Prediction of the Formulation Dependence of the Glass Transition Temperatures of Amine-Epoxy Copolymers Using a QSPR Based on the AM1 Method

Jason A. Morrill,*,† Robert E. Jensen, Phillip H. Madison, and Cary F. Chabalowski

United States Army Research Laboratory, Weapons and Materials Research Directorate, AMSRL-WM-BD, Aberdeen Proving Ground, Maryland 21005-5066

A designer Quantitative Structure−Property Relationship, based upon molecular properties calculated using the AM1 semiempirical quantum mechanical method, was developed to predict the glass transition temperature of amine-cured epoxy resins based on the diglycidyl ether of bisphenol A. The QSPR ($R^2 = 0.9977$) was generated using the regression analysis program, COmprehensive DEscriptors for Structural and Statistical Analysis. By applying an ad hoc treatment based on the elementary probability theory to the quantitative structure−property relationship analysis a method was developed for computing bulk polymer glass transition temperatures for stoichiometric and nonstoichiometric monomeric formulations. A model polymer was synthesized and found to validate our model predictions.

INTRODUCTION

In describing the tendency of a polymer to absorb energy the glass transition temperature (T_g) is a property of fundamental importance. It is the temperature above which thermal motions of atoms about their equilibrium positions in a polymer increase the free volume enough to allow long-range coupled motions of larger groups of atoms along the backbone of the polymer.[1] At the glass transition temperature the increase in the mobility of the polymer backbone gives rise to a change from a glassy state to a more fluid rubber-like state. The change in the backbone rigidity of a polymer consequently leads to a change in a host of the physical properties of the polymer. Among these properties are density, mechanical moduli, dielectric coefficients, viscosity, and gaseous or liquid absorption.[2] Toward designing new polymeric materials and predicting their overall physical performance it is useful to have a means by which the T_g of a polymer can be reliably predicted.

Numerous researchers have found that T_g values for linear polymers can be estimated on the basis of Quantitative Structure−Property Relationships (QSPRs) developed from the structure of the repeat units making up the polymer. Many such approaches have relied on the additive nature of the repeat units in contributing to the T_g value of the polymer and are termed group additivity methods. One of the best-known examples of the group additivity approach is that of van Krevelen.[3] The van Krevelen approach is purely empirical consisting of a weighted sum of scalar quantities associated with functional groups commonly occurring in polymers. This method provides a rapid and computationally inexpensive approach to the estimation of T_g values but has been criticized due to its limited extensibility to systems composed only of functional groups that have been previously investigated.[4]

Bicerano[5] extended the group additivity concept to a more generally applicable model including solubility parameters and topological considerations that are independent of specific functional groups. These quantities were used as descriptors of the cohesive forces and chain stiffness of polymers as these properties relate to glass transition temperatures. Using quantities of this type a relationship with a coefficient of determination of 0.9749 was obtained for a data set consisting of 320 polymers, thus accounting for 95% of the variance in the experimental T_g values. The model overestimated the T_g values for a series of polymers consisting of phthalimide groups. The lack of performance of the model for a specific class of polymers is likely due to it being based on a large and widely varied training set of polymers. This is a well-known phenomenon experienced in QSPR studies in which superior correlations are obtained from a congeneric series of molecular species and has led to the development of so-called "designer" correlation equations.[6−9]

Koehler and Hopfinger extended the group additivity approach to include terms describing intermolecular interactions, such as hydrogen bonding, to relate the contribution of interchain interactions to polymer T_g values.[10] This approach made use of a probe to describe nonbonded electrostatic and steric repulsions, via a molecular mechanics treatment. A linear combination of backbone and side chain flexibility parameters from Boltzmann distributions of conformational energy states was also used to predict T_g values. Several correlations were obtained all with coefficients of determination in excess of 0.9.

A designer correlation equation was developed by Camelio et al.,[9,11] to investigate the role of bulky substituents in determining the glass transition temperatures of substituted methacrylates and acrylates. The Camelio QSPR model, referred to as the Energy, Volume, and Mass (EVM) model, was developed to describe the effects of bulky substituents in terms of energetic barriers to rotation along the polymer backbone as affected by the proximity of the center of mass

---

and volume of substituents to the polymer backbone. To give a more accurate picture of the volume of the "free" space around the polymers Camelio randomized the conformations of the polymer models by a molecular dynamics/molecular mechanics simulation. A standard deviation of Cartesian coordinates of atoms along the polymer backbone was then used to define the volume of an elliptical cylinder around the polymer backbone. The other terms used in the correlation analyses included intramolecular energies, intermolecular van der Waals energies, and repeat unit molecular weights. Correlations obtained using these descriptors gave coefficients of determination greater than 0.90 for several groups of polymers.[8,10]

In contrast to previous models that are based on specifically chosen molecular descriptors, Katritzky and co-workers created a generally applicable QSPR development program, called COmprehensive DEscriptors for Structural and Statistical Analysis (CODESSA).[12] CODESSA is typically used in conjunction with a molecular modeling package to analyze hundreds of computed molecular descriptors and then sort them into multivariate linear equations based on the statistical correlation of the descriptors to a property of interest. This approach has been shown in numerous studies to result in superior correlations for a wide variety of molecules and several physical and biological properties. Katritzky used CODESSA to predict glass transition temperatures for 88 linear homopolymers using a five-parameter equation that gave a coefficient of determination of 0.946.[4] Cao and Lin have tested the same set of 88 linear homopolymers against a correlation involving their own chosen descriptors in an attempt to derive a more physically meaningful QSPR.[13] Their descriptors were derived on the basis of chain mobility (including that of substituents along the polymer backbone) and intermolecular forces. In their model they included a descriptor for the volume of a terminal substituent of a side chain and a topological descriptor for the longest path (largest number of backbone bonds along the side chain less one) of a substituent. Intermolecular interactions were treated for polymers having polar groups by consideration of electronegativities and atomic polarizabilities. The parameters were found to give a QSPR having a coefficient of determination of $R^2 = 0.9056$.

