ATOMISTIC SIMULATION OF MACROMOLECULES

Thesis by

Youyong Li

In Partial Fulfillment of the Requirements for the

degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2004

(Defended November 15, 2004)

ACKNOWLEDGEMENTS

Studying at Caltech is very happy and fruitful. The beautiful campus and the excellent weather make everyone enjoy the life here. But the most important thing to me is that I have benefited a lot from the people around here, especially from my advisor, Bill Goddard.

I would like to acknowledge the constant support and encouragement from Bill. I still clearly remember those mornings that Bill spent with me as early as 5 AM. His intellectual integrity, creativity, enthusiasm and dedication to science not only impressed me, but also will guide me for my whole life and career.

I also would like to thank my thesis committee: Aron Kuppermann, Zhen-Gang Wang, and Bob Grubbs. I benefited not only from their excellent course lectures, but also from the fruitful discussion of my proposals.

I owe a great deal to everyone in MSC (Bill's group), past and present, for his or her friendship in the last four years. My thanks go to Mario Blanco's instructions of CED calculations, discussion of dendrimer studies with Tahir Cagin, Siddharth Dasgupta's advice in initiating nylon study, Peter Meulbroek's help for my walk from C++ to python, Nagarajan Vaidehi's help to my touch of Bio-field, Shiang-Tai Lin's help on VAC calculations. I really enjoy the discussion and even casual chat with Peng Xu, Qingsong Zhang, Weiqiao Deng, Seung Soon Jang, Yun Hee Jang, Valeria Molinero, Prabal Maiti, Lvzheng Zhang, Qing Zhang, Huangzhang Shen, Fangqiang Zhu, Victor Kam, Joyce Peng, Darryl Willick, Shirley Wu, and all the other MSC members.

Last, I dedicate this thesis to my parents and my wife. I would like to express my gratitude and my appreciation to my wife, Jing Leng, for her love, support, encouragement, scientific discussion, and for giving me a happy and complete life.

ABSTRACT

This thesis focuses on the atomistic simulation of polymers/dendrimers material properties and development/applications of Monte Carlo methods for macromolecules. The main topics and their outlines are listed as following.

1) Structures and properties of crystalline polymers from theory. Although crystalline polymers such as nylon are important industrial materials, it is difficult to get the details of the various structures/properties and the conversion between them from the experiment. Using molecular modeling, we successfully predicted the complicated structures/properties and illustrated the process of forming the polymer crystal and conversion mechanism among those structures. (Chapters 1 and 2)

2) Packing mechanism of self-assembly dendrimer balls with soft coronas. Using the vibrational density of state (DoS) derived from molecular dynamic simulations, we investigate the free energy of the liquid crystal formed by soft dendrimer balls. We find that the preferred lattice for soft balls is different from the hard balls and illustrate the mechanism. (Chapter 3)

3) Development of CCBTX Monte Carlo method for polymer and dendrimer. Although computer simulation has developed as a powerful research tool to study polymer/dendrimer materials properties recently, it has been hampered by the difficulties of sampling amorphous polymer/dendrimer configurations efficiently. We develop the efficient Continuous Configurational Biased TX (CCBTX) method to generate high-quality amorphous polymer and dendrimer atomistic structures directly. The code is implemented in C++ and ported in python environment, which provides friendly interface. (Chapter 4)

4) Thermodynamic functions, critical exponents, and theta temperatures of polymer chains from CCBB Monte Carlo method. We examine the thermodynamic properties (entropy, energy, end-to-end distance) of isolated polymer chains with the Monte Carlo method. (Chapter 5)

TABLE OF CONTENTS

Chapter 1

NYLON 6 CRYSTAL STRUCTURES, FOLDS, AND LAMELLAE FROM THEORY*

ABSTRACT

Although polyamide "nylon 6" polymer is an important industrial material, there remain many questions about the details of the various structures and the conversion between them. Using the MSXX force field (developed previously from ab initio quantum calculations), we predict the crystal structures, folds and lamellae of nylon 6, leading to the following results:

(a) Assuming infinite chains and evaluating the free energy of all 112 regular crystal structures, we find three classes of crystal structures: α form, γ form, and δ form. We find that at 300 K the α form is most stable with γ and δ higher by 0.4 and 0.3 kcal/mol/(amide unit), respectively. We calculate the Young's modulus in the chain direction to be 295 GPa for α, 135 GPa for γ, and 253 GPa for δ. These values are above the experimental value of 168 GPa for α form because the experimental system has a finite lamella thickness, disorder in the chain conformation, and imperfections in the crystallinity.

(b) We find the thermostability of α form over other forms arises from Intra H-bonds in the α form, which are dynamically and entropically favored.

(c) We propose five detailed steps in the transition between the α and γ forms. We also identify the structures of the other two experimentally observed metastable forms, β and δ. Our structures explain the available fiber X-ray results.

* Published in Y. Li, W. A. Goddard III, *Macromolecules* **2002,** *35,* 8440-8455.

(d) The H-bond schemes for all regular crystal structures are examined. We find that the γ form has a more linear (stronger) H-bond than α form, which is consistent with the interpretation from solid state NMR.

(e) Considering that nylon forms lamellae with finite thickness in the chain direction we considered all 5 possible loop structures and the two best (of 8) possible stacking schemes for the folded sheets together with the 14 possible sheet displacements. We find that the optimum lamella for α form has the alkane loop fold (one amide per loop) and packs so that adjacent sheets are displaced by ± 3.7 Å (3b/14), which is in good agreement with the conclusion from fiber X-ray. Our amide pocket model explains the observed sheet displacements in nylon 6, nylon 66, and also the progressive shear in nylon 66 and nylon 46.

1. Introduction

The polyamide "nylon 6" polymer is an important polymeric material with applications ranging from carpet and automotive parts to intimate apparel.[1] In addition, nylon 6 stores the a larger amount of iodine[2] than any other polymer and it is claimed to be a good solid-state electrolyte.[2] Nylon 6 has the -[-NH$(CH_2)_5$(CO)-]- repeating group, leading to a structure in which the peptide units (NH-CO) provide hydrogen bonding between polymer chains. Although nylon is highly crystalline, the presence of the crystalline lamella in an amorphous matrix makes it difficult to obtain precise crystallography. Despite the numerous important industrial applications and numerous experimental studies on nylon 6, there remain many questions about the details of the various structures and the conversion between them. As with other highly crystalline polymers (such as polyethylene and PET), the chains in the crystalline regions of nylon 6 tend to be re-entrant to preserve chain-chain bonding and the nature of these folds complicates the optimum spacing and structure of the lamella and the connections between them.

In order to learn about the various factors controlling these structures and their properties, we used the MSXX force field developed by Dasgupta et al.[3] to carry out theoretical

calculations (quantum mechanics, molecular dynamics, and molecular mechanics) to determine the optimum packing and fold structures of nylon 6. Section 2 describes the details of these methods. Section 3 reports the results for infinite chain model structures of nylon 6 and compares to the experimental data. Section 4 discusses the results for fold structures of nylon 6. The summary is in section 5.

2. Calculation details

We used the MSXX FF[3] with the MSC version of PolyGraf (version 3.30, Caltech version) for all calculations. Some quantum mechanics (QM) simulations were performed to validate the results with the MSXX FF. Cerius2 (v4.0) was also used for graphics and manipulations. The electrostatic and van der Waals (vdW) interactions used the accuracy bounded convergence acceleration (ABCA) Ewald technique[4] for computing the nonbond energies of periodic systems. We used an accuracy of 0.001 kcal/mol. All structures were minimized to an rms force on all atoms of 0.01 kcal/(mol Å) for atom and rms stresses of 0.1 kcal/(mol Å) using conjugate gradient method.

2.1. Force field

The MSXX force field for simulation of nylon polymers was derived from *ab initio* QM calculations.[3] Special emphasis was given to the accuracy of the hydrogen bond potential for the amide unit and the torsional potential between the peptide and alkane fragments.

This hydrogen bond potential was derived from MP2/6-31G** calculations of the formamide dimer. Subtracting electrostatic interactions (based on fixed-point charges extracted from QC on the monomers) leads to a repulsive exponential form (Eq. 1) of the short-range hydrogen bond potential[3] with A=0.028 kcal/mol, C=0.251 Å, and R_e=3.017 Å. Instead of the original charge scheme in Ref. 3, we now use the improved charge scheme

from Ref. 5 (see section 2.2). The differences are mainly from the methylene groups, which do not affect the parameters used for hydrogen bond potential.

$$E_{vdW}^{EXP} = A\exp\left[-\frac{(R-R_e)}{C}\right] \quad (1)$$

The full torsion potential between peptide and alkane fragment was calculated by optimizing the geometry (using HF/6-31G**) at each point on the torsional curve and the torsional potential is represented by a Fourier series (Eq. 2) in MSXX force field.[3]

$$E_{torsion} = \frac{1}{2}\sum_{n=0}^{n=6} V_n \cos n\tau \quad (2)$$

where τ is the torsional angle (τ=0 for cis), and V_n is the barrier (energy of cis over trans).

Detailed MSXX force field was described in previous paper.[3]

2.2. Charges

We use potential derived charges (PDQ) based on quantum mechanical calculations (HF/6-31G**) of model systems. These charges are based on calculations for long alkyl chains functionalized with an amide linkage, where a minimum of five carbons to either side was required for charge convergence.[5] Based on a series of calculations for shorter alkane chains functionalized with an amide, the charge perturbation within a long alkane chain due to each functional unit was extracted.[5] The charge scheme for nylon 6 is summarized in Figure 1.

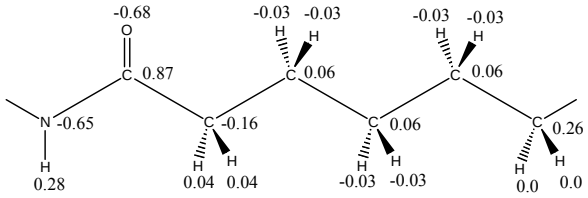


Figure 1. Charge scheme for a monomer fragment in polymer chain of nylon 6

2.3. Vibrational calculations

The analytic second derivative matrix (Hessian) obtained directly from the complete energy expression was used to calculate the vibrational modes and frequencies. This Hessian was also used with additional terms to calculate the elastic constants (including Young's Modulus). To obtain the zero point energy and temperature dependent entropy (S), enthalpy (H), and free energy (F) as a function of temperature, we calculated the vibrational modes for a 3x3x3 mesh of reciprocal lattice vectors, based on the unit cell with four chains each with two formula units. We also used a 5x5x5 mesh for the model structure proposed by Holmes[6] and found a total free energy change of only 0.03 kcal/mol/(amide unit). The methodologies are reported in Ref. 7, 8 and implemented in the VIBRATE, THERMO, and ELASTICA Modules in PolyGraf.

2.4. R-factor calculation

We used the "Diffraction-Crystal" module in Cerius2 4.0 to calculate the fiber x-ray diffraction intensities and to obtain the R-factor against the experimental data. The intensity for each hkl reflection was calculated using

$$I(hkl) = \left\{\sum f_n \cos 2\pi(hx_n + ky_n + lz_n)\right\}^2 + \left\{\sum f_n \sin 2\pi(hx_n + ky_n + lz_n)\right\}^2 \quad (3)$$

where f_n is the scattering factor of atom n and x_n, y_n, z_n are the fractional coordinates of atom n. The summation is over all atoms in the unit cell.

We used the global anisotropic temperature factor to correct the intensities. In this method, the intensity of an hkl reflection is reduced by a factor of the form:

$$\exp\left[-2\left(h^2 B_a + k^2 B_b + l^2 B_c\right)\right] \tag{4}$$

where: $B_a = 2\pi^2 \frac{\langle a^2 \rangle}{a^2}$, $B_b = 2\pi^2 \frac{\langle b^2 \rangle}{b^2}$ and $B_c = 2\pi^2 \frac{\langle c^2 \rangle}{c^2}$ are dimensionless. $\langle a^2 \rangle$, $\langle b^2 \rangle$ and $\langle c^2 \rangle$ are mean squared atomic displacements (Å^2) in a crystal with unit cell dimensions a, b, c (Å). We determined the $\langle a^2 \rangle$, $\langle b^2 \rangle$ and $\langle c^2 \rangle$ from NVT molecular dynamics simulations.

No polarization factor or crystal monochromator factors were applied to the intensity calculation.

The R-factor is defined as following:

$$R = \frac{\sum \left| I_{obs.} - I_{calcd.} \right|}{\sum I_{obs.}} \tag{5}$$

3. Infinite chain model structures of nylon 6: results and discussion

3.1. α form, γ form, and intermediate forms from fiber X-ray experimental results

Two crystalline forms in nylon 6, α and γ, have been well characterized by x-ray crystallography.[6, 9, 25-27] The plane of the amide group and that of the $(CH_2)_5$ group are parallel in the α form, while in the γ form they are approximately perpendicular. H-bonds

are formed between antiparallel full-extended chains in the α form and between parallel pleated chains in the γ form. The α phase is the thermodynamically most stable crystalline form, and can be obtained by slow cooling from the melt. The γ form is obtained by spinning fibers at a high speed or by iodinating nylon 6 in aqueous KI/I_2 treatment followed by removal of the iodine and potassium with sodium thiosulphate. The γ form can be converted into α by melting followed by recrystallization[10], by annealing at 160°C in a saturated-steam atmosphere without any significant loss of orientation,[11] and by applying stress at room temperatures[12-14].

Besides the well-characterized α form and γ form, there exists the intermediate crystalline phases between them, [15-24] which we will discuss in part 3.5.

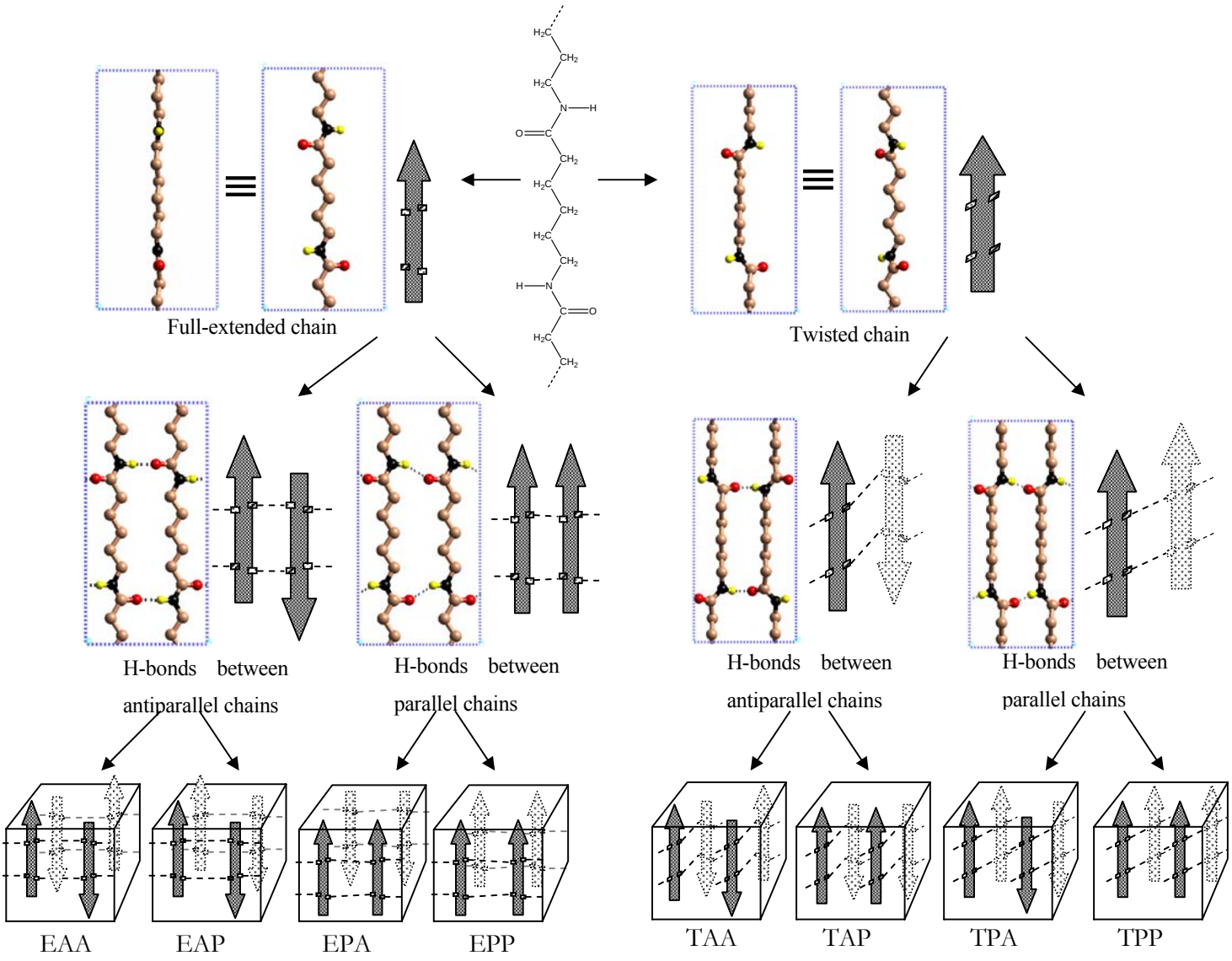


Figure 2. Eight possible classes of crystal structures constructed with different packing schemes

3.2 All possible regular infinite chain crystal structures of nylon 6

There are two distinct different chain conformations of nylon 6: full-extended chain and twisted chain as shown in Figure 2. H-bonds can be formed between antiparallel chains or parallel chains. The corresponding chains in adjacent sheets can be antiparallel or parallel. Thus there are four possible packing schemes of the chains.

From the two types of chain conformations and the four packing schemes, eight possible classes of regular infinite chain crystal structures are constructed as shown in Figure 2.

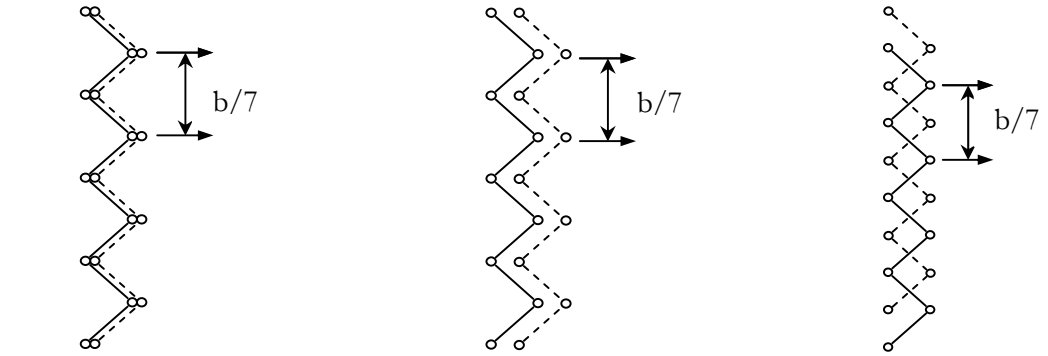
We will classify the structures using a 3-letter index:

The first letter describes whether the chains are full-extended (E) or twisted (T).

The second letter describes the relative direction of the chains *forming H-bonds* parallel (P) or antiparallel (A).

The third letter describes the relative direction of corresponding chains in *adjacent sheets* parallel (P) or antiparallel (A).

The eight classes of crystal structures EPP, EPA, EAP, EAA, TPP, TPA, TAP, and TAA are shown in Figure 2.



a. "uncrossed" chains b. shift "uncrossed" chains c. "crossed" chains

Figure 3. Three stacking schemes of polymer chains in adjacent sheets (b is the cell length of the unit cell in chain direction)

There are three distinct ways to stack adjacent H-bond sheets in nylon 6 crystal, as illustrated in Figure 3. The two chains shown in 3a, 3b or 3c represent the corresponding chains in adjacent H-bond sheets. In Figure 3b, the two chains are shifted in both *b* and *a* direction (chain direction and H-bond direction, respectively). The *shift "uncrossed"-chains* and *"crossed"-chains* stacking schemes as shown in Fig. 3b and 3c are better than *"uncrossed"-chains* stacking scheme as shown in Fig. 3a. For EPP, EPA, EAP and EAA, the unshifted *"uncrossed"-chain* structure in Fig. 3a is 0.7~0.8 kcal/mol/(amide unit) higher than the other two.

In addition, the single H-bond sheet in EPP, EPA, EAP and EAA can slide in the chain direction by an integer number of b/7 with respect to the adjacent sheets, as shown in Figure 3.

Combining the 7 slides with *"crossed"* and *shifted "uncrossed"* leads to a total of 14 possible crystal model structures for each class (14 different sheet displacements).

Multiplying the 8 classes with the 14 stacking schemes (14 sheet displacements) leads to a total of 112 regular infinite chain crystal structures

3.3. Free energy and Young's modulus of all possible regular crystal structures of nylon 6

We specified the space group of the simulated unit cell as[28] $P2_1$ and calculated the potential energy, ZPE (Zero Point Energy), entropy, enthalpy, and free energy of all 112 regular crystal structures after energy minimization in PolyGraf 3.21. The unit cell and the atomic coordinates were optimized simultaneously. To compare the various model structures we consider the free energy at 300K. The results are listed in Table 1 and illustrated in Figure 4. The parameters of the unit cell are listed in Table 2. When the SD (sheet displacement) in Table 1 is an odd integer, the crystal stacks with *"crossed"-chains* as in Figure 3c. When SD is even, the crystal stacks with *shift "uncrossed"-chains* as in Figure 3b.

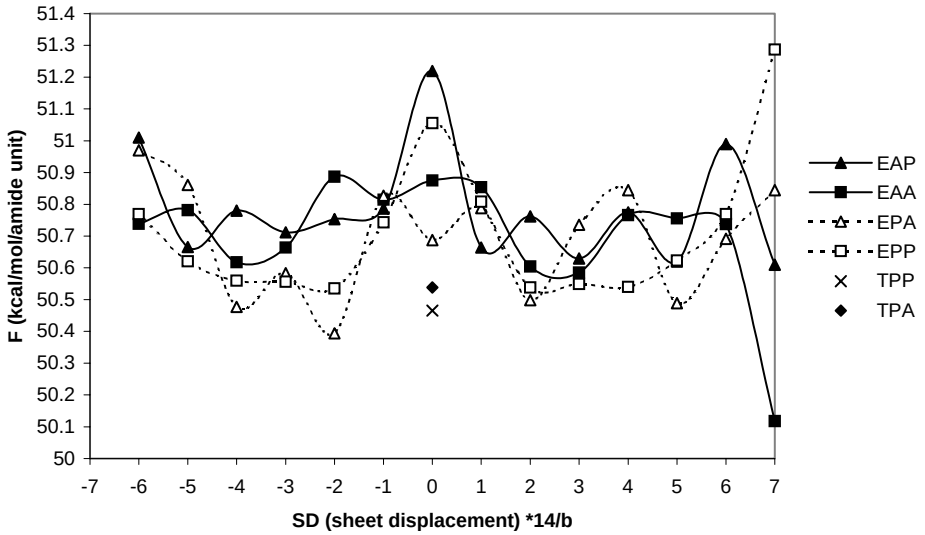


Figure 4. Free energy of different infinite chain model structures at 300 K

3.3.1 EAP and EAA (α form)

First we consider the EAA and EAP classes, each of which has 14 models. All of these are denoted as α form. As shown in Figure 5, EAP and EAA can transform to each other without difficulty. This transformation does not change the chain conformation and does not modify any hydrogen bonds. Thus only vdW interactions are involved in the barrier, which is estimated to be 0.8 kcal/mol/(amide unit). Within the EAA and EAP classes, there are 14 possible model structures differing from each other by the sheet displacement, as shown in Fig. 3. The 14 possible model structures can also transform to each other easily by overcoming the vdW energy barrier of about 0.7 kcal/mol/(amide unit).

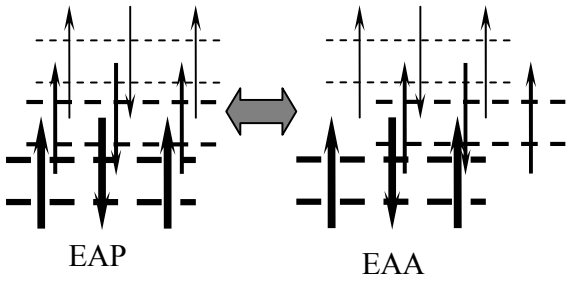


Figure 5. The transition between EAP and EAA

For EAP the lowest free energy model is EAP+7 but there are 2 other models (+3, +5) with energies within 0.01 kcal/mol/(amide unit).

For EAA the lowest energy model is EAA+7 and it is 0.5 kcal/mol/(amide unit) lower than EAP+7. As shown in Fig. 4, EAA+7 is the best model structure and is roughly 0.4 kcal/mol/(amide unit) better than other models. In EAA+7, the coulomb interaction between the amide units of the adjacent H-bond sheets is significantly favored over that of the other models.

The model structure concluded by Holmes[6] from fiber X-ray (corrected by Simon[29]) to account for the α form is EAP±3, which is *not* consistent with the results here. However,

after considering the fold structures in section 4.4, we find the same optimum crystal structure as Holmes[6].

Leon *et al.*[30] reported results of rigid body sliding for infinite chains using the PCSP method to estimate interchain separations (PCSP does not minimize the energy of the conformation or cell for each displacement in our understanding). They calculate that the sheet displacement with the most favorable potential energy is $4b/_{14}$. From relaxing the chains and unit cells we find that the most favorable is 7b/14. This indicates that it is important to relax the structures.

3.3.2 TPA (γ form) and TPP

The pleating in the hydrogen-bonded sheets in TPP and TPA prevents them from sliding by b/7 with respect to the adjacent sheets. The only reasonable model structures are TPA+0 and TPP+0. For other sheet displacements, the TPP and TPA structures relax to EPP or EPA after energy minimization.

The TPA+0 model structure corresponds to the γ form, which is exactly the same as the model structure proposed by Arimoto[9]. The free energy is 0.4 kcal/mol/(amide unit) higher than EAA+7.

The free energy of TPP+0 at 300K is very close to TPA+0 (the difference is only 0.07 kcal/mol/(amide unit). But it is not possible to form a folded structure for TPP, since all chains are parallel to each other. Thus TPP is not a reasonable regular crystal structure in the real fiber.

3.3.3 EPA (δ form) and EPP

We denote EPA as the δ form. We find that EPA fits the x-ray characteristics of the metastable crystalline phase of nylon 6 described by Murthy[23]. He found the fiber-axis diffraction scan of the metastable phase is similar to that of the α crystalline phase, but the equatorial diffraction scan is similar to that of the γ phase. We find that

EPA has the same b length (Table 2) as α form leading then to the same meridional scan,

EPA has very similar a, c, β angle parameters to the γ form (Table 2) leading thus to the same equatorial scan.

This agrees exactly with the experimental observations for the δ form. As discussed in section 3.4 below, EPA (δ form) accounts for an intermediate phase between α form and γ form (more precisely, between β form and γ form). The best model structure of EPA is EPA-2, which is 0.3 kcal/mol/ (amide unit) worse than EAA+7.

The best model structure of EPP is EPP-2 and the free energy of EPP±2, ±3, ±4 are almost undistinguishable. As with TPP, EPP has parallel chains so that it is not possible to form a folded crystal structure for the real fiber.

3.3.4 TAP and TAA

We find that TAP and TAA lead to high energies. Indeed the TAP and TAA unit cells relax to EAP and EAA with energy minimization. Keeping them rigid so that they cannot convert leads to energies 4 to 20 kcal/mol/(amide unit) higher than the other crystal forms. Such high energies make these structures extremely unlikely.

3.3.5 The Young's modulus

We predict a Young's modulus in the chain direction of 295 GPa for α, 135 GPa for γ, and 253 GPa for δ. The value for α is above the experimental value[31] of 168 GPa, because the experimental system has finite thickness lamellae, disorder in the chain conformation, and imperfections.

3.4 *Intramolecular* and *intermolecular* hydrogen bonds and the thermostability of α form over other form

Since the molecular weight of Nylon polymer is large, we consider that each lamella involves the same polymer chain folded repeatedly (of course there may be more than one polymer chain in a lamella and the polymer chain in one lamella may exit the lamella and connect through an amorphous region to an adjacent lamella or back to the same one). Each such covalently connected polymer chain we will consider as one *molecule.* For an isolated molecule, we would expect the chain to fold repeatedly to allow intramolecular hydrogen bonds, denoted hereafter as: *Intra H-bonds*. Folding an isolated polymer chain to form Intra H-bonds necessarily leads the adjacent chain to be in opposite directions. Thus *Intra H-bonds* are always between antiparallel chains.

It is also possible to have crystals in which the hydrogen bonds are between different molecules or between remote parts of the same molecule. We will denote this case of intermolecular hydrogen bonds as *Inter H-bonds*.

The lowest enthalpy for a folded molecule with Intra H-bonds would lead to a 2D sheet with a constant length between folds. Two such sheets could be packed to form a 3D structure where the intermolecular interactions would be dominated by van der Waals (non coulomb) interactions since that H-bonds are all within sheets. As discussed in section 4 (below), this packing leads to the α structure of Nylon, which is the most stable crystal structure for nylon 6.

An alternative packing is to start with the above structure having parallel sheets of intramolecular bonds and rotate the amide groups to make hydrogen bonds to the molecule in the adjacent sheet, leading to *Inter H-bonds*. As discussed in section 4 (below), this packing leads to the γ structure of Nylon-6, which can be formed experimentally from the α structure by iodination. Also we will find that δ form has only *Inter H-bonds.*

Thus for nylon 6, when H-bonds are parallel with the fold direction, the H-bonds are intra-molecular. Otherwise, the H-bonds are intermolecular. Prior to crystallization we would expect intra H-bonds to be favored since this does not require that the motions of different molecules be correlated. Thus we would expect nucleation and growth to be dominated by

molecules containing Intra H-bonds. The energetics of individual molecule for nylon 6 are discussed in section 4.3. The recent simulations by Welch P. and Muthukumar[32] support this role of Intra H-bonds. They conclude that, "Lamella thickening is a highly cooperative process requiring the mobility of all chains in the crystals."[32] Obviously, intra H-bonds favor molecular mobility and formation of the H-bond sheet inside molecule is enthalpically favored. These considerations lead to the conclusion that Intra H-bonds are favored over Inter H-bonds, both energetically and kinetically. These Intra H-bonds account for the thermostability of the α form over all other forms.

3.5 Transition mechanism and intermediate β form, δ form between α form and γ form

3.5.1 Fiber X-ray results for intermediate forms in the literature

After the elucidation of the structures for the α form and γ form of nylon 6,[6,9,33] many studies have examined the crystalline phase intermediate between α and γ.[12,13,16-24] The intermediate phase is normally observed in fibers under stress, but has been observed in relaxed fibers, and it can be transformed into either the conventional α form or γ form using suitable thermo-mechanical treatments. For example, stretching the fibers up to their breaking points[12,13] or boiling in water at temperatures between 100°C and 160°C[12,13,34] transforms the intermediate phase to α form.

Holmes[6] discussed an unstable structure he denoted as the β form. The very obvious difference between β form and α form is that the meridional spot 020 of fiber x-ray (barely visible in the α form) is the strongest spot in the fiber diffraction for the β structure.

Murthy[23] pointed out that the fiber-axis diffraction scan of the metastable phase is similar to that of the α crystalline phase, while the equatorial diffraction scan is similar to that of the γ phase.

Recent fiber X-ray results of Auriemma et al [24] show that the mesomorphic form is made of small mesomorphic aggregates of chains and the chains have disordered conformations with the H-bonds are formed in different directions. They consider the term mesomorphic to refer to the phase intermediate between α form and γ form. As discussed below we consider there to be 3 distinct intermediate phases. The one discussed by Auriemma is the one often referred to as β.

