COMPUTER AIDED POLYMER DESIGN USING MULTI-SCALE MODELLING

K. C. Satyanarayana[1], J. Abildskov[1], R. Gani[1*], G. Tsolou[2] and V. G. Mavrantzas[2]

[1]CAPEC, Department of Chemical and Biochemical Engineering, Phone: +45 4525 2882, Fax +45 45932906,
Technical University of Denmark, DK-2800 Lyngby, Denmark.
E-mail rag@kt.dtu.dk
[2]Department of Chemical Engineering, University of Patras & Institute of Chemical Engineering
and High-Temperature Chemical Processes (FORTH-ICE/HT), Patras GR 26504, Greece.

Abstract - The ability to predict the key physical and chemical properties of polymeric materials from their repeat-unit structure and chain-length architecture prior to synthesis is of great value for the design of polymer-based chemical products, with new functionalities and improved performance. Computer aided molecular design (CAMD) methods can expedite the design process by establishing input-output relations between the type and number of functional groups in a polymer repeat unit and the desired macroscopic properties. A multi-scale model-based approach that combines a CAMD technique based on group contribution^{plus} models for predicting polymer repeat unit properties with atomistic simulations for providing first-principles arrangements of the repeat units and for predictions of physical properties of the chosen candidate polymer structures, has been developed and tested for design of polymers with desired properties. A case study is used to highlight the main features of this multi-scale model-based approach for the design of a polymer-based product.
Keywords: CAMD; Polymer Design; Molecular dynamics; Property prediction; Group contribution methods; Polyisobutylene; Diffusivity; Multi-Scale Modelling.

INTRODUCTION

The search for new polymeric compounds with desired end-user properties is an everyday problem in chemical, pharmaceutical and food industries. Traditional molecular design approaches relying on a series of experiments for synthesizing each product candidate and evaluating its performance are costly, highly time consuming, tedious and capital intensive, and demand considerable resources. On the other hand, the structures of polymers suitable for applications in various chemicals-based products have increased in complexity, further complicating the design problem. Consequently, predictive computational methods (see Figure 1) capable of evaluating candidates for specific applications have received considerable attention in the last years. They can expedite the design process by establishing a link to the microstructure, thereby developing quantitative structure-property relationships that can facilitate the development of new products with a set of tailored pre-specified properties or functions. The objective of this paper is to present an extended CAMD technique for polymer design and to illustrate its use through a case study involving resolution of design issues on macro-, meso- and micro-scales.

Compared to the macroscopic property-based approaches, the computer-aided multi-scale model-based approaches to the polymer design problem are unique in their own way, but do have limitations in their applications. The modelling techniques used, for example, are computationally much more demanding than the correlations used in macroscopic property-based approaches. In this paper, however, the two approaches are integrated into a two-stage design algorithm. For example, in the extended

Computer Aided Molecular Design (CAMD) technique applied to polymer repeat-unit design, the macro-scale property constraints are evaluated using a set of pre-selected predictive property models based on functional groups (van Krevelen, 1990; Sataynarayana et al., 2007), while, atom-connectivities are used for predicting missing group contributions and/or properties (Bicerano, 2002; Satayanarayana et al., 2007) at the meso-scale. Molecular descriptors are then used to evaluate the polymer chain on the micro-scale (Tsolou et al., 2008). This leads to a multi-scale model-based design approach where feasible polymer repeat-units are identified at the group and atom-connectivity scales while the polymer chain length and structure are established on the micro-scale.

For organic chemicals, group contribution based property prediction models (Marrero and Gani, 2001) are sufficiently simple and reliable for use in CAMD, but due to the limited availability of experimental data, there are large gaps in the group parameter tables for polymer properties. At times there can also be a polymer repeat-unit structure that cannot be totally represented by groups when predicting its properties with a specific group contribution method. This is overcome by a meso-scale approach, where an atom-connectivity index method (Gani et al., 2005) is applied to determine the missing groups and predict their contributions automatically without the need for additional experimental data. This integration of the group and atom-connectivity index based models is called the group contribution^{plus} (GC$^+$) model, as proposed by Satyanarayana et al. (2007). Since the atom-connectivity index based property prediction models

have comparatively lower accuracy than the group contribution models, they would be used mainly to fill out the gaps in the parameter tables of the group contribution models.

Bulk polymer models based on the identified basic repeat-units are used as input in molecular simulation algorithms (micro-scale) to identify a polymer chain with the desired number of repeat-units and then to either predict or validate the required set of properties (since some properties can vary with polymer chain length or number of repeat-units). The accuracy of the results (e.g., thermodynamic, structural, conformational, etc.) obtained from molecular simulations depends mainly on the efficiency of the method used to carry out the simulations and the reliability of the force-field employed for the description of the intra- and inter-atomic interactions. Examples of polymers typically addressed with such a method include polystyrene (Spyriouni et al., 2007; Harmandaris et al., 2006; Milano and Müller-Plathe, 2005), polyethylene terephthalate (Kamio et al., 2007), polycarbonates (Abrams and Kremer, 2003), polyphenylene dendrimers (Carbone et al., 2007), and also DNA (Knotts et al., 2007). In this paper, the method developed by Tsolou et al. (2008) is used.

This paper presents the new extended multi-scale model based CAMD approach for polymer design and highlights the important issues, in particular, the micro-scale analysis of polymer repeat-units identified through the macro-meso scale analysis involving the design of a polymeric hermetic seal. First, however, a brief background of the algorithms and property prediction methods associated with the multi-scale model-based CAMD approach is given.
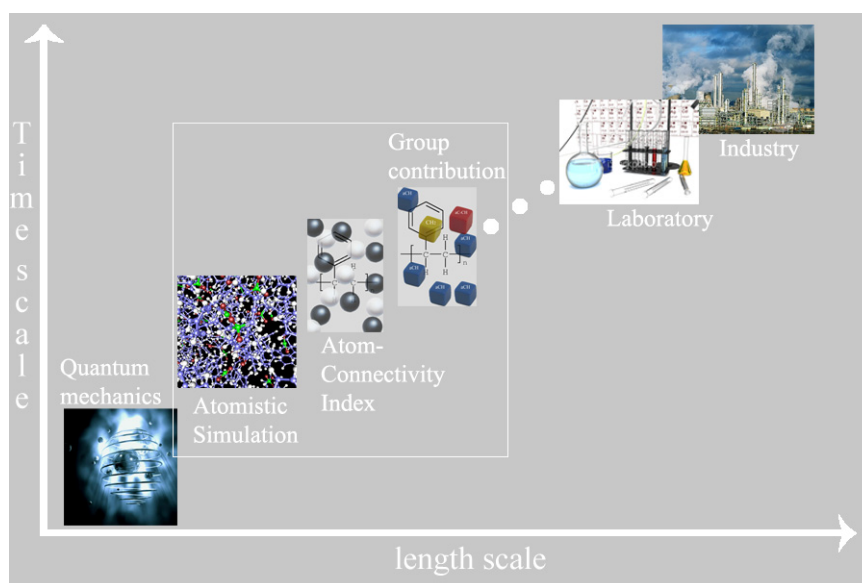