All of the discussed approaches to predicting T_g values were developed on the basis of linear homopolymers despite the importance of cross-linked copolymers in commercial applications. The focus on linear homopolymers is likely due to the complexity associated with constructing an atomistic model of the branching that exists in a networked copolymer. This complexity, when coupled with the added computational cost of modeling a polymer structure large enough to incorporate various degrees of cross-linking, has previously made simulations of the effect of cross-linking on polymer properties such as T_g extremely computationally expensive, even at modest levels of theory. Despite this difficulty, attempts have been made by researchers to model the effect of the extent of cross-linking on T_g. Galy used a group additivity model to predict T_g values for nonstoichiometric amine-cured epoxy networks.[14] Bicerano employed a relation that involved a correction to maximal experimental glass transition temperature values as a function of the number of rotatable bonds and average number and molecular weight of repeat units between cross-links.[15] And, Beloshenko related the mean statistical molecular mass between cross-links and compression moduli (both obtained from experimental data) to the T_g for an amine/anhydride cured epoxy cross-linked copolymer.[16] These approaches all require the existence of experimental data or previously determined group additivity parameters for a polymer of interest in order to make a prediction of T_g for that *same* polymer.

Our interest is in the rational design of amine-cured bisphenol A epoxy copolymers with improved impact and shock resistance. Empirically, it is known that substitution of aromatic backbones of polymers of this type with cycloaliphatic species yields increased localized molecular mobility in the glassy state of the polymer. These short-range relaxation mechanisms yield increased high strain rate impact and shock resistance. However, this approach often leads to inferior performance in other critical polymer properties, such as the coefficient of thermal expansion. For successful application a true balance of properties is necessary, and using traditional empirical methods to discover an optimal copolymer formulation may never be achieved due to the elevated number of formulations that must be probed. Herein we describe our results for the prediction of the behavior of the glass transition temperature with respect to the stoichiometric ratio of monomers in a cross-linked amine-cured model epoxy polymer.

PROCEDURE

General Computational Procedure. In general, our computational procedure involves the development of a designer QSPR for prediction of the glass transition temperatures of amine-cured polymers based on DGEBA. In a second step the QSPR is used to predict the contribution of repeat units, each built to mimic cross-link points having varied degrees of branching in a networked copolymer. The contributions of each of the cross-link points are then multiplied by a weighting factor obtained from a previously developed method[17] to describe the probability of finding a finite chain in a polymer network. Finally, the individual contributions are summed to arrive at a composite predicted T_g for any given stoichiometric ratio. This approach retains the group additivity concept proposed by van Krevelen[3] but is fundamentally different. The group contributions used here are not tabulated values; rather they are predicted solely on the basis of computable molecular descriptors, found to be predictive of T_g, and the appropriate choice of a polymer repeat group.

Construction of Polymer Models. Toward simplifying the polymer structures, representative polymer models were constructed for a set of 13 amine-cured epoxy copolymers for which T_g data exists in the literature. The monomers in the data set consisted of DGEBA cured with one amine curing agent. The curing agents consisted of 11 diamines and 2 monoamines and were formulated in a 1:1 stoichiometry of functional groups, which is defined as one amine N−H bond per DGEBA oxirane ring. Since the literature data corresponded to stoichiometric formulations, it was desirable to build models having relative numbers of monomeric subunits that are equivalent to their respective literature stoichiometries.

914 *J. Chem. Inf. Comput. Sci., Vol. 44, No. 3, 2004*

MORRILL ET AL.

Scheme 1. Repeat Unit (Right) Resulting from the Events Used by Miller and Macosko To Derive a Set of Recursion Relations for Solving the Probability that a Randomly Chosen Monomeric Subunit Leads to a Finite Chain
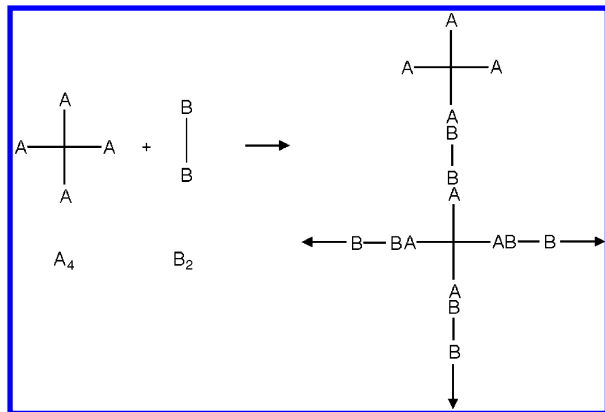
It is generally observed for systems of this type that a maximally cross-linked polymer results from equivalent stoichiometries of functional groups.[18] Further, maximal cross-linking results in an infinitely large and hyperbranched macromolecule.[16,19] Choosing a repeat unit for such a complex polymer is not as straightforward as it is for linear homopolymers or semiregularly repeating linear copolymers. Fortunately, the nature of branching in polymers of the type with which we are concerned has been studied previously, and this work led to our choice of a suitable repeat unit to use for the basis of our QSPR model.

To predict the concentration of elastic chains and cross-link densities for calculating molecular weight averages of hyperbranched polymers the work of Flory[18] was simplified by Miller and Macosko.[16] Their approach retained Flory's ideal network assumptions:

(1) all functional groups of the same type are equally reactive;

(2) all functional groups react independently of one another;

(3) no intramolecular reactions occur in finite species.

Their derivation was used to describe random stepwise copolymerization of a four-functional monomer with a two-functional monomer and is analogous to the copolymerization of a primary diamine (four-functional) and a diglycidyl ether (two-functional). The derivation begins with the reaction of the four-functional monomer (A_4) with a two-functional monomer (B_2) (see Scheme 1).

The probability of A_4 reacting with B_2 such that a finite chain results $P(F_A^{out})$ is given by the law of total probability

$$P(F_A^{out}) = P(F_A^{out}|A_4 \text{ reacts})P(A_4 \text{ reacts}) +$$
$$P(F_A^{out}|A_4 \text{ does not react})P(A_4 \text{ does not react}) \quad (1)$$

where the first term is the conditional probability that A_4 leads to a finite chain given that A_4 has reacted and the second term is the conditional probability that A_4 leads to a finite chain given that A_4 has not reacted. For the copolymer model in this study eq 1 takes the form

$$P(F_A^{out}) = P(F_B^{in})p_A + 1(1 - p_A) \quad (2)$$

where the probability that returning from B_2 is a finite chain, $P(F_B^{in})$, is equivalent to the probability that A_4 leads to a

finite chain given that A_4 has reacted, and p_A is the fraction of A_4 monomers having reacted, or the probability that A_4 reacts. The second term in eq 2 simply states that the probability that A_4 leads to a finite chain given that A_4 has *not* reacted is unity and that the probability A_4 does not react is the fraction of A_4 monomers that do not react ($1 - p_A$). Because B_2 is a two functional monomer the probability of returning from a finite chain at B_2, $P(F_B^{in})$, must be equivalent to the probability of leading out to a finite chain at B_2, $P(F_B^{out})$, thus giving rise to eq 3.

$$P(F_B^{in}) = P(F_B^{out}) \quad (3)$$

The probability that leading out from B_2 results in a finite chain is analogous to leading out from A_4 and finding a finite chain and results in eq 4.

$$P(F_B^{out}) = P(F_A^{in})p_B + 1(1 - p_B) \quad (4)$$

For another A_4 monomer adding to the growing chain, the probability that leading in from that monomer is a finite chain is equivalent to the probability that leading out from that monomer results in a finite chain. Leading out from that monomer depends on its functionality, and if all functional groups are assumed to be equally reactive, then the probability of leading out to a finite chain will be the product of individual probabilities of leading to finite chains at each functional group, thus

$$P(F_A^{in}) = P(F_A^{out})^3 \quad (5)$$

If assumptions are made on the particular values that p_A and p_B are allowed to take, then eqs 2−5 become a system of four equations in four unknowns.