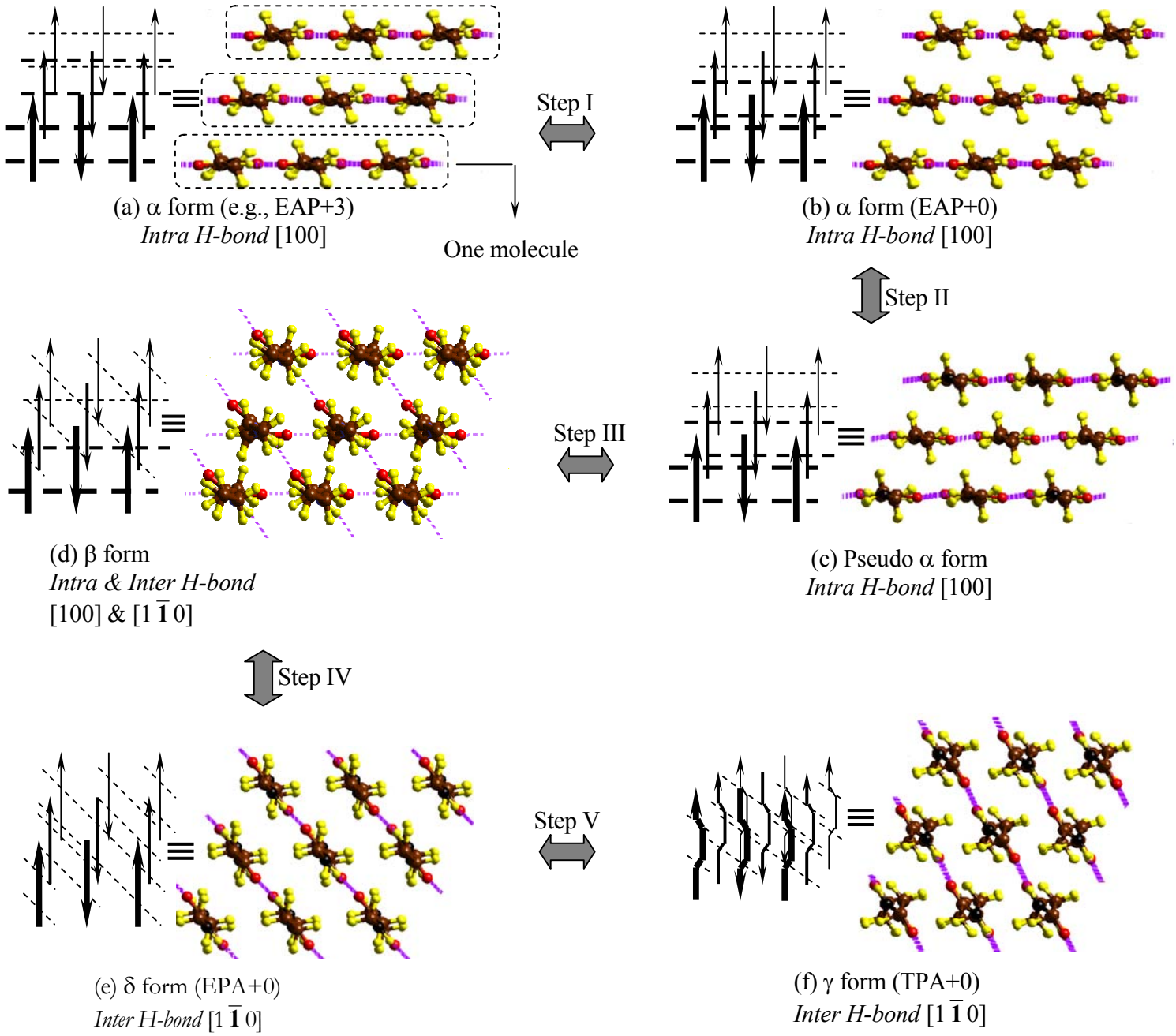


Figure 6. The transition mechanism and metastable forms between α and γ forms

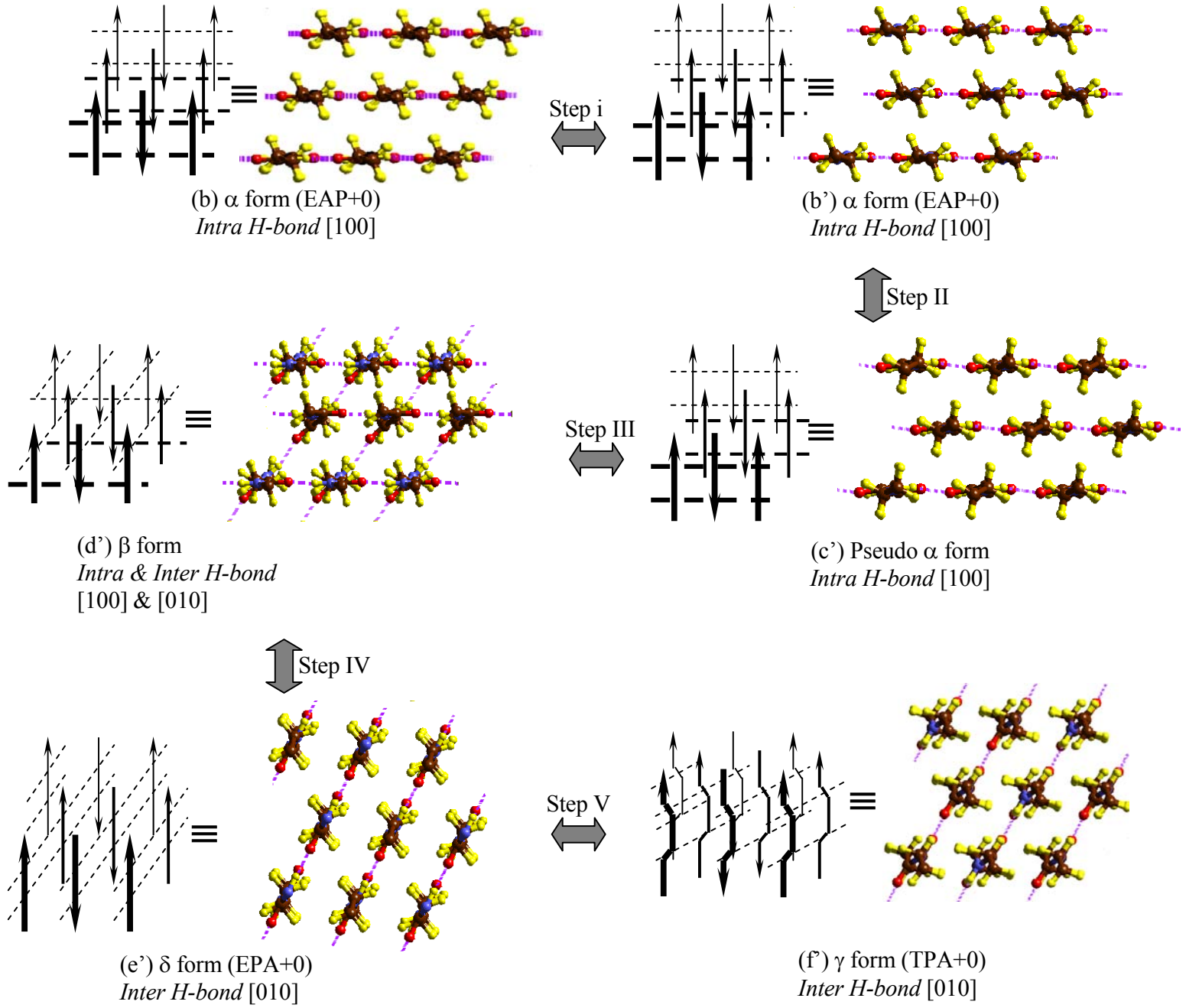


Figure 6'. An alternative of transition mechanism in Fig. 6

3.5.2 Transformation steps between α form and γ form

The transformation between α form and γ form cannot be a direct single-step process because it requires breaking a number of H-bonds, making a number of new H-bonds, and changing chain conformation.

Based on the results in section 3.3 and the fiber X-ray results in section 3.5.1, we propose five detailed transformation steps between α form and γ form as illustrated in Figure 6. Fig. 6, represents the different forms by showing nine chains of nylon 6, consisting of three molecules. The folds connecting chains with each other, are circled in Fig. 6a to help understand the composition of the chains in terms of molecules.

Step I from Fig. 6a to Fig. 6b: The H-bond sheet in the α form slides (by 3b/14) in the chain direction with respect to adjacent H-bond sheet. In Fig. 6a, the amide units in the adjacent H-bond sheets (adjacent molecules) are not on the same height. By overcoming the 0.7 kcal/mol (amide unit) vdW energy barrier (see section 3.3.1), the adjacent H-bond sheet can slide to keep the amide units on the same height as shown in Fig. 6b.

Step II from Fig. 6b to Fig. 6c: All of the amide units are twisted ~10° with respect to the two connected pentamethylene segments while preserving the H-bonds. This structure is similar to α, but has a slightly different chain conformation.

Step III from Fig. 6c to Fig. 6d: Starting from Fig. 6c, we break half of the *Intra H-bonds* in [100] direction and form new *Inter H-bonds* in [1 $\bar{1}$ 0] direction while retaining the other half of the *Intra H-bonds* in [100] direction. This structure is similar to the β form proposed by Auriemma[24] and to the interphase form proposed by Murthy[23].

Step IV from Fig. 6d to Fig. 6e: The remaining *Intra H-bonds* are broken while forming *Inter H-bonds*.

Step V from Fig. 6e to Fig. 6f: The pentamethylene segments rotate as rigid bodies with respect to the amide units so that the chain conformation changes from extended to twisted.

An alternative pathway, which is shown in Fig. 6', from α to γ is in Step i to slide the middle H-bond sheet with respect to the upper and lower ones by a/4 in the hydrogen bond direction to form Fig. 6'b' (Fig. 6'b is the same structure as Fig. 6b). Then in step III the half of the amides that twist do so in the opposite direction ([010] for original α form) to form Fig. 6'd'. In this case the step IV leads to a δ form with the hydrogen bonds in this same [010] direction. Then step V leads to γ, but with the hydrogen bonds in the [010] direction.

3.5.3 β form

We find that the structure in Fig. 6d fits the fiber X-ray results from Holmes[6], Ziabicki[15,16] and Auriemma[24]. Following their notation[6,15,16,24], we name the structure in Fig. 6d(6'd') as the β form. Because the amide units are on the same height in the structure of Fig. 6d(6'd'), they give a periodicity of 0.835 nm (the 010 spot), which is consistent with the fiber X-ray result from Auriemma[24]. We consider this same structure to be responsible for the strong 020 diffraction observed by Holmes[6], who reported a periodicity of 0.862 nm.

The structures of Fig. 6d and 6'd' have H-bonds formed in three distinct directions leading to disorder in the chain conformation. These hydrogen bonds would be in the [100], [1 $\bar{1}$ 0] and [010] directions as discussed above, which is consistent with the conclusion of Auriemma[24] from fiber X-ray results.

3.5.4 δ form

The structure in Fig. 6e(6'e') corresponds to the regular crystal structure EPA+0 discussed in section 3.3.3. We denote this as the δ form because it fits the fiber X-ray results of the intermediate phase observed by Murthy.[23] (see section 3.3.3)

From Fig. 6(6') and section 3.3.2, we can now understand the characteristics of intermediate phase from Murthy[23]. The difference between δ form and α form results from a different H-bonds pattern. The difference between the δ form and the γ form results from

a different chain conformation. Both α form and δ form are composed of extended chains, leading to a Young's modulus much larger than the γ form, which is composed of pleated chains. Thus Table 1 shows that the Young's modulus for the γ form's (TPA) is lower than EPA (the δ form) and lower than EAA, EAP (α form). It is reasonable that stretching the γ form would lead to the full-extended chains of the δ form (EPA), as found by Murthy.[23]

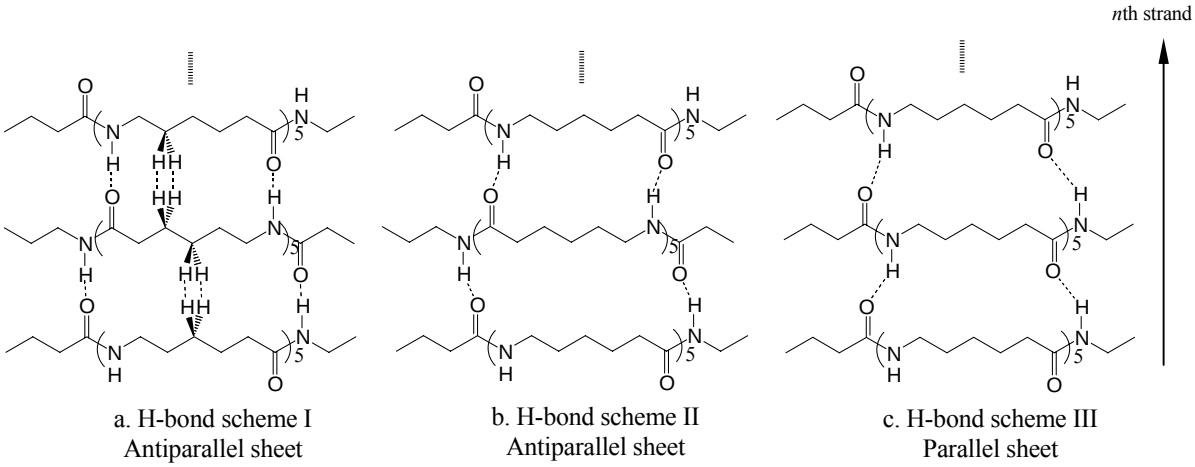


Figure 7. Three H-bond schemes for H-bond sheets. (Illustrated is the three strand case with six formula units in each strand.)

3.6 The H-bond schemes in α form (EAP, EAA), δ form (EPA), and EPP

It is well accepted[6] that for nylon 6, all H-bonds are made perfectly between antiparallel fully extended chains as shown in Figure 7a. For parallel full-extended chains (Figure 7c), the H-bonds cannot be made perfectly due to the bad H-bond angle.[6] Thus the H-bond scheme III in Fig. 7c is worse than the H bond scheme I in Fig. 7a. However, we find that the H-bond scheme I is worse than II and III by 0.6 kcal/mol/(amide unit) (see Table 3).

There are two reasons:

There is a bad contact between the pentamethylene units. Bad contact exists between 10 hydrogen atoms of every repeating unit as shown in Fig. 7a. From Table 3, we see that the vdW part of H-bond scheme I in Fig. 7a is 1.0 kcal/mol/(amide unit) worse than in scheme II and III, as shown in Fig. 7b and 7c.

Although a linear H bond is best, the energy cost of small displacements from linear is small. Thus Table 3 shows that the electrostatic part for H-bond scheme I is 0.4 kcal/mol/(amide unit) better than for Fig. 7b and Fig. 7c because of better H-bonds.

3.6.1 Validation of the FF results

To corroborate these results from the FF, we used semi-empirical QM calculations (AM1) to compare the binding energy of different H-bond schemes in Fig. 7.

Using AM1 we first optimized a single chain containing 6 amides (as shown in Figure 7) with the backbone atoms fixed on the same plane. Then we performed rigid body minimization of the dimer, trimer, and etc. until the 5th strand. This was done for the three H-bond schemes shown in Figure 7 for a three-strand case. Defining $E_{B(n)}$ as the total binding energy with n strands, then we consider:

$E_n = E_{B(n)} - E_{B(n-1)}$ as the incremental binding energy (per amide) for adding the nth strand

$\varepsilon_n = E_n/6$ is the average binding energy increment for each amide of the nth strand.

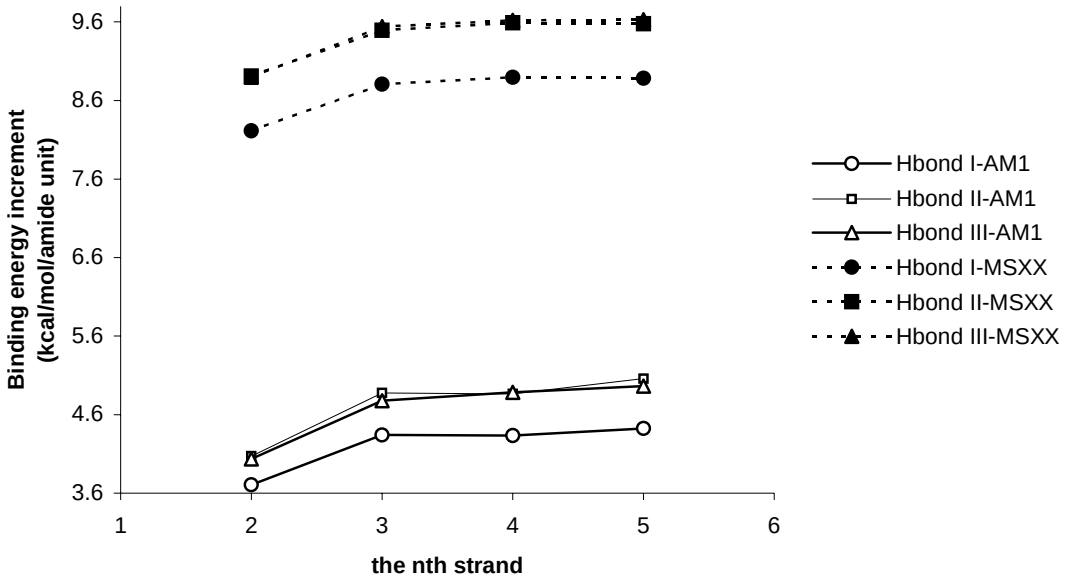
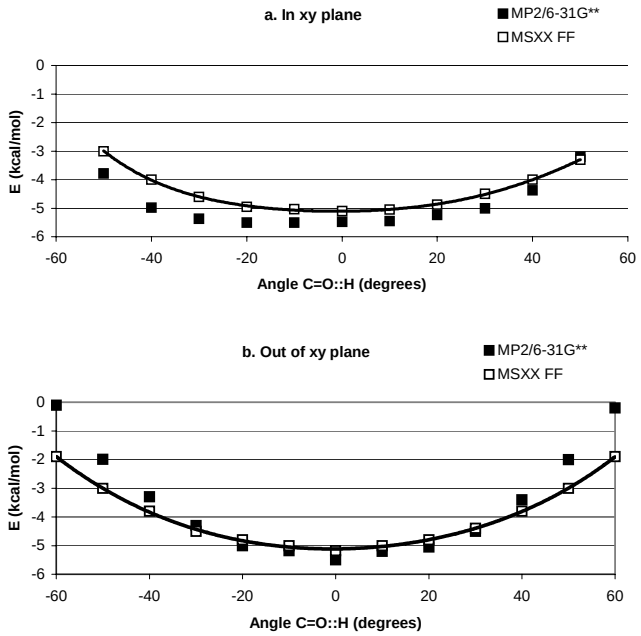
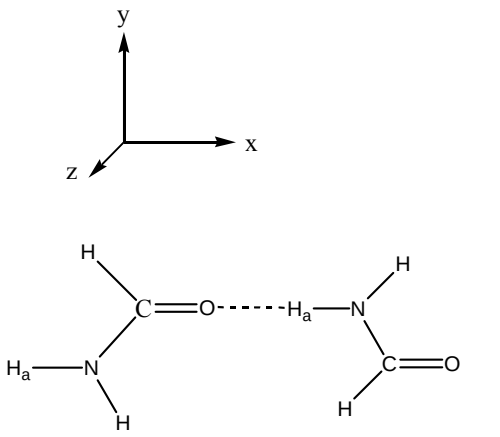


Figure 8. The binding energy increment per amide, $\varepsilon_n = E_n = [E_{B(n)} - E_{B(n-1)}]/6$, for adding the nth strand calculated using AM1 and MSXX FF

Figure 8 shows the ε_n results for each of the three different H-bond schemes. Here we see that ε_n is about the same for $n \geq 3$, but is about 1 kcal/mol weaker for n=2. This suggests a cooperative component to the hydrogen bonding. We see that H bond schemes II and III (with bent hydrogen bonds) are competitive while H-bond scheme I (with linear hydrogen bonds) is 0.5 kcal/mol/residue worse than II and III. This result confirms our MSXX FF result in Table 3.

Similarly, we use MSXX FF to get the binding energy increment as shown in Fig 8. MSXX FF gives us the similar conclusions except that the absolute binding energy is higher than AM1 due to the different charges. By analyzing the energy components of our MSXX FF calculations, we conclude that the cooperative component arises from coulomb interaction.





c. The H-bond formed in Formamide dimer

Figure 9. Dependence of the bond energy Formamide dimer on H-bond angle for translation of one amide unit relative to the other (a) in the y direction (in plane) and (b) in the z direction (out-of-plane) (reproduced from Ref 3)

3.6.2 Dependence of H-bond energy on H-bond angle

In order to understand how the H-bond energy depends on linearity of the bond, we show in Fig. 9 the change in the binding energy of formamide dimer as one formamide is translated in the plane of the dimers (Fig. 9a) or perpendicular to the plane (Fig. 9b). Here we compare the results for the MSXX FF with ab initio QM at the MP2 level. This shows that MSXX FF gives a good description of QM.

Fig. 9a shows that the potential for in-plane sliding (the *y* direction) is quite soft. One can understand this in terms of the sp^2 lone pairs of the carbonyl (see Fig. 10) which lead to significant electronic density extended from the Oxygen in the y direction, making the

interaction energy with the partially positive H favorable for displacements (2 kcal/mol for 50°).

In contrast the electron density at the oxygen drops off quickly in the z direction so that displacement in the out-of-plane *z* direction results in a much stiffer potential (3 kcal/mol for 50°, see Figure 9b).

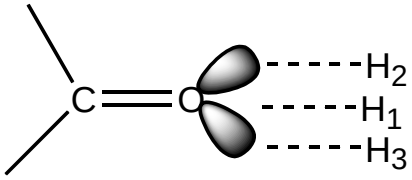


Figure 10. A schematic of the hydrogen bond with different COH angles. (The uniform charge density due to the lone pairs of oxygen in the plane leads to similar hydrogen bond energies for H3, H1, and H2, with H1 best.)

The soft in-plane sliding and the packing effects between methylene units make H-bond schemes II and III ~0.5 kcal/mol better than I in Fig. 7. QM calculations (HF-6-31G*) on optimized glycine β-sheet structures also lead to nonlinear H-bonds as in II and III over the linear ones as in I.[35]

3.7 The H-bond in γ form

It is well accepted that for Nylon 6, the γ form is not stable as α form.[9,23,33] On the other hand, Hatfield et al.[1] concluded from the results of solid-state ^{13}C and ^{15}N NMR experiments, that the H-bond in the γ phase is stronger than in the α phase. We believe that rather than measuring bond strength, their result is related to the shielding of the amide proton, which is probably related to the linearity of the hydrogen bond.

Our simulation results, show that the angle C=O::H in γ phase is more linear than in the α phase. Thus we found this angle to be

170.7^{o} in γ form (TPA+0) and

156 to 158^{o} in EAP (α form) and EAA is (H-bond scheme II in Fig. 7b) and

153^{o} in EPA (δ form) and EPP (H-bond scheme III in Fig. 7c).

A linear H-bond angle should lead to the biggest affect on the charge distribution in the amide units, which should directly affect the chemical shifts observed by solid-state NMR.

3.8. Comparison with fiber x-ray results

We used the "Diffraction-Crystal" module in Cerius^2 4.0 to calculate the fiber x-ray diffraction and to obtain the R-factor with respect to the experimental data.[6,9]

We calculated the R-factor of 28 model structures of EAP and EAA with respect to the intensity data of α form from Holmes[6] The calculated R-factors range from 0.258 to 0.408. In this calculation the mean squared atomic displacements used for the temperature factor are $\langle a^2 \rangle$=0.39 Å^2, $\langle b^2 \rangle$=0.08 Å^2, $\langle c^2 \rangle$=0.46 Å^2, which we derived from NVT molecular dynamics.

The analysis of the critical intensities such as 020 gives conclusion similar to those derived by Holmes[6]. The intensity of 020 becomes close to zero when the sheet displacement is $3/_{14}$b or $4/_{14}$b (2.9A and 3.9A respectively). Otherwise, the intensity/ratios of 020 ranges from 24/2.4% to 61/6.1%, which differs from the observed intensity, 1/0.1%.[6] The calculated intensities/ratios of different sheet displacements for 020 and 040 are listed in Table 4. Holmes suggested that the $3b/_{14}$ sheet displacement is best because the similarly charged polar groups will be uniformly distributed along the b-axis rather than close to each other as they are in nylon 66. Instead we find that the best crystal structure for infinite chains (see Table 1) is EAA+7 and not the EAP+3. In Section 4.4 we discuss the effects of

having finite thickness lamella, where we find the interactions between the folds lead directly to the $3b/_{14}$ sheet displacement.

We find that for infinite chains all EAP and EAA structures lead to a significant discrepancy in the intensities of the 7th layer line, as originally mentioned by Holmes.[6] Displacing the sheet by $b/_7$ as shown in Figure 3 does not affect the intensities of the 7th layer line. However, the "uncrossed" and "crossed" structures shown in Figure 3b and Figure 3c give quite different intensities of the 7th layer line. The "uncrossed" structure gives significant intensities for *h*7*l* with even *l* while the "crossed" structure gives significant intensities for odd *l*. Unfortunately, Holmes[6] listed odd *l* of *h*7*l* only for observed intensities. For this reason the "uncrossed" structure leads to a larger R-factor than the "crossed" structure. We explain the discrepancies for the 7th layer line to the coexistence of "crossed" and "uncrossed" structures.

4. Results and discussion for folded structures (lamella) of nylon 6

4.1 Introduction

It is generally accepted that many linear polymers form crystalline regions or lamella consisting of folded chains. This includes polyethylene, polyoxymethylene, poly(ethylene terephthalate), and nylon. Three models of chain-folded polymer crystals have been proposed[36] as shown in Fig. 11:

Fringed Micelle or Bundle-like Model;

Random Re-entry or "Switchboard" Folded Model;

Adjacent Re-entry Chain-Folded Models (Regular Folding, smooth surface).

Models a and b would lead to total densities substantially low. Model c of lamellae with smooth surfaces would lead to a total density close to the x-ray crystal density. Model d of

lamellae with rough surfaces would be expected to have some voids or bubbles between the lamellae, leading to a density somewhat below the x-ray crystal density.

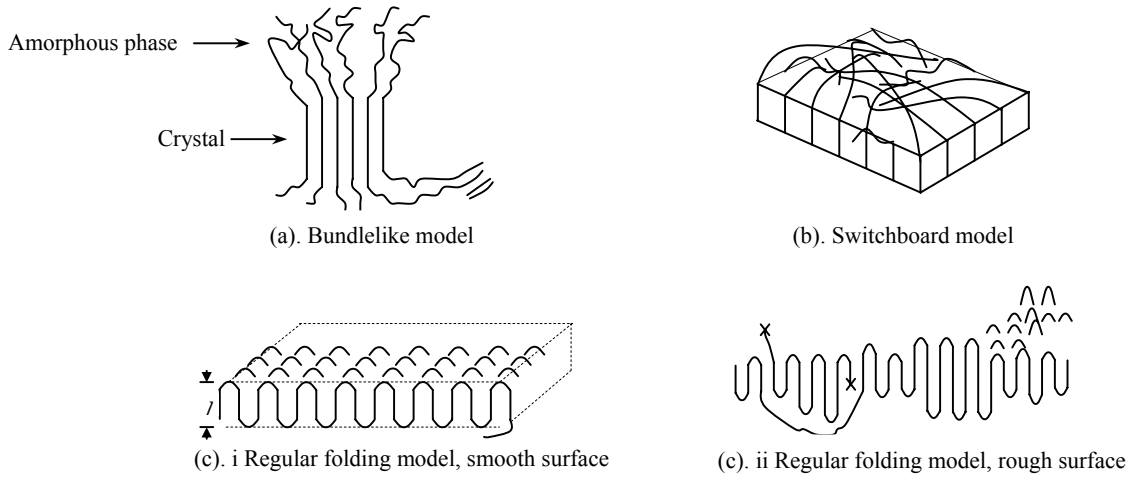


Figure 11. Schematics for suggested fold models of polymer crystals (based on Fig. 2 of Ref. 34)

The measured density of the piece of the monofilament nylon 6 used for the intensity measurements was[6] 1.160 ± 0.001 g/cm^3, which is quite close to the crystal density of 1.233 derived from the x-ray unit cell[6] (a=9.56 Å, b=17.24 Å, c=8.01 Å, β=67.5^o). For the regular lamellae model we constructed, we find a density at 0K of 1.18 ± 0.02 g/cm^3 (the uncertainty depending upon which types of loop). The good agreement with experiment strongly suggests that the experimental system has a regular lamella structure as in Fig. 11c.

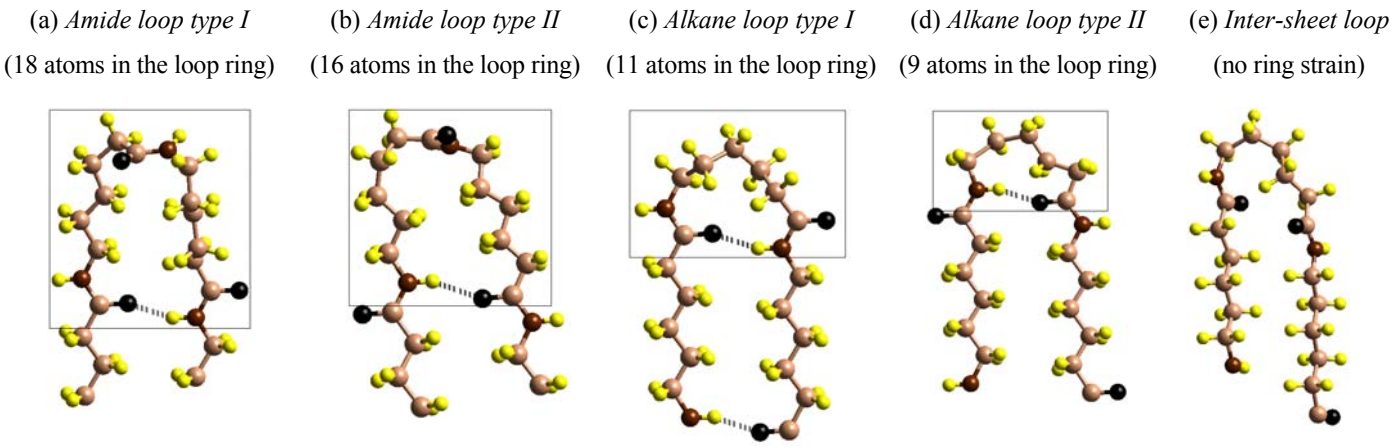


Figure 12. Five types of loop structures for nylon 6

4.2 Different types of loops for nylon 6

There are five types of loops for the lamella with smooth surface shown in Figure 11c. The loop can be formed

with *Intra H-bonds* leading to the four cases in Fig 12 abcd or

with *Inter H-bonds* (*inter-sheet*, Fig. 12e).

For the first four loops must be compatible with forming the H-bonds (dotted lines in Fig. 12 abcd), leading to a ring constraint. This can be done in two ways:

Two amide units to form the loop (Fig. 12a and 12b), denoted as the *Amide loop*

One amide unit to form the loop (Fig. 12c and 12d), denoted as the *Alkane loop*

The loop ring of the *Amide loop* includes

18 atoms (Fig. 12a, type I) or

16 atoms (Fig. 12b, type II),

whereas *Alkane loop* has

11 (Fig. 12c, type I) or

9 (Fig. 12d, type II).

Thus the *Amide loop* leads to less ring strain than the *Alkane loop*, but the *Alkane loop* has one additional amide unit in the stem chain for favorable hydrogen bonding and packing. For both *Amide loop* and *Alkane loop*, there are two cases referred to as and type II. Type II has two less atoms in the loop than type I leading to a stiffer more highly strained loop.

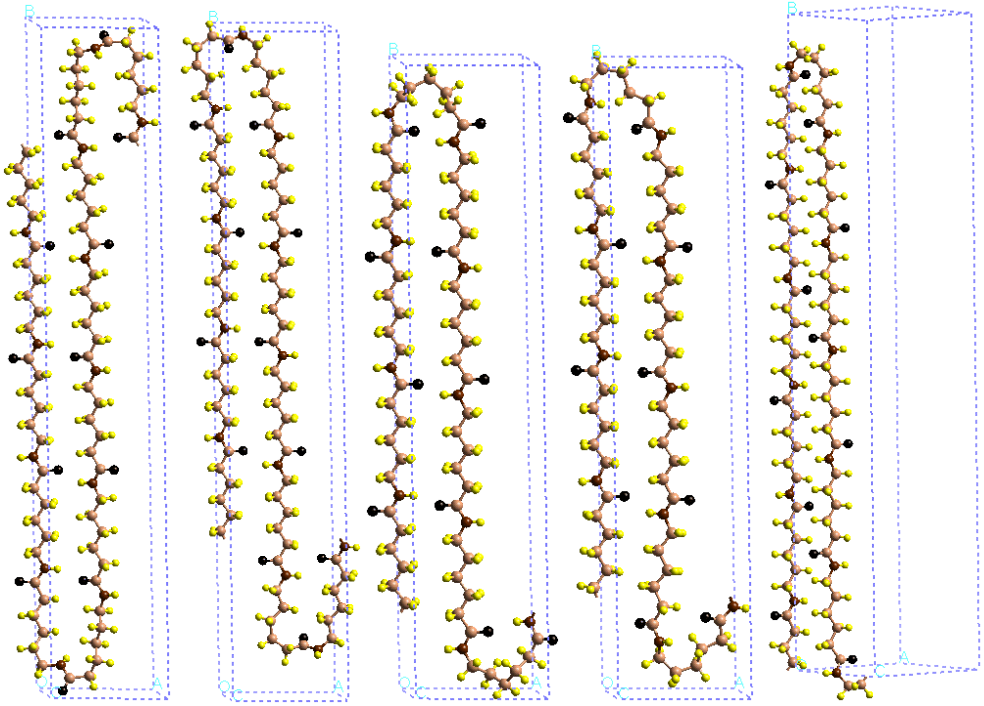
In contrast, the *inter-sheet loop* (between H-bond sheets) leads to very little constraint on the loop ring. In addition, it uses only one amide unit in the loop, just as for the *intra-sheet Alkane loop*. For comparing the *inter-sheet loop* case with the various Intra H-bond cases, we selected the lowest energy inter-sheet H-bond case.

In order to determine the energetics for different types of loops, we performed NVE molecular dynamics (MD) of the model structures shown in Figure 12. MD is necessary to allow the loops to achieve their most favorable shape. In these calculations we fixed the stem regions, allowing only the atoms in the loop to move. To be consistent we included the interactions among the 57 atoms shown in Fig. 12 but allowed only the ones in the boxes to move in the dynamics. We then performed a total of 1ns MD using 1 fs time step at 600K. Then we analyzed the 20 potential energy lowest frames and minimized them to get candidate structures for the optimum structure of each type of loop. MD of whole system will give a better analysis in cost of a lot of computer time. By using the strategy described above, in which MD is performed before energy minimization, we get reasonable results in cost of normal computer time.

