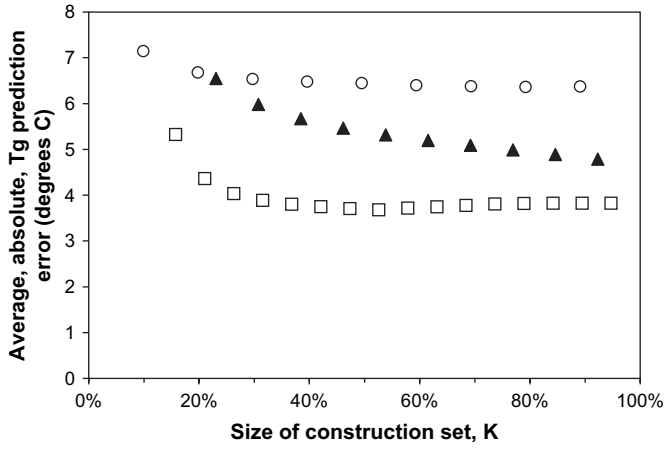


Fig. 6. The evolution of the average, absolute error (ΔT_{av}) in the predicted T_g values as a function of construction set size k for the three polymer classes discussed in this work. (I) Polycarbonates and polyarylates (open circles); (II) poly(DTE-*co*-DT-*co*-PEG carbonate)s (filled triangles); (III) Poly(I_2DTE-*co*-I_2DT-*co*-PEG carbonates) (open squares).

5.5. Structure calculation from target T_g

Eq. (1) can also be used in the reverse manner from the above approach, namely polymer structure prediction for a target T_g value when the fit coefficients A and C are known. Knowledge of the fit coefficients A and C, and the choice of a target T_g now lead to the identification of the polymer with an M/f value that corresponds to one desired composition or a range of polymer compositions in the case of terpolymers, allowing the best suitable material to be chosen. In the case of terpolymers, combination of Eqs. (1) and (2) and the fact that the sum of mass fractions equals unity results in Eqs. (4a) and (4b):

$$x_{M_3} = \left(\frac{\frac{M}{f_{M_1}}}{\frac{M}{f_{M_1}} - \frac{M}{f_{M_3}}}\right) + x_{M_2}\left(\frac{\frac{M}{f_{M_2}} - \frac{M}{f_{M_1}}}{\frac{M}{f_{M_1}} - \frac{M}{f_{M_3}}}\right) - \left(\frac{T_g - C}{A\left(\frac{M}{f_{M_1}} - \frac{M}{f_{M_3}}\right)}\right) \quad (4a)$$

Table 3
The average, absolute error in the T_g values predicted by Eq. (1) for the three polymer classes found in Fig. 1 at $k = 50\%$

Polymer class	Polymers	Set size	k_{min} [a]	ΔT_{av}: average T_g prediction error (K)
I	All polycarbonates and polyarylates	100	7[b]	6.4
II	Poly(DTE-*co*-DT-*co*-PEG carbonate)s	13	7	5.3
III	Poly(I_2DTE-*co*-I_2DT-*co*-PEG carbonate)s	19	7	3.7

[a] The minimum size of the construction set (in number of *X*,*Y* pairs) necessary to obtain the best T_g prediction (lowest ΔT_{av}) in that class.
[b] No asymptotic behavior, arbitrarily chosen value of *k*.

$$x_{M_2} = x_{M_3}\left(\frac{\frac{M}{f_{M_1}}-\frac{M}{f_{M_3}}}{\frac{M}{f_{M1}}-\frac{M}{f_{M_2}}}\right)+\left(\frac{\frac{M}{f_{M_1}}}{\frac{M}{f_{M_1}}-\frac{M}{f_{M_2}}}\right)-\left(\frac{T_g-C}{A\left(\frac{M}{f_{M_2}}-\frac{M}{f_{M_1}}\right)}\right) \quad (4b)$$

where x_{Mi} is the mass fraction of the polymer repeat unit i; *M* is the molecular weight of the repeat unit (in g/mol); *f* is the number of flexible bonds in the repeat unit; $(M/f)_i$ is the mass-per-flexible-bond of the repeat unit; T_g is the targeted glass transition temperature of the polymer (in K); and *A* and *C* are polymer class specific constants.