The necessary reaction steps for setting up these equations have led to a repeat unit for a fully cross-linked amine-cured epoxy copolymer that bears a connectivity like that shown in Scheme 1. To solve the system of four equations Miller and Macosko assumed two cases, one for amine excess (in which the epoxide is fully reacted) and one for epoxide excess (in which the amine is fully reacted). The stoichiometric ratio, r, of monomeric functional groups is defined as

$$r = \frac{p_B}{p_A} \quad (6)$$

For the case of amine excess, all of the epoxide groups are assumed to have reacted, thus $p_B = 1$ and $p_A = 1/r$. Substituting these values and solving the system of equations yields

$$P(F_A^{out}) = \left(r - \frac{3}{4}\right)\frac{1}{2} - \frac{1}{2} \quad (7)$$

For the case of epoxide excess, all of the amine N−H bonds are assumed to have reacted thus, $p_A = 1$ and $p_B = r$. Repeating the substitution and again solving the system of equations gives

$$P(F_A^{out}) = \left(\frac{1}{r} - \frac{3}{4}\right)\frac{1}{2} - \frac{1}{2} \quad (8)$$
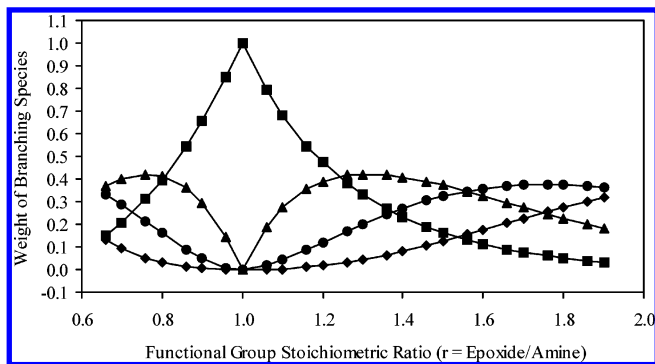
Figure 1. Plot of the probabilities that a randomly chosen monomer will have m out of f of its branches leading to the infinite polymer network as determined from eq 9. The values plotted are for $m = 4$ (■), $m = 3$ (▲), $m = 2$ (●), and $m = 1$ (◆).

These probabilities were then substituted into the binomial distribution function

$$P(X_{m,f}) = \binom{f}{m} P(F_A^{out})^{f-m}[1 - P(F_A^{out})]^m \qquad (9)$$

to arrive at the probability that an f-functional monomer (diamine) chosen at random in the infinite polymer network would have m out of f of its functional groups leading to the infinite network. For the case of amine excess ($r < 1$), eq 7 was substituted into eq 9 and for the epoxide excess case ($r > 1$), eq 8 was substituted into eq 9. For each case the resulting values obtained from the binomial distribution function were plotted in Figure 1.

The values are plotted for $f = 4$ (diamine) and m having values from 4 to 1, representing the fractional portions of structure types ranging from the fully cross-linked repeat units ($m = 4$) to those repeat units having one branch ($m = 1$) leading to the infinite polymer network. As shown in Figure 1 as the formulation deviates from being stoichiometric the portion of the polymer having four branches leading to the infinite polymer network decreases with subsequent increase in the other branched species.

In deriving a relationship for the cross-link density Miller and Macosko used the probabilities obtained from eq 9 to obtain the relative weights of "effective cross-links" of degree m. In their work the distinction is made as to what constitutes a cross-link or branching point in the infinite network. For example, a four-functional monomer is considered an "effective cross-link" if at least three out of four of its functional groups lead to the infinite polymer network. If either one or two of the functional groups lead to the infinite network, then the monomer is considered to be either a point on a dangling end of the polymer or a point between effective cross-links. We are concerned with the relative weights of all repeat unit types contributing to the glass transition temperature of the networked polymer, thus we do not differentiate between effective cross-links and other components.

Molecular Modeling and Derivation of QSPRs. In the CODESSA framework several descriptors depend explicitly upon the three-dimensional conformation of the molecule for which they are being computed. Examples of such descriptors are the principle moments of inertia and the shadow indices, both of which are computed on the basis of the principal axes of inertia. To avoid the introduction of a bias in terms of such descriptors, conformational analysis

of the polymer models was performed according to a randomization procedure using the program Cerius[2].[20] The randomization procedure was conducted by altering the rotatable bonds (single bonds that are not part of a ring or a terminal hydrogen bond) to a randomly chosen value within a user-defined window of ±65°. The models were then subjected to a maximum of 2500 steps of a Universal Force Field[21] (UFF) energy minimization. This random altering and minimization process was applied 100 times for each polymer model, without an energy cutoff criterion applied to high-energy models. The resulting lowest energy polymer model for each amine was then optimized using a parallel version of Gaussian 98[22] implementing the AM1[23] semiempirical quantum mechanical method with the correct AM1 parameters being applied to sulfur atoms.[24] Due to the large size of the polymer models, characterization of the AM1 optimized structures by means of frequency analyses could not be conducted on all models. For those models that were subjected to frequency analyses six near-zero eigenvalues were obtained. For all models the default convergence criterion were used within Gaussian 98. To compute the quantum mechanical descriptors for statistical analysis using CODESSA[25] the program AMPAC *with Graphical User Interface*[24] was used to carry out single point calculations on all preoptimized geometries. The QSPR equation was developed upon these models using the "best multilinear regression analysis" method within CODESSA following a similar procedure to that described by Katritzky.[4] The best correlation equation resulting from this analysis was then used to predict the dependence of T_g on the formulation of comonomers of bis(p-aminocyclohexyl)methane (PACM) and DGEBA.