Figure 1: Multi scales for developing a new polymer product on the basis of the required property.

BACKGROUND

The Group Contribution^plus (GC^+) Model

This approach for property prediction resulted from the integration of the group contribution and the atom-connectivity index methods developed by Gani et al. (2005), with emphasis on properties of organic chemicals. They have been further extended to predict homo-polymer repeat unit properties using the Marrero-Gani group contribution method and the atom-connectivity index method (Sataynarayana et al. (2007)). The Marrero-Gani group contribution method (Marrero and Gani (2001)) has advantages over other group contribution methods, since a larger range of groups, classified as first-, second- and third-order, are available and cover a wider range of organic chemicals. The two main requirements of any group contribution method are:

1. Ability to represent the polymer repeat-unit structure with the groups available in the selected property prediction method.

2. Ability to estimate the necessary property by using the property contributions of the groups used to represent the polymer structures.

However, since the number of reliable data on polymer systems is limited, there are large gaps in the group contribution parameter tables and, consequently, for many polymer repeat-unit structures, their properties cannot be predicted. This is the case, for example, with the polymer repeat-unit structure of poly-bismethoxyphosphazene (https://polymer.nims.go.jp/PoLyInfo/), for which the Marrero-Gani group contribution method (Marrero and Gani, 2001) cannot fully represent its structure (see Figure 2). Such a method alone is therefore not widely applicable because the need of missing group contributions cannot be totally ruled out. To retain the advantages of the Marrero-Gani group contribution methods and to overcome the limitation of their missing contributions, this group contribution method is integrated with the atom-connectivity index method, and the combined model is called the GC^+ model.

The Marrero-Gani Group Contribution Method

The general expression for all the single value pure component property models based on the Marrero-Gani group contribution method (Marrero and Gani, 2001) is given by:

$$f(X) = \sum_i N_i C_i + w \sum_j M_j D_j + z \sum_k O_k E_k \qquad (1)$$

where, C_i is the contribution of the first-order group of type-i that occurs N_i times, D_j and E_k are the contributions of the second-order group of type-j and of the third-order group of type-k that occur M_j and O_k times, respectively. In the first level of estimation (where $w = z = 0$), only first-order groups are allowed; in the second level (where $w = 1$ and $z = 0$), first- and second-order groups are involved; and in the third level (where $w = 1$ and $z = 1$), contributions of groups of all levels are allowed in the calculation. The left-hand side of Eq. 1 is a simple function $f(X)$ of the property X. When selecting this function one tries to satisfy as closely as possible the following criteria: (1) The function has to achieve additivity in the contributions C_i, D_j and E_k (thereby, for example, providing a linear programming problem formulation for CAMD); (2) The function has to enable the best possible fit of the experimental data (thereby providing the necessary accuracy); and (3) The function should provide good extrapolation capability and, therefore, a wide applicability (thereby providing the means for obtaining reliable and innovative solutions). The procedure for property prediction is the same as in the case of organic chemicals except that here the same set of groups describes the polymer repeat-unit structures, i.e., structures having two free attachments. The list of polymer repeat-unit properties modelled through this method is given in Satyanarayana et al. (2007).
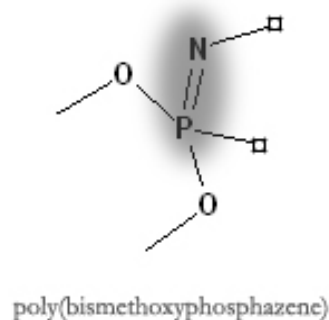

poly(bismethoxyphosphazene)

Figure 2: Incomplete molecular structure representation – the Marrero-Gani method (2001) does not describe the shaded part. The small squares in the above structure denote the free ends of the polymer repeat unit structure.

The Atom-Connectivity Index Method

In this case, the polymer repeat-unit structures are represented by atoms and atom-connectivity indices, as proposed by Gani et al. (2005):

$$f(Y) = \sum_i a_i A_i + b\left(^v x^0\right) + 2c\left(^v x^1\right) + d \qquad (2)$$

where, Y is the sought polymer property, A_i is the contribution of atom i occurring in the molecular structure a_i times, $^v x^0$ and $^v x^1$ the zero- (atom) and first-order (bond) connectivity indices, respectively, as described by Kier and Hall (1986), and b, c and d adjustable parameters. The contributions A_i are also treated as adjustable parameters. Development of the atom-connectivity based models for predicting polymer properties is explained in detail by Satyanarayana et al. (2007).

The CAMD Technique for Polymer Repeat-Unit Design

The CAMD technique (macro-meso scale) used in this work (Sataynarayana et al., 2009) involves three main steps: problem formulation, repeat unit identification and result analysis, similar to the algorithms proposed by Harper and Gani (2000) and Gani et al. (1991). Note that many other CAMD techniques proposed by others (Maranas, 1996, Venkatasubramanian et al., 1994) could, in principle, have also been employed.

Problem Formulation

The design problem here refers to identifying a set of polymer repeat-units with their corresponding chain length size and distribution possessing a specified set of properties. The important variables are the building blocks (groups) that are used to generate and represent the polymer repeat-unit structures, as well as to predict their properties. There are two types of constraints: structural constraints and property constraints. The structural constraints are related to the size and chemical feasibility of the generated structures (e.g., for a representation by groups: how many groups, the maximum and minimum numbers that can be present, how many times a particular group can occur in the repeat-unit, and the rules for combining them). Moreover, any repeat-unit structure that is generated should also satisfy certain structural conditions, such as the condition that there must be two free attachments in a polymer repeat-unit structure.

