

Molecular modelling of polymers:

5. Inclusion of intermolecular energetics in estimating glass and crystal-melt transition temperatures

M. G. Koehler* and A. J. Hopfinger†

*Departments of Chemistry and Medicinal Chemistry and Pharmacognosy,
University of Illinois at Chicago, PO Box 6998, Chicago, IL 60680, USA
(Received 31 March 1988; revised 13 June 1988; accepted 18 June 1988)*

The method of constructing a quantitative structure–property relationship to estimate the glass transition temperature (T_g) from conformational entropy and mass moments has been extended to take into account intermolecular interactions. The intermolecular energies are estimated by driving probe units about torsion angle units. The average interaction energy between a probe and the torsion angle unit is taken as a molecular descriptor used in conjunction with entropy and mass moment to construct the quantitative structure–property relationships (QSPR). Three types of probe units have been considered, CH_3 , O^- , and H^+ . QSPRs have been constructed for T_g using 35 polymers, and for the melt transition temperature (T_m) using 30 polymers. The QSPRs are formulated by performing multidimensional linear regression analysis between observed T_g s (T_{m} s) and the calculated molecular descriptors. A very significant QSPR involving backbone and side chain entropies, backbone mass moments, and the intermolecular energies of the O^- and H^+ probes could be constructed for the glass transition data. It was not possible to formulate a significant QSPR for T_m . This may be due to the spatial anisotropic nature of a polymer crystal *versus* a polymer glass.

(Keywords: glass transition temperature; intermolecular interactions; intermolecular energy)

INTRODUCTION

In a recent paper¹ we were able to develop quantitative structure–property relationships, QSPRs, which relate the glass transition temperature, T_g , to the intramolecular flexibility of the polymer chain. Intramolecular flexibility, in turn, is quantitatively represented by a linear combination of backbone and side chain contributions, respectively, to the intramolecular conformational entropy and mass moment of the chemical repeat unit, normally the monomer. Conformational entropy is computed from the Boltzmann distribution of conformational energy states arising from torsion rotations about backbone and sidechain bonds. The mass moments are currently taken to be the same as the backbone and individual sidechain masses associated with a monomer (chemical repeat) unit. The use of the principal moments of inertia of backbone and sidechains is being considered as an alternative to masses.

The current T_g model does not consider contributions from intermolecular interactions to the glass transition process. The neglect of intermolecular interactions in estimating T_g has been justified by assuming that the polymer geometry of the glassy state is locally non-ordered, leading to a spatially orthotropic intermolecular energy field over the length of any individual polymer chain. However, in one of the QSPRs the T_g of poly(hexamethylene adipamine) is underestimated by

about 77K. We believe this is a direct consequence of not taking into account intermolecular hydrogen bonding that can occur in the glass for this particular polymer. Moreover, the overestimation of the T_g of poly(isobutylene) by 68K for the same QSPR may be due to neglecting poor local packing of chain segments which results in a loss of stabilizing interchain dispersion energy relative to other polymers. Hence, it would be useful to be able to estimate intermolecular interactions as might occur in the glassy state.

We are also interested in determining the extent to which the current formalism developed to estimate T_g can be extended to construct QSPRs for the polymer crystal to melt transition temperature, T_m . The observed relationships between T_g and T_m in many families of polymers² makes this extension of the formalism an obvious goal. However, it is quite clear from lattice packing energy calculations³ that the medium surrounding a polymer chain in a crystal is much less orthotropic with respect to intermolecular interactions, than the glassy matrix of a polymer.

Thus, the major methodology goal of this paper is the development of a simple molecular modelling scheme to estimate intermolecular interactions that will complement intramolecular entropy and mass moment in developing QSPRs for both T_g and T_m .

METHODS

The key component to being able to efficiently estimate intramolecular entropy and mass moment is the representation of the polymer in terms of torsion angle units. A

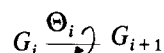
* Present address: Allied-Signal Inc., Engineering Materials Research Center, 50 East Algonquin Road, Box 5016, Des Plaines, IL 60017-5016, USA

† To whom correspondence should be addressed

0032-3861/89/010116-11\$03.00

© 1989 Butterworth & Co. (Publishers) Ltd.

torsion angle unit is schematically defined as:



where G_i and G_{i+1} are structural groups connected by a bond about which the torsion angle Θ_i occurs. The polymer is built up by connecting torsion angle units together such that the 'right' structural group of the i th torsion angle becomes the 'left' structural group of the $i+1$ torsion angle unit. Figure 1 illustrates the structure of a linear polymer in terms of torsion angle units.

Torsion angle units permit a global molecular property of a polymer system to be computed as a scalar sum of the individual molecular properties of the constituent torsion angle units. That is, we have formulated the estimation of global molecular properties in terms of a group additive property (GAP) model². The global molecular properties can be correlated to macroscopic properties of the system, in this case T_g and T_m , to hopefully yield a QSPR.

The major limitation of GAP models is that the requisite GAP parameters to make an estimation of a property are not always available. Thus, while the calculation of the global molecular property is straightforward, it cannot be done because of the missing data. However, the formulation of a GAP model using torsion angle units permits the application of molecular modelling methods to estimate the GAP parameters associated with any torsion angle unit. Thus, one only needs to know, or develop, a scheme to compute the requisite group additive properties using molecular modelling to insure being able to parameterize any torsion angle unit. Moreover, once the GAP properties are computed for a torsion angle unit, they can be stored in a database and used whenever the torsion angle unit turns up in a study. Overall, the marriage of molecular modelling and GAP approaches, and their joint application to define torsion angle unit properties, provides an open-ended, general approach to estimate macroscopic polymer properties.

It should be pointed out that the torsion angle unit used here is one-dimensional (one degree of conformational freedom) in torsion angle, Θ_i , when bond lengths and angles are held constant, which is the case in our formalism. This representation of the torsion angle unit

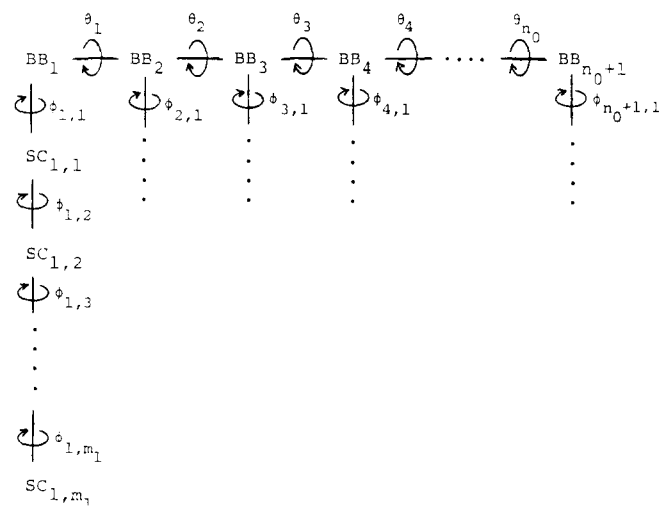
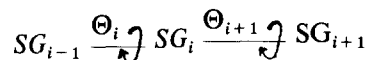