Next we must consider how thick to make the lamellae. Considering oligomers of nylon 6 containing just 10-amide units[37], the Lamella Stacking Periodicity (LSP) from x-ray

diffraction is 4.77 ± 0.05 nm. Such systems will fold just once suggesting a lamella thickness for nylon 6 that is 2.8 times the b length of the original unit cell[6] (17.24 Å). Trifan et al.[38] showed from x-ray experiments that the lamella thickness of nylon 66 is approximately 58 Å (that is 3.3 times the b spacing) and independent of the bath temperature. The optimum lamella thickness involves a balance of thermodynamics and kinetics. Thick lamella leads to increased stability but the kinetics limits the growth rate thick lamella (requiring the long chains to be aligned as the crystal grows). Based on these experimental results we decided to aim at a lamella thickness of ~50 Å.

In order to use just one type of loop for *intra-sheet loop* cases, it is necessary to have an even number of amides in each stem. To achieve a thickness of ~ 50 Å leads to the choice of four amides per stem for *Alkane type* and five amides per stem for *Amide type*.



(a). Amide I (b) Amide II (c) Alkane I (d) Alkane II (e) between sheets

Figure 13. Individual molecules of five different types of loop structures

4.3 Optimum loop type for individual molecules

First we consider the optimum loop type for the individual molecules shown in Fig. 13. In order to eliminate inter-molecule interactions we fixed the distance between layers (molecules) to be 50 Å, and we fix the LSP (b length of the unit cell) to be 100 Å.

The energy cost of forming the various types of loops for individual molecules is shown in Figure 14 and Table 6. The fold energy cost is normalized by the number of folds. The reference energy is a single H-bonded sheet formed between infinite antiparallel chains without a fold (the distance between sheets again fixed at 50 Å). This MSXX FF energy is –35.940 kcal/mol/(amide unit).

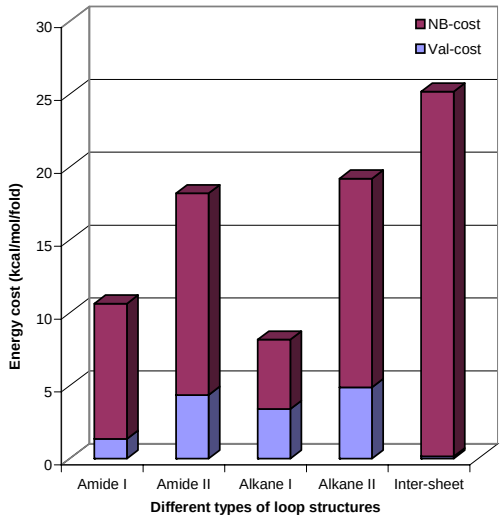


Figure 14. Energy cost of different types of loops in individual molecule

The energy cost of the fold part arises from two terms:

Nonbond part (packing energy cost) and

Valence part (ring constraint cost).

The cost of losing the H-bond is included in the *Nonbond part*. Here we see that

the *Alkane loop type I* is best, with a cost of 8.2 kcal/mol/fold

next is *Amide loop type I* (10.6),

then *Amide loop type II* (18.2), and

then *Alkane loop type II* (19.2).

The difference between *Amide loop type I* and *Alkane loop type I* is 2.4 kcal/mol/fold, which is less than one H-bond. For both *Amide loop* and *Alkane loop*, *Type II* is always worse than *Type I* because of the additional ring constraint in *Type II*.

For the *inter-sheet loop* case, the individual molecule energy is much less favorable (by ~ 17 kcal/mol/fold, see Figure 14) than the *Alkane loop type I* case, because of the lack of H-bonds. We see from Fig. 14 that the energy cost for one molecule of *inter-sheet loop* almost comes from the *Nonbond part* (mainly H-bond part). Thus the *inter-sheet loop* is not favored for individual molecules and is not favored by dynamics.

Jones et al[37] discussed the "*Amide Fold*" and two types of "*Alkane Fold*" for singly-folded finite chains of nylon 6 using the CVFF force field. They concluded that the "*Amide Fold*" and one of the two "*Alkane Folds*" are both good folds for nylon 6. However, they did not provide any details of the fold energy cost.

4.4 The stacking of the fold sheets and the best fold sheet displacement

4.4.1 Two best of eight possible stacking schemes for folded sheets

Before going to 3D lamella crystal structures, we will first consider all possible stacking schemes of the folded sheets.

Because the loop parts of folded sheets are wave-like, there are two ways of stacking the adjacent loops, as shown in Fig. 15. We call them *Loop stacking type* I and *Loop stacking type II*, respectively. *Loop stacking type I* is better than *Loop stacking type II*, because there

is less vdW repulsion energy between the adjacent loops. The energy difference between them varies depending on the fold sheet displacements and the way of stacking the folded sheets. In the following calculations we will consider only *Loop stacking type I*.

(a) *Loop stacking type I* (b) *Loop stacking type II*

Figure 15. Two ways of stacking adjacent loops (The fold structure is projected down the chain axis. The wave line and the small circle represent the loop and the straight stem segment, respectively.)

(a) *Eclipse fold-sheet stacking* (b) *Staggered fold-sheet stacking*

Figure 16. Two ways of stacking adjacent fold sheets

(a) *Eclipse lamella stacking* (b) *Staggered lamella stacking*

Figure 17. Two ways of stacking the adjacent lamellae

Besides the two ways of stacking of adjacent loops shown in Figure 15, there are two ways of stacking folded sheets in the H-bond direction as shown in Figure 16. We call these *Eclipse fold-sheet stacking* and *staggered fold-sheet stacking*, respectively.

In addition, considering the stacking of the folded sheets in the chain direction, there are two ways of stacking the lamella, as shown in Fig. 17. We call these *Eclipse lamella stacking* and *Staggered lamella stacking*. The difference between them and the conversion between them involves only loop vdW interactions. For the folded sheets of *Alkane loop type I* shown in Figure 12c, *Eclipse lamella stacking* is better than *Staggered lamella stacking* by 0.422 kcal/mol/fold (keeping the c length still fixed at 50 Å).

Different lamella stacking type gives different optimum b spacing. For the folded sheets of *Alkane loop type I*, *Staggered lamella stacking* gives 45.3 Å, which is 1.3 Å larger than for *Eclipse lamella stacking.*

To be consistent, all the unit cells simulated below are constructed using *Eclipse lamella stacking*.

From Fig. 15, 16, 17, there are total 2^3=8 possible stacking schemes for folded sheets. We will consider just the two best of them.

4.4.2 Folded sheet displacements

The folded sheets can slide in chain direction with respect to adjacent sheets just as infinite chain model structures discussed in Sections 3.2 and 3.3. Fig. 18 gives the Fold energy cost dependence on the fold sheet displacement. The reference energy here is the energy of EAP+3 infinite chain model structure in Table 1. Thus the “fold energy cost” in Fig. 18 includes both the energy cost of converting stem amide units into loop units and the energy of stem segments sheet displacement in chain direction.

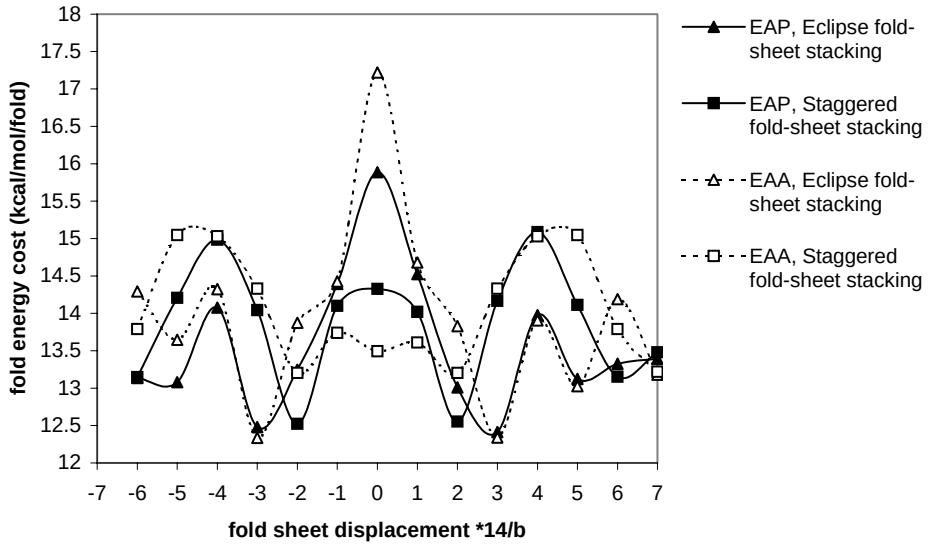
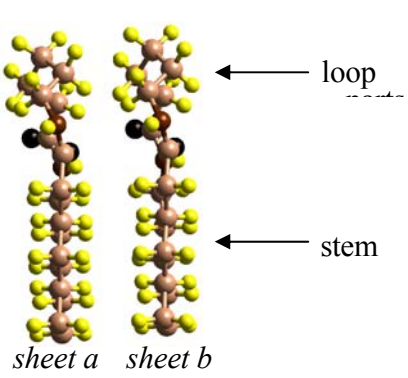
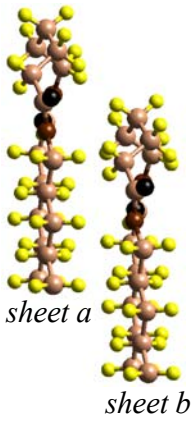


Figure 18. Fold energy cost for different sheet displacements of *Alkane loop type I*



a. The fold sheet displacement is zero. There is strong vdw repulsion between the folded parts, but the packing of the stem parts is good.



b. The fold sheet displacement is $3b/_{14}$. The amide part in *sheet a* provides much space for the folded part in *sheet b*.

Figure 19. Various fold sheet displacements of *Eclipse fold-sheet stacking*

For both EAP and EAA we find that the best displacement for *Eclipse fold-sheet stacking* is $3/_{14}$ b. Two factors affect the fold sheets stacking.

One factor comes from the loop parts (see Fig. 19). The loop parts need more space than the straight stem segment. Thus there exists a strong repulsion energy of the loop parts for zero sheet displacement of *Eclipse fold-sheet stacking*, as shown in Fig. 19a. In addition, the amide part provides much space for the loop of the adjacent sheet, as shown in Fig. 16b.

The second factor is that the stacking between the stem parts is better than the stacking between the loop part and the stem part, which is not significant compared with the first factor.

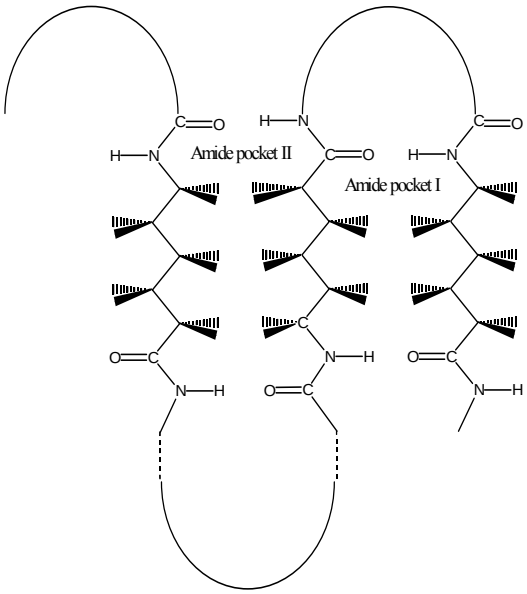


Figure 20 Amide pockets for the loop in adjacent sheet

For *Eclipse fold-sheet stacking*, the best place to accommodate the adjacent loop is the *Amide Pocket I* as shown in Fig. 20. Thus the best fold-sheet displacement is $3/_{14}$b.

For *Staggered fold-sheet stacking*, there is no strong repulsion energy between adjacent loop parts for zero sheet displacement. Here the best place to accommodate the adjacent

loop is *Amide pocket II* instead of *Amide pocket I,* as shown in Fig. 20. Thus the best fold-sheet displacement is $2/_{14}$b. See Fig. 18 for the energy comparisons.

These results of the optimum sheet displacement are consistent with the conclusion from fiber X-ray results.[6]

4.5 Optimum loop type in 3D lamella crystal

Section 4.3 found that the optimum loop type for an individual molecule is *Alkane loop type I*. In the 3D lamella crystal, this might change because of *Inter H-bonds* (including H-bonds formed between adjacent loops) and different ways of stacking of the folded layers.

To simplify the comparisons, we consider only *Eclipse fold-sheet stacking* and *Loop stacking type I* in constructing 3D lamella crystals of all five different loop types.

For *intra-sheet loop type*, the 3D lamella crystal structures were constructed from EAP infinite chain model structures.

For *inter-sheet loop type*, the lamella crystal structures were constructed from EAA infinite chain model structures.

For the fold sheet displacement, we considered only $3/_{14}$b and $1/_{14}$b.

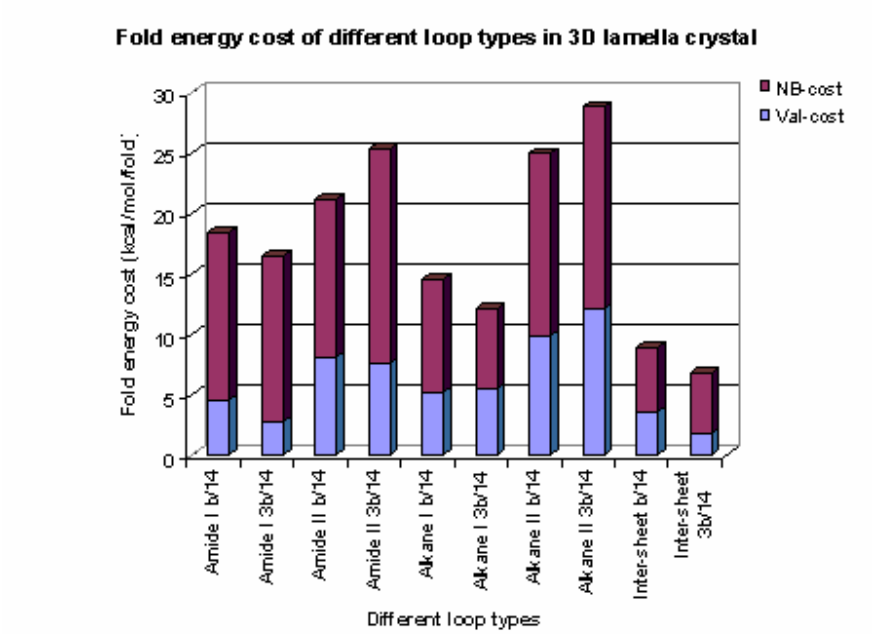


Figure 21. Fold energy cost of different types of loops for 3D lamella crystal

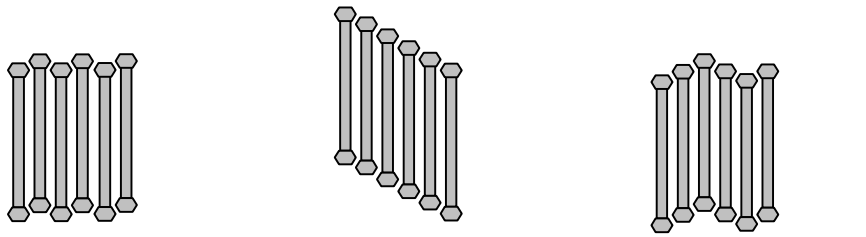
The results in Figure 21 show that, in the 3D lamella crystal structure, the *inter-sheet* loop between H-bond sheets is more stable by 5.3 kcal/mol/fold than the best *intra-sheet* loop (*Alkane loop type I*). The reason for this is the lack of loop ring constraints for the fold *between* H-bond sheets, allowing greatly reduced strain and allowing 100% of the H-bonds to be formed.

The *Inter-sheet* loop leads to *Inter H-bonds*, which considering the dynamics of the crystal growth should be worse than loops with *Intra H-bonds* (See section 3.4).

The four types of loops with *Intra H-bonds* have relative energies in the 3D lamella crystal in the same order as for individual molecules. Thus from best to the worst: *Alkane loop type I*, *Amide loop type I*, *Amide loop type II*, and *Alkane loop type II*. Although the amide

groups in the *Amide loop type I* can form weak H-bonds between adjacent loops, this costs a significant amount of packing energy (involving the methylene group and the amide group from stem to loop). The result is that *Alkane loop type I* remains the best type for *intra-sheet* loops.

Thus *Alkane loop type I* should be the dominant loop type for 3D lamella crystals of nylon 6. However, it is not possible to have *Alkane loop type I* in both sides of the stem if there are an odd number of amide units in the straight stem segment. On the other hand *Amide loop type I*, which is only 4 kcal/mol/fold worse than *Alkane loop type I* is allowed no matter how many amide units are in the straight stem segment. Thus we expect that the lamella in experimental structures will have coexisting *Alkane loop type I* and *Amide loop type I*, but with a majority of *Alkane loop type I*.



a) Monoclinic unit cell (staggered shear)

b) Triclinic unit cell (progressive shear)

c) Complicated (roughly staggered shear)

Figure 22. Various stacking schemes for H-bonded sheets

4.6 Monoclinic versus triclinic

All the folded crystal structures above are monoclinic unit cells (Fig. 22a) in which alternate sheet displacements are EAP+3 and EAP-3. In Fig. 22b, we consider the best

lamella structure of nylon 6 when each sheet displacement is the same (EAP+3), leading to a triclinic unit cell. We find that triclinic is 0.213 kcal/mol/residue more stable than monoclinic. This is explained by our Amide Pocket model. The Amide Pocket effect is not quite as good if the same pocket must provide space for both adjacent loop parts as in Fig. 22a. The monoclinic and triclinic lamella in Fig.22a and 22b are the extreme case of perfectly ordered structures. More likely is that growth kinetics will lead to lamella illustrated in Fig.22c, which has a nearly random combination of 22a and 22b. The lamella in 22c would be entropically favored and is probably close to the real lamella for nylon 6.

From fiber x-ray[6], it has been determined that the lamella in nylon 66 α form is progressive shear (Fig. 22b) whereas nylon 6 α form and nylon 66 β form have staggered shear (Fig. 22a). This difference between Nylon 66 and Nylon 6 arises because it is easier for the amine residue in Nylon 66 to form the loop (due to less loop constraint than in the carbonic acid residue). Also, the amine residue of nylon 66 is bigger than the residue of nylon 6. Thus the Amide Pocket effect in nylon 66 should be more significant than in nylon 6, leading to a stronger preference for progressive shear observed in nylon 66 α.

A progressive shear was also concluded to be favored for nylon 46[39] and explained it from surface H-bonds with an assumption that the fold part including two amide units.[39] Our amide pocket model explains the enthalpy favorable progressive shear without such an assumption. Further work need to be performed for nylon 66 and nylon 46 to clarify the best loop structure and competing factors.

5. Summary

Using the MSXX force field (derived from *ab initio* quantum calculations), we predict the crystal structures and folded (lamella) structures of nylon 6.

(1) Assuming infinite chains and evaluating the free energy for all 112 regular crystal structures, we find three classes of regular crystal structures: α form (with amide bonds parallel to the methylene sheets), γ form (with amide bonds ~ perpendicular to the methylene sheets), and δ form (somewhat intermediate between α and γ). These results for α and γ agree well with experimental data. We find that at 300K the α form is most stable with γ and δ higher by 0.4 and 0.3 kcal/mol (per amide) respectively. The Young's modulus in the chain direction is 295 GPa for α, 135 GPa for γ, and 253 GPa for δ. The only experimental data is 168 GPa for α which is below the calculated value because of the finite thickness of the lamella, the disorder of the chain conformation, and non-perfect crystal in the experimental system.

(2) The concepts of *molecule, Intra & Inter H-bonds* are introduced and we find the thermostability of α form over other forms comes from *Intra H-bonds* in α form, which are dynamically and entropically favored.

(3) Five detailed transition steps between α form and γ form are proposed and the structures of the two metastable crystalline βforms and δ form are identified. The structure of β form, (which has disordered chain conformations and H-bonds in [100], [010], and $[1\,\bar{1}\,0]$ directions) is consistent with the fiber X-ray results from Auriemma et al[24]. The δ form defined here accounts for the metastable crystalline phase between β form and γ form, and fits the characteristics of the metastable crystalline discussed by Murthy.[23] It has a similar fiber-axis diffraction scan as α form and has a similar equatorial diffraction scan as γ form. These results suggest that stretching and relaxing the fibers, would transform between γ and δ.

(4) The H-bond schemes for all regular crystal structures are examined. We find that the γ form has a more linear H-bond than α form, which is consistent with the conclusion from solid NMR[1].

(5) Since Nylon forms lamellae with finite thickness in the chain direction, we considered all five possible loop structures and the two best (of 8) possible stacking schemes for folded sheets together with 14 folded-sheet displacements. The Five types of loop structures of nylon 6 are:

Two with *Intra H-bonds* including two amide units (called *Amide loop type I* and *II*)

Two with *Intra H-bonds* including one amide unit (called *Alkane loop type I* and *II*)

One with *Inter H-bonds* including one amide unit

We find that for individual molecules the *Alkane loop type I* is the best with the others worse by at least 4.5 kcal/mol/fold. However in 3D lamella crystals, the inter-sheet case with loops *between* hydrogen bond sheets (leading to γ form or δ form packing) has an energy 5.3 kcal/mol/fold lower than the best intra -sheet case. This is because there is no ring constraint on the structure of the loops.

We find that the optimum lamella for α form have the alkane loop fold (one amide per loop) and pack so that adjacent sheets are displaced by ± 3.7 Å ($3/_{14}$b), which is in good agreement with the conclusion from fiber x-ray (Holmes[6]). Amide pocket model is proposed to explain the observed sheet displacement in nylon 6, and it can also explain the observed sheet displacement in nylon 66 and progressive shear in nylon 66 and nylon 46.

6. Acknowledgment

We thank Dr. Siddharth Dasgupta for helpful advice in initiating this project. This research was partially supported by grants from NSF (CHE99-85574), DOE ASCI ASAP, and ARO-MURI. The facilities of the MSC are also supported by grants from DOE, NSF (CHE 99-77872), ARO (MURI), ARO (DURIP), NIH, ChevronTexaco, 3M, Seiko-Epson,

Avery-Dennison, General Motors, Kellogg's, Asahi Chemical, Beckman Institute, and Nippon Steel.

7. References:

1. Hatfield, G. R.; Glans, J. H.; Hammond, W. B. *Macromolecules* **1990**, 23, 1654.
2. Murthy, N. S.; Khanna, Y. P. *Chem. Mater.* **1993**, 5, 672.
3. Dasgupta S.; Hammond, W. B.; Goddard III W. A. *J. Am. Chem. Soc*. 1996, 118, 12291.
4. Karasawa, N.; Goddard III, W. A. *J. Phys. Chem*. **1989**, 93, 7320.
5. Dasgupta S.; K. A. Brameld, C. -F. Fan, W. A. Goddard III, *Spectrochimica Acta Part A* 53 1997, 1347.
6. Holmes, D. R.; Bunn C. W. and Smith D. J., *J. Polym. Sci.* 1955, 17, 159.
7. Karasawa, N.; Dasgupta, S.; Goddard III W. A. *J. Phys. Chem*. **1991**, 95, 2261
8. Karasawa, N.; Goddard III, W. A. *Macromolecules* **1992**, 25, 7268.
9. Arimoto H., Ishibashi M., Hirai M. and Chatani Y., *J. Polym. Sci. A* 1965, 3, 317.
10. Abu Isa, I. *J. Polym. Sci., Polym. Chem. Edn* 1971, 9, 199.
11. Murthy, N. S., Szollosi, A. B., Sibilia, J. P. and Krimm, S. *J. Polym. Sci., Polym. Phys.* Edn 1985, 23, 2369.
12. Miyasaka, K. and Makishima, K. *J. Polym. Sci. A1* 1967, **5**, 3017.
13. Miyasaka, K. and Ishikawa, J. *J. Polym. Sci. A2* 1968, 6, 1317.
14. Hiramatsu, N. and Hirakawa, S. *Polym. J.* 1982, 14, 165.
15. Ziabicki, A. *Collect. Czech. Chem. Commun.* 1957, 22, 64.
16. Ziabicki, A. *Kolloid Z* **1959**, 167, 132.
17. Ziabicki, A. and Kedzierska, K. *J. Polym. Sci.* **1959**, 2, 14
18. Avramova, N. and Fakirov, S. Polym. Commun. 1984, 25, 27.
19. Roldan, L. G. and Kaufman, H. S. *Polym. Lett.* **1963**, 1, 603.
20. Illers, K. H.; Haberkorn, H. and Simak, P. Makromol. Chem. 1972, 158, 285.
21. Stepaniak, R. F.; Garton, A.; Carlsson, D. J. and Wiles, D. M. *J. Polym. Sci., Polym. Phys. Edn* **1979**, 17, 987.
22. Startsev, O. V.; Iordanskii, A. L. and Zaikov, G. Ye. Polym. Sci, *USSR* **1988**, 30, 1625.
23. Murthy N. S., *Polymer communications*, 1991, 32, 10, 301.
24. Auriemma, F.; Petraccone, V.; Parravicini, L. and Corradini, P. *Macromolecules* **1997**, 30, 7554.
25. Malta, V., Cojazzi, G., Fichera, A., Ajo, D., Zanetti, R. *Eur. Polym. J.* **1979**, 15, 765.
26. Salem, D. R.; Weigmann, H.-D., *Polym. Commun.* **1989**, 30, 336.
27. Murthy N. S., Minor H, *Polymer communications*, 1991, 32, 10, 297.
28. The energy difference between P1 and $P2_1$ is negligible. We find that $P2_1$ keeps the same shape for the two amide units in each chain in the simulated unit cell while P1 does not. Thus we considered only $P2_1$.

29. Simon, P., Argay, Gy. *J. Polym. Sci., Polym. Phys. Ed.* **1978**, 16, 935.
30. Leon, S.; Aleman, C.; Munoz-Guerra, S. *Macromolecules* **2000**, 33, 5754.
31. Sakurada, I.; Kaji, K. *J. Polym. Sci., C* **1970**, 31, 57.
32. Welch P. and Muthukumar M. *Phys. Rev. Lett.* **2001**, 87, 21: art. No. 218302.
33. Arimoto, H., *J. Polym. Sci. A* 1964, 2, 2283.
34. Murthy, N. S.; Szollosi, A. B.; Sibillia, J. P. and Krimm, S. *J. Polym. Sci., Polym. Phys. Edn* **1985**, 23, 2369.
35. Zhao Y. L. and Wu Y. D., *J. Am. Chem. Soc.* **2002** 124, 1571.
36. Hoffman, J. D. *SPE Transactions* **1964**, 315.
37. Jones, N. A.; Sikorski, P.; Atkins, E. D. T. and Hill M. J., *Macromolecules* **2000**, 33, 4146.
38. Trifan, D. S. and Terenzi, J. F., *J. Polymer Sci.* **1958**, 28, 443.
39. Bermudez, M.; Leon, S.; Aleman, C.; Munoz-Guerra, S.; *J. Polym. Sc. Polym. Phys. Edn* **2000**, 38, 41.

Table 1. The free energy (kcal/mol/amide unit), Young's modulus (GPa) of regular crystal structures of nylon 6 at 300 K

SD*14/b		7	6	5	4	3	2	1	0	-1	-2	-3	-4	-5	-6
EAP	F	50.61	50.99	50.62	50.77	50.63	50.76	50.66	51.22	50.79	50.75	50.71	50.78	50.66	50.01
α	E_Y	295.8	257.5	285.5	294.6	294.6	284.6	267.3	293.1	268.2	283.4	294.6	294.4	285.2	263.4
EAA	F	50.12	50.74	50.76	50.77	50.59	50.60	50.85	50.88	50.82	50.89	50.66	50.62	50.78	50.74
α	E_Y	294.9	295.2	276.1	281.2	300.2	301.1	243.5	245.7	241.7	238.8	300.4	300.7	277.7	279.2
EPA	F	50.84	50.69	50.49	50.84	50.74	50.50	50.79	50.69	50.83	50.39	50.58	50.48	50.86	50.97
δ	E_Y	277.8	282.4	103.4	277.8	290.9	274.2	266.9	137.5	171.7	253.2	282.5	115.5	278.2	228.0
EPP	F	51.29	50.77	50.62	50.54	50.55	50.54	50.81	51.06	50.74	50.54	50.56	50.56	50.62	50.77
	E_Y	299.3	259.1	224.5	226.0	238.1	238.9	220.0	292.8	218.2	238.9	243.8	226.2	233.9	259.0
TAP	F								50.54						
γ	E_Y								135.0						
TPP	F								50.47						
	E_Y								153.6						

[a] There are 8 amide units in the simulation unit cell as shown in Figure 2.
[b] SD is the sheet displacement. For EPA and EAA, the definition of SD is a little tricky. We invert the adjacent sheet first and then define the sheet displacement. b as the length of the unit cell in chain direction.
[c] The free energy at 300K is the sum of the zero point energy (ZPE), the potential energy and entropy at 300K.
[d] Young's modulus (GPa) is in the chain direction.

Table 2. The unit cell parameters for various forms of nylon 6 crystal. Simulated at 0 K and experimental strrucrues at 298 K

Crystal form	A/Å	b/Å	c/Å	β (angle)
EAP-crossed	9.41	17.68	8.11	65.1°
EAP-uncrossed	9.53	17.67	7.87	68.1 °
EAA-crossed	9.49	17.68	8.07	65.0°
EAA-uncrossed	9.53	17.67	8.06	65.1°
Exp. α [6]	9.56	17.24	8.01	67.5°
TPA	4.97	17.33	8.68	127.5°
Exp. γ [9]	4.78	16.88	9.33	121°
EPA-crossed	4.72	17.67	7.91	114.6°
EPA-uncrossed	4.77	17.67	8.75	122.0°
Exp. δ [23]		17.2		

Table 3. The energy for the H-bond schemes shown in Fig. 7 using the MSXX FF[a]

E (kcal/mol/amide unit)	H-bond I	H-bond II	H-bond III
Val	30.644	30.592	30.656
VdW	4.321	3.388	3.217
Electrostatic	-70.005	-69.605	-69.534
Total	-35.040	-35.625	-35.661

[a] The c length of unit cell is restricted to 50 Å.

Table 4. The calculated (020) and (040) intensities/relative ratios for different sheet displacements of EAP, compared to experiment (ref. 6)

SD *14/b	0	-1	-2	-3	-4	-5	-6	Exper
020	60.3/6.1%	49.5/5.1%	25.3/2.6%	3.1/0.3%	2.2/0.2%	23.3/2.4%	47.5/5.0%	1/0.1%
040	20.6/2.1%	8.7/0.9%	0.5/0.1%	17.2/1.8%	17.9/1.9%	1.2/0.1%	6.6/0.6%	6/0.6%
SD *14/b	1	2	3	4	5	6	7	
020	49.0/5.0%	21.6/2.2%	2.9/0.3%	3.9/0.4%	23.9/2.4%	50.5/5.1%	60.7/6.1%	
040	8.0/0.8%	1.8/0.2%	17.8/1.8%	15.6/1.6%	0.9/0.1%	9.5/1.0%	21.6/2.2%	

Table 5. The unit cell parameters of nylon 6 lamella crystal after minimization at 0 K

Fold type of lamella	A/Å	b/Å	c/Å	β (angle)
Amide I	9.42	53.58	8.45	64.5°
Amide II	9.43	54.10	8.03	72.9°
Alkane I	9.38	44.03	8.11	70.0°
Alkane II	9.37	44.20	8.06	73.4°
Inter-sheet	9.47	53.74	8.24	64.8°

Table 6. The energy cost different types of fold in individual molecule and 3D lamella crystal (SD=$3/_{14}$b) (kcal/mol/fold)

Fold type	Amide I	Amide II	Alkane I	Alkane II	Inter-sheet
Individual molecule	10.625	18.194	8.164	19.202	25.173
3D lamella	16.432	25.335	12.107	28.790	6.809

Chapter 2

CRYSTAL STRUCTURE AND PROPERTIES OF N6/AMCC COPOLYMER FROM THEORY AND FIBER XRD*

ABSTRACT

The MSXX force field developed previously from *ab initio* quantum calculations for studies of nylon are used to study the crystal structure and properties of the copolymer of nylon 6 with AMCC (4-aminomethylcyclohexanecarboxylic acid). For the isolated chain conformation of the copolymer, we consider both axial and equatorial connections of the chain with the cyclohexane ring and find that the best is Chair-ee-St, which has equatorial connections on both ends of chair cyclohexane. We consider 12 possible crystal structures for the copolymer (the best four conformations of the isolated chain with the three forms of packing these chains: α form, γ form, and δ form). With 12.5 % of AMCC in the copolymer, we find that γ form with the Chair-ee-St chain structure is the most stable, even though α form is most stable for nylon 6. The calculated X-ray diffraction patterns of the predicted crystal structure fit both equatorial and meridional scans of XRD very well. There are two reasons that make α form less stable for the copolymer. One is the bad contact between the axial hydrogen atoms of the cyclohexane ring and the CH_2 hydrogens. The other is the difficulty of intra-molecular H-bonds in the copolymer.

The predicted chain-axis repeat distance of the copolymer (0K) is 1.4Å smaller than for the α form of Nylon 6, in good agreement with the X-Ray results, which indicates that it is 1.5 Å smaller (at 300K).