The property constraints are those related to the needs for the desired product, expressed in terms of target property values. Mathematically, these constraints are written as $P_l \leq P(n_i) \leq P_u$, where n_i is the number of times the i-th group appears in the polymer; $P(n_i)$ the vector of predicted properties for the polymer; and P_l and P_u the vectors of lower and upper bounds, respectively, specified on the selected target polymer properties.

Repeat-Unit Identification and Screening (Macro- and Meso-Scales)

The employed CAMD technique is of the "generate-and-test" type, where all feasible polymer repeat-unit structures are generated from a set of building blocks and subsequently tested against the property constraints. The adopted multi-level approach of Harper and Gani (2000) has been employed (Sataynarayana et al., 2009). For the polymer repeat-unit design problem, which uses the GC^+ model (Sataynarayana et al., 2007) for predicting the properties of the generated repeat-unit structures, polymer repeat-units are generated in successive hierarchical steps and are screened against a corresponding set of property constraints. The level of structural details increases as the number of promising candidates decreases in each subsequent step.

Result Analysis

If necessary, the screened repeat units satisfying the property constraints are further analyzed using data from literature sources or generated from other more accurate, but non-group contribution (off-line) property prediction methods. Here, a final selection of a set of feasible polymer repeat-units is made.

Micro-Scale Computations

Having identified the basic polymer repeat unit from the CAMD technique does not solve the design problem since the complete information on the structural and conformational properties of the polymer chain is still missing at this stage. This information can be obtained from direct molecular simulations capable of providing the physical properties of the corresponding polymeric material as a function of size (number of repeat units) and operational variables such as the temperature and the pressure. Molecular dynamics (MD) and Monte Carlo (MC) are two such methods capable of providing exact results for statistical mechanics problems (for a given molecular model) in preference to approximate solutions. Monte Carlo, in particular, has developed into a powerful tool for simulating the properties of complex systems such as chain molecules, because of its capability to

accelerate system equilibration through the implementation of large or unphysical moves that do not require the system to follow the natural trajectory. Molecular Dynamics, on the other hand, is attractive because of its simplicity (it relies on the solution of Newton's equations of motion for each atom in the system) and its ability to provide real-time information. With the help of coarse-graining methodologies, in particular, and the rapid increase of computing power, the two methods have developed into state-of-the art tools for capturing the subtleties of structure-property relations relevant to present-day polymer technology and the prediction of key material properties starting from the detailed consideration of chemical constitution and molecular architecture. The method used in this work has been reported by Tsolou et al. (2008) and the application details are given in the case-study.

MULTISCALE MODEL BASED APPROACH FOR POLYMER DESIGN

The schematic work-flow of the proposed multilevel model-based approach for design of polymers with the required target properties is shown in Figure 3. Note that the design problem is divided into two stages – first the polymer repeat unit structures satisfying a sub-set of (macro-scale) properties are determined. Then a polymer repeat-unit is selected for further analysis with respect to the remaining sub-set of (micro-scale) properties. By integrating the different scales, from macro- to micro-scale, the uniqueness of each method is retained and their limitations are avoided to some extent. Therefore, they offer a much more promising avenue to the design of polymers satisfying a set of performance criteria.
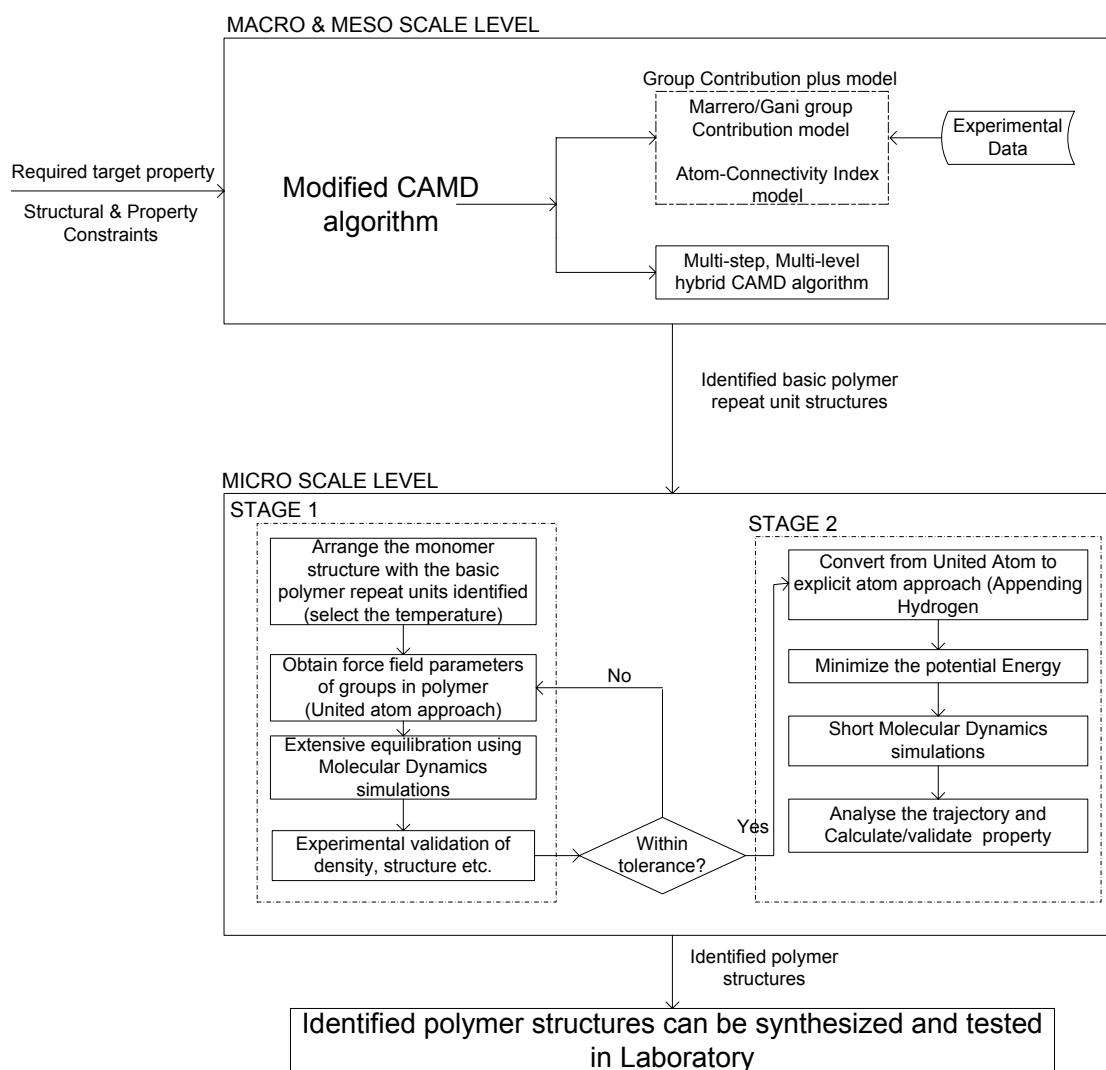