Prediction of T_g for Nonstoichiometric Monomeric Ratios. To predict glass transition temperatures for nonstoichiometric copolymer formulations, polymer models were built to represent various degrees of branching of repeat units in the polymer network. Each of the models for prediction was built on the basis of the equations derived by Miller and Macosko. For all of the diamines in the training set, structures were built with a diamine surrounded by four monomers of DGEBA and another amine built onto one of the distal functional groups of the epoxide monomers. Such a connectivity not only bears resemblance to the structural repeat unit resulting from Miller and Macosko's recursive technique, but it also retains a 1:1 stoichiometric ratio of functional groups of comonomers. For predicting the off-stoichiometric dependence of the T_g, structural types representing different degrees of branching were built by removing, in succession, one of the three epoxides not already having an amine on each end. This resulted in four structure types shown in Figure 2.

These models were then subjected to the same randomization and optimization routine described for the training set of structures. Following the randomization and optimization, descriptors were calculated to predict the contribution of each of the structure types to the glass transition temperature value. These T_g contributions were then multiplied by the corresponding fractional contributions (from eq 9), and the resulting products were summed to arrive at a composite predicted T_g for the given stoichiometric ratios. For the epoxide excess structures, representation of various degrees

Table 1. Data for the Training Set of Amine-Cured Epoxy Polymers from the Best Four-Parameter Correlation Equation ($R^2 = 0.9977$)'

amine curing agent	lit. T_g (°C)	pred T_g (°C)	diff (°C)	ref
3,5,5-trimethylcyclohexylamine (TMCA)	84	83.6	−0.4	28
aniline	85	86.2	1.2	32
1,3-bisaminomethylcyclohexane (13BAC)	150	149.4	−0.6	33
isophorone diamine (IPD)	165	161.4	−3.6	34
4,4′-methylenebis(2,6-diethylaniline) (MDEA)	165	163.4	−1.6	35
2,2′-bis(4-(4-aminophenoxyl)phenyl)propane (BAPP)	169	170.5	1.5	36
meta-phenylene diamine (MPDA)	169.2	170.9	1.7	37
4,4′-methylenebis(3-chloro-2,6-diethylaniline) (MCDEA)	172	171.0	−1.0	29
trimethylene glycol di-*p*-aminobenzoate (TMAB)	180	183.6	3.6	38
4,4′-diaminodiphenyl methane (DDM)	183	183.0	0.0	39
bis(4-(4-aminophenoxyl)phenyl) sulfone (BAPS)	183	181.5	−1.5	30
3,3′-diaminodiphenyl sulfone (33DDS)	184	185.3	1.3	32
4,4′-diaminodiphenyl sulfone (44DDS)	221	220.3	−0.7	32
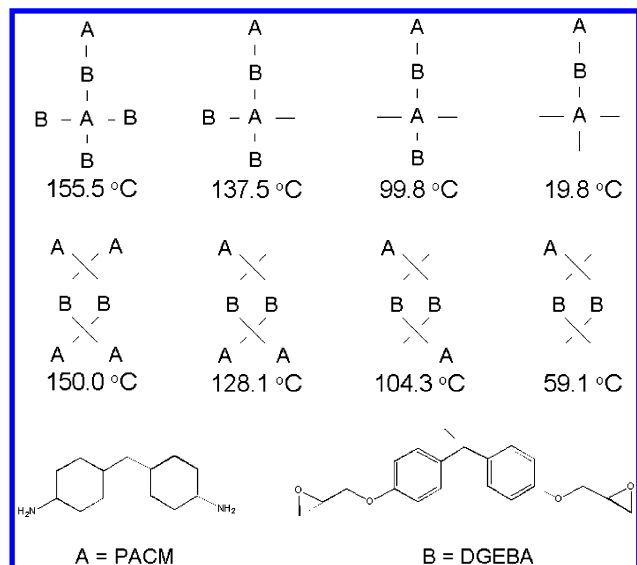
Figure 2. Branching patterns used to predict the stoichiometric dependence of T_g for PACM with predicted contributions to the polymer T_g. The top row represents the epoxy (**B**) excess structures, and the middle row represents the amine (**A**) excess structures.

of branching by the successive removal of one of the epoxide subunits from the repeat unit was intuitive. It is well-known that for polymers of this type branching is decreased for stoichiometric ratios of less than one. Representing various degrees of branching for amine excess repeat units involved a less intuitive process of trial and error. After several attempts at employing various representations of amine excess structures, the suggestion of Miller and Macosko that an entanglement of two polymer chains can behave like a cross-link was used (see Figure 2). The entanglements were built by starting with two linear trimers each consisting of an epoxide monomer with an amine monomer on each of its two ends. The two oligomers were then placed along perpendicular lines at a distance such that the Connolly surfaces of the center of the epoxide subunits were in approximate contact. The Connolly surfaces were calculated using a zero probe radius and plotted using the Cerius² graphical user interface. The distance between the center carbon atoms of the epoxides was then constrained, and the system was treated with the same randomization and molecular mechanics optimization routine described above. The system was then optimized without constraints using the AM1 method within Gaussian 98 and descriptors calculated using AM1 within AMPAC and CODESSA as described above.

Validation of T_g Predictions for Nonstoichiometric Monomeric Ratios. As with the polymers of the training set the epoxy resin was based upon DGEBA (Shell Chemical Company EPON 828, epoxy equivalent weight (EEW) = 190 g/equivalent). The DGEBA epoxy resin was cured with bis(*p*-aminocyclohexyl)methane (PACM, EEW = 52.5 g/equivalent), which was acquired from Air Products and Chemical, Incorporated. The stoichiometry of the DGEBA-PACM polymerization reaction was varied to yield epoxy/amine ratios ranging from approximately 2:1 through 0.6:1 in regular intervals, including the stoichiometric ratio of 1:1. The DGEBA-PACM mixtures were then oven cured at 160 °C for a period of 2 h. The T_g values of the reacted samples were measured using a TA Instruments 2980 Differential Scanning Calorimeter (DSC). The DSC was ramped to 225 °C at a constant heating rate of 15 °C/minute through 2 cycles to ensure complete cure and matching thermal histories of all of the samples studied. The T_g values used to validate the predictive model were then taken from the second heat.