Figure 1 Schematic representation of a linear polymer monomer unit in terms of backbone, BB_j , and sidechain, $SC_{i,j}$, structure groups which, in turn, define the component torsion angle units of the monomer

is not a restriction. Higher dimensional torsion angle units can be defined if needed. For example,



defines a two-dimensional torsion angle unit. Obviously, the computational effort needed to compute the molecular properties increases with dimensionality. However, the coupling of structural units composing the polymer chain also increases which should enhance the estimation of molecular properties. The polymer chain can also be decomposed into arbitrary sequences of torsion angle units of different dimensions.

Computational procedures to determine intramolecular entropy and mass moments of torsion angle units have been formulated and carried out¹. Table 1 contains the intramolecular entropies and mass moments for many torsion angle units.

A molecular modelling scheme to estimate intermolecular energies for torsion angle units is developed as part of this paper. The estimation of intermolecular energetics follows from the estimation of potential energy fields about a molecule⁴, which is a molecular mechanics based generalization of the electrostatic potential field of a molecule⁵.

The geometry for computing intermolecular energetics of the one-dimensional torsion angle unit is shown in Figure 2. An intermolecular probe unit is placed a perpendicular distance d^* from the geometric centre of the bond, C_i , between the two structure groups, G_i and G_{i+1} . The angular location of the probe relative to C_i is equivalent to Θ_i , and is initially taken to be arbitrary. However, d^* always corresponds to the minimum distance from C_i such that no steric violations occur between the probe and the atoms of the torsion angle unit. Once this distance, d^* , is identified by systematically increasing the radial distance, d , from C_i at Δd increments, the energy between the probe and the torsion angle unit is minimized with respect to d starting at d^* . The probe and the torsion angle unit energetics is taken to be the sum of the non-bonded steric and electrostatic components of a molecular mechanics force field⁶ as given by equation (1).

$$E_p(\Theta_i, d) = \sum_{i=1}^N \left[\left(\frac{-A_{ip}}{r_{ip}^6} + \frac{B_{ip}}{r_{ip}^{12}} \right) + \left(\frac{Q_i Q_p}{\epsilon r_{ip}} \right) \right] \quad (1)$$

In equation (1), Θ_i is fixed, and $E(\Theta_i, d)$ is minimized to d_{\min} as a function of d . The first two terms in parentheses in equation (1) constitute the non-bonded steric potential

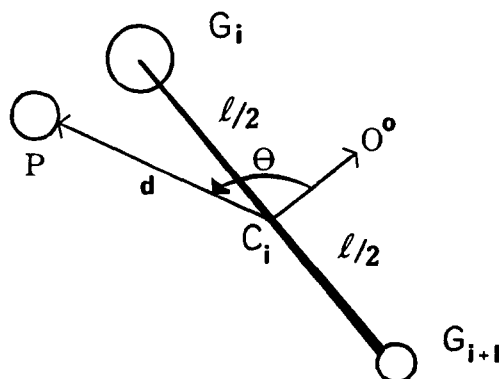


Figure 2 Geometry used to compute the intermolecular energies of torsion angle units. See text for definition of symbols

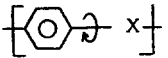
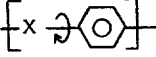
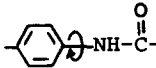
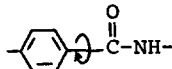
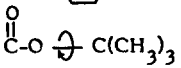
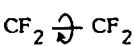
Table 1 Torsion angle units and corresponding mass moments, $M(\theta)$, conformational entropies, $S(\theta)$, and intermolecular probe energies $\langle E_D \rangle$, $\langle E_+ \rangle$, and $\langle E_- \rangle$

Torsion angle unit $\text{-}\overset{\text{X}}{\text{C}}\text{-}\overset{\text{R}}{\text{C}}\text{-CH}_2\text{-}$ and $\text{-CH}_2\text{-}\overset{\text{R}}{\text{C}}\text{-}\overset{\text{X}}{\text{C}}\text{-}$	$M(\theta)$ (a.m.u.)	$S(\theta)$ (cal K ⁻¹ mol ⁻¹)	$\langle E_D \rangle$ (kcal mol ⁻¹)	$\langle E_+ \rangle$ (kcal mol ⁻¹)	$\langle E_- \rangle$ (kcal mol ⁻¹)
X = -CH ₂ -	14.0	4.30	-1.39	0.45	-0.26
-O-	15.0	3.33	-0.86	-1.53	0.69
	21.0	0.78	-1.66	-2.19	-0.36
	29.0	2.17	-2.08	-2.69	-0.49
	29.0	2.17	-2.16	-2.55	0.30
-S-	23.0	2.28	-1.89	-0.89	-0.42
-SO ₂ -	39.0	0.68	-2.75	-3.25	-1.08
	28.5	1.16	-2.16	-4.45	-4.16
	28.5	1.30	-2.16	-4.22	-4.39
	45.0	0.67	-3.44	-0.96	-0.48
	45.0	0.76	-3.18	-0.88	-0.51
-CH(OH)-	22.0	1.48	-1.83	-0.80	-1.88
-CHF-	24.0	1.69	-1.52	-0.81	-0.38
-CHCl-	31.0	0.80	-1.93	-0.90	-0.46
-CF ₂ -	32.5	1.48	-1.69	-1.27	-0.16
-CCl ₂ -	48.0	0.84	-3.00	-1.19	-0.20
-CFCl-	39.5	0.90	-2.20	-1.36	-0.38
X = CH(R)					
R =	23.6	2.38	-1.83	-2.93	0.66
-CH ₃	14.0	3.06	-1.51	0.51	-0.37
-CH ₂ -	12.0	2.92	-1.44	0.46	-0.29
-O-	14.5	3.29	-0.89	-1.72	0.83
	18.3	1.17	-1.57	-2.33	-0.46
	34.3	1.00	-3.26	-1.09	-0.59
	34.3	1.03	-3.31	-0.69	-0.45
-S-	19.7	3.41	-1.75	-0.77	-0.49
-SO ₂ -	30.3	0.57	-2.58	-3.46	-1.37
	23.3	1.46	-2.03	-2.67	-2.05
	23.3	2.03	-1.95	-1.87	-2.43
X =	20.5	0.68	-1.60	-1.39	-0.44