Figure 3: Schematic work flow of multilevel modelling for predicting polymer repeat unit structure with the required target property.

After the integration of macro and meso-scales for the fragments to be used in the CAMD technique, the next step is to decide on the type of the CAMD technique. Different types have been proposed for homo-polymer repeat-unit structure identification for a specified (desired) set of properties: techniques based on mixed integer linear programming (MILP) using group contribution property models by Maranas (1996); a genetic algorithm based technique also employing group contribution property models by Venkatasubramanian et al. (1994); and the multi-step group contribution-atom connectivity based technique, adopted from the multi-step, the multi-level hybrid CAMD algorithm of Harper and Gani (2000) for molecular design by Sataynarayana et al. (2009). The latter, which suppresses the combinatorial explosion problem, has been modified to use the GC$^+$ models of Satyanarayana et al. (2007) and has been further extended in this paper to add the micro-scale analysis option.

The identified basic polymer repeat-units then undergo analysis with the micro-scale approach, which involves two further stages (as proposed by Tsolou et al., 2008). In the first stage, a pre-selected sub-set of polymer repeat-units is used to build polymer chains using molecular modelling methods. United atom model approach is used to build and simulate the polymer chains in order to obtain well equilibrated configurations. Equilibrated structures are used in the second stage, where structures are converted from united atom to all atom models by appending hydrogen atoms to them, followed by the energy minimization procedure to overcome the steric overlaps. A short molecular dynamics simulation (to provide real-time information) can be performed either to calculate/validate whether the generated polymer chain (with the considered number of repeat units) has the required property or not. Dependency of the property on the varying chain length can also be studied in this micro-scale approach.

The obtained polymer configurations that have the required set of properties could then be tested through laboratory experiments to verify that they indeed satisfy the desired performance criteria. This methodology can reduce the need for laboratory scale synthesize-and-test of thousands of polymers, thereby significantly reducing the time and cost of developing a new polymer based product. Note, however, that although the micro-scale simulations are time consuming, they are applied only to a selected few polymer repeat-units.

CASE STUDY

As a test case a problem involving the design of a polymeric hermetic stopper has been considered. Hermetic seal based stoppers have wide applications in food, chemical and pharmaceutical industries, as they intend to secure against the entry of air or microorganisms and to maintain the safety and quality of the contents stored in their containers. There are a number of materials that are used as sealants such as natural rubber and cork. Natural rubber has a disadvantage for use as a stopper for medical (or drug) bottles as it can cause latex allergies in patients (http://www.sciencedaily.com). Cork stoppers also have disadvantages as they may at times spoil the stored contents due to their varying pore size (United States Patent 6528152 (2003)). It would be interesting to design a polymeric material for use as a hermetic stopper (sealant) by retaining the advantages of existing stopper materials and overcoming their disadvantages. Designing a hermetic seal to be used in common in the food, chemical and pharmaceutical industries could be too ambitious, as each of these areas may have different types of materials that need to be stored. Narrowing the problem definition here, our case study will be confined to the design of a polymeric material that is suitable for use as a hermetic seal mainly in pharmaceutical industries. In other words, it is required to design a polymeric material that can be used as a medical container stopper, which is a hermetic sealant.

Several examples highlighting the application of the model-based approach for design of polymeric products at the repeat unit level has been reported earlier by the authors (Sataynarayana et al., 2009)). Another example of only the micro-scale analysis for an identified polymer repeat-unit (polyethylene) as a polymeric membrane for gas separation has also been previously reported (Hostrup et al., 2006)).

Application: Macro-Meso-Scale Approach

The design problem to be solved through this methodology is to identify first the basic polymer repeat-unit structures that can be used as hermetic stoppers by matching the macro-scale property targets. This requires specifying initially a set of structural and macro-scale property constraints.

Property Constraints

For the case at hand, the constraints refer to: (a) glass transition temperature, (b) dielectric constant, and (c) low permeation to moisture and air.

(a) The polymeric stopper is required to be elastic over a wide range of temperature, and it is preferred to be made of a rubbery amorphous polymer so that its glass transition temperature is less than or equal to the room temperature (300 K). The upper limit for glass transition temperature can often be relaxed to 350 K. By doing this, the scope of missing some candidates that have the required property (but are not identified due to inconsistency of the experimental data reported and also considered while developing the property prediction models) is minimized to some extent.

(b) Some contents stored in the medical containers need to be constantly agitated with electrical agitators before use. It is therefore required that the polymeric stopper should not conduct electricity and electrolyze the stored contents in the medical containers. This implies a low dielectric constant as a property target. Here, we impose a maximum constraint of 2.5 on the value of the dielectric constant.

(c) The stopper should protect the contents stored from moisture and air, and exhibit very low water (or moisture) absorption. As the GC$^+$ model for predicting water absorptivity has not yet been developed, the method of Van Krevelen (1990) was used to predict this property (to verify whether the identified repeat-units have low water absorptivity or not) for the identified repeat-units that satisfy both the glass transition temperature and dielectric constant constraints. Moreover, low permeation to air implies low permeation to nitrogen (N_2) and oxygen (O_2). Since there is no group contribution method using the same groups to study the solubility and diffusivity (the key factors determining permeation) in the polymer, and, since the polymer chain length is known to affect these properties, micro-scale modelling is employed for determining these properties (along with the arrangement of repeat units to form polymer chains).