The Young's modulus in the chain direction is calculated to be 93 GPa for the copolymer (at 0K), which compares to 135 and 295 GPa predicted for γ form and α form nylon 6,

* Published in Y. Li, W. A. Goddard III, N. S. Murthy, *Macromolecules*, **2003**, *36*, 900-907.

respectively. The indroduced cyclohexane ring locates between the two amide pockets of the adjacent hydrogen bond sheets and has two major effects on the properties of the copolymer:

i) It causes twisted conformations, which decreases the Young's modulus of the copolymer in chain direction.

ii) It makes the chain rigid, which likely is responsible for the decrease in sensitivity of the copolymer to moisture

1. Introduction

Copolymers provide a robust and economical route to enhance the performance range of a polymer. "Isomorphous" replacement of ε-aminocaproic acid residues in nylon 6 by 4-aminomethyl-cyclohexancecarboxylic acid (AMCC) has been reported by several authors.[1-5] The nylon 6 lattice can accommodate less than 30 mole-% AMCC residues before a new structure appears.[5] Cis-AMCC isomerizes to the more thermodynamically stable trans isomer during copolymerization with caprolactam (CL) and also during homopolymerization.[5]

Nylons crystallize in two crystalline forms commonly referred to as α and γ. In the α form the hydrogen bonds are between antiparallel chains while in the γ form they are between parallel chains.[6] The α form is stable in nylon 4 while the γ form is stable for nylon 8 and above.[7] Nylon 6 is unique in that it is observed to crystallize easily in either crystalline forms and can be transformed from one to another.[8]

In this paper we examine the effect of geometric configuration on the ability of 4-aminomethylcyclohexanecarboxylic acid (AMCC) to "isomorphously" replace ε-aminocaproic acid residues in nylon 6. Here we present detailed studies of the structure for

the copolymer using both molecular modeling and X-ray diffraction. We conclude that just 12.5% of AMCC in the copolymer is sufficient to change the energetic preference of nylon 6 from the α form to the γ form. We will use the theory to analyze why this change occurs.

2. Calculation details

We use the MSXX FF[7] with the MSC version of PolyGraf (version 3.30, Caltech version) for all calculations. Cerius2 (v4.0) was also used for graphics and manipulations. The electrostatic and van der Waals (vdW) interactions use the accuracy bounded convergence acceleration (ABCA) Ewald technique[9] for computing the nonbond energies of periodic systems. We use an accuracy of 0.001 kcal/mol. All structures are minimized to an rms force on all atoms of 0.01 kcal/(mol Å) for atom and rms stresses of 0.1 kcal/(mol Å) using conjugate gradient method.

2.1. Force field

The MSXX force field for simulation of nylon polymers was derived from *ab initio* QM calculations.[7] Special emphasis was given to the accuracy of the hydrogen bond potential for the amide unit and the torsional potential between the peptide and alkane fragments.

This hydrogen bond potential was derived from MP2/6-31G** calculations of the formamide dimer. Subtracting electrostatic interactions (based on fixed-point charges extracted from QC on the monomers) leads to a repulsive exponential form (Eq. 1) of the short-range hydrogen bond potential[7] with A=0.028 kcal/mol, C=0.251 Å, and R_e=3.017 Å. Instead of the original charge scheme in Ref. [7], we now use the improved charge scheme for nylon 6 from Ref. [10] (see section 2.2). The difference of the two charge schemes comes from the methylene groups, which does not affect the parameters used for hydrogen bond potential.

$$E_{vdW}^{EXP} = A\exp\left[-\frac{(R-R_e)}{C}\right] \quad (1)$$

The full torsion potential between peptide and alkane fragment was calculated by optimizing the geometry (using HF/6-31G**) at each point on the torsional curve and the torsional potential is represented by a Fourier series (Eq. 2) in MSXX force field.

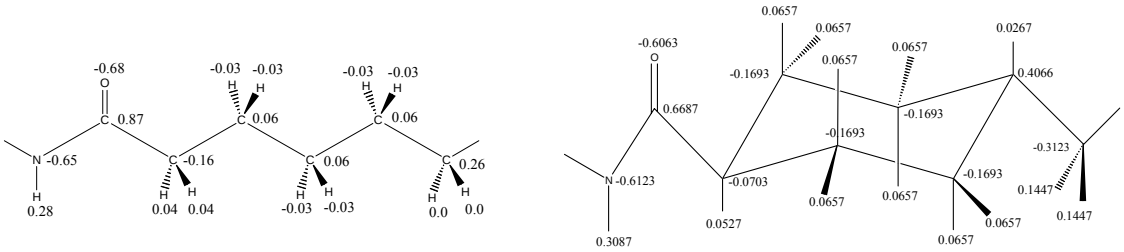
$$E_{torsion} = \frac{1}{2}\sum_{n=0}^{n=6} V_n \cos n\tau \quad (2)$$

where: τ is the torsional angle (τ=0 for cis), and V_n is the barrier (energy of cis over trans).

Detailed MSXX force field was described in previous paper.[7]

2.2. Charges

We use potential derived charges (PDQ) based on quantum mechanical calculations (HF/6-31G**) of model systems. These charges are based on calculations for long alkyl chains functionalized with an amide linkage, where a minimum of five carbons to either side was required for charge convergence.[10] Based on a series of calculations for shorter alkane chains functionalized with an amide, the charge perturbation within a long alkane chain due to each functional unit was extracted.[10] We use the same method to get the potential derived charges (PDQ) for AMCC monomer by fitting the total charges and the dipole moment fragment. The charge schemes for nylon 6 and AMCC are summarized in Figure 1a and 1b.



a. nylon 6 monomer b. AMCC monomer

Figure 1. Charge schemes for nylon 6 monomer and AMCC monomer fragments in copolymer chain

2.3. Vibrational calculations

The analytic second derivative matrix (Hessian) obtained directly from the complete energy expression was used to calculate the elastic constants (including Young's Modulus). The methodologies are reported in ref [11,12] and implemented in the VIBRATE, THERMO, and ELASTICA Modules in PolyGraf.

2.4 X-ray diffraction calculation

We use the "Diffraction-Crystal" module in Cerius2 4.0 to calculate the fiber x-ray diffraction intensities and to obtain diffraction pattern to compare with the experimental data. The intensity for each hkl reflection was calculated using:

$$I(hkl) = \left\{\sum f_n \cos 2\pi(hx_n + ky_n + lz_n)\right\}^2 + \left\{\sum f_n \sin 2\pi(hx_n + ky_n + lz_n)\right\}^2 \quad (3)$$

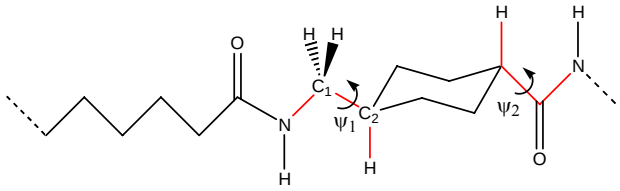
where f_n is the scattering factor of atom n and x_n, y_n, z_n are the fractional coordinates of atom n. The summation is over all atoms in the unit cell.

No polarization factor, crystal monochromator factor or temperature factor are applied to the intensity calculation.

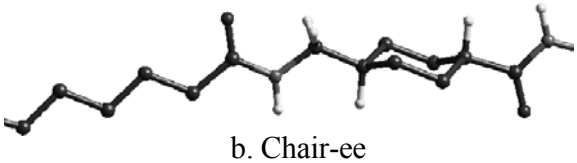
3. Results and discussions

3.1 Isolated chain conformation of copolymer

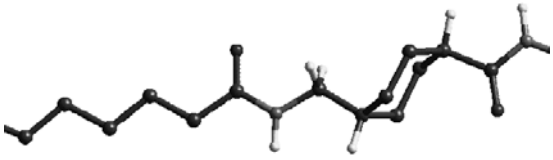
Before discussing the crystal structure of copolymer, we first analyze the isolated chain conformation of copolymer.



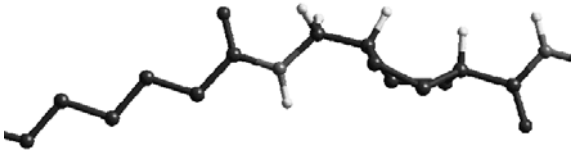
a. Copolymer of nylon 6 with AMCC with definitions of chain-ring torsions Ψ_1 and Ψ_2 (Here Ψ_1=0 and Ψ_2=0)



b. Chair-ee



c. Chair-aa



d. Boat-ee

Figure 2. Three ring conformations of cyclohexane in the copolymer chain

To maintain the roughly straight chain for the copolymer required for good crystal packing, there are three important ring conformations for the cyclohexane as shown in Figure 2bcd.

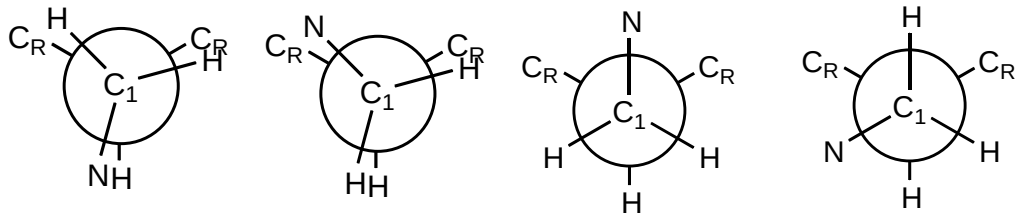
"Chair-ee" as shown in Fig. 2b: The cyclohexane ring of chair conformation uses two equatorial bonds to connect chain ends.

"Chair-aa" as shown in Fig. 2c: The cyclohexane ring of chair conformation uses two axial bonds to connect chain ends.

"Boat-ee" as shown in Fig. 2d: The cyclohexane ring of boat conformation uses two equatorial bonds to connect chain ends. The minimized structure has a twist boat conformation instead of regular boat. Twisted boat is local minimum, while regular boat is a transition state. See Figure 1.9 in Ref 13.

In addition we could construct mixed axial-equatorial structures. However all of these lead to the chains leaving the cyclohexane ring at right angles, making it less likely to pack well.

Besides the three possible ring conformations of the cyclohexane, the two bonds, which the cyclohexane ring uses to connect chain ends, may have different torsion values. Those two torsions are indicated as chain-ring torsions Ψ_1 and Ψ_2 as indicated in Figure 2a.



a. E1 $\Psi_1=0, \Psi_2=0$

b. E2 $\Psi_1=-120^\circ, \Psi_2=120^\circ$

c. Sg $\Psi_1=-180^\circ, \Psi_2=180^\circ$

d. St $\Psi_1=-60^\circ, \Psi_2=60^\circ$

Figure 3. The four possible sets of chain-ring torsion Ψ_1 (Ψ_1 and Ψ_2 are varied simultaneously and $\Psi_1=-\Psi_2$)

To keep the chain segment as shown in Figure 2 a good repeat unit, we vary the torsions Ψ_1 and Ψ_2 simultaneously. For instance, changing Ψ_1 by an angle such as $+60^o$ will make Ψ_2 changed by -60^o simultaneously. Figure 3 shows the four possible chain-ring torsions of Ψ_1. "E", "S", "g", and "t" stand for "Eclipsed", "Staggered", "gauche", and "trans", respectively.

Combining the three ring conformation with four chain-ring torsions of Ψ_1 leads to a total 12 isolated chain conformations of copolymer.

We set the chain segment as shown in Figure 2 as a repeat unit in the unit cell and fixed the distance between two adjacent chains at 50Å to avoid chain interactions. Torsions except Ψ_1 and Ψ_2 in the chain are fixed at 180^o. Figure 4 and Table 1 show the minimized relative energy of 12 chain conformations from MSXX force field.

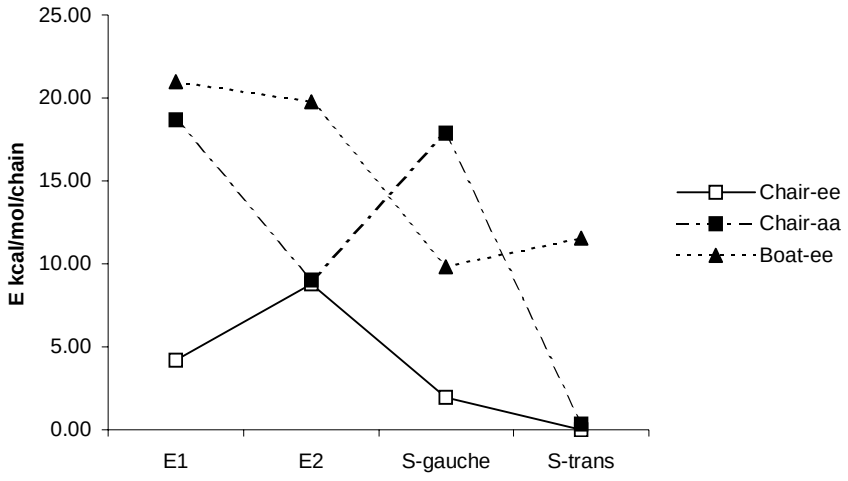


Figure 4. Relative energy of 12 isolated chain conformations

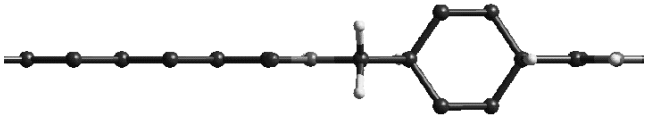
By analyzing the energy component, we find that the energy difference among those conformations mainly comes from the torsion part. We list the torsion part energy in Table 1.

In addition, the repulsion between axial groups of cyclohexane decreases the stability of isolated chain. The repulsion in Chair-aa-Sg becomes significant, making its energy to high by 17.89 kcal/mol.

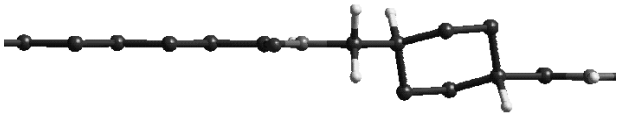
We find that with the chains attached energy difference between chair and twisted boat is 9.82 kcal/mol/chain (see Table 1), which is bigger than the value of 6.9 kcal/mol for cyclohexane.[14] This increase arises from steric interactions of the substituted groups, which is consistent with the conclusion in Ref [13].

The chain-axis repeat distance of 12 isolated chain conformations are listed in Table 1. Chair-ee-E1 has the longest chain and similar chain shape as full-extended chain in α form of nylon 6. Chair-aa-Sg has the shortest chain and it causes the great repulsion among axial groups.

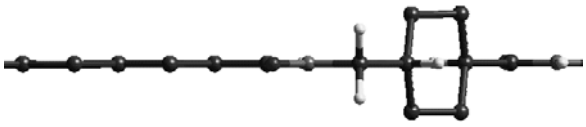
In conclusion, the four best conformations among all 12 possibilities of isolated chain are Chair-ee-St, Chair-ee-Sg, Chair-aa-St, and Chair-ee-E1. Although Chair-ee-E1 is not energy favorable compared with the other three, it keeps the longest chain and is most compatible with the full-extended nylon 6 chains. We will consider it in the following discussion. Those four conformations of copolymer chain and their relative energy are shown in Figure 5.



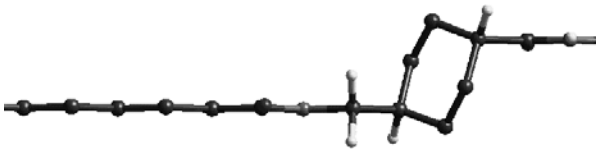
a. Chair-ee-E1(E=4.18kcal/mol/amide unit)



b. Chair-ee-St(E=0.0kcal/mol/amide unit)



c. Chair-ee-Sg(E=1.97kcal/mol/amide unit)



d. Chair-aa-St(E=0.32kcal/mol/amide unit)

Figure 5. The four best chain conformations of N6/AMCC copolymer

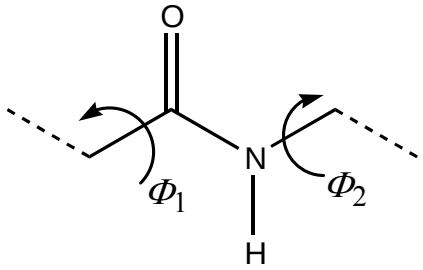


Figure 6. Definitions of amide torsion parameters: torsion Φ_1 [C-C(amide)] and torsion Φ_2 [N(amide)-C]

3.2 Crystal structure of copolymer from molecular modeling

The single bond torsion potentials, C-C(amide) torsion Φ_1 and N(amide)-C torsion Φ_2 as defined in Figure 6, are particularly important for nylon. The primary difference between

the α form and γ form of nylon 6 comes from these two dihedrals. We did systematic study[8] of crystal structures of Nylon 6 before this study of copolymer and we found three types of regular crystal structures: α form, γ form, and δ form. Similarly as the study of nylon 6, we denote the definitions of α form, γ form, and δ form of copolymer as following:

α form: $\Phi_1 \approx \Phi_2 \approx \pm 166^o$, and hydrogen bonds are formed between antiparallel chains.

γ form: $\Phi_1 \approx \Phi_2 \approx \pm 127^o$, and hydrogen bonds are formed between parallel chains.

δ form: $\Phi_1 \approx \Phi_2 \approx \pm 166^o$, and hydrogen bonds are formed between parallel chains.

The Φ_1 and Φ_2 of the same chain have opposite signs (e.g. $\Phi_1 \approx +127^o$, $\Phi_2 \approx -127^o$) to keep the straight chain

One copolymer chain of two amide units as shown in Figure 5 and three chains of nylon 6 are used to construct crystal structure of copolymer, whose mole fraction is 12.5%. We will consider α form, γ form, and δ form together with four conformations of copolymer chain as shown in Figure 5. There are totally 12 possible crystal structures.

Unlike nylon 6 homopolymer, whose hydrogen bond sheet can slide easily over the adjacent sheet,[8] the hydrogen bond sheet in copolymer cannot slide due to the large size of cyclohexane ring. As shown in Figure 11, the cyclohexane ring requires significant space and wants to be adjacent to the amide unit of the adjacent hydrogen bond sheet. (See the Amide pocket model in Ref.[8]) The favored energy is 0.8 kcal/mol/(amide unit) to fit the cyclohexane ring in the Amide Pocket of adjacent hydrogen-bond sheet.

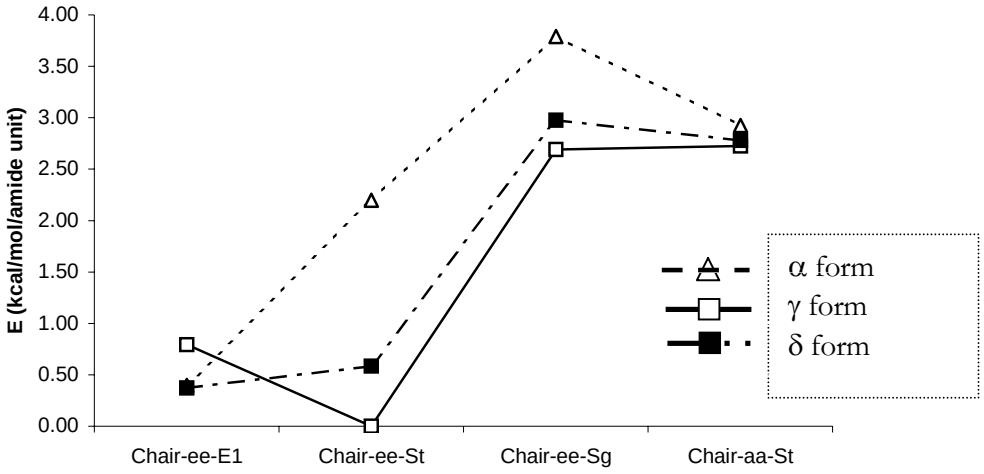


Figure 7. Relative energy (kcal/mol/amide unit) of the 12 crystal structures. For the three crystal forms (α,γ,δ) with four chain conformations (Chair-ee-E1, Chair-ee-St, Chair-ee-Sg, Chair-aa-St)

For the three crystal forms (α,γ,δ) with four chain conformations (Chair-ee-E1, Chair-ee-St, Chair-ee-Sg, Chair-aa-St)

Figure 7 and Table 2 show the relative energy of 12 crystal structures. The chain-axis repeat distances are listed in Table 2.

In general, α form is worse than δ form and δ form is slightly worse than γ form. The reasons are discussed in section 3.5.

The relative packing energies of Chair-ee-St, Chair-aa-St, and Chair-ee-Sg have the same trend as for the isolated chain conformations. However, the unstable isolated chain conformation Chair-ee-E1 has a good packing energy relative to other isolated chain conformations. As mentioned previously, Chair-ee-E1 leads to the longest chain conformation with a chain shape similar to the full-extended chain of nylon 6, and hence is anticipated to be most compatible in nylon 6 matrix.

In conclusion, the isolated chain conformation of Chair-ee-St is the best (see section 3.1) and the γ form of Chair-ee-St is the best crystal structure for the copolymer.

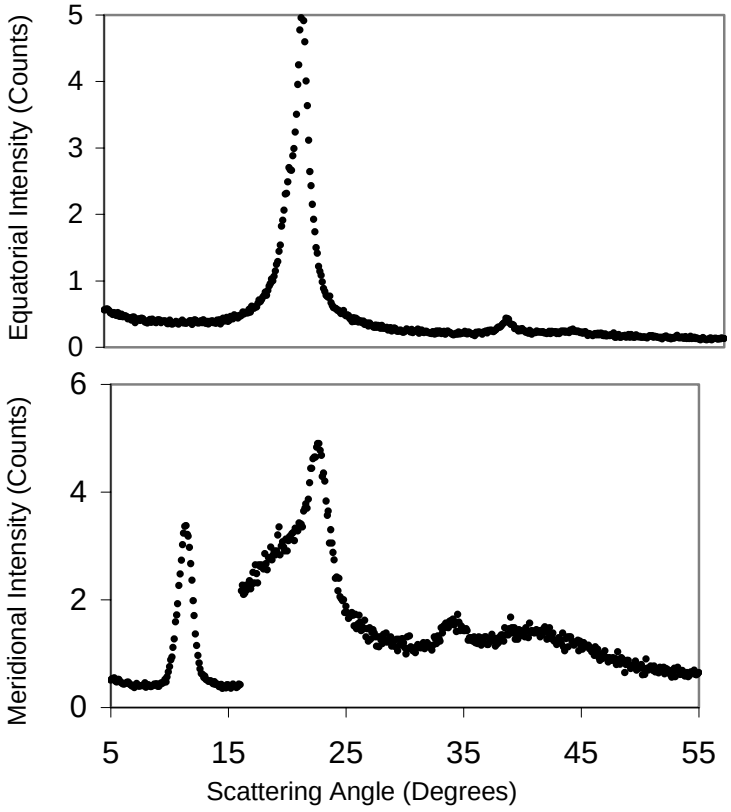


Figure 8. Experimental Equatorial and Meridional XRD scans of copolymer fiber

3.3 Crystal structure of copolymer from XRD

Figure 8 shows XRD scans from drawn/heat-set (annealed) fibers from the copolymer. The equatorial scan of the copolymer fiber is typical of the γ crystalline form. For nylon 6, the unit cells of the α and γ forms are different, and therefore the equatorial reflections of the α and γ forms appear at slightly different angles. The two intense reflections of the α and γ forms are at, respectively, 21° (200, α_1) and 24° (002 + 202, α_2), and 22° (100, γ_1) and 23°

(201 + 200, γ_2).[15] See XRD scans of nylon 6 in Figure 2 in Ref. [15]. The equatorial scans of the undrawn fibers of N6 homopolymer and N6/AMCC copolymer are quite similar,[16] indicative of the presence of γ form in both the fibers. While N6 transforms into the α form upon heat-setting, the copolymer remains in the γ form,[16] which is consistent with the molecular modeling result. Detailed discussion is in part 3.5.

The meridional scan of the copolymer fiber is similar to the N6 homopolymer. (Compare Figure 8 with Figure 2 in Ref. [15]) The significant difference is that the (0 k 0) reflections show clearly that the chain-axis repeat is shorter in the copolymer than in N6. The chain-axis repeat distance is 15.7 Å in the copolymer compared to 17.2 Å for α and 16.8 Å for the γ crystalline forms of nylon 6. See part 3.6 for further discussion.

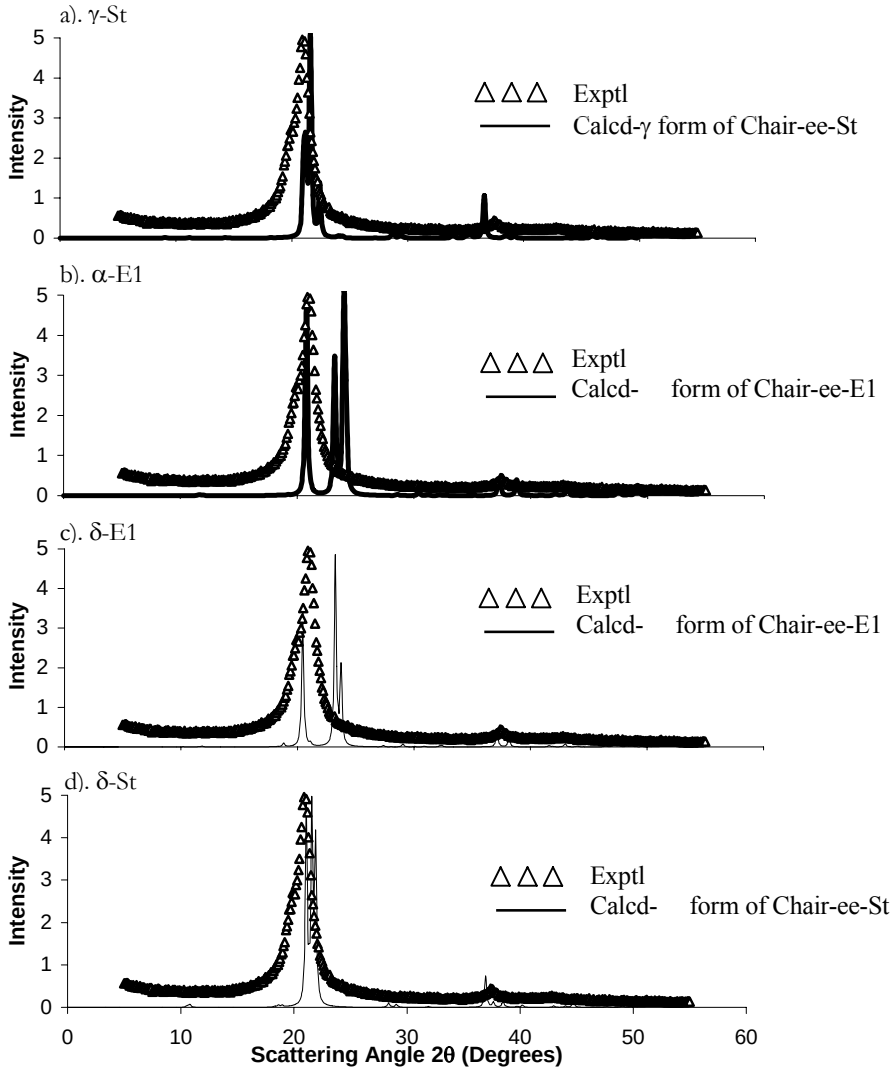


Figure 9. Calculated equatorial scan of 12.5% N6/AMCC copolymer Compared with experimental equatorial scan of 10% N6/AMCC annealed fiber; (a) Calculated γ form of Chair-ee-St, (b) Calculated α form of Chair-ee-E1, (c) Calculated δ form of Chair-ee-E1, (d) Calculated δ form of Chair-ee-St (Only γ-St and δ-St fit the experimental equatorial scan pattern)

3.4 Calculated X-ray diffraction patterns

From the results in Section 3.2, the best crystal structure is the γ form of Chair-ee-St. The second best is the δ form of Chair-ee-E1, while the third is the α form of Chair-ee-E1 and the fourth is the δ form of Chair-ee-St. These four best structures include the best structure in each of α, γ, δ forms. We calculated the X-ray diffraction patterns of these four structures to compare with the experimental diffraction patterns.

Figure 9 shows the calculated X-ray diffraction patterns of equatorial scan for these four structures of 12.5% N6/AMCC copolymer. (Note that the maximum intensity of calculated has been rescaled to be the same as experimental. See part 2.4 for calculation details.) The calculated equatorial scan of γ form Chair-ee-St and δ form Chair-ee-St both fit the observed fiber XRD equatorial scan well, while Chair-ee-E1 α form differs significantly. Indeed, the equatorial scans from the α and γ forms of N6/AMCC copolymer are similar to those from the α and γ forms of homopolymer nylon 6. (Compare Figure 9 with Figure 2 in Ref. [15]).

The equatorial scan does not distinguish between the γ form Chair-ee-St and the δ form Chair-ee-St. Both have similar equatorial scans as the experimental diffraction pattern. Consequently, we calculate the meridional scan for those structures to make further comparision with experiment. For ideally oriented fibers in which the *c** axis coincides with the fiber axis, the reciprocal-lattice points with indices (00*l*) (for monoclinic unit cell) cannot intersect the sphere of reflection. Thus no reflections appear on the meridian of the diffraction pattern unless the fiber is somewhat inclined with respect to the direct beam. However, Buerger-type precession cameras have the advantage of including the meridional reflections, which in principle are not accessible to the stationary flat-film techniques. Consequently, we calculate the precession pattern instead of stationary flat-film pattern.

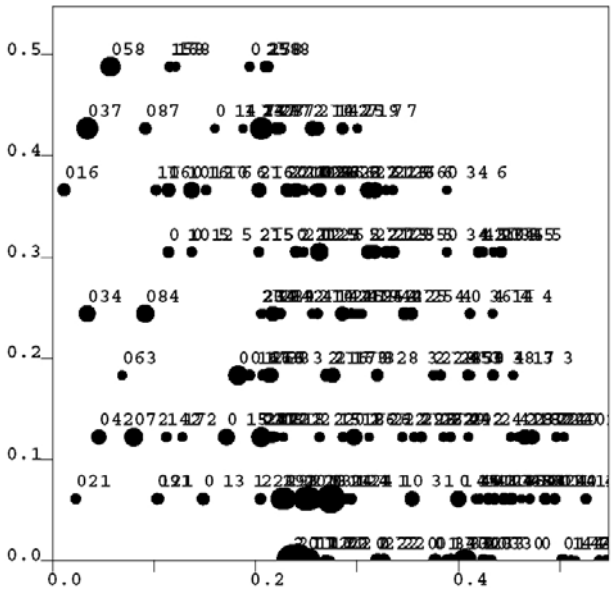


Figure 10. The quadrant precession diffraction pattern of γ form of Chair-ee-St calculated from Cerius2. (The intensities are indexed from the reindexed unit cell shown in Figure 12.)

Figure 10 shows the quadrant precession diffraction pattern of the γ form of Chair-ee-St calculated from Cerius2. Figure 10 shows that the absolute meridional scan from precession pattern is still zero. The calculated precession patterns of the other three structures, which are not shown here, are similar and the absolute meridional scans are zero. The reason is that our minimized crystal structures are triclinic instead of monoclinic. There are 44 chains in the re-indexed monoclinic unit cell and they neutralize the intensities (00*l*) to zero. In the experiment, the actual fibers frequently possess a sufficiently wide spread of orientations about the ideal fiber axis to generate meridional reflections. From this point, we estimate a toleration of 0.05 (reciprocal angstroms) along equatorial direction and we assume that the diffraction intensities inside 0 ± 0.05 (reciprocal angstroms) along equatorial direction are observable in the experimental meridional scan. In this way, we get the "meridional" intensities of these four structures as shown in Figure 11. (The maximum intensity of calculated has been rescaled to experiment as in Figure 9.)