Table 1 Continued

Torsion angle unit $\text{-X-}\overset{\curvearrowright}{\text{C}}\text{H}_2\text{-}$ and $\text{-CH}_2\overset{\curvearrowright}{\text{C}}\text{-X-}$	$M(\theta)$ (a.m.u.)	$S(\theta)$ (cal K ⁻¹ mol ⁻¹)	$\langle E_D \rangle$ (kcal mol ⁻¹)	$\langle E_+ \rangle$ (kcal mol ⁻¹)	$\langle E_- \rangle$ (kcal mol ⁻¹)
$\text{>C}_b\text{H}\overset{\curvearrowright}{\text{C}}\text{O-}$	14.5	2.17	-0.92	-1.43	-0.67
$\text{>C}_b\text{H}\overset{\curvearrowright}{\text{C}}\overset{\text{O}}{\parallel}\text{C-O-}$	28.5	0.83	-1.83	-1.59	-0.27
$\text{>C}_b\text{H}\overset{\curvearrowright}{\text{C}}\text{O}\overset{\text{O}}{\parallel}\text{C-}$	28.5	0.96	-1.70	-1.45	-0.56
$\text{>C}_b\text{H}\overset{\curvearrowright}{\text{C}}\text{CH}_2\text{-}$	13.5	1.45	-1.26	+0.29	-0.40
$\text{>C}_b(\text{CH}_3)\overset{\curvearrowright}{\text{C}}\overset{\text{O}}{\parallel}\text{C-O-}$	27.5	0.40	-1.72	-2.89	-0.72
$\text{-O}\overset{\curvearrowright}{\text{C}}\text{CH}_3^{(b)}$	15.5	4.72	-0.95	-1.43	0.46
$\text{CH}_2\overset{\curvearrowright}{\text{C}}\text{CH}_3$	14.5	4.33	-1.30	0.40	-0.31
$\text{-}\overset{\text{O}}{\parallel}\text{C-O}\overset{\curvearrowright}{\text{C}}\text{CH}_3$	29.5	3.00	-2.13	-2.51	0.35
$\text{-CH}_2\overset{\curvearrowright}{\text{C}}\text{C}(\text{CH}_3)\text{X-}$ and $\text{-C}(\text{CH}_3)\text{X}\overset{\curvearrowright}{\text{C}}\text{CH}_2\text{-}$					
X = -CH_3	28.0	1.22	-2.50	+0.77	-0.66
$\overset{\text{O}}{\parallel}\text{-C-O-}$	28.3	0.76	-2.83	-1.89	-0.52
$\overset{\text{O}}{\parallel}\text{-C-}$	23.3	1.17	-2.59	-1.58	-0.63
$\left[\text{C}_6\text{H}_4\overset{\curvearrowright}{\text{C}}\text{X} \right]$ and $\left[\text{X}\overset{\curvearrowright}{\text{C}}\text{C}_6\text{H}_4 \right]$					
X = -O-	46.0	0.65	-2.89	-1.09	0.17
$\overset{\text{O}}{\parallel}\text{-C-}$	52.0	0.39	-3.51	-2.69	-1.05
$\text{-O}\overset{\text{O}}{\parallel}\text{C-C-}$	68.0	0.48	-4.09	-3.05	-0.23
$\text{-C}(\text{CH}_3)_2\text{-}$	59.0	0.57	-4.16	+0.53	-0.82
$\text{-O}\overset{\text{O}}{\parallel}\text{C-}$	60.0	0.58	-3.78	-2.86	-0.65
$\overset{\text{O}}{\parallel}\text{-C-O-}$	60.0	0.58	-3.46	-2.98	-0.79
$\text{-CH(F)}\overset{\curvearrowright}{\text{C}}\text{CH(F)-}$	32.0	1.06	-1.36	-1.03	-0.52
$\text{-CF}_2\overset{\curvearrowright}{\text{C}}\text{CF(Cl)-}$	58.0	1.12	-2.18	-1.92	-1.08
$\text{C}_6\text{H}_4\overset{\curvearrowright}{\text{C}}\text{NH}\overset{\text{O}}{\parallel}\text{C-}$	59.5	0.66	-4.08	-2.05	-2.54
$\text{C}_6\text{H}_4\overset{\curvearrowright}{\text{C}}\overset{\text{O}}{\parallel}\text{C-NH-}$	59.5	0.60	-3.89	-2.60	-2.11

Table 1 Continued

Torsion angle unit	$M(\theta)$ (a.m.u.)	$S(\theta)$ (cal K ⁻¹ mol ⁻¹)	$\langle E_D \rangle$ (kcal mol ⁻¹)	$\langle E_+ \rangle$ (kcal mol ⁻¹)	$\langle E_- \rangle$ (kcal mol ⁻¹)
					
and 					
	59.5	0.58	-4.05	-2.07	-2.39
	59.5	0.53	-3.95	-2.62	-2.08
	50.5	1.37	-3.52	-1.06	+0.28
	31.0	0.91	-1.96	-2.13	-0.80

(a) These torsion angle units define backbone-sidechain branching units

(b) These torsion angle units terminate sidechains

function with the A_{ip} and B_{ip} parameters taken from the set proposed by Hopfinger⁶. The last term in parentheses in equation (1) is the electrostatic potential with the Q_x the partial charges on the atoms and probe. The Q_i were computed from CNDO/2 calculations⁷. The molecular dielectric, ϵ , was set equal to 3.5 (ref. 6).

Equation (1) does not necessarily give proper accounting to intermolecular hydrogen bonding. However, the electrostatic term does very much dominate the favourable intermolecular energetics when atom i is a hydrogen bond donor/acceptor and the probe, p , is an acceptor/donor.

When d_{\min} , and correspondingly, $E_p(\Theta_i, d_{\min})$, are determined for a given Θ_i , Θ_i is increased by $\Delta\Theta_i$, and the energy minimization of $E_p(\Theta_i, d)$ repeated as a function of d . This cycle, in turn, is carried out over the 360° available to Θ_i . The resulting set of $E_p(\Theta_i, d_{\min})$ is then used to compute the average intermolecular energy, $\langle E_p \rangle_i$, for p which is taken to be a component descriptor for the intermolecular energy.

$$\langle E_p \rangle_i = \frac{\sum_{j=1}^N E_p(\Theta_i, d_{\min})_j \exp(-E_p(\Theta_i, d_{\min})_j/RT)}{\sum_{j=1}^N \exp(-E_p(\Theta_i, d_{\min})_j/RT)} \quad (2)$$

In our analyses, $\Delta\Theta_i = 30^\circ$ so that $N = 12$; T was set to room temperature, 298K. In principle, the T in equation (2) should be treated as a variable in estimating T_g and T_m . However, conformational energy is independent of temperature for the fixed valence geometry approximation applied in this model. Moreover, the range in RT for the T_g and T_m values considered is only about 0.8 kcal mol⁻¹. Since RT appears in the same functional form in the numerator and denominator of equation (2) it has a very modest effect upon $\langle E_p \rangle_i$ over a range of only 0.8. Lastly, the $\langle E_p \rangle_i$ are correlation, as opposed to thermodynamic, properties. The $\langle E_p \rangle_i$ are designed to reflect relative composite average values of intermolecular interactions. Thus, there is not a requirement for computational precision of these quantities with respect to temperature to yield a significant QSPR.