Structural Constraints

Based on the property constraints, it is more efficient to pre-select a basis set of groups (or building blocks) that are needed to build the basic polymer repeat-unit structures. In this way, the combinatorial size of the problem is reduced and redundant structures are eliminated a priori. Since the purpose is to design a rubbery amorphous polymer (as stated above) with low permeability to moisture, the polymer repeat units should not have hydrogen bonding properties. Therefore, the polar and hydrogen bond forming functional groups can be removed, a priori, from the basis set. The reason is that they often enhance chain-chain attractions (as they make possible dipole-dipole and hydrogen bonding intermolecular forces), which tend to enhance crystallinity and tensile strength. Considering that the polymer repeat-unit will not contain any polar or hydrogen bonding groups, only the groups forming hydrocarbon polymers are first pre-selected. Note that according to the combination rules, acyclic, cyclic (aromatic) and cyclic (non-aromatic) structures are generated separately. Since it is also known that the double bonds in polymers render them poor for weather resistance, thermal resistance, oxygen resistance and ozone resistance (United States Patent 5994477 (1999)), the functional groups forming these structures are also removed. Consequently, starting from a possible 86 functional groups available as building blocks, considering the structure-property relationships, we are limited to the CH_3, CH_2, CH and C groups for generating acyclic non-hydrogen bonding polymer repeat-units. Note that even with these 5 functional groups, thousands of polymer repeat-unit structures can be generated.

In addition to deciding the groups of the basis set, it is also required to decide the minimum and maximum number of groups that can occur in a basic polymer repeat-unit structure and also the number of times each group can be repeated. Given the side chain forming groups (CH and C) that are included in the basis set and in order to avoid too many side chains in the basic polymer repeat-unit, the number of times each group can occur in the repeat unit is limited to 2 implying that the maximum number of groups that can occur with the combination of all the groups included in the basis set is 6. Thus;

$$2 \leq \sum_j n_j \leq 6 \qquad (3)$$

where, n_j is the number of groups of type j.

For our case study, the structural groups considered in the basis set was classified on the basis of their valency (number of free attachments) and the general octet rule was followed while generating structures based on these building blocks. As this case study aims at generating a basic homo-polymer repeat-unit that has two free ends, the criterion of having two free ends in each generated structure is also specified. Thus a simple relation for structural feasibility of a collection of groups with this constraint is;

$$\sum_j (2 - v_j) n_j = 2(m-1) \qquad (4)$$

where, n_j, v_j are the number and valency, respectively, of groups of type j and m =1, 0 or -1 for acyclic, monocyclic and bicyclic groups, respectively. Here, since for the groups considered, only acyclic polymer repeat units are generated, therefore m takes the value of 1.

Structural and property constraints were given as input to the CAMD software. Using the automated algorithm, the basic polymer repeat-units were designed based for the above mentioned structural criteria and their properties were estimated using the group contribution based property models. The estimated properties for the glass transition temperature and dielectric constant for each repeat-unit were checked to verify whether they met the specified property constraints (mentioned above). The repeat-unit structures satisfying both the structural and the macro-scale property constraints were obtained as the desired candidates for possible use as hermetic stoppers.

Consequently, from the specified input to the CAMD software, 112 candidates of basic polymer repeat-unit structures were generated and 106 candidates were selected from them based on the given constraints (including some isomers of the generated repeat units). These 106 structures can be generated from 12 unique structures that satisfy all chosen constraints and these are listed in Table 1. It can be noted that the repeat-unit structure of polyisobutylene (PIB) is among those obtained from the CAMD algorithm. PIB is known to be an important elastomer with several interesting properties such as low glass transition temperature (it is classified as the least fragile polymer in Angell's terminology (Angell, 1995) and high friction coefficient. From different literature sources (Kamio et al., 2007; Abrams and Kremer, 2003), it is also reported that this polymer presents markedly low permeability properties to small gas molecules in comparison to other elastomers due to the presence of the two bulky pendant methyl groups on each basic repeat unit structure. The results from the macro-scale stage therefore confirm known data and gives confidence that the methodology is able to find potentially good candidates.

Permeability is described as the product of solubility and diffusivity. Lower solubility and diffusivity values always correspond to low permeability. Because it is highly time consuming to study all of the identified basic polymer repeat-unit structures at the micro-scale level (at this point of time), only PIB was selected to illustrate the concepts and, at the same time, determine the final polymer design.

Table 1: Identified basic polymer repeat units from the CAMD algorithm.

Structural Constraints			
Basis set CH3, CH2, CH, C Number of groups	6 (Max)	2 (Min)	
Property Constraints	**Min**	**Max**	
Glass Transition temperature (K) Dielectric Constant		350 2.5	
Basic polymer repeat units	T_g **(K)**	**E**	W_{abs} **(g of H_2O/g of polymer)**
-[-CH(CH$_3$)-]n-	302.51	2.48	0.00004
-[-CH$_2$-CH$_2$-]n-	168.72	2.47	0.00004
-[-CH$_2$-CH(CH$_3$)-]n-	257.91	2.37	0.00004
-[-CH$_2$-C(CH$_3$)$_2$-]n-	342.1	2.23	0.00003
-[-CH(CH(CH$_3$)$_2$)-]n-	303.13	2.28	0.00004
-[-CH(CH$_3$)-CH(CH$_3$)-]n-	302.51	2.26	0.00004
-[-CH((CH$_2$)$_2$CH$_3$)-]n-	235.61	2.25	0.00004
-[-C(CH$_3$)((CH$_2$)$_2$CH$_3$)-]n-	307.43	2.12	0.00003
-[-CH(CH(CH$_3$)CH$_2$CH$_3$)-]n-	275.75	2.14	0.00004
-[-CH$_2$-CH(CH(CH$_3$)$_2$)-]n-	276.24	2.17	0.00004
-[-CH(CH(CH$_3$)((CH$_2$)$_2$CH$_3$))-]n-	257.91	2.03	0.00004
-[-CH((CH$_2$)$_2$-CH(CH$_3$)$_2$)-]n-	258.32	2.05	0.00004