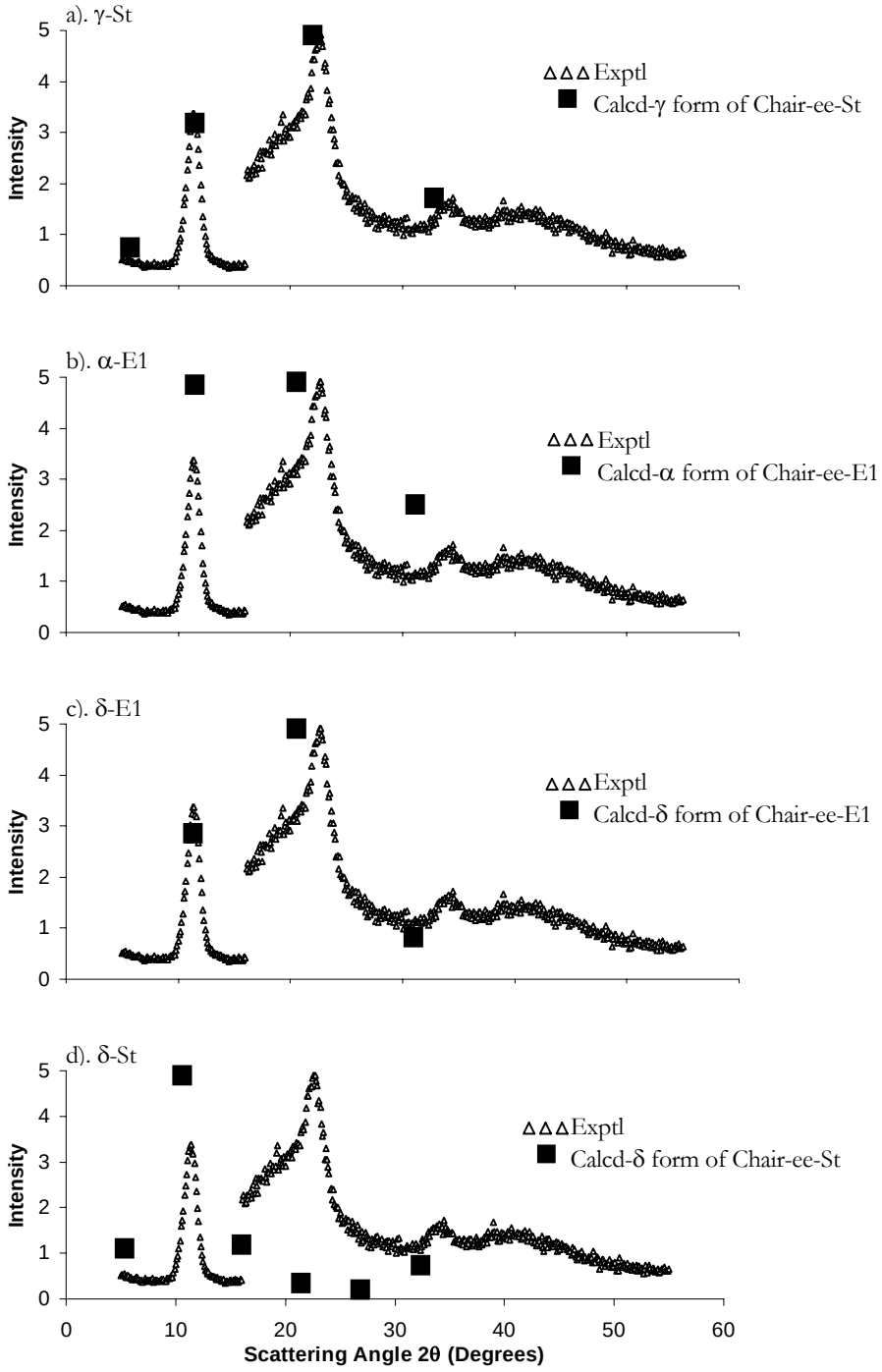
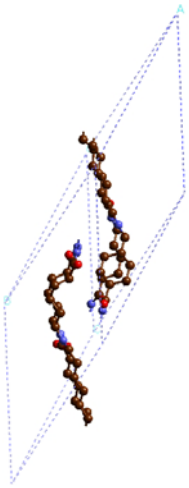


Figure 11. Calculated meridional intensities of 12.5% N6/AMCC copolymer Compared with experimental equatorial scan of 10% N6/AMCC annealed fiber; (a) Calculated γ form of Chair-ee-St, (b) Calculated α form of Chair-ee-E1, (c) Calculated δ form of Chair-ee-E1,

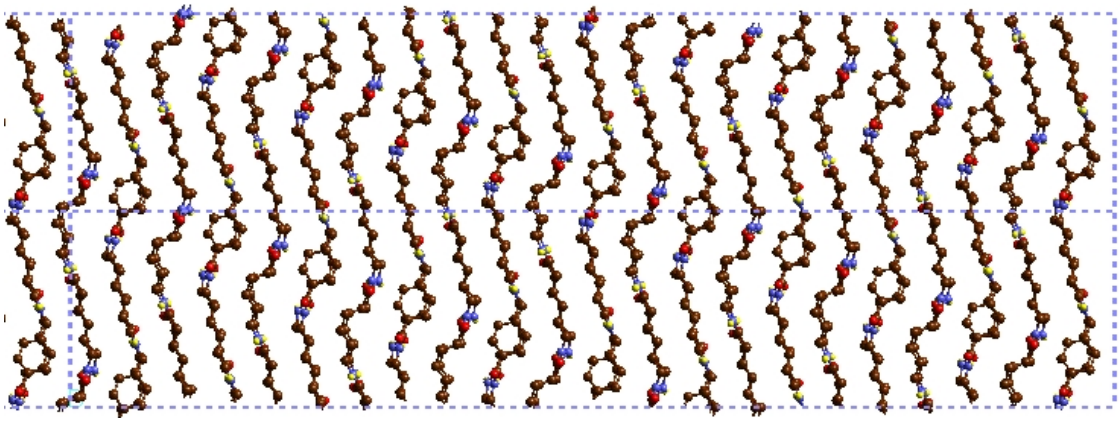
(d) Calculated δ form of Chair-ee-St (Only γ form of Chair-ee-St gives meridional intensities fit XRD pattern)

From Figure 11, we see that only the γ form of Chair-ee-St fits the experimental meridional scan. Thus although the calculated equatorial scan of δ form of Chair-ee-St fits the experimental equatorial scan very well, its calculated meridional intensities differ from the experimental meridional scan significantly.

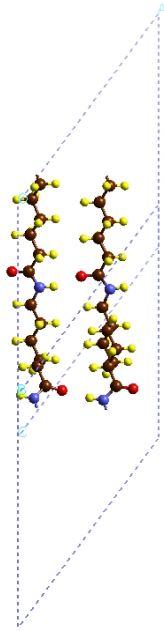
In conclusion, by comparing the calculated equatorial scan and meridional intensities of the best four candidate structures γ form of Chair-ee-St, α form of Chair-ee-E1, δ form of Chair-ee-E1, and δ form of Chair-ee-St, we find only γ form of Chair-ee-St fits the experimental patterns very well, while the other three differ substantially. This confirms our calculated energy results in Part 3.2.



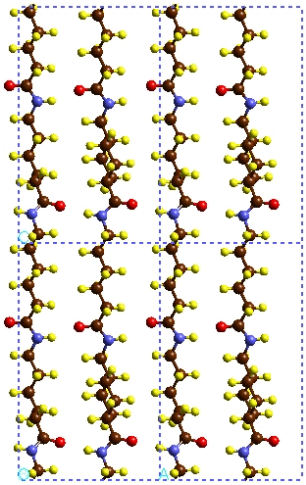
12*a*. Original triclinic unit cell projected in H-bond direction



12*c*. Reindexed monoclinic unit cell projected in H-bond direction



12*b*. One H-bond sheet in the original triclinic unit cell

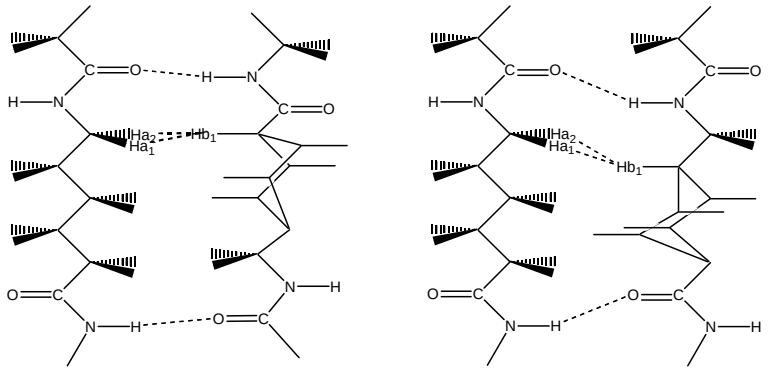


12*d*. One H-bond sheet in the reindexed monoclinic unit cell

Figure 12. Crystal structure of γ form of Chair-ee-St (*a* and *b* are visualizations of original triclinic unit cell, whose cell parameters are *a*= 18.25 Å, *b*=15.56 Å, *c*=16.37 Å, α=30.57°, β=136.88°, γ=147.50°; *c* and *d* are visualizations of re-indexed monoclinic unit cell, whose cell parameters *a*= 9.81 Å, *b*=87.12 Å, *c*=16.37 Å, γ=88.80°) (See comment [18])

3.5 Relative stability of different crystalline forms of copolymer

The homopolymer of nylon 6 can be produced either in the α form, a mixture of α and γ crystalline forms, or in the γ form depending on the spin/draw conditions.[15] In contrast, the copolymer exists only in the γ form, and retains the short chain-axis repeat distance of this γ form even after annealing in an autoclave.[16] IR spectra confirm that conformation in the copolymer in the copolymer is the same as that in the γ form in the N6 homopolymer.[16]



a. H-bond between antiparallel chains

b. H-bond between parallel chains

Figure 13. Hydrogen bonds between antiparallel chains or parallel chains

For the copolymer of N6/AMCC, the α crystalline form is not favorable, even though it is the most favorable for the nylon 6 homopolymer. There are two reasons:

1. In the α form, the hydrogen bonds are in the plane of the trans alkane segment. The hydrogen bonds in the α form are formed between antiparallel chains (Fig. 13a). In the γ and the δ forms, the hydrogen bonds are formed between parallel chains (Fig. 13b). As discussed in part 3.6 of ref.[8], unlike in the γ and δ form, there are unfavorable contacts between the hydrogen atoms in α form of N6. This situation is worsened for the copolymer because of the large size of the cyclohexane ring; the hydrogen atoms on the

cyclohexane ring are in contact with the CH2 atoms on the adjacent chains, and this is not desirable. From Figure 12*a* and 12*c*, we can see that the cyclohexane ring in γ form of Chair-ee-St locates between two amide pockets (See Ref 8 for definition) avoiding the bad Hydrogen contacts.

2. Intra-molecular H-bonds are hard to form due to large loop strain of the copolymer chain.

We consider a covalently connected polymer chain as one molecule.[8] As discussed in section 3.4 of Ref.[8], for the α form, hydrogen bonds are formed *intra* molecularly, while in γ or δ forms, they are formed *inter* molecularly. *Intra* Molecule H-bonds in the α form are entropically favored and dynamically favored from solution.[8] Assuming 90% hydrogen bonds are formed in amorphous region and most of them are *inter* Molecule H-bonds, the γ form is kinetically favored from the melt. Annealing the γ form of nylon 6 transfers it to the thermo-stable α form with *Intra* Molecule H-bonds. Thus as-spun fibers of N6 are in the γ form, while drawn/heat-set (annealed) fibers are in the α form.[15,16]

There is a ring constraint in the loop part of Intra-molecular H-bonds, while there is not such a ring constraint in Inter-molecular H-bonds.[8] This competing factor favors Inter-molecular H-bonds in nylon 6 and becomes more significant in N6/AMCC copolymer due to the rigidity of copolymer chain.

The thermal characteristics and the effect of moisture[16] on the copolymer suggests that the copolymer chain is *more* rigid than the N6 homopolymer chain. It is easy to understand that a cyclohexane ring is more rigid than pentamethylene. Thus in the copolymer, intra-molecular H-bonds are much harder to form due to large loop strain of the copolymer and rigidity of cyclohexane than N6 homopolymer, which makes the α form of copolymer unfavorable.

3.6 Short repeat distance for copolymer

The chain-axis repeat of the copolymer is found to be 15.7 Å from fiber X-ray diffraction, (See Figure 8) shortest ever observed for nylon 6. The reported chain-axis repeats for the α, γ, and the complex of iodine/N6 are 17.2, 16.8 Å, and 15.8 ~ 16.0 Å, respectively.[16,17] This short repeat is observed at co monomer concentration as low as 10-15 mol %.

Based on the results of parts 3.1 and 3.2, we attribute this unusual short chain-axis repeat distance to the twisted chain conformation of copolymer. The best isolated chain conformation of copolymer is Chair-ee-St, which is *not* fully extended (due to the eclipsed bonds of the cyclohexane and the adjacent chain) as shown in Figure 5b. In addition, the torsions Ψ_1 and Ψ_2 are twisted and packed as γ form (See part 3.2). As shown in Figure 12, the N6 chains adjacent to the cyclohexane ring adjust their shapes to get good packing energy.

We calculate (0K) a chain-axis repeat distance of 16.3 Å for the best structure, γ form of Chair-ee-St of copolymer, which is 1.4 Å shorter than our calculation for α form N6 and 1.0 Å shorter than our calculation for γ form N6. From Table 3, we see that chain-axis repeat distance of copolymer from fiber X-ray is 1.5 Å shorter than experimental α form N6 and 1.1 Å shorter than experimental γ form N6. There exist systematic differences between the experimental chain-axis repeat distance from XRD (300K) and the theoretical value (0K). This systematic difference arises from disorder in the chain conformations at room temperature (experiment) compared to the 0K of theory. The fact that the actual fibers frequently possess a sufficiently wide spread of orientations about the ideal fiber axis will also decrease the observed chain repeat distance from the meridional scan of XRD. In addition imperfections and folding in the crystal experiment will generally decrease the chain-axis repeat distance. Table 3 shows that the chain-axis repeat distance from theory is consistent with experimental result.

3.7 Young's modulus for copolymer

The Young's modulus in chain direction is evaluated by the method described in part 2.3. The results are listed in Table 2. The Young's modulus of the best structure, γ form of Chair-ee-St is calculated to be 93 GPa, which is lower than the α form and γ form of N6 homopolymer (295 GPa and 135 GPa, respectively[8]). As discussed in parts 3.1, 3.2 and 3.6, there is a slight twist in the chain to accommodate the eclipsing of CH bonds, which gives the chains in the copolymer a slightly unextended character. This twisted chain conformation decreases the Young's modulus of the copolymer in chain direction.

This is considerable less for the α,δ forms of N6/AMCC copolymer. We note that the chain axis repeat for α,δ are 7% longer than for γ. Thus that extending the chains in the chain direction by tension might stabilize α,δ. However as discussed in part 3.5 and Ref 8, the loops in the α form are not comparable as in γ and δ and the situation is worsen when cyclohexane is introduced. This suggests that the γ->α transition of copolymer is harder than of pure nylon 6.

The introduce of AMCC core into nylon 6 matrix leads directly to two results:

The chain conformation in the favorable packing structure becomes twisted instead of straight, which decreases the Young's modulus of the copolymer in chain direction.

It makes the chain rigid, which likely is responsible for the decrease in sensitivity of the copolymer to moisture. Thus the wet-modulus of the copolymer is higher than that of the homopolymer, even though dry modulus of the copolymer is a little lower than the homopolymer. (See Table 1 in Ref.[16])

4. Summary

The MSXX force field developed previously from *ab initio* quantum calculations for studying nylon is used here in conjunction with and X-ray diffraction to determine the

crystal structure of copolymer of nylon 6 with AMCC (4-aminomethylcyclohexanecarboxylic acid).

We calculate the 12 crystal structures formed from all 4 plausible isolated chain conformations of copolymer with the three packing forms observed in nylon (γ, α, and δ). We predict (Fig. 7) that the best structure is the γ form of Chair-ee-St, which has the chains equatorial to chair cyclohexane. Bad contacts between the axial hydrogen atoms of the cyclohexane and the CH_2 of the nylon 6 on adjacent chains together with the difficulty of intra-molecular H-bonds in the copolymer make the α form unfavorable. Indeed, this prediction is confirmed by the fiber X-ray diffraction experiment, which is in good agreement with the predicted patterns (Fig. 9 and 11).

We predict that the stable structure (γ form of Chair-ee-St) of AMCC/N6 copolymer has a chain repeat distance 1.4Å shorter than the predicted distance of α form of nylon 6. This is conformed by the experiments, which give the difference 1.5 Å between AMCC (15.7 Å) and nylon 6 α form (17.2 Å). The twisted chain conformation caused by torsion Ψ_1, Ψ_2, Φ_1, and Φ_2 accounts for the decrease in the chain repeat distance.

The Young's modulus in the chain direction is calculated to be 93 GPa for the copolymer (at 0K), which compares to 135 and 295 GPa for GPa for γ form and α form nylon 6, respectively.

The introduction of cyclohexane into nylon 6 has two major effects to its properties:

It causes twisted conformations, which decreases the Young's modulus of the copolymer in chain direction.

It makes the chain rigid, which decrease the sensitivity of the copolymer to the moisture.

5. Acknowledgment

This research was partially supported by grants from NSF (CHE99-85574), DOE ASCI ASAP, and ARO-MURI. The facilities of the MSC are also supported by grants from DOE, NSF (CHE 99-77872), ARO (MURI), ARO (DURIP), IBM-SUR, NIH, ChevronTexaco, 3M, Seiko-Epson, Avery-Dennison, General Motors, Kellogg's, Asahi Chemical, Beckman Institute, and Nippon Steel.

6. Reference

(1) E. W. Pietrusza, Nesty, G. A., R. Pinter. In *U. S. Pat.*, 1962; Vol. 3,037,002.
(2) M. Levine, Temin, S. C. *J. Polym. Sci.* **1961**, *49*, 241.
(3) M. N. Bogdanov, Kudryavtsev, G. I., F. M. Mandrosova, I. A. Spirina, D. E. Ostromogol'skii. *Vysokomol. Soedin.* **1961**, *3*, 1326.
(4) Temin, S. C.; Levin, M. U.S. Pat. 2,910,457
(5) F. R. Prince, Pearce, E. M., R. J. Fredericks. *Journal of polymer science: part A-1* **1970**, *8*, 3533-3541.
(6) Kohan, M. I. *Nylon Plastics Handbook*; Hanser: New York, 1995.
(7) S. Dasgupta, Hammond, W. B., W. A. Goddard III. *J. Am. Chem. Soc.* **1996**, *118*, 12291.
(8) Y. Li, W. A. G. I. *Macromolecules* **2002**, *35*, 8440.
(9) N. Karasawa, Goddard, W. A., III *J. Phys. Chem.* **1989**, *93*, 7320.
(10) S. Dasgupta, Brameld, K. A., C. -F. Fan, W. A. Goddard III. *Spectrochimica Acta Part A* **1997**, *53*, 1347.
(11) N. Karasawa, S. D., W. A. Goddard III. *J. Phys. Chem.* **1991**, *95*, 2261.
(12) N. Karasawa, W. A. Goddard. I. *Macromolecules* **1992**, *25*, 7268.
(13) Juaristi, E. *Conformational behavior of six-membered rings*; VCH Publishers, Inc.: New York, 1995.
(14) D. A. Dixon, Komornicki, A. *J. Phys. Chem.* **1990**, *94*, 5630.
(15) Murthy, N. S. *Polym. Commun.* **1991**, *32*, 301.
(16) N. S. Murthy, Cooke, R. S., R. G. Bray. *Macromolecules* **Submitted**.
(17) Murthy, N. S. *Macromolecules* **1987**, *20*, 309.
(18) We clarify the source of the calculated meridional and equatorial intensities (to specify the value of *hkl*) by re-indexing the unit cell. The original minimized unit cell is triclinic as shown in Figure 12*ab*. To facilitate the analysis, we re-index the unit cell as shown in Figure 12*cd*. The monoclinic unit cell shown in Figure 12*cd* is 11 times the original unit cell shown in Figure 12*ab* in *b* direction (*b* is the stacking direction of H-bond sheets, the *a* direction is the H-bond direction and the *c*

direction is the chain direction). The intensities shown in Figure 10 are indexed from the monoclinic unit cell. A careful study of Figure 10 reveals that we do *not* include (037) and (058) intensities in Figure 11a. The reason is that the copolymer we studied using theory assumes that cyclohexane ring is distributed uniformly in the matrix, while in the experimental study, the cyclohexane ring is distributed randomly in the matrix. Thus the intensity (037) of 7^{th} layer and (058) of 8^{th} layer would like to be negligible in the XRD meridional scan, and they are relatively strong in our calculated meridional intensities. From our re-indexed values of *hkl*, we can understand why the intensities inside the region of 0 ± 0.05 (reciprocal angstroms) along the equatorial direction are observed in the meridional scan. All these intensities have h=0 and k a small number. Because the monoclinic unit cell is 11 times the original unit cell and the b length of the re-indexed unit cell is 87.1 angstroms, so that it is reasonable that 0kl intensities are in approximately directions of 00l

Tables

Table 1. Relative total energy E, torsion energy $E_{torsion}$, (kcal/mol/chain) and Chain-axis repeat distance b (Å) of all 12 possible isolated chain conformations
(3 ring conformations with 4 ring-chain torsions)
(For each ring conformation, the best ring-chain torsion is shown in bold. In the italics, we show the four best conformations among all 12 possibilities, which we will analyze in the periodic case.)

	E1			E2			Sg			St		
	E	$E_{torsion}$	b	E	$E_{torsion}$	b	E	$E_{torsion}$	b	E	$E_{torsion}$	b
Chair-ee	*4.18*	*3.52*	*17.0*	8.79	7.72	16.5	*1.97*	*2.92*	*16.0*	***0.00***	***0.00***	***16.5***
Chair-aa	18.66	7.63	16.2	9.01	4.61	16.6	17.89	4.43	14.1	***0.32***	***1.34***	***16.1***
Boat-ee	20.96	16.59	16.6	19.77	13.28	16.6	**9.82**	**7.07**	**16.5**	11.53	8.13	16.9

Table 2. Relative energy E (kcal/mol/amide unit), Young's modulus in chain direction E_Y (GPa), and Chain-axis repeat distance b (Å) of the 3 types of crystal structures (α,γ,δ) formed by 4 chain conformations (Chair-ee-E1, Chair-ee-St, Chair-ee-Sg, Chair-aa-St)
(The best one for each type (α,γ,δ) is shown in bold.)

	form			form			form		
	E	E_Y	b	E	E_Y	b	E	E_Y	b
Chair-ee-E1	**0.40**	**237.85**	**17.4**	0.79	73.03	16.9	**0.37**	**165.20**	**17.4**
Chair-ee-St	2.20	123.43	16.9	**0.00**	**92.74**	**16.3**	0.58	50.48	16.6
Chair-ee-Sg	3.79	50.49	16.3	2.69	29.01	15.8	2.98	34.02	16.2
Chair-aa-St	2.93	84.79	16.3	2.73	128.69	16.2	2.78	105.73	16.3

Table 3. Young's modulus in chain direction E_Y (GPa) and Chain-axis repeat distance b (Å) of nylon 6 form, form and copolymer from theory (γ form of Chair-ee-St) and experiment

	Nylon 6 form		Nylon 6 form		N6/AMCC	
	E_Y	b	E_Y	b	E_Y	b
Exp.	168	17.2		16.8		15.7
Theory	295	17.7	135	17.3	93	16.3

Chapter 3

EFFICIENCY OF VARIOUS LATTICES FROM HARD BALL TO SOFT BALL: THEORETICAL STUDY OF THERMODYNAMIC PROPERTIES OF DENDRIMER LIQUID CRYSTAL FROM ATOMISTIC SIMULATION*

ABSTRACT

Self-assembled supramolecular organic liquid crystal structures at nanoscale have potential applications in molecular electronic, photonics, and porous nanomaterials. Most of such structures are formed by aggregation of soft spherical supramolecules, which have soft coronas and overlap each other in the packing process. Our main focus here is to study the possible packing mechanisms via molecular dynamics simulations at atomistic level. We consider the relative stability of various lattices packed by the soft dendrimer balls, first synthesized and characterized by Percec et al. (*J. Am. Chem. Soc.*, **1997**, 119, 1539) with different packing methods. The dendrons, which form the soft dendrimer balls, have the character of a hard aromatic region from the point of the cone to the edge with C_{12} alkane "hair". After packing into a sphere, the core of the sphere has the hard aromatic groups while the surface is covered with the C_{12} alkane "hair". In our studies we propose three ways to organize the hair on the balls: ***Smooth/Valentino*** balls, ***Sticky/Einstein*** balls, and ***Asymmetric/Punk*** balls, which lead to three different packing mechanisms: ***Slippery***, ***Sticky***, and ***Anisotropic,*** respectively. We carry out a series of molecular dynamics (MD) studies on three plausible crystal structures (A15, FCC and BCC) as a function of density and analyze the MD based on vibrational density of state (DoS) method to extract enthalpy, entropy, and free energies of these systems. We find that anisotropic packed A15 is favored over FCC, BCC lattices. Our predicted X-ray intensities of the best structures are in excellent agreement with experiment. "Anisotropic ball packing" proposed here plays

* Published in Y. Li, S-T Lin, W. A. Goddard III, *J. Am. Chem. Soc*. **2004**, 126, 1872-1885.

an intermediated role between the enthalpy-favored "disk packing" and entropy-favored "Isotropic ball packing", which explains the phase transitions at different temperatures. Free energies of various lattices at different densities are essentially the same, indicating that the preferred lattice is not determined during the packing process. Both enthalpy and entropy decrease as the density increases. Free energy change with volume shows two stable phases: condense phase and isolated micelle phase. The interactions between the soft dendrimer balls are found to be lattice dependent when described by a 2-body potential because the soft ball self adjusts its shape and interaction in different lattices. The shape of the free energy potential is similar to the "square shoulder potential". A model explaining the packing efficiency of ideal soft balls in various lattices is proposed in terms of geometrical consideration.

1. Introduction

The liquid crystal phase formed by self-assembled supramolecules with 3D nanoscale periodicity has been researched extensively[1-13], which has potential applications in molecular electronics[2], photonics[5], and porous nanomaterials[6]. In particular, Percec and coworkers have advanced a rational design and synthesized monodendrons that self-assemble through various molecular recognition mechanisms into rod-like[14], cylindrical[15], and spherical[16] supramolecular dendrimers, which self-organize into column lattice[15] or cubic lattice[1,4,16,17]. Wedge-shaped dendrons such as **I**, **II** and **III** depicted in Scheme 1, have so far been found to form either columnar or cubic phases. In the former, dendrons assemble like flat pizza slices into disks, which then stack into columns, which eventually pack to form a hexagonal array.[15] Dendrons with more alkyl chains are cone-shaped and assemble into supramolecular spheres. So far, these spherical aggregates have been known to pack on three types of lattices, Cub $Pm\bar{3}n$,[16] Cub $\mathrm{Im}\bar{3}m$,[17] and Tet $P4_2/mnm$.[1] The preferred formation of the $Pm\bar{3}n$ lattice in thermotropic spherodic cubic mesophases is a fact which surprises, especially as related ordered assemblies of block copolymers form the

BCC ($\mathrm{Im}\bar{3}m$) lattice. In addition, the $Pm\bar{3}n$ lattice was not only found for dendritic molecules, it was also reported for the thermotropic mesophases of amphiphilic molecules,[18,19] star shaped molecules,[20] and amphiphiles with perfluorinated chains.[21]

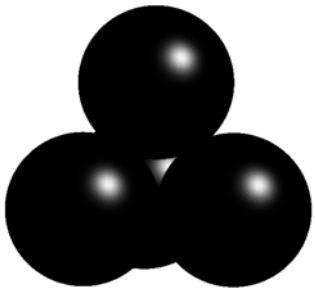


Figure 1. Close-packed hard balls

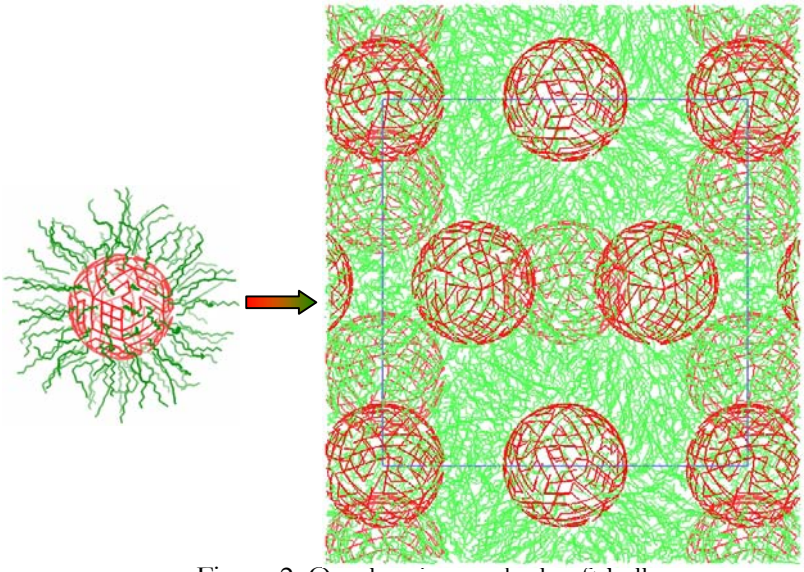


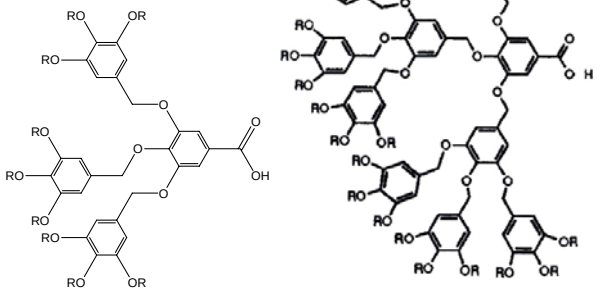
Figure 2. Overlapping packed soft balls

One of the most important principles to pack ***hard*** spheres into periodic lattice is minimization the interstitial volume (Figure 1), which leads to the face-centered cubic (cubic close-packed) structure and hexagonal close-packed structure. The packing of hard spheres (billiard balls and noble gases) into infinite two-dimensional arrays leads to close packing with each ball having 6 equally spaced neighbors in a plane with the planes packed into periodic arrays in three dimensions that lead to 12 nearest neighbors for each ball. Stacking the close packed layers as ABCABC leads to the face-centered cubic (cubic close-packed, denoted FCC) structure, while ABABAB stacking of close packed layers leads to the hexagonal close-packed structure (denoted HCP). Indeed the stable crystal structure for all noble gases and the favored structures for most metals is FCC, HCP or DHCP (hexagonal with ABACABAC packing). Another popular structure with metals is body centered cubic (BCC) in which atoms are at the corners and center of a cube, leading to 8 nearest neighbors. In addition, symmetric molecules such as CH_4 and C_{60} fullerene crystallize into structures that are slightly distorted FCC at low temperature and fully FCC

at higher temperature. Furthermore, it has been shown that the face-centered-cubic packing maximizes the total entropy.[22]

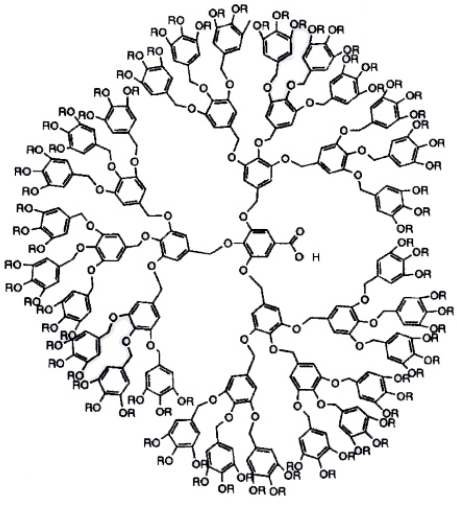
Obviously, the simple principles that explain the structure approproiate for hard balls such as noble gas atoms, CH_4, and C_{60} don't apply to ***soft*** spheres, which instead of a simple ball-ball surface contact, have flexible hair that can overlap each other to achieve more complex packing as in Figure 2. Almost all self-assembled supramolecular aggregates are soft spheres and can be approximated by hard spherical cores and soft aliphatic coronas, as shown in Figure 2.

What would be the principle to pack soft balls together? Here we propose three packing mechanisms of soft balls and use classical atomistic molecular dynamic simulations to investigate the efficiency of various lattices for soft balls from determining their thermodynamic properties. We focus on the three compounds **I**, **II** and **III** (scheme 1), which have been reported to form spheres with nearly integer number of dendrons.[16] In section 2 we describes the details of our calculations. Section 3 reports the results and the discussion. The summary is presented in Section 4.



I

II



III

Scheme 1. R=$C_{12}H_{25}$, **I**: G2 (12G2-AG) **II**: G3 (12G3-AG), **III**: G4 (12G4-AG). We denote these as G2, G3, and G4 through the text. Ref.[16] denotes them as 12G2-AG, 12G3-AG, and 12G4-AG respectively, which are 12G2-AG: 3,4,5-Tris[3',4',5'-tris(n-dodecan-1-yloxy)benzyloxy)]benzoic Acid; 12G3-AG: 3,4,5-Tris{3',4',5'-tris[3",4",5"-tris(n-dodecan-1-yloxybenzyloxy)benzyloxy]benzyloxy}benzoic Acid, and 12G4-AG: 3,4,5-Tris(3',4',5'-tris{3",4",5"-tris[3'",4'",5'"-tris(n-dodecan-1-yloxybenzyloxy) benzyloxy] benzyloxy}benzyloxy)benzoic Acid.

2. Methods

2.1 Force field

The methylene (CH_2) and methyl (CH_3) groups in $C_{12}H_{25}$ were treated as united atoms C_32 and C_33, that is, each CH_2 or CH_3 unit was treated as a single neutral pseudoatom. The van der Waals interaction for this coarse-grained polyethylene model is taken from the SKS (Siepmann-Karanorni-Smit) force field[23-25], which was developed to describe the thermodynamic properties of n-alkanes. The bond-stretching force constant, which is not presented in the original SKS force field, is taken from the AMBER force field[26,27]. The torsion potential of the SKS force field had been taken from the OPLS (Optimized Potentials for Liquid Simulation) force field of Jorgensen.[28] The parameters for all other atoms including hybrid terms with C_32, C_33 are taken from the generic Dreiding force field.[29] There is no coulomb term in the setup energy of the simulation.

The force field for the coarse-grained polyethylene model uses valence terms of the form

$$E^{VDW}(R) = D_0\left\{\left(\frac{R_0}{R}\right)^{12} - 2\left(\frac{R_0}{R}\right)^{6}\right\} \tag{1a}$$

$$E^{bond}(R)^a = \frac{1}{2}K_R(R-R_0)^2 \tag{1b}$$

$$E^{angle}(\theta) = \frac{1}{2} K_\theta (\theta - \theta_0)^2 \quad (1c)$$

$$Etorsion(\phi)^b = \sum_n \frac{1}{2} V_n \left[1 - d_n \cos(n\phi)\right] \quad (1d)$$

where the parameters are summarized in Table 1.

Table 1. Force filed parameters used for coarse-grained alkane segments. The functional forms are given in Eq. (1).[ab] (all parameters from SKS united atom force field unless otherwise indicated.)

E^{vdW}	CH_2	R_0[c]	4.4113	D_0[d]	0.09339		
	CH_3	R_0[c]	4.4113	D_0[d]	0.2265		
E^{bond}	CH_2-CH_2	R_0[c]	1.54	K_b[e]	520		
	CH_2-CH_3	R_0[c]	1.54	K_b[e]	520		
E^{angle}	CH_2-CH_2-CH_2	θ_0[f]	114	K_θ[g]	124.19		
	CH_2-CH_2-CH_3	θ_0[f]	114	K_θ[g]	124.19		
$E^{torsion}$	CH_2-CH_2-CH_2-CH_2	$V_1(d_1)$[d]	1.4109(-1)	$V_2(d_2)$[d]	-0.271(1)	$V_3(d_3)$[d]	2.787(-1)
	CH_2-CH_2-CH_2-CH_3	$V_1(d_1)$[d]	1.4109(-1)	$V_2(d_2)$[d]	-0.271(1)	$V_3(d_3)$[d]	2.787(-1)

[a]The force constants for bond-stretching potential function was introduced from the AMBER force field[26,27] because the original SKS force field uses fixed bond distance.