The major difficulty in using the formalism reported here for estimating intermolecular interactions is in the selection of a probe, p , which meaningfully reflects the dominant type of intermolecular interactions at play in

the polymer system of interest. There does not seem to be any definite solution to this problem. Thus, we were forced to take advantage of the statistical nature of constructing a QSPR. This is the same approach that was successfully employed in developing quantitative structure-activity relationships, QSARs, using potential energy fields for a set of 2,4-diaminobenzylpyrimidine inhibitors of dihydrofolate reductase⁴. A probe to reflect non-bonded/dispersion interactions—a unified atom representation of a CH₃ group⁸ has been used in the intermolecular energy calculations. Probes to reflect electrostatic interactions with a positively charged group—a hydrogen atom with a unit positive charge, and with a negatively charged group—an oxygen atom with a negative formal charge, have also been considered. The average interaction energies $\langle E_D \rangle$, $\langle E_+ \rangle$ and $\langle E_- \rangle$, respectively, of a torsion angle unit with these three probes have been determined. In turn, $\langle E_D \rangle$, $\langle E_+ \rangle$, and $\langle E_- \rangle$ have each been considered as independent correlation terms, along with the intramolecular conformational entropy and mass moments, to establish the QSPR.

Multidimensional linear least-square regression analysis^{9,10} has been used to construct the correlation equation representing the QSPR. A variety of functional representations of the selected independent variables have been considered for the QSPR. The optimum functional representation, based upon the statistical significance of fit for the regression analyses explored, has been found to contain only linear terms in the independent variables

$$T_x = \alpha S_B + \beta M_B + \sum_i [\delta_i S_S(i) + r_i M_S(i)] + \epsilon \bar{E}_D + \rho \bar{E}_+ + \lambda \bar{E}_- + \omega \quad (3)$$

In equation (3), x can be g or m for the glass or crystal melt transition, respectively. S_B and $S_S(i)$ are the backbone and the i th sidechain contributions to the monomer conformational entropy. M_B and $M_S(i)$ are the corresponding mass moments of the monomer backbone and the i th sidechain of the monomer unit (Figure 1). \bar{E}_D , \bar{E}_+ and \bar{E}_- are the dispersion, positive electrostatic, and

negative electrostatic intermolecular energies for the complete monomer unit.

The greek letters in equation (3) are the regression coefficients which indicate the relative importance of each molecular property in specifying T_x . Moreover, comparison of the equivalent regression coefficients of the T_g and T_m QSPR equations provide a means of assessing the relative importance of the molecular properties upon the two transition processes.

The backbone entropy, S_B , and mass moment, M_B , are 'normalized' with respect to the number of torsion angle units, n_0 , composing the backbone of the monomer. It is to be noted that S_B and M_B include one-half contributions from each of the two torsion angle units joining a monomer unit to adjacent monomers on the left and right. The value of n_0 includes these two half torsion angle unit contributions and, thus, has a value one greater than the actual number of torsion angle units composing a monomer.

$$S_B = \frac{1}{n_0} \sum_{i=1}^{n_0} s_i(\Theta) \quad (4a)$$

$$M_B = \frac{1}{n_0} \sum_{i=1}^{n_0} m_i(\Theta) \quad (4b)$$

The terms $s_i(\Theta)$ and $m_i(\Theta)$ are the i th torsion angle unit entropy and mass moment, respectively, as reported in Table 1.

The entropy and mass moment of the i th sidechain of the monomer are also normalized by dividing each of these total respective quantities by the number of torsion

angle units, n_i , making up the i th sidechain.

$$S_S(i) = \frac{1}{n_i} \sum_{j=1}^{n_i} s_{i,j}(\Theta) \quad (5a)$$

$$M_S(i) = \frac{1}{n_i} \sum_{j=1}^{n_i} m_{i,j}(\Theta) \quad (5b)$$

The terms $s_{i,j}(\Theta)$ and $m_{i,j}(\Theta)$ are the same as in equations (4a) and (4b) and obtained from Table 1.

The three intermolecular energies, \bar{E}_D , \bar{E}_+ , and \bar{E}_- are computed as the respective sums of $\langle E_D \rangle$, $\langle E_+ \rangle$, and $\langle E_- \rangle$ for all backbone and sidechain torsion angle units composing the monomer divided by the number of such torsion angle units.

$$E_Y = \frac{1}{n_0} \sum_{i=1}^{n_0} \langle E_Y \rangle_i + \sum_{i=1}^{n_0} \frac{1}{n_i} \sum_{j=1}^{n_i} \langle E_Y \rangle_{i,j} \quad (6)$$

In equation (6), Y can be D, +, or - and $\langle E_Y \rangle_i$ and $\langle E_Y \rangle_{i,j}$ are taken from Table 1.

The normalization of the entropies, mass moments, and intermolecular energetics scales each of these properties to a single typical torsion angle unit characteristic of the polymer. This scaling, in turn, permits a meaningful comparison among polymers and the generation of QSPRs.