Application: Micro-Scale Approach

Recently, Tsolou et al. (2008) undertook a systematic study of model PIB systems ranging in molecular length from 8 up to 80 repeat (monomeric) units per chain at different temperature conditions (from 600 K down to 300 K). A key feature of this work was the development and implementation in the molecular simulations of a new united-atom (UA) force field for PIB, which reproduces faithfully the conformational, structural and thermodynamic properties of the polymer. In the UA model representation, hydrogen atoms are lumped together with carbon atoms to define a new CH_x type of united atoms, implying a significant reduction in the total number of degrees of freedom of the system; simulations with the united-atom model are therefore fast and very efficient. Guided by Tsolou et al. (2008), the first stage of the micro-scale approach involved the execution of atomistic MD simulations with a C_{320} PIB system (i.e., a system in which each chain contains 80 monomers or 320 carbon atoms) at several temperatures. We chose to work with the C_{320} PIB system because it is long enough to exhibit density characteristics typical of high molecular weight PIB, as shown in Figure 4 (note that the previously calculated data from Tsolou et al. (2008) have been confirmed in this work by repeating the simulations). The outcome of these simulations was an ensemble of well-relaxed configurations, which were used next as input to all-atom (AA) simulations (simulations in which the atomistic detail is restored with the insertion of the missing hydrogen atoms along each PIB chain) for the accurate prediction of the solubility and diffusivity properties of PIB to small molecules such as O_2 and N_2 of interest here. Through such an approach, Tsolou et al. (2008), for example, obtained excellent predictions of O_2, N_2, He and Ar solubility in PIB as a function of temperature. In the present study, the approach was extended in order to calculate the diffusivity of O_2 and N_2 in PIB. Overall, the steps undertaken in the micro-sale approach included the following:

1. A 3-chain PIB system with 80 monomers per chain was subjected to long atomistic MD simulations in the NPT ensemble using the parallel LAMMPS code and the new UA model introduced by Tsolou et al. (2008).

2. Well-relaxed configurations from these UA simulations at four different temperatures (350 K, 450 K, 500 K and 550 K) were selected and converted to AA representations by re-inserting the hydrogen atoms that had been neglected in the UA simulations.

3. After a short energy minimization of the chosen PIB system, five O_2 molecules were randomly inserted at different positions inside the matrix at the temperature of interest, and the energy of the system was minimized again. Extra thermal re-equilibration through a short NVT molecular dynamics run for a short time (up to 500 ps) followed.

4. Long MD runs (up to 3 ns, depending on the temperature) in the NPT and NVT statistical ensemble were carried out, which allowed us to monitor the time evolution of the 5 oxygen molecules. In all simulations with the all-atom model, the very detailed and accurate COMPASS force field was used.
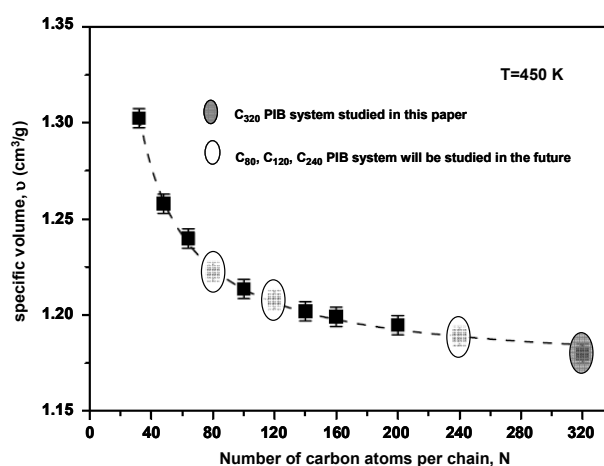

Figure 4: Specific volume versus the number of carbon atoms per chain (data reported by Tsolou et al. (2008) has been repeated and new data marked with circle have been added).

The diffusivity of oxygen molecules in the PIB matrix was studied in terms of the time evolution of their mean square displacement (MSD) in the course of the MD simulation. The diffusion coefficient D can then be readily computed from the linear part of the curve through the Einstein relation (Karayiannis et al. (2004)):

$$D = \frac{1}{6N_\alpha} \lim_{t\to\infty} \frac{d}{dt} \sum_{i=1}^{N_\alpha} \left\langle \left[\mathbf{r}_i(t) - \mathbf{r}_i(0) \right]^2 \right\rangle \qquad (5)$$

where, r_i denotes the position vector of the center of mass of the diffusant and angular brackets denote the ensemble average over all five (i.e., $N_a=5$) O_2 molecules and over all possible time origins in the course of the trajectory. Typical plots of MSD versus time at the four different temperatures for NPT addressed here (350, 450, 500, and 550 K) are shown in Figure 5. Clearly, as the temperature decreases,

the diffusivity of O_2 in PIB slows down considerably. The resulting diffusivity values are shown in Figures 6-7 in the form of Arrhenius plots, which are very helpful since, through them, we can quantify their temperature dependence. Indeed, by plotting log (D) versus 1/T, a linear plot is obtained, whose slope provides an estimate for the activation energy of the diffusive process, defined as:

$$E_a^{app} = -\frac{R}{0.434} \frac{\partial(\log D)}{\partial\left(1/T\right)} \qquad (6)$$

The value obtained from the data of Figure 6 is 5.7 kcal/mol and Figure 7 is 7.6 kcal/mol, while the reported experimental value is 11.9 kcal/mol reported by van Amerongen (1950) from diffusivity data measured over a range of lower temperature values than the ones covered here by MD. In fact, van Amerongen (1950) reports that the activation energy is not constant but increases with decreasing temperature; this indicates that the relatively lower value of E_a^{app} predicted by our micro-scale approach is reasonable.
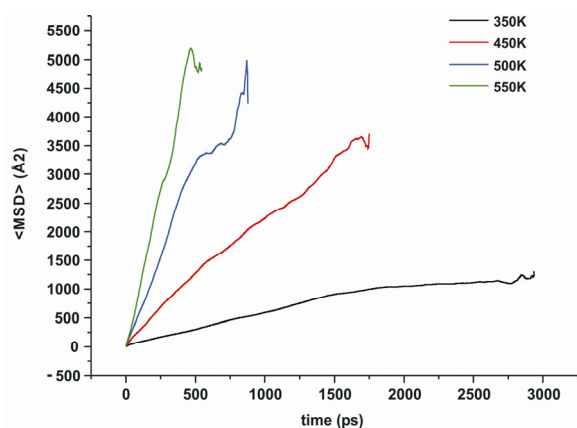