[b]The torsion potential of the SKS force field had been taken from the OPLS (Optimized Potentials for Liquid Simulation) force field.[28]

[c]In Å. For vdW, SKS uses $E^{VDW}(R) = 4\varepsilon\left\{(\sigma/R)^{12} - (\sigma/R)^6\right\}$. Thus $R_0 = \sqrt[6]{2}\sigma$.

[d]In kcal/mol. For vdW, D_0(kcal/mol)=$k\varepsilon$ (SKS), where k is the Boltzmann constant.

[e]In kcal/mol/Å^2. [f]In degrees. [g]In kcal/mol/rad^2.

2.2 The vibrational density of states (DoS) from the velocity autocorrelation (VAC) function

To obtain the vibrational density of states (DoS) $S(\upsilon)$ as a function of the frequency υ for a given density and temperature, we start with the mass weighted velocity autocorrelation (VAC) function

$$C(t) = \sum_{i=1}^{N}\sum_{\alpha=1}^{3} m_i c_{i\alpha}(t) \tag{2}$$

where m_i is the mass of atom i; and $c_{i\alpha}(t)$ is the α component (α=x, y, and z) of velocity autocorrelation of atom i

$$c_{i\alpha}(t) = \lim_{\tau\to\infty} \frac{\int_{-\tau}^{\tau} \mathrm{v}_{i\alpha}(t'+t)\mathrm{v}_{i\alpha}(t')dt'}{\int_{-\tau}^{\tau} dt'} = \lim_{\tau\to\infty} \frac{1}{2\tau}\int_{-\tau}^{\tau} \mathrm{v}_{i\alpha}(t'+t)\mathrm{v}_{i\alpha}(t')dt' \tag{3}$$

Taking the Fourier transform of C(t) then leads to the vibrational density of states(DoS)

$$S(\upsilon) = \frac{2}{kT}\lim_{\tau\to\infty}\int_{-\tau}^{\tau} C(t)\exp(-i2\pi\upsilon t)dt \tag{4}$$

Integrating $S(\upsilon)$ gives the total degrees of freedom of the system, i.e.,

$$\int_0^{\infty} S(\upsilon)d\upsilon = 3N \tag{5}$$

Generally, MD simulations for a condensed system remove the center of mass translations (3 degrees of freedom) since the energy must be independent of the origin. Thus $S(\upsilon)$ is renormalized such that the integration of eq. 5 gives the 3N-3.

From various test calculations we found that having the velocities every 4 fs is sufficient to obtain an accurate description of the high frequency DoS and we found that a total time span (after equilibration) of 20~40 ps is generally adequate to give the low frequencies modes sufficient for accurate entropies (see section 2.5).

2.3 Thermodynamic properties from molecular dynamics

Given the vibrational density of states for a given V and T, we can calculate the partition function $Q(V,T)$ by treating the continuous DoS as a continuum of uncorrelated harmonic oscillators

$$\ln Q = \int_0^\infty d\upsilon S(\upsilon) \ln q_{HO}(\upsilon) \tag{6}$$

where

$$q_{HO}(\upsilon) = \frac{\exp(-\beta h\upsilon/2)}{1 - \exp(-\beta h\upsilon/2)} \tag{7}$$

is the partition function of a harmonic oscillator with vibrational frequency v ,Here β=1/kT, h is Planck's constant, and $S(v)$ dv is the number of modes between frequencies v and v+dv. Given the partition function, the thermodynamic properties are determined as

$$E = V_0 + T\beta^{-1}\left(\frac{\partial \ln Q}{\partial T}\right)_{N,V} = V_0 + \beta^{-1}\int_0^\infty d\upsilon S(\upsilon) W_E(\upsilon) \tag{8a}$$

$$S = k\ln Q + \beta^{-1}\left(\frac{\partial \ln Q}{\partial T}\right)_{N,V} = k\int_0^\infty d\upsilon S(\upsilon) W_S(\upsilon) \tag{8b}$$

$$A = V_0 - \beta^{-1} \ln Q = V_0 + \beta^{-1} \int_0^\infty d\upsilon S(\upsilon) W_A(\upsilon) \tag{8c}$$

where

$$W_E(\upsilon) = \frac{\beta h\upsilon}{2} + \frac{\beta h\upsilon}{\exp(\beta h\upsilon) - 1} \tag{9a}$$

$$W_S(\upsilon) = \frac{\beta h\upsilon}{\exp(\beta h\upsilon) - 1} - \ln[1 - \exp(-\beta h\upsilon)] \tag{9b}$$

$$W_A(\upsilon) = \ln \frac{1 - \exp(\beta h\upsilon)}{\exp(-\beta h\upsilon/2)} \tag{9c}$$

are weighting functions and V_o is a reference energy. Therefore, within the assumption that the system is ergotic within the time scale of the calculation, all thermodynamic properties are determined. We need to provide only the reference energy V_o and the vibrational density of states distribution $S(v)$ from the molecular dynamics simulations.

We choose the reference energy such that in the classical limit (h→0) the energy evaluated from eq. 8a is equivalent to the total energy E^{MD} (kinetic plus potential) obtained from molecular dynamics simulation. For a system of harmonic oscillators in the classical limit, the energy is $3NkT$ (the equipartition theorem), where $3N$ is the total degrees of freedom of the system. Thus we write the reference energy as

$$V_0 = E^{MD} - \beta^{-1} 3N \tag{10}$$

where N is the total number of atoms in the system. For condensed systems, the total degrees of freedom of the system are 3N-3, and Eq. (10) becomes $V_0 = E^{MD} - \beta^{-1}(3N-1)$. It is useful to note that the energy (and free energy) determined this way includes the zero point energy (first term on the RHS of eq. 9a). This contribution is important when quantum effects are significant in the system.

2.4 Molecular dynamics simulation

Molecular dynamics software Lammps[30] is used to perform NVT molecular dynamics simulations with the time step set to 1 fs. For each structure, we start by assigining the initial velocities (Gaussian distribution) on each atom to give a system temperature of 20 K, and then perform MD simulations for 10 ps at 10 K to equilibrate the structure. The temperature of the system is then increased from 10 K to 277 K (recrystallization temperature) steadily over a period of 2 ps (the temperature was increased by 267/2000 K every time step), followed by equilibration runs at 277 K for 8 ps. Then a 40 ps NVT molecular dynamics is performed with the atomic velocities, system energy, temperature, pressure saved every 4 fs. This trajectory information is later used in the velocity autocorrelation analysis to obtain the vibrational spectrum from which the thermodynamic properties were calculated. The use of 40 ps is found to be sufficient for obtaining a converged free energy as shown in section 2.5.

Molecular dynamics are performed on a Linux cluster of 80 dual-processor Dell PowerEdge 2650s (P4 Xeon 2.2/2.4GHz, 2G Memory, 54G HD) at MSC. Each dynamics takes 1~5 days depending on the system size (4,000~25,000 atoms).

2.5 Convergence of the free energy calculation

In order to compare the efficiency of various lattices for dendrimer balls consistently, we build various lattice structures from the same initial isolated dendrimer ball (see section 3.3 below). However, there may exist deviations in the free energy evaluated from the molecular dynamics, especially for the entropy that is dominated by the low frequency modes in the system. To access the convergence of our calculated thermodynamic properties, we run three independent molecular dynamics from the same initial structure but assigning different initial velocity distributions (same overall system temperature). From the three different MD trajectories, we determine the RMS deviation of free energy,

enthalpy, entropy as shown in Figure 7 of section 3.4. To discuss how the deviation depends on the correlation time, we select the point with a large deviation in Figure 7A (sticky G2 Ball in BCC lattice) and calculate the deviations in free energy using different correlation times: 5ps, 10ps, 20ps, 40ps. Figure 3 shows that using a longer correlation time leads to a better free energy evaluation and 40ps is sufficient to give satisfying deviation to compare various lattices as shown in Figure 7.

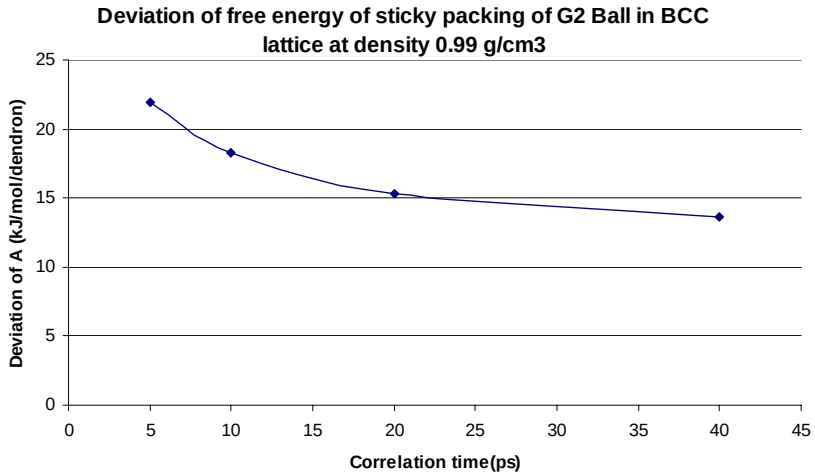


Figure 3. Convergence of the free energy from vac

3. Results and discussion

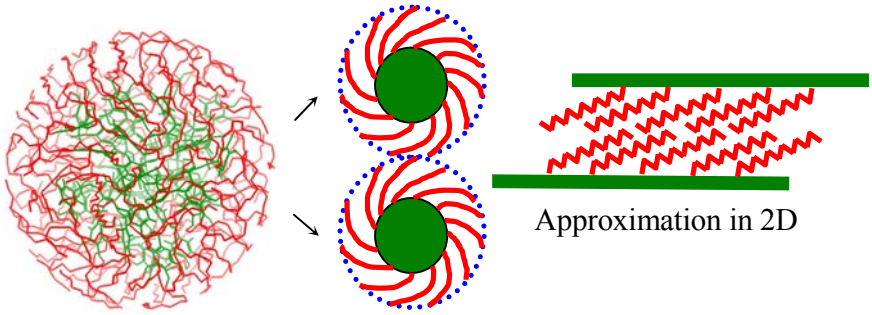
3.1 Three proposed packing mechanisms of the soft balls with polyethylene tails

The liquid crystal phase formed by self-assembled supramolecules with 3D nanoscale periodicity has been researched extensively.[1-13] For example, studies of electron density profiles and histograms computed from the X-ray diffraction data[16] demonstrate that compounds **I**, **II**, and **III** are self-assembled in supramolecular dendrimers resembling spherical micelles, which self-organize in a 3-dimensional cubic $Pm\bar{3}n$ lattice. This

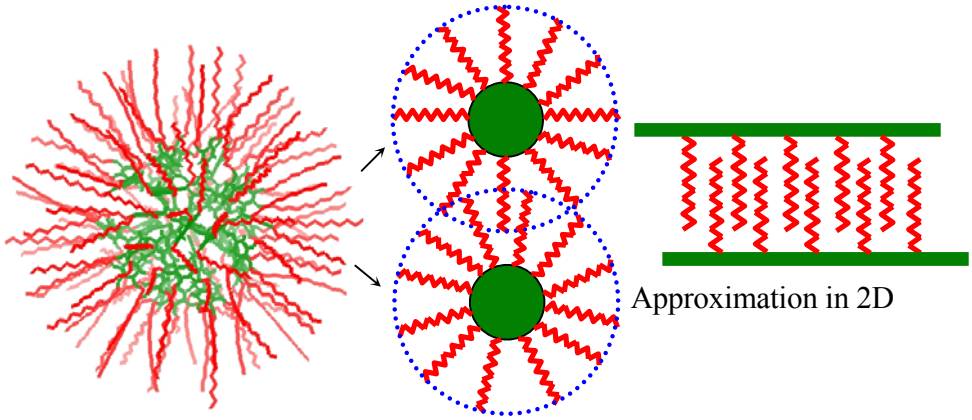
thermotropic liquid crystal phase is similar to that of the lyotropic $Pm\bar{3}n$ phase found in some amphiphile/water systems. These supramolecular dendrimers contain a poly (benzyl ether) core dispersed in an aliphatic matrix of nearly uniform density, which is made up of the melted terminal long alkyl chains of the monodendrons.

Geometrical aspects of liquid crystal phases have been studied in considerable detail for lyotropic systems[31-33] with interfacial curvature being recognized as the key factor for determining the phase type.[34,35] However, it is very difficult to investigate the interfacial curvature from experiment because of the nearly uniform density of the aliphatic matrix.

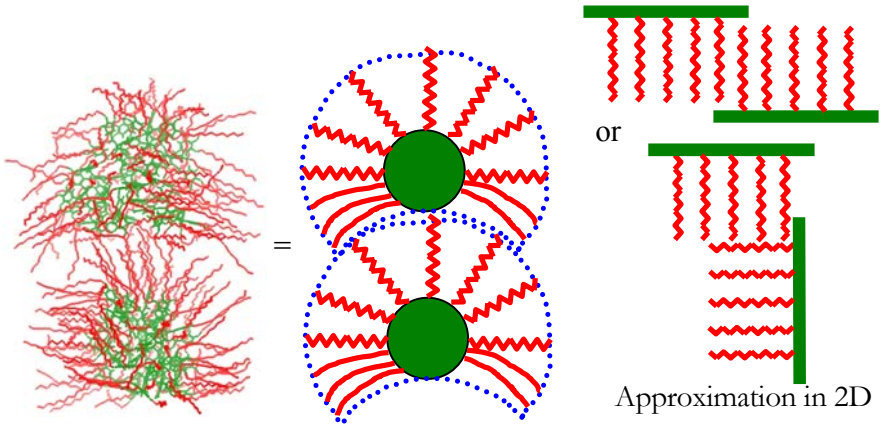
Based on our simulations in atomistic scale, we propose three packing mechanisms of soft balls with soft corona composed by polyethylene as shown in Figure 4.



A. ***Slippery packing*** of ***Smooth/Valentino balls***; this provides favorable intra-ball vdW but with increased unfavorable guache dihedras



B. ***Sticky packing*** of ***Sticky/Einstein balls***; this provides favorable inter-ball vdW, while retaining favorable trans torsions of the alkyl chains



C. ***Anisotropic packing*** of ***Asymmetric/Punk balls***, this balances the favorable inter-ball vdW and intra-ball vdW while minimizing unfavorable guache dihedrals

Figure 4. Three proposed packing mechanisms of soft balls with soft corona composed by polyethylene chains

For the soft balls having soft corona composed of long alkyl chains, we considered three classes of structures:

Smooth/Valentino balls shown in Fig. 4A, which leads to ***Slippery packing mechanism***. Here the C_{12} alkane "hair" is slicked down on each ball**.** This maximizes the favorable van der Waals attraction between intra-ball chains and is optimum for separate balls. However the best intra-ball interactions may lead to an increase in the number of gauche dihedrals and hence less favorable torsional interactions. This likely leads to a minimum in the favorable inter-ball interactions between the chains.

Sticky/Einstein balls shown in Fig. 4B, which leads to ***Sticky packing mechanism***. Here the alkyl hairs stick out perpendicular to the surface to provide maximum surface area for adjacent balls to interact, maximizing the favorable intermolecular van der Waals attraction between chains of different balls. This likely leads to a minimum in the favorable intra-ball interactions between chains.

Asymmetric/Punk balls shown in Fig. 4C, which leads to ***Anisotropic packing mechanism***. Here ~ half the surface has Einstein type hair and the other half has Valentino hair. This was motivated by the anisotropic positions on the faces of $Pm\bar{3}n$ lattice, where the balls have two close neighbor balls and other neighbor balls at regular distance. In this packing mechanism, the core of the soft ball deforms to become nonspherical. The chains are slicked down (Valentino) to provide the most favorable by intra-ball VDW while accommodating the short distances between the closest balls on the faces and they are extended (Einstein) to provide the best inter-ball VDW with the other more distant neighbors. Thus in this Anisotropic packing mechanism, the packing of polyethylene chains is a balance between intra-ball type and inter-ball type. In this packing mechanism, the chain shape can be kept as nearly all-trans to minimize the torsional cost which optimizing the intra-ball and inter-ball interactions.

3.2 Preparation of dendrimer balls of different shapes

We use Cerius2[36] to construct the three-dimensional structures including three types of dendrimer balls: Smooth/Valentino (Figure 4a), Sticky/Einstein (Figure 4b), Asymmetric/Punk (Figure 4c).

From the isolated dendrons of compounds G2(**I**), G3(**II**), G4(**III**), we first constructed the Sticky/Einstein balls for G2, G3, G4. The analysis of the experimental X-ray result suggested that 11.3 dendrons self-assemble into the soft balls in G2, 5.8 dendrons assemble into the soft balls in G3, and 1.9 assemble to form the balls for G4 (see Table 3 of Ref. [16]). We construct the G2, G3, G4 balls to have 12, 6, 2 dendrons respectively. This leads to 108 C12 alkyl chains for G2 and 162 for both G3 and G4. Figure 5 illustrates the preparation of the G2 balls.

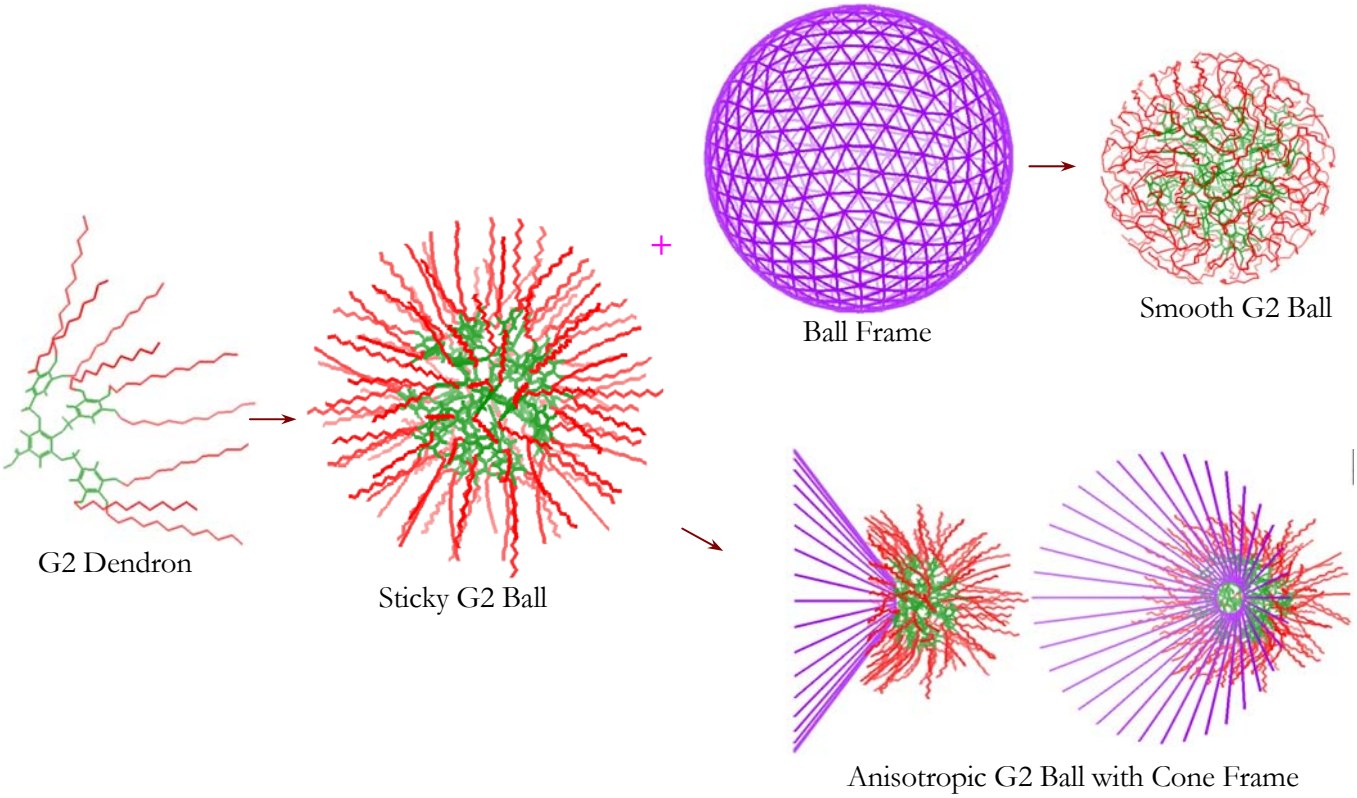


Figure 5. Preparation of the three shapes of G2 dendrimer ball

From the isolated G2 dendron, we first construct the "Sticky/Einstein G2 Ball" composed of 12 dendrons, with the polyethylene chains initially all-trans, pointing outwards uniformly from the ball center. An energy minimization of 500 steps is then performed to optimize the structure.

Next we constructed the structure of the "Ball Frame" shown in Figure 5. This frame is prepared using 500 C_R atoms (the type for C in aromatic assemblies including graphite) uniformly distributed on a sphere. The uniform distribution of C_R atoms is obtained by first placing a point charge on each atom and then minimizing (1000 steps) the Coulombic energy, with a constraint of keeping the C_R atoms a fixed distance from the center. The "Ball Frame" is a tool to be used to compress the "Sticky/Einstein ball" to "Smooth/Valentino ball" by starting with a radius outside of the atoms for the sticky ball and steadily reducing the radius forcing the alkyl chains to deform and eventually for the structure of the Smooth ball, as described next.

To select an initial radius for the "Ball Frame", we note that the surface atoms (C_33) of the "Sticky G2 Ball" lead to a RMS radius of 27.4 Å, whereas the closest distance of soft balls in various lattices (A15, BCC, and FCC) at the expected density 0.99 g/cm^3 is 34.8 Å. Thus we started the size of "Ball frame" at 27.4Å + 4Å = 31.4Å and contracted it to 34.8Å/2 + 4Å = 21.4Å in 10 steps. Here 4Å is the favored distance between C_R and C_33 atoms. In each step, the size of "Ball Frame" decreases 1Å followed by 200 steps of energy minimization on the dendrimer. The final result is the "Smooth G2 Ball" shown in Figure 5.

By rotating a line of 12 C_R atoms 36 times, we get the "Cone Frame" as shown in Figure 5. After compressing the cone angle to almost zero degree, we align the "Cone Frame" together with "Sticky G2 Ball" as shown in Figure 5. Then we expand the cone angle step by step followed by 200 steps of energy minimization unless there is no overlap between the two same Asymmetric/Punk balls at the distance of 34.8 Å as shown in Figure 4c.

3.3 Packing balls into various lattices

From the dendrimer balls prepared in section 3.2, except for the dendrimer ball with smooth shape, we first pack them together in various lattices (A15, FCC, BCC) at low density (0.20 g/cm^3) as shown in Figure 6B, 6D, 6F. We then shrink the unit cell steadily by 1Å each step followed by 500 steps of minimization to allow the polyethylene chains of neighboring balls to pack together. In each step, the ball center is translated to be at the correct position for the new decreased cell length, but the structure within each ball remains unchanged. During these minimizations, we fix the acid hydrogen atoms (denoted as atom type H___A) of the carboxylic acid at the core of each dendron in order to retain the position of the ball center. With this procedure, we get the A15, FCC, BCC lattices at the target density (0.99 g/cm^3) as shown in Figure 6C, 6E, 6G.

For the A15 structure, there are two crystallographically inequivalent types of positions: the face position (2 atoms per face, 6 per cube) and the corner/body center positions (one each per cube), as indicated by black cycles and white cycles in Figure 6. The face positions lead to the closest contacts with neighboring balls and hence we place the anisotropic balls (Figure 4c) at these face positions. For the corner/body center positions of A15 we started with sticky balls. We denote this structure as the "Anisotropic A15" structure.

For the dendrimer ball with smooth shape, we place the balls on the desired positions of various lattices at the target density (0.99 g/cm^3). Our preparation procedure for smooth balls (detailed in section 3.2) guaranteed that the balls would not touch each other when placed into the structures at density 0.99 g/cm^3. Thus the initial structure for the smooth balls led to significant void volume. Then we perform energy minimization of 500 steps, followed by the molecular dynamics runs as described section 2.3. We find that the tail chains relax to fill in the void volume that has been left in the initial structure during energy minimization and molecular dynamics.

In the MD simulations we use 8 independent dendrimer balls for A15 leading to 15840 independent atoms per cell for G2, 24336 for G3 and 24528 for G4. Note that there are no hydrogen atoms in the alkyl chains as the united atom force field is used. For the FCC structure we use 4 independent balls, leading to half the number of atoms as for A15. For the BCC structure we use 2 independent balls, leading ¼ the number of atoms as for A15.

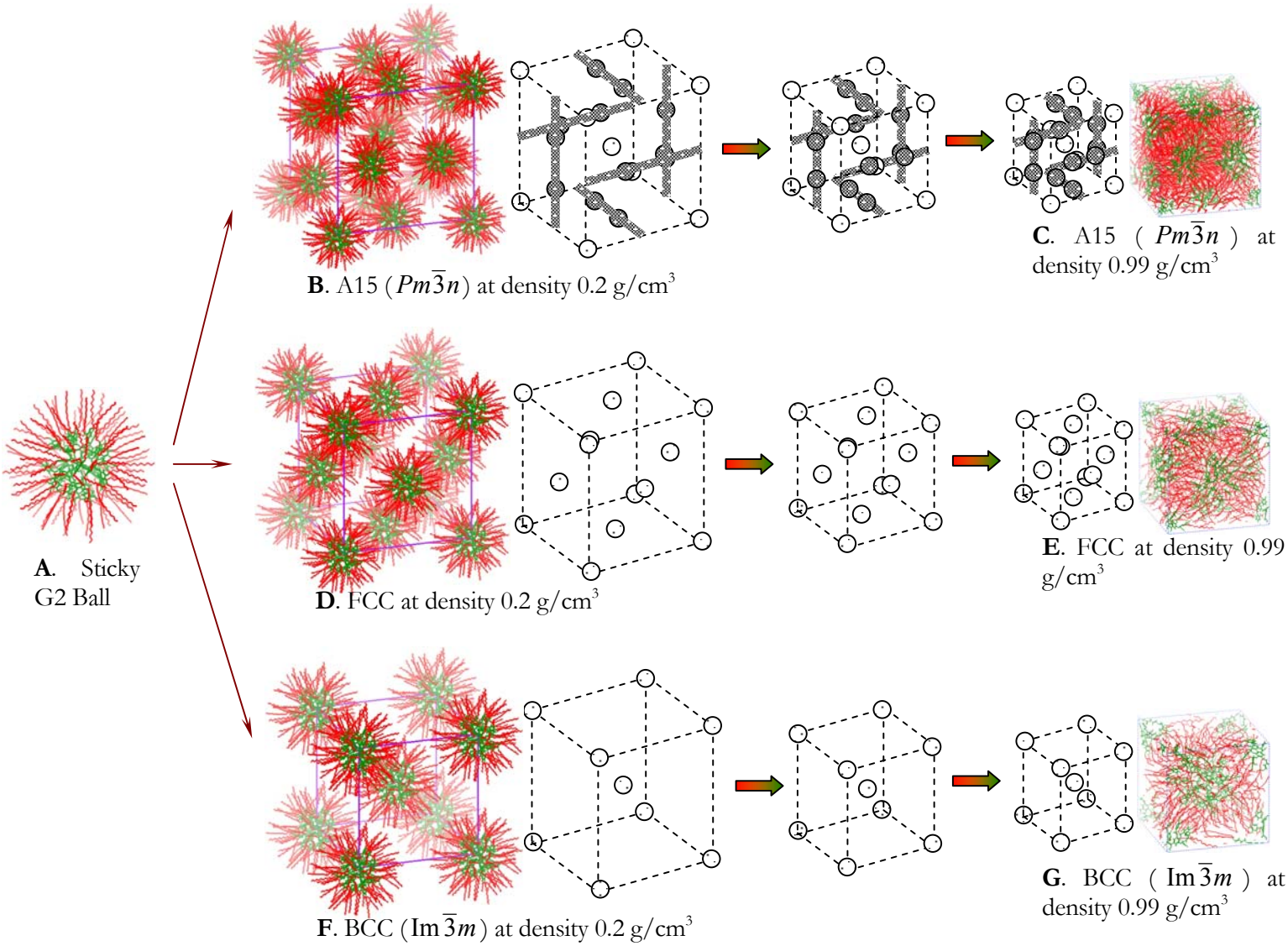
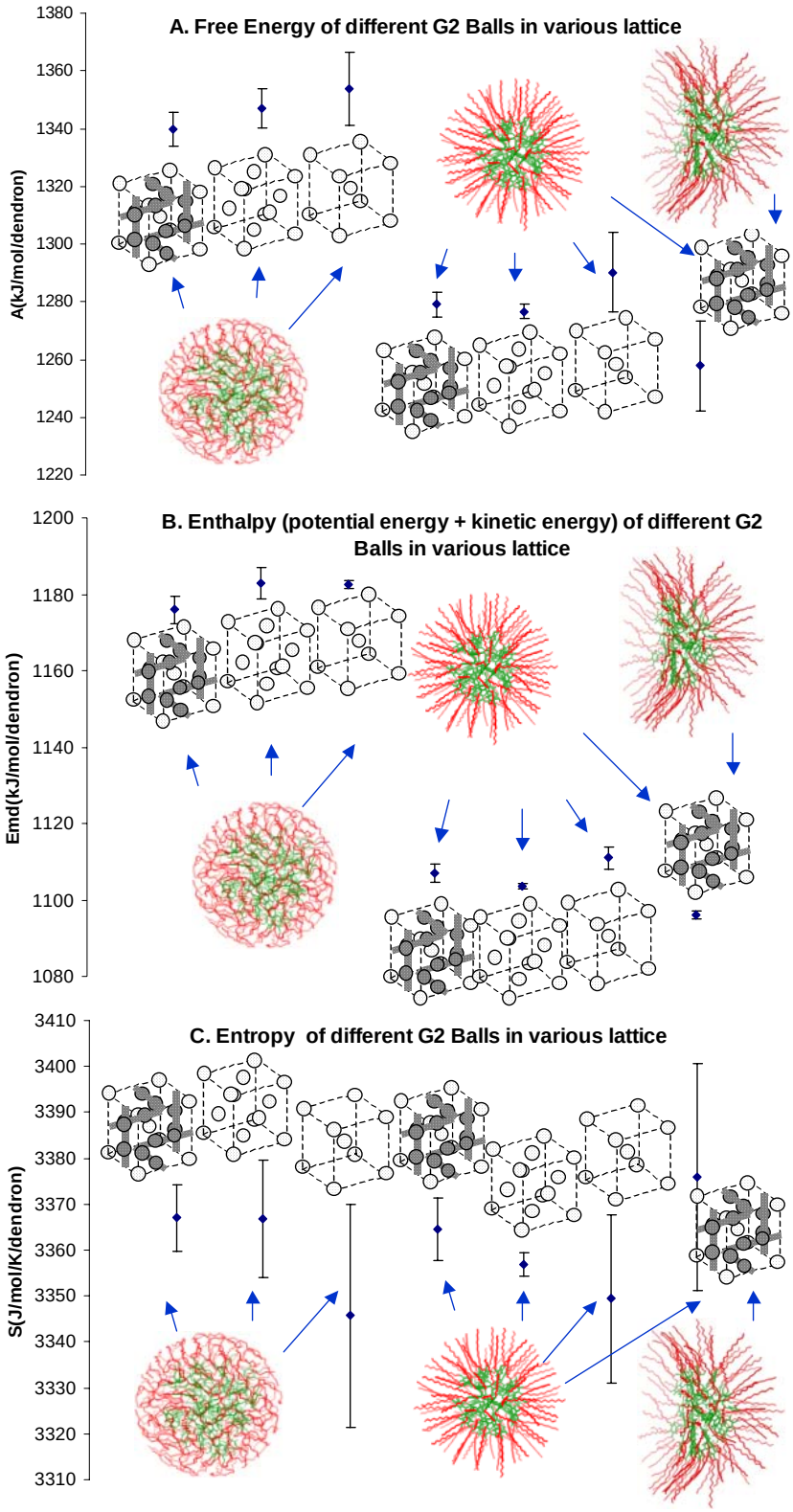
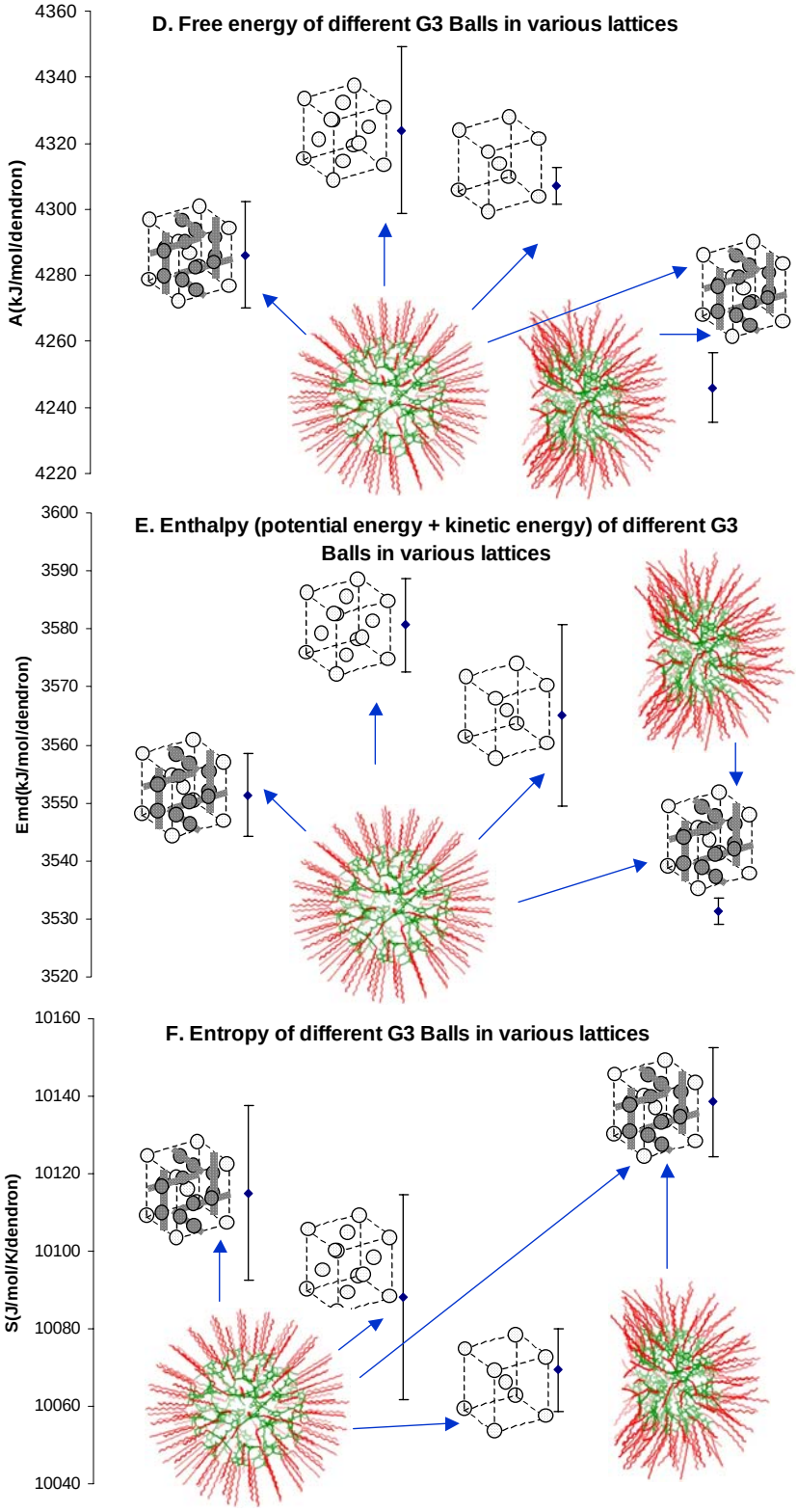


Figure 6. Packing dendrimer balls into A15, FCC, BCC at low density (0.2g/cm^3) followed by compressing in multiply small steps until the target density (0.99g/cm^3) is achieved.