RESULTS

Table 2 contains one of the sets of polymers investigated along with the observed T_g s², conformational entropies, mass moments, and the three average intermolecular

Table 2 Polymers, corresponding entropy, mass moment, and intermolecular energy descriptors, and the observed and calculated (using eq. (10)) T_g values as well as the differences, ΔT_g , in observed and calculated values

No.	Polymer	S_B (cal mol ⁻¹ K ⁻¹)	M_B (a.m.u.)	S_S (cal mol ⁻¹ K ⁻¹)	M_S (a.m.u.)	\bar{E}_D (kcal mol ⁻¹)	\bar{E}_+ (kcal mol ⁻¹)	\bar{E}_- (kcal mol ⁻¹)	T_g (obs.) (K)	T_g (calc.) (K)	ΔT_g (K)
1	-CH ₂ -O-	3.33	15.0	-	-	-0.86	-1.53	0.69	188/243	220.9	22.1
2	-(-CH ₂) ₂ -O-	3.65	14.7	-	-	-1.04	-0.87	0.37	206/246	221.5	-15.5
3	-(-CH ₂) ₃ -O-	3.82	14.5	-	-	-1.13	-0.54	0.22	195/228	210.6	17.4
4	-(-CH ₂) ₄ -O-	3.91	14.4	-	-	-1.18	-0.34	0.12	185/194	204.7	-10.7
5	-(-CH ₂)-	4.30	14.0	-	-	-1.39	0.45	-0.26	143/250	176.9	-33.9
6	-CH ₂ -CH(CH ₃)-	1.93	14.0	-	-	-1.51	0.51	-0.37	238/299	249.0	-11.0
7	-CH ₂ -CH(ϕ)-	1.00	34.3	-	-	-3.26	-1.09	-0.59	353/380	334.3	18.7
8	-CH ₂ -CH(F)-	1.70	24.0	-	-	-1.52	-0.81	-0.38	253/314	307.8	6.2
9	-CH ₂ -CF ₂ -	1.48	32.5	-	-	-1.69	-1.27	-0.16	238/286	325.6	-39.6
10	-CH(F)-CH(F)-	1.06	32.0	-	-	-1.36	-1.03	-0.5	323/371	329.4	-6.4
11	-CH ₂ -CH(Cl)-	0.80	31.0	-	-	-1.93	-0.90	-0.46	247/354	341.1	12.9
12	-CF ₂ -CF(Cl)-	1.12	58.0	-	-	-2.18	-1.92	-1.08	318/373	377.0	-4.0
13		1.96	38.4	-	-	-2.59	-2.10	-0.22	346	351.2	-5.2
14		0.58	56.5	-	-	-3.77	-2.00	-0.36	393/420	397.2	-4.2
15		0.53	63.5	-	-	-4.13	-1.26	-0.53	414/423	381.1	32.9
16	-CH ₂ -C(CH ₃)- COOCH ₃	0.76	28.3	1.70	28.5	-2.38	-2.29	-0.41	378	353.8	24.2

Table 2 Continued

No.	Polymer	S_B (cal mol ⁻¹ K ⁻¹)	M_B (a.m.u.)	S_S (cal mol ⁻¹ K ⁻¹)	M_S (a.m.u.)	E_D (kcal mol ⁻¹)	E_+ (kcal mol ⁻¹)	E_- (kcal mol ⁻¹)	T_g (obs.) (K)	T_g (calc.) (K)	ΔT_g (K)
17	$-\text{CH}_2-\text{C}(\text{CH}_3)-$ $\text{COOCH}_2\text{CH}_3$	0.76	28.3	2.29	23.7	-2.15	-1.79	-0.56	338	319.3	18.7
18	$-\text{CH}_2-\text{C}(\text{CH}_3)-$ $\text{COO}(\text{CH}_2)_{15}\text{CH}_3$	0.76	28.3	3.95	15.7	-1.59	-0.14	-0.34	288	258.6	29.4
19	$-\text{CH}_2-\text{C}(\text{CH}_3)-$ $\text{COOC}(\text{CH}_3)_3$	0.76	28.3	0.89	39.0	-2.72	-1.93	-0.42	380	373.3	6.7
20	$-\text{CH}_2-\text{C}(\text{CH}_3)-$ $\text{COO}-\text{C}_6\text{H}_5$	0.76	28.3	0.49	43.7	-2.71	-2.41	-0.69	385	393.9	-8.9
21	$-\text{CH}_2-\text{C}(\text{CH}_3)_2-$	1.22	28.0	-	-	-2.50	0.77	-0.66	198/243	272.6	-29.6
22	$-\text{CH}_2-\text{CH}(\text{OH})-$	1.48	22.0	-	-	-1.83	-0.80	-1.88	343/372	332.1	10.9
23	$-\text{CH}_2-\text{CH}(\text{C}_2\text{H}_5)-$	2.92	12.0	2.89	14.0	-1.36	0.40	-0.32	228/249	218.2	9.8
24	$-\text{CH}_2-\text{CH}(\text{C}_3\text{H}_7)-$	2.92	12.0	3.36	14.0	-1.37	0.41	-0.31	221/287	213.0	8.0
25	$-\text{CH}_2-\text{CH}(\text{OCH}_3)-$	3.29	14.5	3.45	15.0	-0.91	-1.58	0.36	242/260	242.2	-0.2
26	$-\text{CH}_2-\text{CH}(\text{OCH}_2\text{CH}_3)-$	3.29	14.5	3.28	14.7	-0.97	-1.20	0.27	231/254	234.2	-3.2
27	$-\text{CH}_2-\text{CH}-$ COOCH_3	2.38	23.6	1.92	29.0	-1.91	-2.49	0.35	279/282	318.1	-36.1
28	$-\text{CH}_2-\text{CH}-$ $\text{COOCH}_2\text{CH}_3$	2.38	23.6	2.44	24.0	-1.77	-1.95	0.05	251	283.3	-32.3
29	$-\text{CH}_2-\text{CH}-$ $\text{COOC}(\text{CH}_3)_3$	2.38	23.6	1.10	39.5	-2.25	-2.13	0.33	314	305.3	8.7
30	$-\text{CH}_2-\text{CH}-$ $\text{COO}(\text{CH}_2)_3\text{CH}_3$	2.38	23.6	3.19	20.0	-1.66	-1.26	-0.04	219	252.1	-33.1
31	$-(\text{CH}_2)_{10}-\text{O}-\text{C}(=\text{O})-\text{C}_6\text{H}_4-\text{C}(=\text{O})-\text{O}-$	3.40	23.4	-	-	-1.84	-0.54	-0.31	268/298	240.5	27.5
32	$-\text{NH}-\text{C}_6\text{H}_4-\text{NHC}(=\text{O})-\text{C}_6\text{H}_4-\text{C}(=\text{O})-$	0.59	59.5	-	-	-3.97	-2.34	-2.25	411/428	438.4	-27.4
33	$-\text{NH}-(\text{CH}_2)_6-\text{NHC}(=\text{O})-(\text{CH}_2)_4-\text{C}(=\text{O})-$	3.27	18.8	-	-	-1.65	-1.16	-1.58	318/330	277.2	40.8
34	$-\text{CH}_2-\text{CH}-$ C_6H_{15}	2.92	12.0	3.83	14.0	-1.38	0.43	-0.29	208/228	207.8	0.2
35	$-\text{CH}_2-\text{CH}-$ $\text{O}-(\text{CH}_2)_5\text{CH}_3$	2.92	12.0	3.86	14.3	-1.28	0.02	-0.21	196/223	216.9	6.1