RESULTS AND DISCUSSION

QSPR Derivation. The literature and predicted T_g values, and their differences, resulting from the correlation equation used to predict the formulation dependence of T_g of an amine-cured DGEBA epoxy resin copolymer, are listed in Table 1. Also listed in Table 1 are the abbreviations used for the names of the polymers in the training set. The coefficient of determination corresponding to these data is $R^2 = 0.9977$, and the leave-one-out cross-validated coefficient of determination is $R^2_{cv} = 0.9951$. These values indicate that there exists a high degree of correlation between the predicted and experimental values for the training set of structures and that the training set of structures is insensitive to removal of any one particular structure. Given the small size of the data set the adjusted coefficient of determination, R^2_{adj}

$$R^2_{adj} = 1 - \left(\frac{\sum_{i=1}^{N}(y_{oi} - y_{pi})^2}{\sum_{i=1}^{N}(y_{oi} - \bar{y})^2} \right) \frac{N-1}{N-M} \quad (10)$$

was calculated. In eq 10, N is both the number of experimental (y_o) T_g values from the literature and the number of predicted (y_p) values, M is the number of descriptors, and $\bar{y}$

AMINE-EPOXY COPOLYMERS

J. Chem. Inf. Comput. Sci., Vol. 44, No. 3, 2004 **917**

Table 2. Correlation Equation Parameters and Associated Significance Values Appearing in the Best Four-Parameter Correlation ($R^2 = 0.9977$)

descriptor/intercept	coefficient	t-test	$R^{2\,a}$	$R^{2\,b}$
av nucleophilic reactivity index for a C atom ($\bar{N}_C$)	-1.4314×10^5	-39.11	0.80	0.62
av bond order of an O atom (O_{BO})	-2.1660×10^3	-23.25	0.24	0.24
weighted negative charged partial surface area (*WNSA2*)	1.0567×10^{-3}	13.08	0.44	0.17
max. electron−nuclear attraction for a C−N bond (CN'_{en})	5.0311	5.26	0.43	0.55
intercept	7.8202×10^2	2.25		

a Single descriptor correlation coefficient. *b* Correlation coefficient for correlation with all other descriptors.

is the mean of experimental values from the literature. The value of the adjusted coefficient of determination is $R^2_{adj} = 0.9970$ indicating a high degree of correlation for the training set of polymers even after applying a penalty for the relative number of descriptors and structures used to derive the correlation equation. The correlation is also characterized by a Fisher value of $F = 864.0$; this value exceeds the 99.9% confidence level[26,27] indicating a correlation that accounts well for the variance in the experimental data.

The parameters appearing in the correlation equation, including the coefficients to the descriptors and the intercept, and coefficients of determination associated with the individual descriptors are listed in Table 2. The values of the coefficients of determination for each descriptor indicate that each descriptor is moderately to highly predictive of T_g for the study set of monomers. The values of the coefficients of determination for each descriptor correlated with the set of remaining descriptors in the equation indicate a moderate to low degree of collinearity.

The most significant descriptor appearing in the correlation equation, according to the t-test, was the average Fukui nucleophilic reactivity index for a carbon atom ($\bar{N}_C$).[28] This descriptor is computed according to the following equation

$$\bar{N}_C = \frac{1}{k} \sum_{j=1}^{k} \sum_{i \in C}^{n} \frac{c_{ij}^2}{(1 - E_{HOMO})} \tag{11}$$

In eq 11, c_{ij} is the ith coefficient to an atomic orbital centered on the jth carbon atom, which contributes to the highest occupied molecular orbital, where there are n such atomic orbitals and k such carbon atoms. Also in eq 11, E_{HOMO} is the energy of the highest occupied molecular orbital. The original intent of the Fukui nucleophilic reactivity indices was to describe the tendency of a molecule to react as a nucleophile. This descriptor depends on the magnitude of the energy of the highest occupied molecular orbital and the extent to which that orbital is localized on the carbon atoms of the polymer model. Analysis of E_{HOMO} with respect to T_g resulted in a coefficient of determination of $R^2 = 0.2272$. This suggests that most of the correlation of this descriptor is contained within the atomic orbital coefficients. Multiplication of $\bar{N}_C$ by the quantity, $1-E_{HOMO}$, appearing in the denominator of eq 11 and analyzing the product with respect to T_g results in a coefficient of determination of $R^2 = 0.7914$. The difference between these two correlations corroborates the idea that the correlation lies in the atomic orbital coefficients. Further, the coefficients appearing in the equations resulting from the regressions of $\bar{N}_C$ and the described product with T_g both bear a negative sign. This indicates that the HOMO is delocalized from the carbon atoms to the heteroatoms with increasing T_g values. The shift in contribu-

tion to the HOMO from the carbon atomic orbitals to the heteroatoms suggests that the rotable bonds involving these atoms are made more rigid by delocalizing the HOMO among them, resulting in stabilization of the polymer backbone and increased T_g values. Thus, we propose that $\bar{N}_C$ is a measure of the rigidity of the bonds along the backbone of the polymers of the training set.

By the t-test the second most significant descriptor appearing in Table 2 is the average bond order of an oxygen atom (O_{BO}). The AM1 bond orders were calculated within AMPAC from a Coulson population analysis of the density matrix elements. For this descriptor the matrix elements correspond to the oxygen atoms in the polymer model. According to Table 2 the coefficient of determination for this descriptor in a single parameter equation is $R^2 = 0.24$. While all polymer models contain oxygen atoms, analysis of Table 1 shows that the three largest T_g values correspond to amines that contain sulfone functional groups that have two doubly bonded oxygen atoms, thus raising the average bond order for these models. We believe that the high significance indicated by the t-test coupled with the low predictive quality indicated by the single parameter coefficient of determination is a consequence of the small size and composition of the training set of polymers. This descriptor suggests a donor−acceptor property for the models containing the sulfone groups as well as for TMAB, which contains a 4-aminobenzoate moiety. The donor−acceptor effect for these polymers would arise from the sulfone and carbonyl functionalities drawing electron density from the amino group by induction. The transfer in electron density is similar in effect to the delocalization described above. Here again, the increase in electron density around the rotable bonds neighboring the aromatic rings would strengthen these bonds, consequently raising the amount of energy required for rotation about the bond, thus increasing the T_g for these polymers. Further support for this interpretation can be found in the difference between T_g values for 33DDS and 44DDS listed in Table 1. Although the described inductive effects could contribute to a higher value for the T_g of 33DDS, through some degree of p-π delocalization, the amino groups for this diamine are in a *meta* position with respect to the sulfone, and this delocalization is probably not nearly as strong as that in 44DDS. The *para* orientation for the functional groups exhibited in 44DDS could lead to further stabilization through more delocalization of the nitrogen lone pair into the aromatic ring, resulting in an even higher T_g value for this copolymer.