A. Free Energy of different G2 Balls in various lattice
A(kJ/mol/dendron)
1380
1360
1340
1320
1300
1280
1260
1240
1220
B. Enthalpy (potential energy + kinetic energy) of different G2 Balls in various lattice
Emd(kJ/mol/dendron)
1200
1180
1160
1140
1120
1100
1080
C. Entropy of different G2 Balls in various lattice
S(J/mol/K/dendron)
3410
3400
3390
3380
3370
3360
3350
3340
3330
3320
3310


D. Free energy of different G3 Balls in various lattices
A(kJ/mol/dendron)
4360
4340
4320
4300
4280
4260
4240
4220
E. Enthalpy (potential energy + kinetic energy) of different G3 Balls in various lattices
Emd(kJ/mol/dendron)
3600
3590
3580
3570
3560
3550
3540
3530
3520
F. Entropy of different G3 Balls in various lattices
S(J/mol/K/dendron)
10160
10140
10120
10100
10080
10060
10040

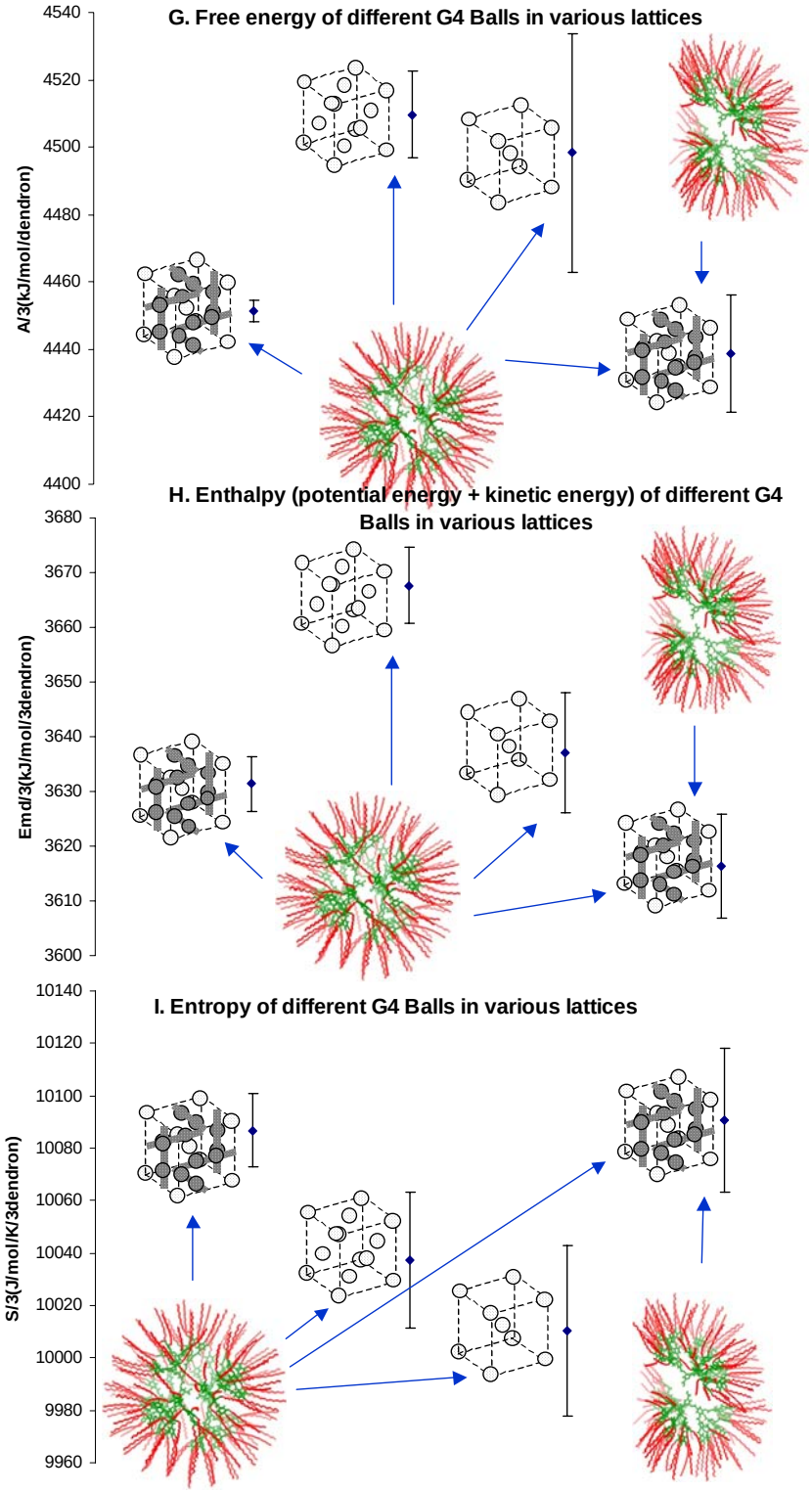

G. Free energy of different G4 Balls in various lattices
A/3(kJ/mol/dendron)
4540
4520
4500
4480
4460
4440
4420
4400
H. Enthalpy (potential energy + kinetic energy) of different G4 Balls in various lattices
Emd/3(kJ/mol/3dendron)
3680
3670
3660
3650
3640
3630
3620
3610
3600
I. Entropy of different G4 Balls in various lattices
S/3(J/mol/K/3dendron)
10140
10120
10100
10080
10060
10040
10020
10000
9980
9960

Figure 7. Helmholtz free energy (*A*=*E*md-*ZPE*-*TS*), enthalpy (*E*md=potential energy from MD including kinetic energy but not zero point energy), and entropy (*S*) of different shape balls in various lattices. Each MD simulation is carried out with fixed Volume using a Nose-Hoover thermostat (denoted s NVT) at the density 0.99 g/cm^3. These MD simulations are carried out at a temperature of 277K, which is recrystallization temperature in Ref. [16]. Shown here are the average and the RMS uncertainties, obtained from three independent MD simulations from using independent sets of initial velocities.

3.4 Free energy of various lattices

Figure 7 shows the free energy (A), energy (Emd), and entropy (S) of different shape balls packed into various lattices. The comparison of generations 2, 3, 4 are shown in Figure 7ABC, DEF, GHI respectively.

Comparing Figure 7A with 7B, 7D with 7E, 7G with 7H, we see that free energy of various lattices is in the same trend as the enthalpy of various lattices. The best lattice is the "Anisotropic A15" lattice for all Generations (2, 3, and 4).

For generation 2, A15, FCC and BCC composed by sticky balls are essentially not distinguishable. We have the same conclusion for A15, FCC and BCC lattices by smooth G2 Balls. However the lattices composed by sticky balls are found to be much more stable than the lattices composed by smooth balls.

From Figure 7C, F, I, we can see that the deviations of entropy is pretty big compared with those of enthalpy in Figure 7B, E, H, which makes the difference among those lattices not distinguishable. However, we see the same trend for all generations: Anisotropic A15 > sticky A15 > sticky FCC > sticky BCC, indicating that A15 is slightly favorable in terms of entropy consideration.

For generations 3 and 4, the differences among various lattices are larger than those for generation 2. We can see that the best lattice is still Anisotropic A15 lattice, followed by A15, BCC, and FCC in the sequence.

In conclusion, anisotropic A15 lattice gives the best (lowest free energy) packing of the G2, G3, and G4 balls composed by compounds **I**, **II** and **III**. For the three packing mechanisms proposed in section 3.1, anisotropic packing is the best. This is because it allows a good balance between inter-ball vdW and intra-ball vdW and accommodates a variety of ball-ball distances without sacrificing the all-trans intra-chain conformation of the alkyl chains. In contrast, slippery packing with smooth ball shape leads to bad enthalpy, although it has the optimal intra-ball vdW.

Furthermore, Asymmetric/Punk ball shape can be viewed as the intermediate between the disk shape and isotropic ball shape as shown in Figure 8. Dendrons with tapered fan shape assemble flat pizza-like slices into disks, which then stack into columns that eventually form a hexagonal array.[15] Dendrons with more alkyl chains are cone-shaped and assemble into supramolecular spheres. So far, these spherical aggregates have been known to pack on three lattices, Cub $Pm\bar{3}n$,[16] Cub $\mathrm{Im}\bar{3}m$,[17] and Tet $P4_2/mnm$.[1] By tuning the number of alkyl chains, there exists the intermediate shape between disk shape and isotropic ball shape as shown in Figure 8B. In fact, from the analysis in Figure 7, this type of Asymmetric/Punk ball shape makes the A15 structure superior to other structures.

For a specified compound,[1] experiments find the following phase sequences at different temperatures:

glass < 110°C < $\mathrm{Col_h}$ ($\mathrm{Col_h}$ = hexagonal columnar) < 140°C < Cub $Pm\bar{3}n$ < 153°C < Tet $P4_2/mnm$ (tetragonal distortion of cubic packing)<163°C < Iso(isotropic liquid).

This indicates that the dendron exhibits tapered fan shape at low temperature, while it adopts cone shape at high temperature. In other words, the packing of disk shape dendrimers in Figure 8A is enthalpicaly preferred and the packing of ball shape dendrimers

in Figure 8C is entropically favored. The results for the Asymmetric/Punk ball shape in Figure 8B, suggests a way to understand the above phase sequence. The Col_h structure is composed by 100% disks as shown in Figure 8A. The Cub $Pm\bar{3}n$ structure has 75%(6/8) Asymmetric/Punk balls (coordination number 14) and 25%(2/8) isotropic balls (coordination number 12). The Tet $P4_2/mnm$ structure has 67%(20/30) Asymmetric/ Punk balls (16 of them have coordination number 14 and 4 of them have coordination number 15) and 33%(10/30) isotropic balls (coordination number 12).[1] Our assumption that tapered fans is favored enthalpically is consistent with the observed ordering of the phases with Col_h lowest, Cub $Pm\bar{3}n$ = next, and Tet $P4_2/mnm$ at the highest temperature.

In the Cub $Pm\bar{3}n$ liquid crystal phase formed by compounds G2(**I**), G3(**II**), and G4(**III**), it is found that nearly integer number of dendrons self-assemble into sphere. Specifically, the packing numbers of dendrons in each sphere of G2, G3, and G4 are 11.3, 5.9, and 1.9.[16] We use integer number 12, 6, 2 in the simulation. Consider an isolated ball formed by aggregation of dendrons, the core part of the ball is compact involving interactions between the phenyl groups. However, the soft corona, which is composed by polyethylene chains, is quite loose. Thus the packing number of dendrons into a ball is determined by the core part of the dendrimer ball. The average radius gyrations of the core part of the G2, G3, G4 balls from our best anisotropic A15 structure (equilibrated after 60ps dynamics) are 11.02 Å, 13.39 Å, and 13.88 Å respectively. Assuming that the density is uniform in the ball core part, we derive the volume occupied by the core from the relation of $Rg = \sqrt{3/5}r$. Dividing this core volume by the volume of the phenyl group of each dendron at density 0.99 g/cm^3, we get the maximum packing number of dendrons in each G2, G3, G4 ball: 12.48, 7.15, 2.62. This result indicates that the packing of dendrons into ball is not yet saturated. The "anisotropic packing" as shown in Figure 4C results in the non-spherical balls and unsaturated packing. Table 2 lists the physical properties of the dendrimer balls formed by 12G2-AG, 12G3-AG, 12G4-AG.

Table 2. Physical properties of the spherical supramolecule formed by 12Gn-AG

Generation	μ(model)[a]	N(Model)[b]	Rg/Å[c]	Rg(Core)/Å[c,d]	μ(Predicted)[e]	μ(exptl)[f]	Saturation degree[g]
2	12	4680	17.75	11.02	12.48	11.3	90.5%
3	6	7092	20.24	13.39	7.15	5.8	81.1%
4	2	7116	20.97	13.88	2.62	1.9	72.5%

[a]We use 12, 6, 2 dendrons to form the dendrimer ball of Generation 2, 3, 4, which are close to the experimental values: 11.3, 5.8, 1.9.[16]

[b]There are totally 4680, 7092, 7116 atoms in each dendrimer ball of 12G2-AG, 12G3-AG, 12G4-AG constructed for simulation. However, the hydrogen atoms in the $C_{12}H_{25}$ tail chains are implicit during molecular dynamics (see section 2.1).

[c]The radius gyration of each ball or core of the ball is evaluated from the best liquid crystal structure at density 0.99 g/cm3 after 60ps NVT molecular dynamics and averaged from the 8 independent balls in the unit cell.

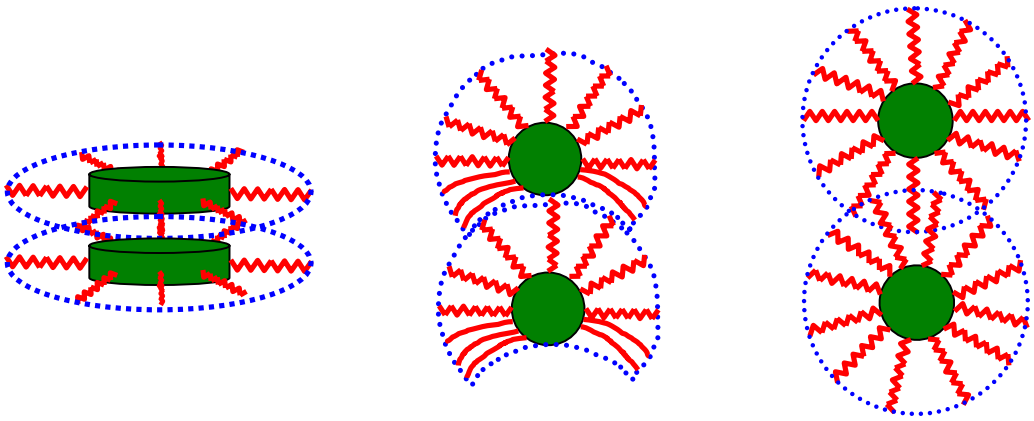
[d]This is the radius gyration of the core part of each ball (exclude the $C_{12}H_{25}$ tail chains)

[e]Based on the relation of $Rg = \sqrt{3/5}r$, we determine the volume occupied by core part of each ball from Rg(Core). Then we divide it by the volume of the core part of each dendron at the density 0.99 g/cm3, we get the maximum packing number of dendrons in each G2, G3, G4 ball.

[f]This packing number was determined from X-ray analysis and listed in the Table 3 of Ref. [16].

[g]Saturation degree is defined as μ(exptl)/ μ(Predicted).

Figure 9 shows the density of states (power spectrums) of various G2 lattices, which includes anisotropic A15 lattice and FCC, BCC lattices composed by sticky shape G2 balls. Other lattices, which are not shown here, have similar power spectrums. From Figure 9, we can see that the power spectra of various lattices are almost indistinguishable, even in the zoom-in Figure 9B. Indeed, the analysis of Figure 7 shows that there is no significant difference in the entropies of various structures.



A. Disk shape composed by tapered fan shape dendron

B. Asymmetric/Punk ball shape as intermediate shape between disk shape and cone shape

C. Isotropic ball shape composed by cone shape dendron

Figure 8. Packing of disk shape, Asymmetric/Punk ball shape, and isotropic ball shape dendrimers

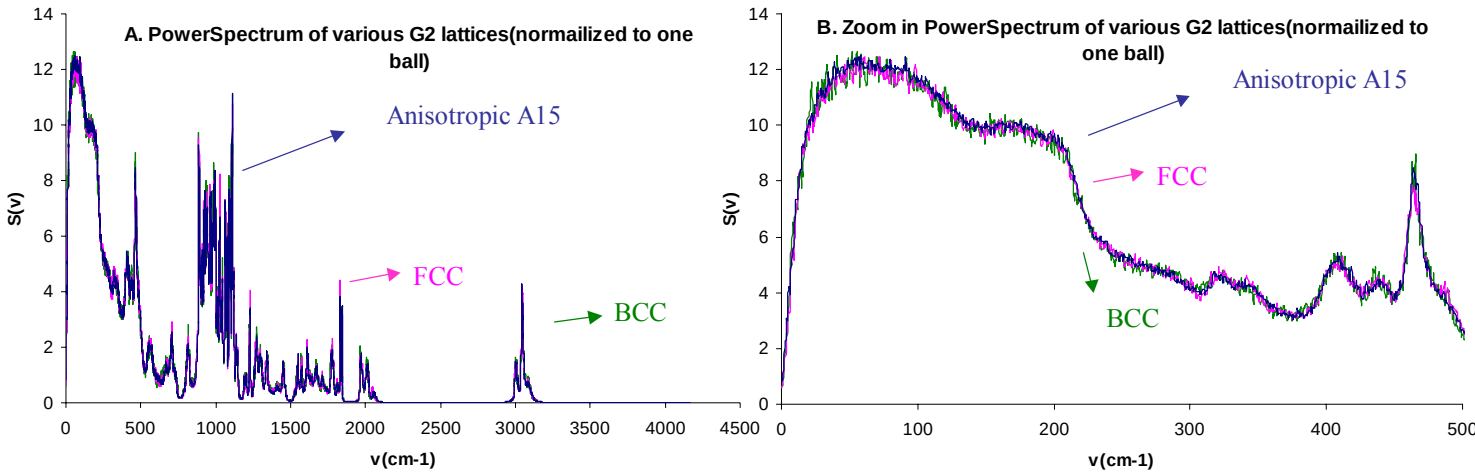


Figure 9. Power spectrum of various G2 structures using the best case for each structure (symmetric sticky balls for FCC and BCC, and asymmetric sticky balls for A15). Each power spectrum is from one of the three MD runs. There results are normalized to one ball; separate colors (blue for A15, purple for FCC, and green for BCC) are used for various structures but the differences are negligible.

3.5 Comparison with experimental X-ray

Figure 10 shows the X-ray intensities of 12G2-AG, 12G3-AG from experiment[16] and prediction from the best structure "anisotropic A15 structure". The predicted intensities fit the experimental intensities very well. This confirms our predicted structure, free energy analysis here and the electron density profile analysis in Ref. [16].

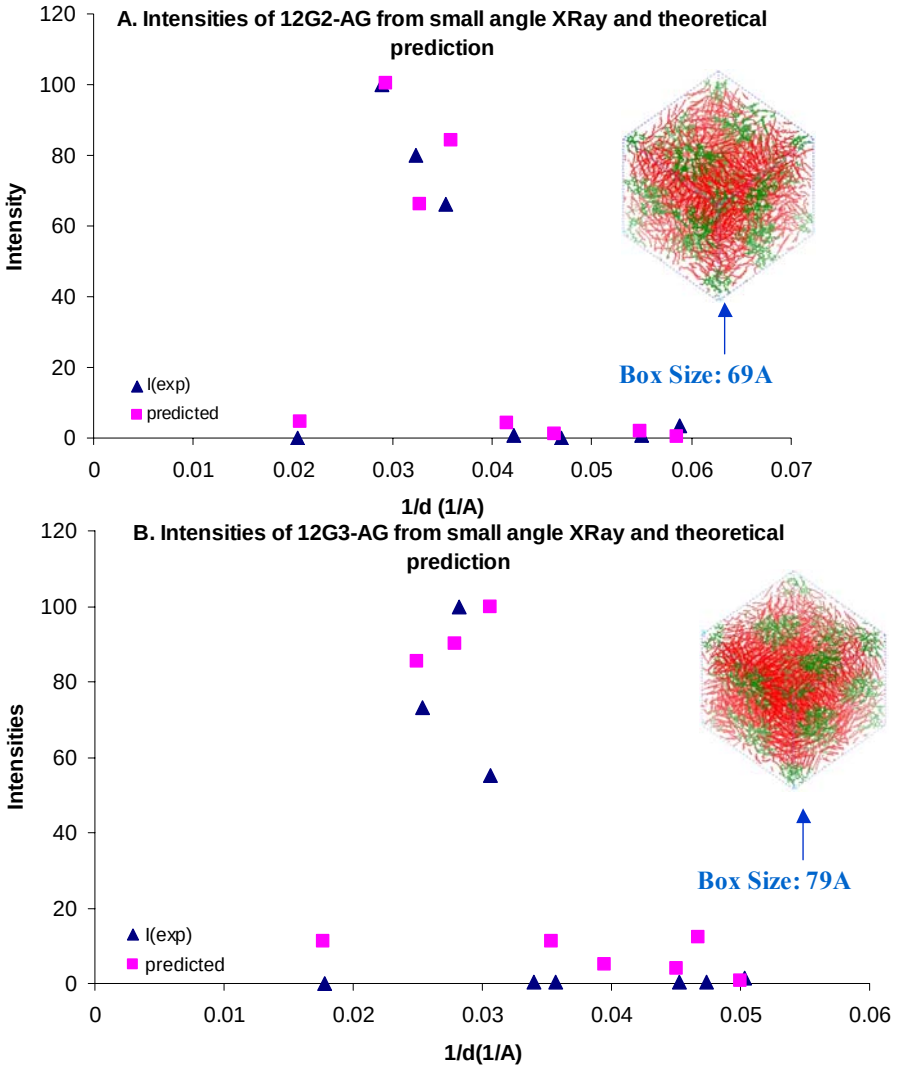
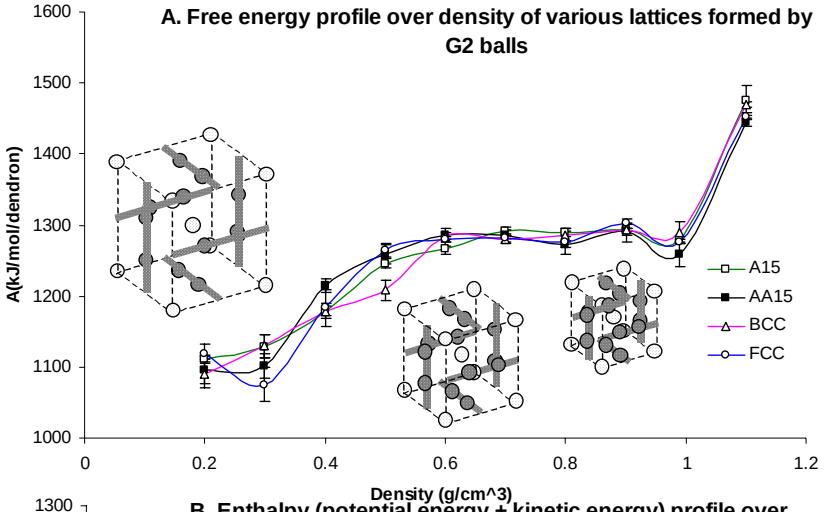
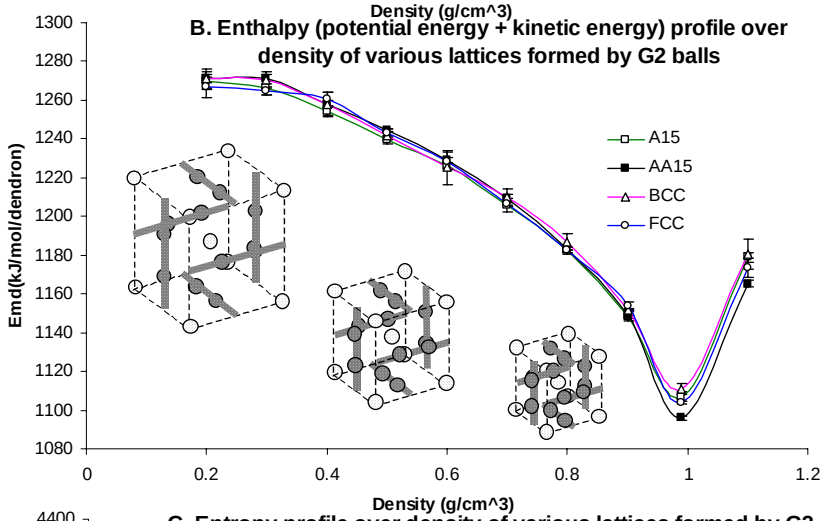


Figure 10. Predicted X-ray intensities of 12G2-AG, 12G3-AG dendrimers from the best structure, Anisotropic A15 structure, compared with experiment[16]. The X-ray diffraction intensities are calculated using the "Diffraction-Crystal" module in Cerius2 4.0. No polarization factor, crystal monochromator factor or temperature factor is applied to the intensity calculations. These intensities are from a single snapshot from the MD trajectory (after 60ps). We apply the factor 0.98 to the predicted intensities of 12G2-AG, which arises from power (11.3/12, 1/3), where exptl has 11.3 dendrons, our structure has 12 dendrons in each ball. The hydrogen atoms of polyethylene chains are not considered during intensities calculation


A. Free energy profile over density of various lattices formed by G2 balls
A(kJ/mol/dendron)
Density (g/cm^3)
A15
AA15
BCC
FCC


B. Enthalpy (potential energy + kinetic energy) profile over density of various lattices formed by G2 balls
Emd(kJ/mol/dendron)
Density (g/cm^3)
A15
AA15
BCC
FCC

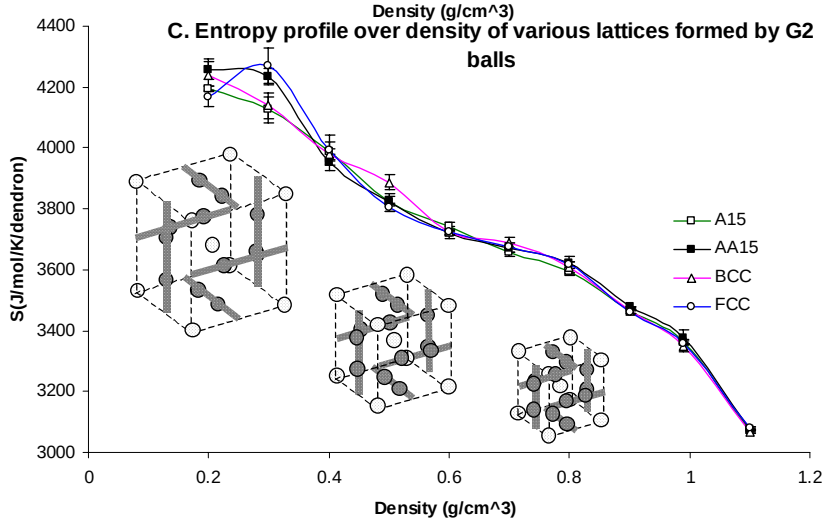

C. Entropy profile over density of various lattices formed by G2 balls
S(J/mol/K/dendron)
Density (g/cm^3)
A15
AA15
BCC
FCC

Figure 11. Free energy, enthalpy, and entropy as a function of density for four cases of G2 dendrimers. AA15 is the A15 structure composed of anisotropic balls on the face and isotropic balls on the corner and body center. The other three structures: A15, FCC, and BCC are based on isotropic balls only. MD simulation details are the same as Figure 7.

3.6 Free energy profile over density of various lattices

In order to understand the packing process of the soft balls, we calculate the free energy, enthalpy, and entropy of various lattices at different densities as shown in Figure 11. The structures are prepared as described in section 3.3 and Figure 6. For the densities lower than 0.99 g/cm3, we fix the central atom H___A in the core of each dendron to keep the ball centers at the desired positions in the lattice during molecular dynamics.

From Figure 11B, we can see that the enthalpies of various lattices all have a minimum at ~ 0.99 g/cm^3. The best structure AA15 (Anisotropic A15) is only favored at density above 0.90 g/cm^3. The enthalpy decreases monotonically during the packing process. The inter-ball vdW interaction dominates the packing energy. It decreases slowly in the early packing stage due to the less interaction among the soft balls at the low density.

The entropies for all structures decrease with increasing density, as shown in Figure 11C. That is because the overlap between the neighboring soft balls constrains the conformations of the chains of the balls, leading to a short-ranged repulsive interaction. The differences among various lattices are negligible above density of 0.6g/cm^3. However, entropy evaluation has a big deviation at the low density as shown in Figure 11C. For nearly isolated soft balls at the low density there is much more flexibility, leading to the large differences.

The free energy profiles are shown in Figure 11A. The free energy profiles of various lattices show the same trend, which can be approximately described as three levels. The lowest level is at the density from 0.2~0.3 g/cm^3. From 0.5~0.99 g/cm3, the free energy shows a flat stage profile as the middle level. The free energy is the worst (highest) at the high density: 1.10 g/cm^3.

Free energy profile over volume of the best Anisotropic A15 lattice at 277K as shown in Figure 12 indicates 2 stable phases: a. condense phase at the density 0.99 g/cm^3; b. isolated Micelle phase at the density 0.30 g/cm^3. This figure is derived from Figure 11A and essentially is A-V plot of the micelle from condense phase to isolated phase. The isolated phase is evaluated in vacuum without solvent and is unphysical. The critical pressure determined from the two stable phases is 0.033Gpa, which is bigger than 1 atm. This implies that in the recrystallization process, the unfavorable solvent at 277K is the driving force to form the condense phase.

In the BCC and FCC lattices, the ball interacts equally with its neighbors. Thus we can derive a 2-body potential from Figure 11. Figure 13 shows the 2-body free energy/enthalpy potential in BCC and FCC lattices. Obviously, the 2-body potential of the soft dendrimer balls is lattice dependent. The reason is that the soft dendrimer balls self adjust its shape and interaction in different lattices. However, the 2-body potential from BCC and FCC lattices share the same trend. The 2-body free energy potential in BCC and FCC lattices can be explained as so called "square shoulder potential".[37-40] In this potential, the soft balls do not interact each other when they are far apart, and they interact via a constant repulsive value at short distance. This soft shoulder arises from the entropy repulsive interaction. The shoulder positions in BCC and FCC are different due to the different neighbor distribution.

The 2-body Emd (enthalpy) potential in BCC and FCC are similar as Lennard Jones potential. And the optimum distances in BCC and FCC are different (See Figure 13B). From the optimum point (at the density 0.99 g/cm^3), we can derive the Lennard Jones 12 6 potentials as shown in Figure 11B (Solid points). The 2-body enthalpy potential of soft dendrimer balls has a much sharper well than the Lennard Jones potential. The interaction of the soft balls increases dramatically when bringing them together.