This reverse approach will be illustrated with two specific examples: the poly(DTE-*co*-DT-*co*-PEG_{1000} carbonate)s and the poly(I_2DTE-*co*-I_2DT-*co*-PEG_{1000} carbonate)s. First we adapt Eqs. (4a) and (4b) by replacing the subscripts 1, 2 and 3 by DTE, DT and PEG_M, respectively. Next, we substitute appropriate *M*/*f* values of the monomers from Table 1 and the fit coefficients A and C from Table 2 for these two polymer classes. We then get Eqs. (5a) and (5b) for the class II polymers, poly(DTE-*co*-DT-*co*-PEG_{1000} carbonate)s, and Eqs. (6a) and (6b) for the class III polymers, poly(I_2DTE-*co*- I_2DT-*co*-PEG_{1000} carbonate)s,

$$x_{PEG_{1000}} = 0.21294x_{DT} + 1.8978 - \left(\frac{T_g + 66.374}{227.48}\right) \quad (5a)$$

$$x_{DTE} = 1 - x_{DT} - x_{PEG_{1000}} \quad (5b)$$

$$x_{PEG_{1000}} = 0.24113x_{I_2DT} + 1.3333 - \left(\frac{T_g + 130.02}{399.40}\right) \quad (6a)$$

$$x_{I_2DTE} = 1 - x_{I_2DT} - x_{PEG_{1000}} \quad (6b)$$

where x_{I_xDTE} is the mass fraction of the I_xDTE monomer in the terpolymer, x_{I_xDT} the mass fraction of the I_xDT monomer, $xPEG_{1000}$ the mass fraction of PEG_{1000}, and T_g is the main glass transition temperature of the terpolymer (K). Eqs. (5)–(6) were used to calculate the different terpolymer compositions that correspond to a target T_g. A number of polymer compositions for a target T_g of 60 °C are listed in Table 4. In this specific example it is thus possible to smoothly change the polymer composition from 86/0/14 to 0/71/29 (in mass% DTE/DT/PEG_{1000}) for poly(DTE-*co*-DT-*co*-PEG_{1000} carbonate)s and from 83/0/17 to 0/67/33 (in mass% I_2DTE/I_2DT/PEG_{1000}) for the poly(I_2DTE-*co*-I_2DT-*co*-PEG_{1000} carbonate)s without affecting the T_g of 60 °C. The fact that the stiffer monomer, I_2DTE, increases the T_g of the terpolymer is illustrated by the fact that this class can have a higher PEG_{1000} content than the corresponding non-iodinated polymer while maintaining its T_g of 60 °C. This makes it possible to tune the degradation rate, water uptake, protein adsorption and the cell–material interactions independent of the glass transition temperature by proper choice of polymer type and composition within these two polymer classes.

Table 4
Selected terpolymer compositions (mass%) of poly(I_xDTE-*co*-I_xDT-*co*-PEG_{1000} carbonate)s that have a predicted T_g of 60 °C with $T_g = 60$ °C (mass%)

DTE	DT	PEG_{1000}	I_2DTE	I_2DT	PEG_{1000}
85.85	0	14.15	82.63	0	17.37
73.72	10	16.28	70.22	10	19.78
61.59	20	18.41	57.81	20	22.19
49.46	30	20.54	45.41	30	24.59
37.33	40	22.67	33.00	40	27.00
25.20	50	24.80	20.59	50	29.41
13.07	60	26.93	8.18	60	31.82
0	70.78	29.22	0	66.59	33.41

It should be noted that the predicted values are dry T_gs, plasticization by water of the hydrophilic PEG blocks is not taken into account, and it is likely that the extreme polymer compositions are beyond the predictive capacity of the current method. However, it is evident that this method has broad applications in polymer and (bio)material science in general, as it allows for the decoupling of the thermal (and perhaps mechanical) behavior of a terpolymer from its chemistry and surface energy.