The next most significant descriptor appearing in Table 2 is the weighted negative charged partial surface area (*WNSA2*).[29] This descriptor is calculated according to the following formula

$$WNSA2 = \frac{(PNSA2)\,TMSA}{1000} \qquad (12)$$

where *TMSA* is the total molecular surface area and *PNSA2* is the negative charged partial surface area of the polymer model and is calculated according to the following equation

$$PNSA2 = (\sum_i SA_i^-)\sum_i Q_i \qquad (13)$$

where SA_i^- is the nonoccluded surface area of the *i*th atom bearing a negative charge from a Coulson population analysis and Q_i^- is the value of the negative charge centered on the atom. The surface area of this descriptor is computed as the nonoccluded solvent accessible surface area after increasing the van der Waals radii of the atoms in the molecule by 1.5 Å. This descriptor belongs to a class of Charged Partial Surface Area (CPSA) descriptors that were created by Stanton and Jurs to describe the polar interactions between molecules that give rise to their physical and biological properties. Within the context of T_g *WNSA2* is indicative of associative effects between polymer chains. The Pearson correlation coefficient for this descriptor is $r = -0.6644$ demonstrating a moderately strong negative correlation with T_g. The strong negative correlation indicates an inversely proportional relationship suggesting that localization of negative charge onto a smaller surface area of the polymers results in more polar interactions between the polymer chains for the higher T_g polymers. The presence of a bond order descriptor in the derived QSAR equation also suggests that this descriptor is a measure of increased charge localization resulting in shorter stronger bonds within the polymers exhibiting higher T_g values. Evidence supporting this reasoning can be found in Table 3, in which are listed the descriptors appearing in five equations obtained by a heuristic[12] analysis of all descriptors. These equations are characterized as being the five most correlated to T_g (within the heuristic method) based first on the F statistic and then on the coefficient of determination. The minimum bond order for a nitrogen atom appears in three out of five of the equations as do other related descriptors.

The last descriptor appearing in the correlation equation is the maximum electron–nuclear attraction for a carbon–nitrogen bond. This descriptor is computed within the constructs of the NDDO approximation of semiempirical molecular orbital theory according to the following equation

$$CN'_{en} = \sum_{\mu,\nu \in C} P_{\mu,\nu} \left\langle \phi_\mu \left| \frac{Z_N}{R_{i,N}} \right| \phi_\nu \right\rangle \qquad (14)$$

where $P_{\mu,\nu}$ represents the density matrix elements corresponding to the carbon atom and $\langle \phi_\mu | Z_N/R_{i,N} | \phi_\nu \rangle$ is the electron–nuclear attraction integral for the *i*th valence electron centered on the carbon atom and interacting with the nitrogen atom, which bears a charge of Z_N. This descriptor is related to the strength of the carbon–nitrogen bonds in the backbone of the polymer models. The lack of significance of this descriptor as evidenced by the *t*-test is likely due to the location of the bonds corresponding to this descriptor. Visualization of this descriptor within CODESSA by plotting it upon the corresponding bonds of the polymer models indicates that these bonds are at the terminal end of the

Table 3. Descriptors Appearing in Equations Having the Highest Coefficients of Determination and *F* Values Following a Heuristic Analysis

| descriptor/intercept	*t*-test	$R^{2\,a}$	$R^{2\,b}$
$R^2 = 0.9942\ F = 344.9$			
number of N atoms	28.85	0.82	0.13
min. bond order (>0.1) of an N atom	−10.10	0.35	0.60
av valency of an O atom	−5.54	0.26	0.40
relative number of O atoms	−4.51	0.28	0.62
intercept	12.75		
$R^2 = 0.9925\ F = 264.4$			
max. total interaction for an N−H bond	9.66	0.80	0.69
max. e−e repulsion for an N atom	13.47	0.73	0.41
min. e−n attraction for a C−C bond	5.16	0.11	0.32
max. 1-electron reactivity index for a C atom	−4.33	0.38	0.58
intercept	−6.51		
$R^2 = 0.9910\ F = 220.0$			
max. total interaction for an N−H bond	8.29	0.80	0.71
max. e−e repulsion for an N atom	12.05	0.73	0.50
min. nucleophilic reactivity index for a C atom	−3.78	0.22	0.50
min. e−n attraction for a C−C bond	3.67	0.11	0.26
intercept	−5.00		
$R^2 = 0.9908\ F = 215.9$			
number of N atoms	21.82	0.82	0.24
min. bond order (>0.1) of an N atom	−7.71	0.35	0.37
av valency of an O atom	−4.44	0.26	0.43
total dipole of the molecule	−3.13	0.27	0.42
intercept	11.90		
$R^2 = 0.9905\ F = 207.5$			
number of N atoms	20.83	0.82	0.23
min. bond order (>0.1) of an N atom	−10.24	0.35	0.20
max. electrophilic reactivity index for a C atom	−3.11	0.25	0.18
min. e−n attraction for a C−C bond	3.08	0.11	0.22
intercept	−2.79		

a Single descriptor correlation coefficient. *b* Correlation coefficient for correlation with all other descriptors.

terminal diamine (see Figure 2). Although the descriptor is indicative of bond strength it is difficult to draw a further conclusion since the presence of the terminal amine is a consequence of the chosen connectivity of the polymer models in the training set.

External Validation of QSPR. We believe that the high degree of correlation exhibited by the leave-one-out cross-validated and adjusted coefficients of determination suggest that the correlation equation that we obtained should be applicable to polymers outside the set of polymers serving as a basis for the correlation. However, the true test of the extensibility of any such model is comparison to experimental results. To test the extensibility of the derived QSPR as well as our model for predicting the formulation dependence of T_g we carried out predictions of T_g and compared these values with those from our DSC determinations for the PACM/DGEBA formulations described above. As described, the predictions involved building polymer models exhibiting various degrees of cross-linking. Once these models were built and optimized, descriptors were computed for each of them, and then the derived QSPR was used to predict their individual contribution to the bulk glass transition temperature for the polymer (see Figure 2). This is, in effect, a glass transition temperature prediction for each of the fragment models corresponding to a different degree of cross-linking. These predicted values were then multiplied by probabilistic weighting factors, obtained by the method of Miller and Macosko, that corresponded to the levels of cross-linking that the models represent. The resulting prod-

AMINE-EPOXY COPOLYMERS

J. Chem. Inf. Comput. Sci., Vol. 44, No. 3, 2004 **919**

Table 4. Summary of the Calculation of a Composite Predicted T_g for a Stoichiometric Ratio $r = 0.833$ for PACM/DGEBA