energies. In many cases, two observed T_g s are reported for a particular polymer. The T_g value which best fits the QSPR equation was used in each QSPR analysis. In our initial investigation to predict T_g^1 , in which intermolecular energetics were neglected, the following QSPR was established for 30 of the polymers in Table 2;

$$T_g = -35.18S_B + 1.55M_B - 18.26S_S + 1.35M_S + 327.30$$

$$N = 30 \quad R = 0.929 \quad \text{SD} = 17.9 \quad F = 39.5 \quad (7)$$

In equation (7) N is the number of compounds, R the correlation coefficient, SD the standard deviation of fit, and F the statistical significance of fit. The entropy terms

in equation (7) account for more than 70% of the variance in the T_g values. Thus, our initial thought was to delete the mass moment terms in equation (7), and to replace them with the intermolecular energy descriptors. This resulted in the following QSPR equation:

$$T_g = -21.46S_B - 9.83S_S - 17.24\bar{E}_D - 29.97\bar{E}_+ - 18.12\bar{E}_- + 263.84 \quad (8)$$

$N = 35 \quad R = 0.912 \quad SD = 20.2 \quad F = 28.8$

The relative magnitude of the regression coefficients of equation (8) suggest that no descriptors can be deleted without compromising the integrity of the QSPR. Moreover, a comparison of equations (7) and (8) suggest that the intermolecular energy terms do not fully compensate for the deletion of the mass moments.

The decision to delete mass moments was based upon the statistical significance of the terms in equation (7), not conceptual understanding of the glass transition process. Conceptually, conformational entropy is associated with the thermodynamic component of the glass transition process, while the mass moments represent an attempt to account for the kinetic nature of the glass transition. Equation (8) only contains thermodynamic descriptors, and its marginal significance, relative to equation (7), indicates that mass moments are needed in the QSPR. Thus, we carried out a multidimensional linear regression analysis using all seven descriptors (S_B , M_B , S_S , M_S , $\langle E_D \rangle$, $\langle E_+ \rangle$, and $\langle E_- \rangle$) for the 35 compounds whose T_g values are reported in Table 2.

$$T_g = -25.3S_B - 7.66S_S + 1.40M_B - 0.14M_S - 0.77\bar{E}_D - 31.6\bar{E}_+ - 24.4\bar{E}_- + 275.17 \quad (9)$$

$N = 35 \quad R = 0.959 \quad SD = 14.9 \quad F = 44.0$

An analysis of the coefficients of the independent variables (molecular properties) in equation (9) indicates that the normalized mass of the side chain torsion angle units, and the intermolecular dispersion energy, make minimal contributions to the variance of T_g over the set of polymers given in Table 2. Thus, we carried out a multidimensional linear regression analysis without these two terms. This led to the most significant QSPR for T_g ,

$$T_g = -27.3S_B - 10.1S_S + 1.07M_B - 29.3\bar{E}_+ - 15.1\bar{E}_- + 288.83 \quad (10)$$

$N = 35 \quad R = 0.954 \quad SD = 15.6 \quad F = 58.9$

The predicted T_g s and differences with the observed values using equation (10) are listed in Table 2. The choice of the observed T_g for each compound used to construct equation (10) is obvious. Figure 3 is a plot of observed versus predicted T_g s based upon equation (10).

The optimum QSPR equation to estimate T_m contains the same set of molecular properties as equation (10).

$$T_m = -32.6S_B - 22.1S_S - 2.51M_B - 50.5\bar{E}_+ - 109.8\bar{E}_- + 493.7 \quad (11)$$

$N = 30 \quad R = 0.907 \quad SD = 31.6 \quad F = 22.4$

However, it is clear from the R , SD , and F values that equation (11) is inferior for predicting T_m as compared with equation (10) for predicting T_g . The set of polymers, corresponding molecular descriptors, and the observed², predicted, and differences in T_m values are reported in Table 3. An inspection of the observed, predicted and

difference T_m values suggests that there is no pattern to the errors in estimating T_m . Indeed, only 11 predicted T_m values are within $\pm 20K$ of the observed T_m measurements. Eight predicted T_m values are more than $\pm 50K$ from the measured values. Figure 4 is a plot of observed versus predicted T_m values, and reflects the relatively high random scatter inherent to the predicted and observed T_m values.

DISCUSSION

Equation (10) suggests that it is possible to meaningfully estimate T_g s in terms of conformational entropy, mass moments, and electrostatic intermolecular interactions. However, these molecular properties do not contribute uniformly to the specification of T_g . Thus, multidimensional linear regression T_g equations were generated for

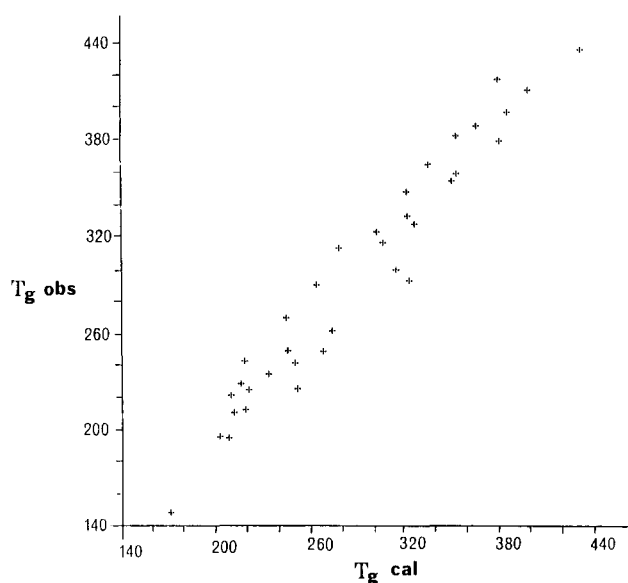


Figure 3 Plot of observed versus calculated (eq. (10)) T_g values for the polymers in Table 2