6. Conclusions

The mass-per-flexible-bond principle was used for the first time to devise a semi-empirical T_g prediction method that is applicable to polymers with any number of comonomers. A self-consistent set of rules for the assignment of bond flexibility, applicable to any polymer structure, were given, as well as formulae for the calculation of the *M*/*f* values of co- and terpolymers. The method was validated by applying it to a 132-member L-tyrosine derived polymer library containing a wide range of polymer structures. A plot of the experimental T_g vs. calculated *M*/*f* values yielded straight lines when the polymers were divided into different classes. This division was explained in terms of the different densities, steric hindrances and intermolecular interactions in each polymer class. An a posteriori predictive accuracy evaluation showed that average, absolute error between predicted and measured T_g ranged from 6.4 to 3.7 K for the three polymer classes investigated in this work. This accuracy is very good considering the data scatter and the large variety of functional groups present in this library, and it is better than most *ab initio*, molecular mechanics, or (semi-)empirical prediction methods found in

the literature. Furthermore, the current method has the advantage of being considerably simpler to use than many of the existing methods. We showed that as little as seven T_g values can be used to accurately predict the T_gs of other polymers in that class. Additionally, it was shown that the proposed method can be used in the reverse manner: to calculate polymer structures that match a certain target T_g value based on the linear fit coefficients of a polymer class. This enables keeping the thermal (and perhaps the mechanical) behavior of a terpolymer constant while independently choosing its chemistry and surface energy.

In summary, we provided a simple, accurate, and universal method for both T_g prediction and polymer structure calculation from a few initial experimental T_g values. This method is likely to have broad applications in polymer and (bio)material science.

Acknowledgments

We thank Dr. Anna Gubskaya and Dr. Paul Holmes for their helpful discussions and Mr. Ross Zimnisky for editing the manuscript. Funding for this work was provided by (i) RESBIO – The National Resource for Polymeric Biomaterials funded by the National Institutes of Health (NIH grant EB001046), (ii) by a collaboration with Professor Roth, funded by NIH grant GM065913, and (iii) by support from REVA Medical, Inc., San Diego, CA.