Figure 5: Mean square displacement of O_2 versus time in a PIB-C_{320} matrix (using NPT ensemble).
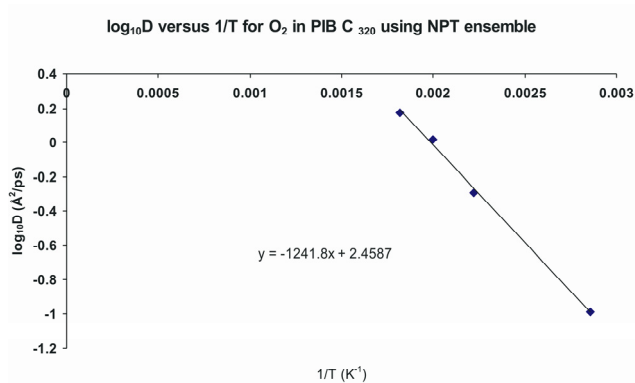

Figure 6: Log(D) versus 1/T for O_2 in a PIB-C_{320} matrix using NPT ensemble (for new improved correlation, contact the corresponding author).
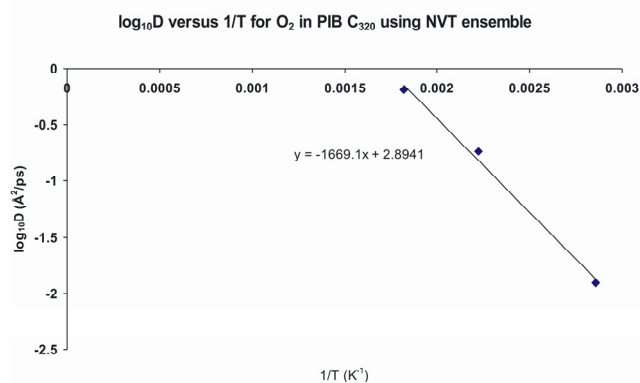

Figure 7: Log(D) versus 1/T for O_2 in a PIB-C_{320} matrix using NVT emsenble (for new improved correlation, contact the corresponding author).

DISCUSSION OF RESULTS

At this point, we should also mention that one of the above predictions for the diffusivity of oxygen in PIB was obtained through NPT simulations with the very accurate COMPASS force-field. It turns out though that, despite its complexity, COMPASS under-estimates the density of PIB (using NPT) by approximately 5-6% compared to the experimental value at the same temperature (see Table 2). Given the sensitivity of the diffusivity values to density (small density variations can have a strong effect on D), new simulations where the MD runs have been executed in the NVT statistical ensemble with the system density fixed to its experimental value (at the given temperature) have been made. Data from these simulations is highlighted in Figure 7. By comparing the results from Figures 6-7, it can be noted that E_a^{app} has a higher value with the NVT ensemble than with the NPT ensemble, which clearly shows how sensitive the diffusivity values are to the change in density of the system. Despite this, the present case study has justified the success of the potential application of a multilevel approach to the design of a polymer with a set of pre-specified properties. Using the macro-meso-scale approach, a set of basic polymer repeat-units promising to have the required target property (for the case study chosen here) is first obtained. Water absorptivity values (estimated using van Krevelen's group contribution method (van Krevelen, 1990)) for the identified basic polymer repeat units were comparatively lower, as shown in Table 1. Alternatively, we could have used all the functional groups to generate millions of structures, and then screen them for the three macro-scale properties and screen out most of the non-hydrocarbon structures. By pre-selecting the functional groups, we started with a smaller search space.

Table 2: Table showing the change in density (from initial density) after NPT runs for PIB-C$_{320}$.

Temperature	Initial Density	Density after NPT runs
550 K	0.798907	0.7562
500 K	0.826817	0.7781
450 K	0.8508	0.8152
350 K	0.905498	0.8658

On the other hand, the macro-meso-scale approach cannot alone decide about the optimal arrangement of the repeat units or the chain length; in the context of the proposed multi-level approach, this is the task of the micro-scale (atomistic or molecular modelling) component. Unfortunately, it is usually too CPU time demanding to study all the repeat units identified from the macro-meso-scale approach under the micro-scale one. This implies that the structures that should be subjected to detailed molecular modelling should have been carefully and thoughtfully chosen so as to provide the best information possible. In the present case, for example, we chose to simulate PIB with C$_{320}$ (polymer chain with 80 repeat units) and compute its permeability properties to small molecules like O$_2$ and N$_2$. Although there is still considerable work to be done at the micro-scale, the first results (together with simulation data from a recent study addressing the solubility of these molecules in PIB) seem to verify the low permeability of O$_2$ in the designed polymer. Then, given that the molecular size of O$_2$ is smaller than N$_2$, we can safely conclude that PIB will also be characterized by low permeability to N$_2$. In a broader sense, this polymer will have a low permeability to air. Moreover, other PIB systems with varying carbon number (C$_{80}$, C$_{120}$ and C$_{240}$, as shown in Figure 4) will also be studied in the future to verify which system has the lowest permeability to air. In conjunction with its low glass transition temperature (Kamio et al., 2007; Abrams and Kremer, 2003), PIB seems to be an excellent candidate for use as a hermetic stopper in medical applications. The next step is to recommend the polymer for laboratory scale synthesis, which will provide the experimental validation of the properties claimed here.