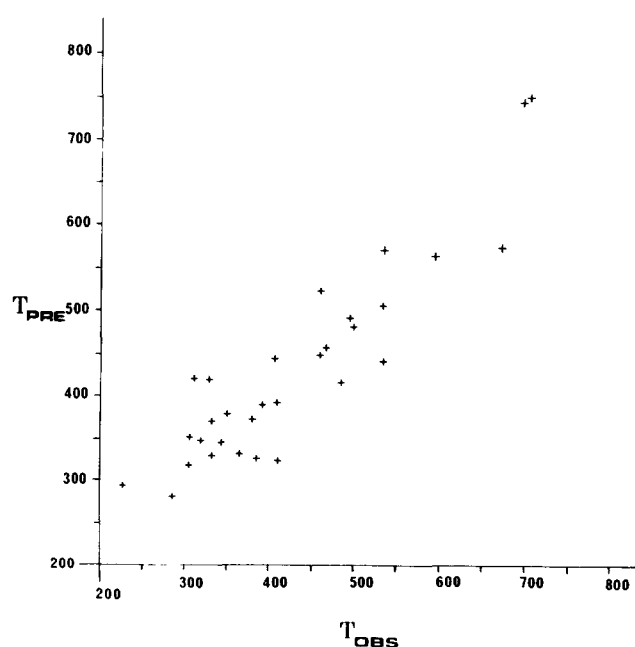


Figure 4 Plot of observed versus calculated (eq. (11)) T_m values for the polymers in Table 3

Table 3 Polymers used to construct the QSPR given by equation (11). The descriptor terms have the same meaning as in Table 2. T_m is the crystal-melt transition temperature





No.	Polymer	S_B (cal mol ⁻¹ K ⁻¹)	M_B (a.m.u.)	S_S (cal mol ⁻¹ K ⁻¹)	M_S (a.m.u.)	\bar{E}_D (kcal mol ⁻¹)	E_+ (kcal mol ⁻¹)	\bar{E}_- (kcal mol ⁻¹)	T_m (obs.) (K)	T_m (calc.) (K)	ΔT_m (K)
1	-CH ₂ -O-	3.33	15.0	-	-	-0.86	-1.5	0.69	333/473	349.3	-16.3
2	-(CH ₂) ₂ -O-	3.65	14.7	-	-	-1.04	-0.87	0.37	335/349	341.4	-6.4
3	-(CH ₂) ₃ -O-	3.82	14.5	-	-	-1.13	-0.54	0.22	308	336.2	-28.2
4	-(CH ₂) ₄ -O-	3.91	14.4	-	-	-1.18	-0.34	0.12	308/333	334.3	-26.4
5	-(CH ₂)-	4.30	14.0	-	-	-1.39	0.45	-0.26	410	324.5	85.5
6	-CH ₂ -CH(CH ₃)-	1.93	14.0	-	-	-1.51	0.51	-0.37	385/481	410.7	70.3
7	-CH ₂ -CH(ϕ)-	1.00	34.3	-	-	-3.26	-1.09	-0.59	498/523	495.0	3.0
8	-CH ₂ -CH(F)-	1.70	24.0	-	-	-1.52	-0.81	-0.38	473	460.9	12.1
9	-CH ₂ -CF ₂ -	1.48	32.5	-	-	-1.69	-1.27	-0.16	410/511	445.8	-35.8
10	-CH ₂ -C(CH ₃) ₂ -	1.22	28.0	-	-	-2.50	0.77	-0.66	275/317	417.4	-100.4
11	-CH ₂ -CH(Cl)-	0.80	31.0	-	-	-1.93	-0.90	-0.46	485/583	485.9	-0.9
12	-CF ₂ -CF(Cl)-	1.12	58.0	-	-	-2.18	-1.92	-1.08	483/533	527.4	5.6
13	-CH ₂ -C(CH ₃)- COOCH ₃	0.76	28.3	1.70	28.5	-2.38	-2.29	-0.41	433/473	521.2	-48.3
14	CH ₂ -CH(OCH ₃)-	3.29	14.5	3.45	15.0	-0.91	-1.58	-0.36	417/423	393.5	23.5
15	-CH ₂ -CH(OCH ₂ CH ₃)-	2.92	12.0	2.89	14.0	-1.36	0.40	-0.32	359	319.8	39.2
16	-CH ₂ -CH- COO(CH ₂) ₃ CH ₃	2.38	23.6	3.19	20.0	-1.66	-1.26	-0.04	275/317	354.8	-37.8
17	-(CH ₂) ₁₀ -O-C(=O)-  -C(=O)-O-	3.40	23.4	-	-	-1.84	-0.54	-0.31	396/411	385.7	10.3
18	-CH ₂ -CCl ₂ -	0.84	48.0	-	-	-3.00	-1.36	-0.38	463/483	456.5	6.5
19	-CF ₂ -CF ₂ -	0.91	31.0	-	-	-1.96	-2.13	-0.80	292/672	581.8	90.2
20	-CH ₂ -CH- CH ₂ -CH ₃	2.92	12.0	2.89	14.0	-1.36	0.40	-0.32	379/415	319.8	59.2
21	-CH ₂ -CH- (CH ₂) ₅ CH ₃	2.92	12.0	3.83	14.0	-1.38	0.43	-0.29	235	294.2	-59.2
22	CH ₂ -CH O=C-O-(CH ₂) ₂ -CH ₃	2.38	23.6	2.91	21.5	-1.71	-1.55	0	388/435	371.2	16.8
23	-(CH ₂) ₂ -NHC(=O)-  -C(=O)-NH-	1.56	38.0	-	-	-2.72	-2.47	-2.59	728	756.8	-28.8
24	-(CH ₂) ₂ -O-C(=O)-  -C(=O)-O-(CH ₂) ₁₀ -O-	3.43	21.3	-	-	-1.72	-0.48	-0.10	338	363.9	-25.9
25	-CH ₂ -CH- O (CH ₂) ₉ CH ₃	3.29	14.5	4.02	14.2	-1.23	-0.18	-0.05	280	276.2	3.8
26	-(CH ₂) ₂ -O-C(=O)-  -C(=O)-O-	2.35	34.3	-	-	-2.39	-1.67	-0.23	533/537	440.9	92.1

Table 3 Continued

No.	Polymer	S_B (cal mol ⁻¹ K ⁻¹)	M_B (a.m.u.)	S_S (cal mol ⁻¹ K ⁻¹)	M_S (a.m.u.)	\bar{E}_D (kcal mol ⁻¹)	\bar{E}_+ (kcal mol ⁻¹)	\bar{E}_- (kcal mol ⁻¹)	T_m (obs.) (K)	T_m (calc.) (K)	ΔT_m (K)
27		3.24	21.5			-1.75	-1.10	-0.24	332	416.3	-84.3
28		3.27	18.8			-1.65	-1.16	-1.58	523.545	572.2	49.2
29		2.39	29.2			-2.30	-1.20	-1.40	606.613	557.0	49.0
30		3.87	17.0			-1.53	-0.17	-0.27	344.358	363.4	-19.4

all possible combinations of molecular properties in Table 2. The minimum number of molecular properties was sought which yielded $R > 0.92$, $SD < 20$ and $F > 50.0$. It was found that an equation involving only S_B , \bar{E}_+ and \bar{E}_- could satisfy these conditions.

$$T_g = -35.3S_B - 29.9\bar{E}_+ - 29.3\bar{E}_- + 321.7 \quad (12)$$

$$N = 35 \quad R = 0.923 \quad SD = 18.2 \quad F = 59.5$$

One might dismiss equation (10) in favour of equation (12) for estimating T_g s as well as inferring mechanisms of molecular action in the T_g process. That is, one might not consider mass moments or side chain conformational entropy as important. This, however, is dangerous to do in that the size of our T_g data base is small—35 compounds. Moreover, in the previous investigation to model T_g^{-1} it was found that the T_g s of a set of homologous polyacrylates and a set of homologous polymethacrylates could be explained solely in terms of side chain entropy and mass moments. Many more polymers need to be considered in generating a reliable universal T_g QSPR. However, it is probably fair to say that S_B , \bar{E}_+ and \bar{E}_- are the dominant molecular properties correlating with T_g .