$P(X_{m,f})$	weighting factor	pred T_g (°C)	product (°C)
$P(X_{1,4})$	0.01653	59.1	0.9769
$P(X_{2,4})$	0.1204	104.3	12.5577
$P(X_{3,4})$	0.3895	128.1	49.8950
$P(X_{4,4})$	0.4727	150.0	70.9050
		composite T_g	134.3
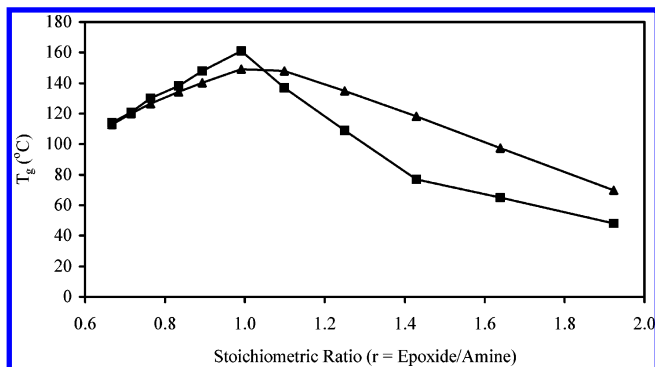
Figure 3. Predicted (▲) and experimental (■) T_g values for PACM cured DGEBA with varied stoichiometric ratios of monomers.

ucts were then summed to arrive at the composite predicted T_g values for each stoichiometric ratio. For example, the weighting factors that correspond to the amine excess stoichiometry $r = 0.833$ are given in Table 4. Also shown in Table 4 are the corresponding predicted glass transition temperature contributions from Figure 2. To arrive at the composite predicted T_g for $r = 0.833$ one simply multiplies across the rows of Table 4 and then takes the sum of the last column. This procedure can be repeated for any similar polymer provided that predicted values for the fragment contribution to the overall T_g are obtained. The results for all experimental and calculated stoichiometric ratios of PACM/DGEBA are plotted in Figure 3.

As shown in Figure 3, the trend in T_g vs stoichiometric ratio is accurately reproduced with a maximum corresponding to a 1:1 stoichiometry of functional groups. These results compare well with those reported by Williams[30] whose work involved predicting T_g values for nonstoichiometric epoxy-amine polymer formulations. In attempting to describe the influence of stoichiometry on T_g Williams used the Nielsen[31] equation

$$T_g = T_{g0} + 3.9 \times 10^4 \nu \qquad (15)$$

where ν is the concentration of elastic chains and T_{g0} is the glass transition temperature of the un-cross-linked polymer. To arrive at the concentration of elastic chains Williams used the method of Miller and Macosko.[16] The value for T_{g0} was obtained by using a value for ν in the Miller and Macosko analysis that corresponds to a 1:1 stoichiometric ratio and a corresponding experimental value for T_g. The difference between the work of Williams and that of our own is that Williams used a T_g value for a given polymer to predict the formulation dependence of T_g for that *same* polymer. In contrast, our approach uses T_g values from similar polymers to arrive at an equation that can be used to predict formulation dependent T_g values for *any* similar polymer. While our results exhibit a greater degree of error than those of Williams they are qualitatively identical. The greater error

in our predictions with respect to experimental results is likely due to the fact that they do not directly involve corrections to experimental values. Rather, our model involves only computed quantities and as such should be expected to exhibit greater error.

CONCLUSIONS

We have derived a designer QSPR capable of predicting the glass transition temperature of amine-cured resins of bisphenol A diglycidyl ether. The descriptors appearing in the correlation equation lend themselves to intuitive interpretation based on the molecular level events that occur when a polymer exhibits a change from a glassy state to a rubbery state. For the study set of polymers, the T_g values appear to be reflective of the strengths of bonds along the backbone of the polymer. The strongly predictive nature of the quantum mechanically derived descriptors presented here is a testimony to the role that semiempirical methods play in studying such large molecular systems. The extent to which our results uniquely depend on the AM1 semiempirical method is not apparent in our results. Descriptors such as the maximum electron−nuclear attraction energy for a carbon−nitrogen bond would not have appeared in methods that predate NDDO methods. However, as suggested by the equations listed in Table 3, it is possible that other descriptors would have substituted, giving equations with lower degrees of correlation. Nevertheless, we feel that the good agreement in the trend between our predictions and the experimental T_g values is reflective of the utility of the AM1 method in predicting experimental trends.

Using an ad hoc probabilistic treatment based on the work of Miller and Macosko we have developed a model that accurately predicts the dependence of T_g on the stoichiometric ratio of comonomers in cross-linked copolymers of amine-cured resins of bisphenol A diglycidyl ether. This is the first such model to do so without requiring an experimental value for the polymer of interest and although our model does not exhibit high quantitative accuracy, our model qualitatively predicts the correct trend. The UFF molecular mechanics method is an all-periodic table method developed for nearly all atom types, and the AM1 semiempirical quantum mechanical method has been parametrized for the elements that comprise nearly all polymers in common use. Further, the treatment of branched chain polymers of Miller and Macosko suggests that a similar methodology would be suitable to other cross-linked polymers. Because of the general utility of the computational methods we anticipate that our approach may be used to generate predictive models for other cross-linked polymers provided that suitable experimental data exists.

REFERENCES AND NOTES

(1) Rosen, S. L. *Fundamental Principles of Polymeric Materials*; John Wiley & Sons: New York, 1993.

(2) Andrews, R. J.; Grulke, E. A. *Glass Transition Temperatures of Polymers*. In *Polymer Handbook*, 4th ed.; Brandrup, J., Immergut, E. H., Grulke, E. A., Eds.; John Wiley & Sons: New York, 1999.

(3) van Krevelen, D. W. *Properties of Polymers Their Estimation and Correlation with Chemical Structure*; Elsevier: Amsterdam, 1976.

(4) Katritzky, A. R.; Sild, S.; Lobanov, V.; Karelson, M. Quantitative Structure−Property Relationship Study of Glass Transition Temperatures of High Molecular Weight Polymers. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 8(2), 300−304.

(5) Bicerano, J. *Prediction of Polymer Properties*; Marcel Dekker: New York, 1996.

(6) Hansch, C.; Fujita, T. In *Classical and Three-Dimensional QSAR in Agrochemistry*; Hansch, C., Fujita, T., Eds.; ACS Symposium Series 606; American Chemical Society: Washington, DC, 1994; pp 1−12.

(7) Vaz, R. J. A QSPR for the T_g of polymers: the Koehler-Hopfinger approach using the Tripos 5.2 Force Field. *Makromol. Chem. Makromol. Symp.* **1993**, *65*, 261−268.

(8) Hamerton, I.; Howlin, B. J.; Larwood, V. Development of Quantitative Structure−Property Relationships for Poly(arylene ether)s. *J. Mol. Graph.* **1995**, *13*, 14−17.