CONCLUSIONS

A multi-level model-based approach for the design of polymers with a set of required properties, which can help to decrease substantially the time for the development of new products (which traditionally has been based on the synthesis of a large number of compounds in the laboratory and exhaustive testing), has been presented. This approach relies first on the application of GC$^+$ property models within the CAMD technique (software) for the identification of the polymer repeat-unit structure, thereby enhancing the range of problems that can be handled. Next, the model-based approach incorporates a micro-scale analysis since the macro-scale CAMD algorithm is not able to give complete information on how the basic polymer repeat-units should be arranged for the resulting polymeric structure to satisfy all the desired target properties. A link is therefore provided for simulations at the micro-scale, which is capable of providing reliable estimates of the desired target (barrier) properties based on the principles of quantum and statistical mechanics and statistical thermodynamics. The application of this multi-scale model-based approach has been illustrated through a case study aimed at designing a polymer that can be used as a hermetic stopper (sealant). One of the basic polymer repeat-units identified from the macro-meso-scale approach (adapted CAMD technique) for such an application was polyisobutylene. In comparison to other elastomers and based on literature data, this polymer is characterized by a low glass transition temperature and have markedly low permeability to O$_2$ and N$_2$ (thus also to air). Building a two-stage approach (development of a united-atom model that can provide reliable predictions of the volumetric, structural and conformational properties of PIB over an extended range of temperature conditions, followed by conversion to an all-atom model for the subsequent permeability studies), the micro-scale approach helped to verify the low levels of O$_2$ and N$_2$ solubility and their small diffusion coefficients (compared, for example, to amorphous PE) in bulk PIB matrices over an extended range of temperatures. It was thus concluded that PIB would be an excellent candidate for constructing hermetic stoppers for use in medical applications.

Obviously, the micro-scale analysis is computationally expensive and, for this reason, the data generated from these simulations are being used currently to develop a higher-order model to predict the micro-scale properties as a function of the chain length parameters. Once a sufficiently large parameter table is obtained, the micro-scale analysis can be made as simple and as computationally fast as the macro-meso scale analysis. Barrier properties for gases and hydrocarbon polymers can then be handled faster and efficiently through the new models with their regressed parameters.

REFERENCES

Abrams, C. F., Kremer, K., Combined coarse-grained and atomistic simulation of liquid bisphenol A-polycarbonate: Liquid packing and intramolecular structure. Macromolecules, 36, 260 (2003).

Angell, C. A., Formation of Glasses from Liquids and Biopolymers. Science, 267, 1924 (1995).

Carbone, P., Negri, F., Müller-Plathe, F. A., Coarse-grained model for polyphenylene dendrimers: Switching and backfolding of planar three-fold core dendrimers. Macromolecules, 40, 7044 (2007).

Bicerano, J., Prediction of Polymer Properties. New York: Marcel Dekker Inc. (2002).

Drug Bottles Containing Natural Rubber Stoppers May Place Latex Allergic Patients At Risk for Reactions. ScienceDaily, June 15, 2001. (http://www.sciencedaily.com)

Gani, R., Harper, P. M., Hostrup, M., Automatic creation of missing groups through connectivity index for pure-component property prediction. Industrial Engineering Chemistry Research, 44, 7262-7269 (2005).

Gani, R., Nielsen, B., Fredenslund, Aa., Group contribution approach to computer-aided molecular design. AIChE Journal, 37, 1318-1332 (1991).

Harper, P. M., Gani, R., A multi-step and multi-level approach for computer aided molecular design. Computers and Chemical Engineering, 24, (2-7), 677 (2000).

Harmandaris, V. A., Adhikari, N. P., van der Vegt, N. F. A., Kremer, K., Hierarchical modelling of polystyrene: From atomistic to coarse-grained simulations. Macromolecules, 39, 6708 (2006).

Hostrup, M., Harper, P. M., Moen, Ø., Muro-Sune, N., Soni, V., Abildskov, J., Gani, R., Computer aided polymer design using group contribution techniques, Computer Aided Chemical Engineering, 22, 257-299 (2006).

Kamio, K., Moorthi, K., Theodorou, D. N., Coarse grained end bridging Monte Carlo simulations of poly (ethylene terephthalate) melt. Macromolecules, 40, 710 (2007).

Karayiannis, N. C., Mavrantzas, V. G., Theodorou, D. N., Detailed atomistic simulation of the segmental dynamics and barrier properties of amorphous poly(ethylene terephthalate) and poly(ethylene isophthalate). Macromolecules, 37, 2978 (2004).

Kier, B., Hall, H. L., Molecular connectivity in Structure Activity Analysis. New York: John Wiley & Sons (1986).

Knotts, IV, T. A., Rathore, N., Schwartz, D. C., de Pablo, J. J., A coarse grain model for DNA. J. Chemical Physics, 126, 084901 (2007).

Manaras, C. D., Optimal computer-aided molecular design: A polymer design case study. Industrial Engineering Chemistry Research, 35, 403 (1996).

Marrero, J., Gani, R., Group-contribution based estimation of pure component properties. Fluid Phase Equilibria, 183-184, 183-208 (2001).

Milano, G., Müller-Plathe, F., Mapping atomistic simulations to mesoscopic models: A systematic coarse-graining procedure for vinyl polymer chains. J. Physical Chemistry B, 109, 18609 (2005).

PolyInfo, National Institute for Material Science, (a polymer database), Japan. https://polymer.nims.go.jp/PoLyInfo/. Accessed (August, 2009).

Satyanarayana, K. C., Gani, R., Abildskov, J., Polymer property modelling using grid technology for design of structured products. Fluid Phase Equilibria, 261, 58-63 (2007).

Satyanarayana, K. C., Abildskov, J., Gani, R., Computer-aided polymer design using group contribution plus property models. Computers and Chemical Engineering, 33, 1004-1013 (2009).

Spyriouni, T., Tzoumanekas, C., Theodorou, D. N., Müller-Plathe, F., Milano G., Coarse-grained and reverse-mapped united-atom simulations of long-chain atactic polystyrene melts: Structure, thermodynamic properties, chain conformation, and entanglements. Macromolecules, 40, 3876 (2007).

Tsolou, G., Mavrantzas, V. G., Makrodimitri, Z. A., Economou, I. G., Gani, R., Atomistic simulation of the sorption of small gas molecules in polyisobutylene. Macromolecules, 41, 6228 (2008).

United States Patent 5994477, Selective hydrogenation method of living polymer having olefinic double bond. (1999).

United States Patent, 6528152, Stopper of foamed thermoplastic synthetic material. (2003).

Van Amerongen, G. J., Influence of structure of elastomers on their permeability to gases. Journal of Polymer Science, 5, (3), 307 (1950).

Van Krevelen, D. W., Properties of Polymers. Amsterdam: Elsevier (1990).

Venkatasubramanian, V., Chan, K., Caruthers, J. M., Computer-Aided Molecular Design using Genetic Algorithms. Computers and Chemical Engineering, 18, 833 (1994).