There is no obvious explanation why T_m cannot be significantly correlated against some combination of S_B , M_B , S_S , M_S , \bar{E}_D , \bar{E}_+ , and \bar{E}_- . If polymers in Table 3 having an absolute difference in predicted and observed T_m s greater than 85K are deleted in constructing a QSPR, the following correlation is achieved

$$T_m = -28.9S_B - 17.5S_C - 1.54M_B - 37.4\bar{E}_+ - 114.7\bar{E}_- + 459.4 \quad (13)$$

$$N = 26 \quad R = 0.959 \quad SD = 24.7 \quad F = 45.9$$

Equation (13) is quite an improvement over equation (11). Unfortunately, the four polymers deleted in construction of equation (13) do not share any common features which might explain why it is more difficult to accurately predict T_m compared with T_g . The four deleted compounds are polyethylene, poly(isobutylene), polytetrafluoroethylene, and poly(ethylene terephthalate). Thus, it is not possible to propose what additional molecular features need to be considered in order to successfully predict T_m s.

One surprising finding is that T_m is predicted to decrease as the normalized mass of the backbone

monomer unit increases in both equations (11) and (13). This is the case since the regression coefficients are negative. The opposite is true for T_g . Intuition suggests that both T_g and T_m should increase with normalized monomer mass. Since the regression coefficients for M_B have a small absolute value, the mass terms make relatively small contributions to specifying T_g and T_m . Hence, this discrepancy may not be meaningful and an artifact of the statistical fit.

A comparison of the regression coefficients of equations (10) and (11) suggests that sidechain entropy and the electrostatic intermolecular energies are more important (the absolute values of the coefficients are larger) for specifying T_m than T_g . The reference point is the backbone entropy whose regression coefficient does not change very much between equations (10) and (11). In particular, the coefficient of \bar{E}_+ changes by more than a factor of seven. It is not obvious why these specific changes are seen, but intuition would support intermolecular interactions being more important in maintaining a crystal than a glass. Moreover, sidechain entropy in a crystal might reflect the onset of disruption of both intrachain and interchain order since sidechains are in more intimate contact in a crystal than in a glass. This could explain the increased significance of S_C in equation (11) as compared with equation (10).

A probable reason for not being able to predict T_m s as well as T_g s may be due to the anisotropic geometric environment inherent to a polymer crystal, in contrast to the orthotropic medium of a polymer glass. The molecular properties we have computed using one-dimensional torsion angle units better model the orthotropic environment of a glass, than the anisotropic geometry surrounding a torsion angle unit in a polymer crystal. There is no allowance in the current model to consider the effects of direction-dependent intermolecular interactions. Specific, periodic lattice interactions which stabilize the crystal are not accounted for in the current model. We have seen in previous polymer crystal structure calculations³ that the electrostatic interactions play a crucial role in specifying lattice geometry, especially the setting angles of the chains. This finding from lattice packing calculations can be taken as evidence that larger structural units, which build anisotropic molecular interactions into the model, are needed to generate significant T_m QSPRs.

Thus, we plan to consider a follow-up QSPR analysis of the polymers in *Table 3* using two, and higher dimensional, torsion angle units to estimate the entropies, mass moments, and intermolecular energetics.

In so far as the data sets are large enough to impart self-consistency in the regression equations, the approach taken here might be used to resolve which of two, or more, reported transition temperatures for a polymer is likely to be correct. That has been done in the derivation of QSPR equations. When two observed T_g , or T_m , values are given, the observed value which maximizes the regression fit is used. Thus, molecular modelling may be of use in resolving conflicting experimental measurements.

Lastly, there is certainly some uncertainty in each of the observed values of the T_g s and T_m s regardless of whether one or two values are reported. Hence, it may be instructive to see how sensitive the QSPRs are to observed T_g and T_m values. This could be accomplished by repeatedly randomly perturbing the observed values for the transition temperatures some average amount, and seeing how sensitive the resultant correlation equations are to the perturbation. This is another study that will be considered in the near future.

ACKNOWLEDGEMENTS

This work was, in part, supported by the Laboratory of

Computer-Aided Molecular Modelling and Design at the University of Illinois at Chicago. All calculations were carried out using the CHEMLAB-II molecular modelling package.

REFERENCES

- 1 Hopfinger, A. J., Koehler, M. G., Pearlstein, R. A. and Tripathy, S. K. *J. Polym. Sci. Polym. Phys. Edn.* in press
- 2 Van Krevelen, D. W. and Hoftyzer, P. J. 'Properties of Polymers', Elsevier, Amsterdam, 1976, and references cited therein
- 3 Tripathy, S. K., Hopfinger, A. J. and Taylor, P. L. *J. Phys. Chem.* 1981, **85**, 1371
- 4 Hopfinger, A. J. *J. Med. Chem.* 1983, **26**, 990
- 5 Politzer, P. and Truhlar, D. G. (Eds.) 'Chemical Applications of Atomic and Molecular Electrostatic Potentials', Plenum Press, New York, 1981
- 6 Hopfinger, A. J. 'Conformational Properties of Macromolecules', Academic Press, New York, 1973
- 7 Pople, J. A. and Beveridge, D. C. 'Approximate Molecular Orbital Theory', McGraw-Hill, New York, 1970
- 8 Mabilia, M., Pearlstein, R. A. and Hopfinger, A. J. in 'Molecular Graphics and Drug Design' (Eds. A. S. V. Burgen, G. C. K. Roberts and M. S. Tute), Elsevier, Amsterdam, 1986, p. 157
- 9 Wold, S., Hellber, S. and Dunn III, W. J. *Acta Pharmacol. Toxicol.* 1983, **52**, 158
- 10 Dunn III, W. J. and Wold, S. *Bioorg. Chem.* 1980, **9**, 505