References

[1] Hoogenboom R. Expanding the polymer toolbox: high-throughput experimentation, microwave irradiation, and grid-like metal complexes. Eindhoven: Technical University of Eindhoven; 2004. p. 216.
[2] Meier MAR, Schubert US. J Mater Chem 2004;14:3289–99.
[3] Weber N, Bolikal D, Bourke SL, Kohn J. J Biomed Mater Res 2004;68(A):496–503.
[4] Kholodovych V, Smith JR, Knight D, Abramson S, Kohn J, Welsh WJ. Polymer 2004;45:7367–79.
[5] Smith JR, Kholodovych V, Knight D, Kohn J, Welsh WJ. Polymer 2005;46(12):4296–306.
[6] Smith JR, Kholodovych V, Knight D, Welsh WJ, Kohn J. QSAR Comb Sci 2005;24:99–113.
[7] Smith JR, Knight D, Kohn J, Rasheed K, Weber N, Kholodovych V, et al. J Chem Inf Comput Sci 2004;44:1088–97.
[8] Smith JR, Seyda A, Weber N, Knight D, Abramson S, Kohn J. Macromol Rapid Commun 2004;25:127–40.
[9] Sharma RI, Kohn J, Moghe PV. J Biomed Mater Res 2004;69A:114–23.
[10] Zeltinger J, Schmid E, Brandom D, Bolikal D, Pesnell A, Kohn J. Biomater Forum 2004;26(1):8–24.
[11] Tangpasuthadol V, Pendharkar SM, Kohn J. Biomaterials 2000;21:2371–8.
[12] Tangpasuthadol V, Pendharkar SM, Peterson RC, Kohn J. Biomaterials 2000;21:2379–87.
[13] Painter PC, Coleman MM. Fundamentals of polymer science. Basel: Technomic Publications AG; 1994.
[14] Di Marzio EA. Polymer 1990;31:2294–8.
[15] Schneider HA. J Res Natl Inst Stand Technol 1997;102(2):229–48.
[16] Schneider HA. Polymer 2005;46:2230–7.
[17] Schneider HA, Di Marzio EA. Polymer 1992;33(16):3453–61.
[18] Schneider HA. J Appl Polym Sci 2003;88:1590–9.
[19] Mark J, Ngai K, Graessley W, Mandelkern L, Samulski E, Koenig J, et al. Physical properties of polymers. Basel: Technomic Publications AG; 2004.
[20] Tangpasuthadol V, Shefer A, Yu C, Zhou J, Kohn J. J Appl Polym Sci 1997;63(11):1441–8.
[21] Katajisto J, Linnolahti M, Pakkanen TA. Theor Chem Acc 2005;113:281–6.
[22] Ton TMT, Rode BM. Macromol Theory Simul 1996;5(3):467–75.
[23] Paul W. AIP Conf Proc 1992;256:145–54.
[24] Wittkop M, Hoelzl T, Kreitmeier S, Goeritz D. J Non-Cryst Solids 1996;201:199–210.
[25] Boyd RH. Trends Polym Sci 1996;4(1):12–7.
[26] Hamerton I, Heald CR, Howlin BJ. J Mater Chem 1996;6(3):311–4.
[27] Hamerton I, Howlin BJ, Klewpatinond P, Shortley HJ, Takeda S. Polymer 2006;47:690–8.
[28] Zhang M, Sundararaj U, Choi P. Polym Mat Sci Eng Prepr 2004;91:1031–2.
[29] Abu-Sharkh BF. Comput Theor Polym Sci 2001;11:29–34.
[30] Becker R, Neumann G. Plast Kautsch 1973;20(11):809–15.
[31] Brown WM, Martin S, Rintoul MD, Faulon J-L. J Chem Inf Model 2006;46:826–35.
[32] Camacho-Zuniga C, Ruiz-Trevino FA. Ind Eng Chem Res 2003;42:1530–4.
[33] Carro AM, Campisi B, Camelio P, Phan-Tan-Luu R. Chemom Intell Lab Syst 2002;62:79–88.
[34] Cypcar CC, Camelio P, Lazzeri V, Mathias LJ, Waegell B. Macromolecules 1996;29:8954–9.
[35] Garcia-Domenech R, de Julian-Ortiz JV. J Phys Chem B 2002;106:1501–7.
[36] Hopfinger AJ, Koehler MG, Pearlstein RA. J Polym Sci Part B Polym Phys 1988;26:2007–28.
[37] Joyce SJ, Osguthorpe DJ, Padgett JA, Price GJ. J Chem Soc Faraday Trans 1995;91(16):2491–6.
[38] Katritzky AR, Rachwal P, Wah Law K, Karelson M, Lobanov VS. J Chem Inf Comput Sci 1996;36(4):879–84.
[39] Kreibich UT, Batzer H. Angew Makromol Chem 1979;83:57–112.
[40] Mattioni BE, Jurs PC. J Chem Inf Comput Sci 2002;42:232–40.
[41] Van Krevelen DW. Properties of polymers. Amsterdam: Elsevier; 1976.
[42] Weyland HG, Hoftijzer PJ, Van Krevelen DW. Polymer 1970;11(2):79–87.
[43] Zhang L, Zhao D, Huang Y. Chin J Polym Sci 2002;20(2):25–30.
[44] Yu X, Wang X, Li X, Gao J, Wang H. Macromol Theory Simul 2006;15:94–9.
[45] d'Acunzo F, Kohn J. Macromolecules 2002;35:9360–5.
[46] Ertel SI, Kohn J. J Biomed Mater Res 1994;28:919–30.
[47] Fiordeliso J, Bron S, Kohn J. J Biomater Sci Polym Ed 1994;5(6):497–510.
[48] Kohn J. Trends Polym Sci 1993;1(7):206–12.
[49] Pulapura S, Kohn J. Biopolymers 1992;32:411–7.
[50] Yu C, Kohn J. Biomaterials 1999;20(3):253–64.
[51] Brocchini S, James K, Tangpasuthadol V, Kohn J. J Am Chem Soc 1997;119(19):4553–4.
[52] Fox TG, Flory PJ. J Appl Phys 1950;21:581.
[53] Fox TG, Flory PJ. J Polym Sci 1954;14:315.
[54] Sousa A, Schut JA, Kohn J, Libera M. Macromolecules 2006;39(21):7306–12.
[55] Jaffe M, Ophir Z, Collins G, Recber A, Yoo S, Rafalko J. Polymer 2003;44:6033–42.
[56] Jaffe M, Ophir Z, Pai V. Thermochim Acta 2003;396:141–52.
[57] Jaffe M, Pai V, Ophir Z, Wu J, Kohn J. Polym Adv Technol 2002;13:926–37.