(9) Cypcar, C. C.; Camelio, P.; Lazzeri, V.; Mathias, L. J.; Waegell, B. Prediction of the Glass Transition Temperature of Multicyclic and Bulky Substituted Acrylate and Methacrylate Polymers Using the Energy, Volume, Mass (EVM) QSPR Model. *Macromolecules* **1996**, *29*, 8954−8959.

(10) Koehler, M. G.; Hopfinger, A. J. Molecular modeling of polymers: 5. Inclusion of intermolecular energetics in estimating glass and crystal-melt transition temperatures. *Polymer* **1989**, *30*, 116−126.

(11) Camelio, P.; Cypcar, C. C.; Lazzeri, V.; Waegell, B. A novel approach toward the prediction of the glass transition temperature: Application of the EVM model, a designer QSPR equation for the prediction of acrylate and methacrylate polymers. *J. Polym. Sci. Polym. Chem.* **1997**, *35*, 2579−2590.

(12) Katritzky, A. R.; Lobanov, V. S.; Karelson, M.; Murugan, R.; Grendze, M. P.; Toomey, J. E. Comprehensive Descriptors for Structural and Statistical Analysis. 1. Correlations between structure and physical properties of substituted pyridines. *Rev. Roum. Chim.* **1996**, *41*(11−12), 851−867.

(13) Cao, C.; Lin, Y. Correlation Between the Glass Transition Temperature and Repeating Unit Structure for High Molecular Weight Polymers. *J. Chem. Inf. Comput.* **2003**, *43*, 643−650.

(14) Won, Y.; Galy, J.; Pascault, J.; Verdu, J. Prediction of the Glass Transition Temperature of Cycloaliphatic Amine-Epoxy Networks. *J. Polym. Sci. Polym. Phys.* **1991**, *29*, 981−987.

(15) Bicerano, J.; Sammler, R. L.; Carriere, C. J.; Seitz, J. T. Correlation between Glass Transition Temperature and Chain Structure for Randomly Cross-linked High Polymers. *J. Polym. Sci. Polym. Phys.* **1996**, *34*, 2247−2259.

(16) Kozlov, G. V.; Beloshenko, V. A.; Varyukhin, V. N.; Lipatov, Y. S. Application of Cluster Model for the Description of Epoxy Polymer Structure and Properties. *Polymer* **1999**, *40*, 1045−1501.

(17) Miller, D. R.; Macosko, C. W. A New Derivation of Post Gel Properties of Network Polymers. *Macromolecules* **1976**, *9*(2), 206−211.

(18) Lee, H.; Neville, K. *Handbook of Epoxy Resins*; McGraw-Hill: New York, 1967.

(19) Flory, P. J. *Principles of Polymer Chemistry*; Cornell University, Ithaca, NY, 1953.

(20) Cerius², Accelrys, Inc., 9685 Scranton Rd., San Diego CA, 92121-3752.

(21) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, A Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114*(25), 10024−10035.

(22) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A4; Gaussian, Inc.: Pittsburgh, PA, 1998.

(23) Dewar, M. J. S.; Zoebisch, E. G.; Healey, E. F.; Stewart, J. J. P. AM1: A New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*(13), 3902−3909.

(24) Dewar, M. J. S.; Yuan, Y. C. AM1 Parameters for Sulfur. *J. Am. Chem. Soc.* **1990**, *29*(19), 3881−3890.

(25) Semichem Inc. Box 1649, Shawnee, KS 66222.

(26) The 99.9% confidence level for 4 degrees of freedom in the numerator and 8 degrees of freedom in the denominator is 14.39 according to the CRC Handbook of Chemistry and Physics.

(27) *CRC Handbook of Chemistry and Physics*, 79th ed.; Lide, D. R., Ed.; Chemical Rubber: Boca Raton, FL, 1998.

(28) Fukui, K. *Theory of Orientation and Stereoselection*; Springer-Verlag: Berlin, 1975.

(29) Stanton, D. T.; Jurs, P. C. Development and use of charged partial surface area structural descriptors in computer assisted quantitative structure−property relationship studies. *Anal. Chem.* **1990**, *62*, 2323−2329.

(30) Vallo, C. I.; Frontini, P. M.; Williams, R. J. J. The glass transition temperature of nonstoichiometric epoxy-amine polymer networks. *J. Polym. Sci. Part B: Polym. Phys.* **1991**, *29*, 1503−1511.

(31) Nielsen, L. E. Cross-linking effect on physical properties of polymers. *J. Macromol. Sci.-Rev. Macromol. Chem.* **1969**, *C3*, 69−103.

(32) Wise, C. W.; Cook, W. D.; Goodwin, A. A. Chemico-diffusion kinetics of model epoxy-amine resins. *Polymer* **1997**, *38*, 3251−3261.

(33) Barral, L.; Cano, J.; Lopez, A. J.; Nogueira, P.; Ramirez, C. Isothermal cure of an epoxy/Cycloaliphatic amine system. Vitrification and gelation. *Polymer. Int.* **1995**, *38*, 353−356.

(34) Won, Y. G.; Galy, J.; Pascault, J. P.; Verdu, J. Prediction of the glass transition temperature of cycloaliphatic amine-epoxy networks. *J. Appl. Polym. Sci. Part B: Polym. Phys.* **1991**, *29*, 981−987.

(35) Eloundou, J. P.; Ayina, O.; Nga, H. N.; Gerard, J. F.; Pascault, J. P.; Boiteux, G.; Seytre, G. Simultaneous kinetic and microdielectric studies of some epoxy-amine systems. *J. Polym. Sci. Part B: Polym. Phys.* **1998**, *36*, 2911−2921.

(36) Grillet, A. C.; Galy, J.; Gerard, J. F.; Pascault, J. P. Mechanical and viscoelastic properties of epoxy networks cured with aromatic diamines. *Polymer* **1991**, *32*, 1885−1891.

(37) Wei, J.; Hawley, M. C.; Demeuse, M. T. Kinetics modeling and time−temperature-transformation diagram of microwave and thermal cure of epoxy resins. *Polym. Eng. Sci.* **1995**, *35*, 461−470.

(38) Wang, X.; Gillham, J. K. Competitive primary amine/epoxy and secondary amine/epoxy reactions: effect on the isothermal time to vitrify. *J. Appl. Polym. Sci.* **1991**, *43*, 2267−2277.

(39) Zukas, W. X. Torsional braid analysis of the aromatic amine cure of epoxy resins. *J. Appl. Polym. Sci.* **1994**, *53*, 429−440.