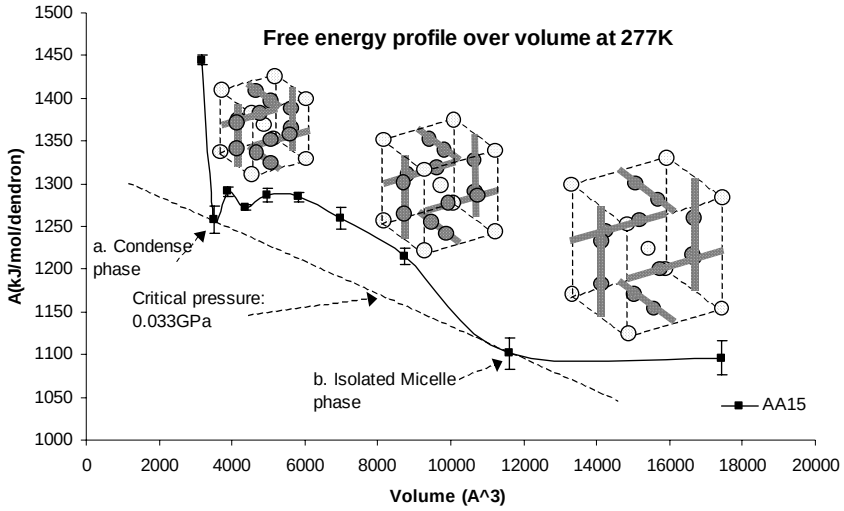


Figure 12. Free energy profile over volume of Anisotropic A15 lattice at 277K

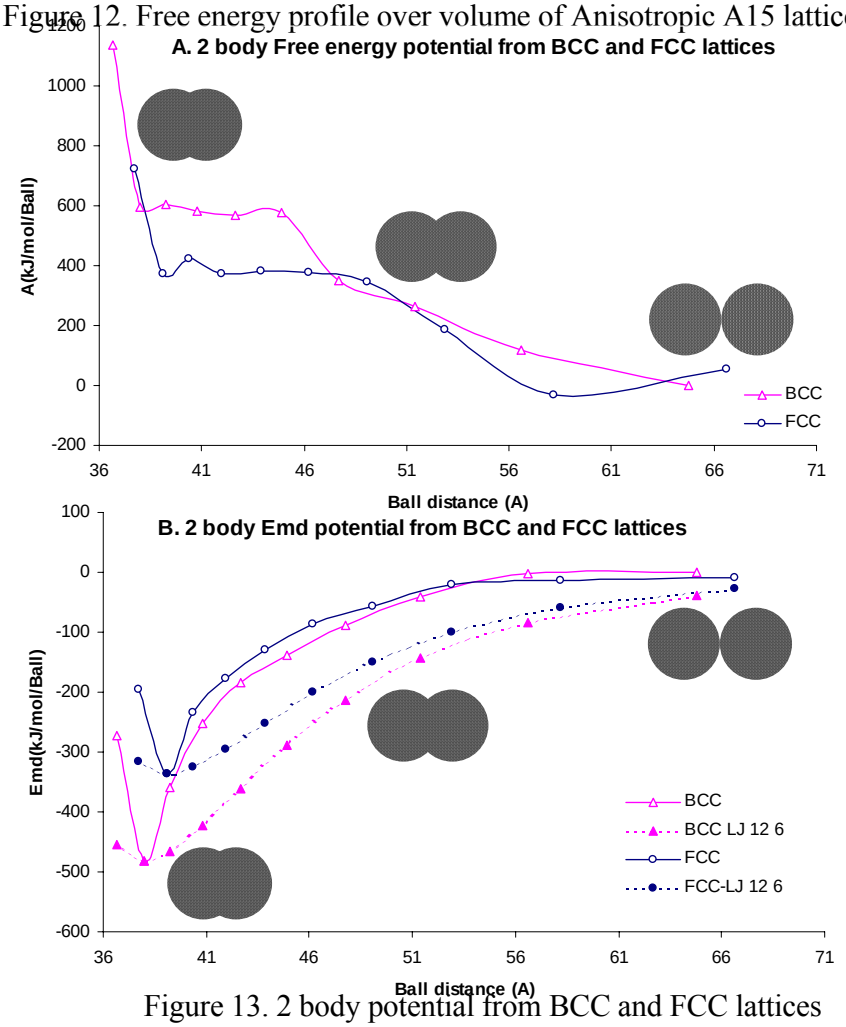


Figure 13. 2 body potential from BCC and FCC lattices

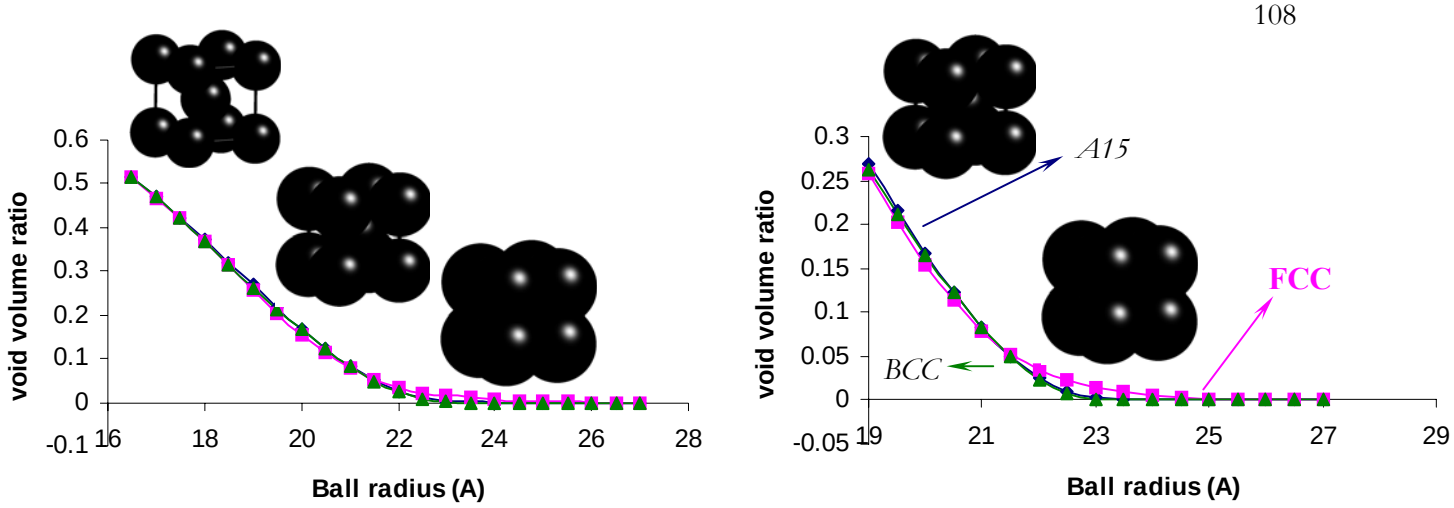


Figure 14. Packing efficiency of ideal soft balls (which overlap each other). The curves are obtained from numerical results.

3.7 Packing efficiency of ideal soft balls (which overlap each other) and some general discussions

In order to investigate the geometric difference of various lattices to pack soft balls, we use the A15, FCC, and BCC lattice boxes of generation 2 above (at the same density: 0.99 g/cm^3), and put ideal balls into desired positions. By increasing the ball radius and let the balls overlap each other, the void volume in the lattice decreases as shown in Figure 14.

When the ball is small enough and there is no overlap between balls, A15, FCC, BCC give the same void volume. By increasing the ball size from 18 Å to 21 Å, the balls in A15 overlap first, then BCC, followed by FCC. That is because FCC is the closest packing lattice for hard balls. In the range of 18 Å ~21 Å, FCC has the least void volume. However, if we increase the ball size from 21 Å to 25 Å, FCC shows the biggest void volume. BCC has the least void volume with A15 is comparable. Further increasing the ball size makes all of the space filled by the soft balls and the void volume is zero for all lattices.

Least void volume means the soft balls use the space in the most efficient way and the overlap between the balls has been kept least. Considering the entropy repulsive

interaction between the soft balls discussed in section 3.6, we conclude that the less overlap in the lattice, the better for the free energy. In view of this, FCC is the lest efficient lattice for ideal soft ball, while BCC is the most efficient lattice.

From the analysis in the previous sections, we can see that dendrimer balls are not ideal soft balls, which can self adjust their shape anisotropiclly to get the best packing of the soft corona. Indeed, they adopt the intermediate shape between disk and isotropic ball to achieve a balance between enthalpy and entropy, which assemble into A15 lattice.

Kamien et al.[41] analyzed the stability of different lattices (A15, BCC, FCC) by considering the tension in the *AB* interface of a diblock copolymer and the stretching of the polymers. They concluded that BCC lattice minimizes the stretching part and A15 lattice minimizes the tension in the interface. They argued that A15 lattice should be favored as the blocks become more symmetric and corroborated this through SCFT.

Indeed, they argued that BCC lattice minimizes the stretching of the polymers, which is consistent with our analysis. We concluded that there exists entropic repulsive interaction between soft balls and BCC lattice minimizes it in terms of ideal soft balls. From an analysis of the competition between enthalpic and entropic interactions in soft balls, we find that the stretching of the polymers (deformation from the sphere cell to the Voronoi / Wigner-Seitz cell) arises from the entropic effects.

However, the tension in the *AB* interface analyzed by them might not be directly applicable to the dendrimer liquid crystal system we studied here. The authors claimed that when the volume fraction of *A*-type monomers (ϕ) is large enough, the *AB* interface takes on the shape of the Voronoi cell. Based upon that, they concluded that A15 minimizes the interfacial tension.

We consider that the packing mechanism of the polyethylene segments of the corona of the soft ball, but not the interfacial tension between aromatic phase and aliphatic phase, is the key factor to determine the efficiency of various lattices. Indeed, Percec's group found that

the density of the polyethylene segments could be used to tune the phase of the dendrimer liquid crystal. Our simulation results indicate that the dimmer balls on the face of A15 structure prefer a non-spherical shape, leading to a good packing of the polyethylene segments. We believe that the interfacial tension is less important between aromatic phase and aliphatic phase, although it might be crucial for the diblock copolymer phases.

Kamien's group also proposed some guidelines for considering the principle to pack colloids.[42,43] They simplified the aliphatic phase in the dendrimer liquid crystal as a dodecyl bilayer, which is dominated by a repulsive free energy that scales as the inverse of the layer thickness, d^{-1}. They argued that A15 is the most stable structure because the entropy is maximized through the minimization of the interfacial area in the bilayer.

We consider the dendrimer liquid crystals to be different from colloidal crystals, which contain solvent. The interaction among colloidal micelles will always be repulsive, even in the crystal. However, in the dendrimer liquid crystal, where there is no solvent, the aliphatic segments of the neighboring balls can pack each other favorably.

Furthermore, we consider that the bilayer approximation may not be accurate enough to distinguish various lattices. Our atomistic simulation results indicate that the various lattices formed by the same type of dendrimer balls are not distinguishable in terms of free energy. (e.g. Figure 7A)

The VAC approach we used here assumes harmonic assumption. By using the same ball to pack into various lattices, we believe this harmonic assumption is unbiased to compare various lattices. There is not any difference of the chemical structures among the various lattices.

The evaluation of the free energy is not completely accurate. There are many other factors influence the absolute free energy of the systems studied here, such as the configurational

entropy part. However, in terms of the relative free energy, the consistent compare gives us the reliable results.

The VAC approach here derives the conformational entropy (translational, rotational, and vibrational) very well. The dendrimer liquid crystal studied here is complicated. Generally we believe that the system cannot transform from one packing approach (Figure 4) to another one easily on the equilibrium density. The means the system is not ergodic. The configuration entropy is not important here and we find that the system only prefers the best configuration.

4. Conclusions

Self-assembled supramolecular organic liquid crystal structures at nanoscale have potential applications in molecular electronic, photonics, and porous nanomaterials. Most of them are aggregated by soft spheres, which have soft coronas and overlap each other in the packing process.

Based on our simulations in atomistic scale, we propose three packing mechanisms of soft balls: "Sticky packing", "Slippery packing", and "Anisotropic packing" and use the vibrational density of state (DoS) derived from classical molecular dynamic simulations to investigate the efficiency of various lattices for soft balls from simulation.

By focusing on the three compounds reported in *J. Am. Chem. Soc.* **1997**, *119*, 1539, which form spheres with nearly integer number of dendrons, we compare the efficiency of various lattices and different packing methods. For the soft spheres with aliphatic corona composed by polyethylene chains, "Sticky packing" is better than "Slippery packing". Anisotropic packed A15 is favored over FCC, BCC lattices. Predicted X-ray intensities of the best structures fit the experiments very well.

"Anisotropic ball packing" proposed here plays an intermediated role between the enthalpy-favored "disk packing" and entropy-favored "Isotropic ball packing", which explains the phase transitions at different temperatures.

Free energy profiles over density of various lattices are essentially the same, which indicates that the preferred lattice is not determined during the packing process. Both enthalpy and entropy decrease as the density increases. Free energy profile over volume shows two stable phases: condense phase and isolated micelle phase. The 2-body potential of the soft dendrimer ball is lattice dependent, because it self adjusts its shape and interaction in different lattices. The shape of the free energy potential is similar to the "square shoulder potential".

A model explaining the packing efficiency of ideal soft balls in various lattices is proposed in terms of geometrical consideration. BCC has the least void volume for the ideal soft ball while FCC has the biggest.

5. Acknowledgments

This paper is dedicated to Paul Miklis, who initiated the study of these Percec systems shortly after the first publication. His early demise was a loss to all of us. In addition to the early stimulation by Paul, we also thank Dr. Tahir Cagin and Virgil Percec for many helpful discussions about these dendrimers and we thank Dr. Mario Blanco with help using Lammps.

This research is based in part on work supported by the U. S. Army Research Laboratory and the U. S. Army Research Office under grant number DAAG55-97-1-0126 (MURI-program officer Doug Kiserow). The facilities of the MSC used in this research have been supported by grants from ARO (DURIP), ONR (DURIP), NSF (MRI), and IBM (SUR). Other support to the MSC is provided by NIH, NSF, DOE, Chevron-Texaco, General Motors, Seiko Epson, Asahi Kasei, Beckman Institute and Toray Corp.

6. References

(1)Ungar, G.; Liu, Y.; Zeng, X.; Percec, V.; Cho, W. D. *science* **2003**, *299*, 1208.

(2) Percec, V.; Glodde, M.; Bera, T.K.; Miura, Y.; Shiyanovskaya, I.; Singer, K. D.; Balagurusamy, V. S. K.; Heiney, P.A.; Schnell, I.; Rapp, A.; Spiess, H. W.; Hudson, S. D.; Duan, H *nature* **2002**, *419*, 384.

(3) Percec, V.; Ahn, C.-H.; Ungar, G.; Yeardley, D. J. P.; Moller, M.; Sheiko, S. S. *nature* **1998**, *391*, 161.

(4) Hudson, S. D.; Jung, H.-T.; Percec, V.; Cho, W.-D.; Johansson, G.; Ungar, G.; Balagurusamy, V. S. K. *science* **1997**, *278*, 449.

(5)Lopes, W. A.; Jaeger, H. M. *nature* **2001**, *414*, 735.

(6) Attard, G. S.; Goltner, C. G.; Corker, J. M.; Henke, S.; Templer, R. H. *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 1315.

(7)Jenekhe, S. A.; Chen, X. L. science 1999, 283, 372.

(8)Zubarev, E. R.; Pralle, M. U.; Li, L.; Stupp, S. I. science 1999, 283, 523.

(9)Stupp, S. I.; Braun, P. V. science 1997, 277, 1242.

(10) Stupp, S. I.; LeBonheur, V.; Walker, K.; Li, L. S.; Huggins, K. E.; Keser, M.; Amstutz, A. science 1997, 276, 384.

(11Orr, G. W.; Barbour, L. J.; Atwood, J. L. science 1999, 285, 1049.

(12Harada, A.; Kataoka, K. science 1999, 283, 65.

(13Muthukumar, M.; Ober, C. K.; Thomas, E. L. science 1997, 277, 1225.

(14)Percec, V.; Chu, P.; Ungar, G.; Zhou, J. J. Am. Chem. Soc. 1995, 117, 11441.

(15)Percec, V.; Johansson, G.; Ungar, G.; Zhou, J. J. Am. Chem. Soc. 1996, 118, 9855.

(16) Balagurusamy, V. S. K.; Ungar, G.; Percec, V.; Johansson, G. J. Am. Chem. Soc. 1997, 119, 1539.

(17) Yeardley, D. J. P.; Ungar, G.; Percec, V.; Holerca, M. N.; Johansson, G. J. Am. Chem. Soc. 2000, 122, 1684.

(18) Borisch, K.; Diele, S.; Goring, P.; Tschierske, C. Chem. Commun. 1996, 2, 237.

(19) Borisch, K.; Diele, S.; Goring, P.; Muller, H.; Tschierske, C. Liquid Crystals 1997, 22, 427.

(20) Cheng, X. H.; Diele, S.; Tschierske, C. Angew. Chem. Int. Ed. Engl. 2000, 39, 592.

(21) Cheng, X. H.; Das, M. K.; Diele, S.; Tschierske, C. Langmuir 2002, 18, 6521.

(22) Mau, S.-C.; Huse, D. A. Phys. Rev. E 1999, 59, 4396.

(23) Siepmann, J. I.; Karaborni, S.; Smit, B. nature 1993, 365, 330.

(24) Smit, B.; Karaborni, S.; Siepmann, J. I. J. Chem. Phys. 1995, 102, 2126.

(25) Martin, M. G.; Siepmann, J. I. J. Am. Chem. Soc. 1997, 119, 8921.

(26) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. J. Am. Chem. Soc. 1984, 106, 765.

(27)Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. J. Comput. Chem. 1986, 7, 230.

(28) Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. J. Am. Chem. Soc. 1984, 106, 6638.

(29) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. J. Phys. Chem. 1990, 94, 8897.

(30) In Large-Scale Atomic/Molecular Maasively Parallel Simulator, version 5.0; CRADA collaboration, Sandia National Laboratory, USA, 1997.

(31) Sadoc, J. F.; Charvolin, J. J. Phys. France 1986, 47, 683.

(32) Israelachvili, J. Intermolecular and Surface Forces; 2nd ed.; Academic Press: London, 1992.

(33) Kratzat, K.; Finkelmann, H. J. Colloid. Interface Sci. 1996, 181, 542.

(34) Gruner, S. M. J. Phys. Chem. 1989, 93, 7562.

(35) Tate, M. W.; Eikenberry, E. F.; Turner, D. C.; Shyamsunder, E.; Gruner, S. M. Chem. Phys. Lip. 1991, 57, 147.

(36) Accelrys_Inc Cerius2 Modeling Environment, Release 4.0; Accelrys Inc: San Diego.

(37) Blohuis, P.; Frenkel, D. J. Phys. C 1997, 9, 381.

(38) Rascon, C.; Velasco, E.; Mederos, L.; Navascues, G. J. Chem. Phys. 1997, 106, 6689.

(39) Lang, A.; Kahl, G.; Likos, C. N.; Lowen, H.; Watzlawek, M. *J. Phys. C* **1999**, *11*, 10143.

(40) Velasco, E.; Mederos, L.; Navascues, G.; Hemmer, P. C.; Stell, G. *Phys. Rev. Lett.* **2000**, *85*, 122.

(41)Grason, G. M.; Didonna, B. A.; Kamien, R. D. **2003**, *91*, art. no. 058304.

(42)Ziherl, P.; Kamien, R. D *Phys. Rev. Lett.* **2000**, *85*, 3528.

(43)Ziherl, P.; Kamien, R. D *J. Phys. Chem. B* **2001**, *42*, 10147.

Chapter 4

THE CONTINUOUS CONFIGURATIONAL BIASED TX METHOD FOR GENERATING AMORPHOUS POLYMER AND DENDRIMER ATOMISTIC STRUCTURES

ABSTRACT

Although computer simulation has developed as a powerful research tool to study polymer/ dendrimer materials properties recently, it has been hampered by the difficulties of sampling amorphous polymer/ dendrimer configurations efficiently. Here we develop the efficient Continuous Configurational Biased TX (CCBTX) method to generate high quality amorphous polymer and dendrimer atomistic structures. When sampling i-th segment in dense multiple chains polymer or dendrimer, we don't know the positions of the atoms in the future segments. The future segments include not only the future segments in the same chain/branch of i-th segment, but also the future segments in the other chains/ branches. We find that it improves the CCB method efficiently to put TXS atoms at the end of each chain/ branch to represent the future segments of that chain/branch. We call the new developed CCB method with TXS atoms in torsion sampling as CCBTX method. The method is much more efficient than the available commercial software module. The code is implemented in C++ and ported in python environment, which provides friendly interface.

1. Introduction

Molecular simulation of bulk polymers provides an atomistic level description of the amorphous material, necessary to understand its structure and to formulate the mechanisms of material behavior, which can be applied to study properties important to polymer applications, such as thermodynamic properties, glass transition, small-molecule sorption and diffusion, plastic and elastic deformation, and molecular mobility mechanisms.

Until recently, the computer simulation of polymer in dense phases or dendrimer has been hampered by the difficulties associated with sampling configuration space efficiently.[1] Molecular dynamics simulation of a simple liquid can start from some artificially prepared, sometimes ordered, configuration.[2] Because of the short relaxation times it does not take long for the system to 'forget' it, so that the calculated properties of the liquid are independent of this initial configuration. However, compared with the simple molecular liquid, the complex molecular architecture and disordered structure that lacks lattice symmetry make amorphous polymers structurally incredibly complicated. It has been observed that the polymer structures and simulation results obtained are very sensitive to the way initial configurations are prepared.[3,4] This is a natural consequence of the long relaxation times peculiar to the polymer/dendrimer molecules. The longest relaxation time of the polymer chain increases as a squared chain length and can reach a tenth of a second.[5] A high quality initial configuration for amorphous polymer /dendrimer is critical for the success of the molecular dynamics.

Efforts have been made to improve the amorphous polymer structures from Monte Carlo method by cutting and re-growing pieces of the polymer chain.[6-9] There is still no efficient way to generate the uncorrelated representative amorphous polymer / dendrimer structures. We develop here a highly efficient variant of the Monte Carlo method to generate high quality amorphous polymer / dendrimer directly.

2. CCBTX method

2.1 Simple sampling direct Monte Carlo (SS-DMC)

For polymer systems described with rigid constraints on bond lengths and angles, SS-DMC is preceded by a random sampling of complete set of torsion angles. In the conventional direct Monte Carlo (DMC) method, which is used for estimating the free energy, an isolated polymer chain is generated by random step-by-step sampling of torsion angles. SS-DMC has been extensively used for self-avoiding walk studies on a lattice.[10] However,

because of drastic sampling attrition, SS-DMC is not applicable to long polymer chains, even for lattice systems.

2.2 Independent rotational sampling

The sampling efficiency of SS-DMC is improved by applying rotational biased sampling, in which torsions are sampled using a weighting function based on the Boltzmann factor of the torsion energy. This is denoted as independent rotational sampling (IRS).[11] For IRS the normalized torsion weighting function (TWF), W_{IRS} is defined as

$$W_{IRS}(\phi) = \frac{g_{IRS}(\phi)}{z_{IRS}} \tag{1}$$

where

$$z_{IRS} = \int_0^{2\pi} g_{IRS}(\phi) d\phi \tag{2}$$

$$g_{IRS}(\phi) = \exp[-\beta E_t(\phi)] \tag{3}$$

2.3 NonBond bias corrected TWF

In independent rotational sampling (IRS), torsion angles are generated in accordance with the weight factor in Eq. (1). With the use of W_{IRS}, IRS effectively excludes high torsion energies throughout the MC sampling. However, spatial overlaps between nonbonding atoms are inevitable, leading to high configurational energies. In order to exclude these overlaps, which dominate the energy of the amorphous polymer /dendrimer, information about the spatial environment in the vicinity of the growing chain end should be introduced into the TWF. The resulting form of the TWF, W^*, is given by

$$W_i^*(\phi_i;\phi_4,...,\phi_{i-1}) = \frac{g_i^*(\phi_i;\phi_4,...,\phi_{i-1})}{z_i^*(\phi_4,...,\phi_{i-1})} \tag{4}$$

where

$$z_i^*(\phi_4,...,\phi_{i-1}) = \int_0^{2\pi} g_i^*(\phi_i;\phi_4,...,\phi_{i-1})d\phi_i \tag{5}$$

$$g_i^*(\phi_i;\phi_4,...,\phi_{i-1}) = g_{IRS}(\phi_i)\exp\left[-\beta\sum_{j=1}^{i-4}E_{LJ}(r_{ij})\right] \tag{6}$$

*W** must be calculated at every step since it depends on all previous steps. The computation time for this TWF is approximately proportional to the step number, *i*; therefore, this sampling method becomes too expensive for systems containing a large number of atoms.

2.4 Continuous configurationally biased (CCB) direct Monte Carlo

To remedy the above problem with rotational sampling, a cutoff length for non-bonding interactions is introduced into the TWF calculation and the implemented off-lattice method is called continuous configurational biased (CCB) direct Monte Carlo method.[11-14] The length of R_c should be taken larger than $l+\sigma$ (l is bond length and σ is Lennard –Jones 12-6 potential parameter) in order to ensure that all possible atomic overlaps are checked. Boltzmann factors for the nonbonding energy between *i*th atom and all other atoms inside the cutoff sphere are included in TWF, W_{CCB}, as

$$W_i^{CCB}(\phi_i;\phi_4,...,\phi_i) = \frac{g_i^{CCB}(\phi_i;\phi_4,...,\phi_{i-1})}{z_i^{CCB}(\phi_4,...,\phi_{i-1})} \tag{7}$$

where

$$z_i^{CCB}(\phi_4,...,\phi_{i-1}) = \int_0^{2\pi} g_i^{CCB}(\phi_i;\phi_4,...,\phi_{i-1})d\phi_i \tag{8}$$

$$g_i^{CCB}(\phi_i;\phi_4,...,\phi_{i-1}) = g_{IRS}(\phi_i)\exp\left[-\beta\sum_{j=1}^{i-4}\Theta(R_C - r_{ij})E_{LJ}(r_{ij})\right] \tag{9}$$

and $\Theta(R)$ is the Heavyside step function

$$\Theta(R)=0, \text{ if } R<0; \ \Theta(R)=1, \text{ if } R\geq 0 \tag{10}$$

The computation time for W_{CCB} is almost independent of i because the only nonbonding atoms considered are those in the local vicinity of a growing chain end. In addition, the list of atoms inside the cutoff circle for the ith atom is automatically available since all the necessary atomic distances were calculated to obtain the energy at the just previous step.

2.5 Continuous configurationally biased TX (CCBTX) direct Monte Carlo

Based on CCB method developed previously,[11,15] we extend it to CCBTX method in order to generate condensed state of multichain amorphous polymer and dendrimer systems. Previously, CCB method has been focused on simulation of united atomistic model of polyethylene[11,15] here we develop CCBTX method for generalized purpose, which can be used to generate high quality dense polymer/dendrimer systems with any detailed atomistic architecture.

$$W_i^{CCBTX}(\phi_i;\phi_4,...,\phi_i)=\frac{g_i^{CCBTX}(\phi_i;\phi_4,...,\phi_{i-1})}{z_i^{CCBTX}(\phi_4,...,\phi_{i-1})}$$

$$z_i^{CCBTX}(\phi_4,...,\phi_{i-1})=\int_0^{2\pi} g_i^{CCBTX}(\phi_i;\phi_4,...,\phi_{i-1})d\phi_i$$

$$g_i^{CCBTX}(\phi_i;\phi_4,...,\phi_{i-1})=g_{IRS}(\phi_i)\exp\left[-\beta\sum_{j=1}^{i-4}\Theta(R_C-r_{ij})E_{LJ}(r_{ij})-\beta\sum_{k=1}^{NumTX}\Theta(R_C-r_{ik})E_{LJ}(r_{ik})\right]$$

To implement the continuous static Monte Carlo method, we first dissect the sampling chain into segments based on the rotatable bonds. Then the segments are added one by one as the desired sequence to construct a complete polymer, dendrimer, or polypeptide. In each step to add one segment, first we evaluate the torsion weighting function (TWF) on the torsion between the current segment and the parent segment. The normalized torsion

weighting function (TWF) W_{CCBTX} is obtained from the above equations. Then we get the auxiliary distribution $P_{\mathrm{CCB}}(\phi)$ is defined as

$$P_{CCB}(\phi) = \int_0^{\phi} W_{CCB}(\phi')d\phi'$$

A random number ξ, uniformly distributed in the interval [0,1] is generated and torsion ϕ is obtained by requiring

$$P_{CCB}(\phi) = \int_0^{\phi} W_{CCB}(\phi')d\phi' = \xi$$

Here we implement the new code with full atomistic description instead of united atom model. During the implementation, we found that it is necessary to improve the CCB sampling scheme to utilize it for full atomistic description.

In the united atom model as shown in Figure 1a, only one TX atom's position is changed when we rotate the torsion. And the torsion weighting function evaluated is appropriate to be used to put one more united atom in the next step by Monte Carlo scheme.

However, for the full atomistic model as shown in Figure 1b, when we rotate the torsion, all the three atoms: TX atom and two hydrogen atoms will be moved. And these three atoms play the similar role in terms of the van der Waals interactions. The most important factor in the torsion weighting function is the spatial overlapping of the atoms. Because there will be more atoms attached to TX atom later, the torsion weighting function obtained in Figure 1b doesn't describe the situation appropriately. Thus we implement it with TXS scheme as shown in Figure 1c. Because we don't know the next torsion angle, we cannot determine the positions of the atoms will be attached to TX atom. We use a phenyl ring instead of the regular three atoms to attach to TX atom. In this way, the torsion weighting function will be better than it evaluated as in Figure 1b.

In addition, different from the united atom model, there are hydrogen atoms connected to the backbone in the full atomistic model. And it is not reasonable to fix all the bond

lengths and bond angles in the Monte Carlo sampling procedure. Thus we combined the Monte Carlo sampling with the energy minimization. After each segment added to the chain, we perform 20 energy minimization steps to relax the system.

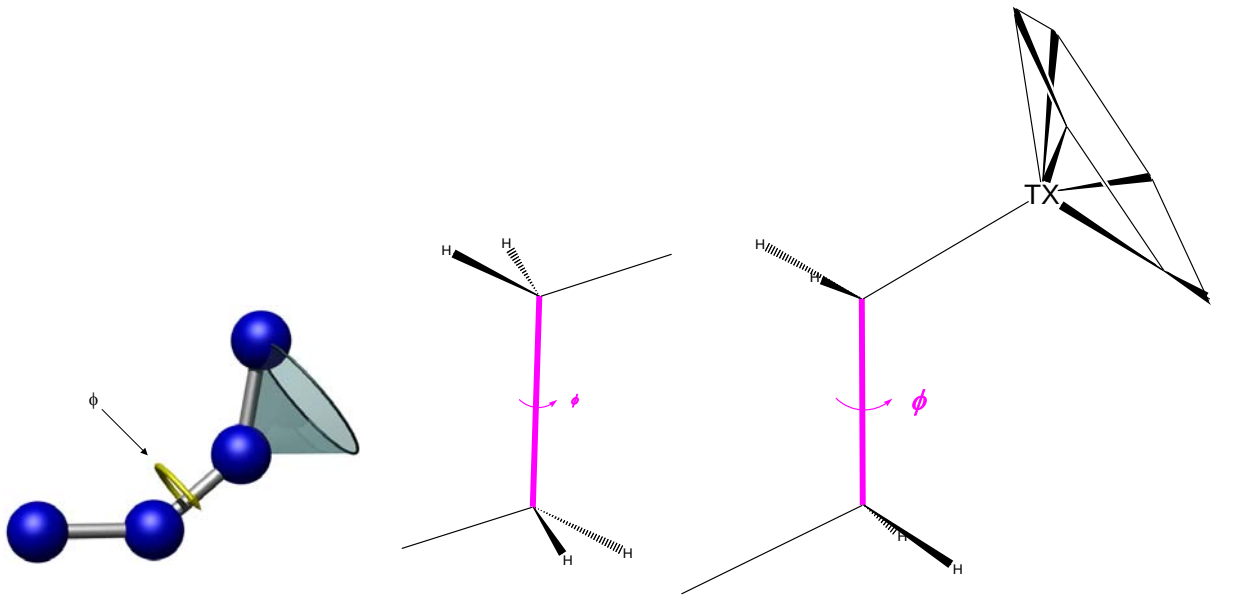


Figure 1. Difference between united atom model and full atomistic model in CCB sampling

3. Distance restraint implementation

In order to utilize our CCBTX method in predicting protein fold structure or cyclic polymer chains, we implement the distance restraint. An efficient restraint technique ensures that the chain conformations are consistent with a set of user-defined inter-monomer distance restraints. Consider using a buildup procedure to construct a chain where segment *j* and

segment k are less than 6 Å apart ($j<k$). The simplest approach is to randomly enumerate all possible conformations of segment j through segment k and discard the "dead end" conformations that do not satisfy this restraint. Unfortunately, this approach becomes prohibitively expensive as the sequential distance between j and k increases, since the detection of a dead end occurs after the construction of segment k.

An algorithm that can determine if a conformation is a dead end prior to the addition of segment k yields a vast improvement in efficiency. The longest distance traversed by each segment addition step is a single bond length, l=1.5 Å. Thus, it is impossible to place segment k within 6 Å of segment j if segment i ($j<i<k$) is greater than 6 + 1.5(k-i) angstroms from residue j. Thus, it is possible to predict at step i if a conformation must eventually result in a dead end at step k.

The geometry restraint method incorporated into the current code is slightly more complex in that it also considers the angle between segments j, i, and i-1. Figure 2 shows the possible positions for segment i+4 in the chain model in the chain model with 120^o bond angle when ϕ_{i+2}, ϕ_{i+3}, and ϕ_{i+4} =0^o or 180^o. Consider a cylindrical coordinate system where the z-axis travels through the bond between segment i-1 and segment i, and the z-axis origin is at segment i-1. The radial axis, ρ, represents the perpendicular distance to the z-axis. In the figure, the solid line around the perimeter traces the maximum radial distance that segment i+4 may be from the z-axis for a given value of z. Hence, this solid line represents the most extreme position in (z,ρ) space that segment (i-1+5) may be placed from segment i-1 and i. Similar diagrams lead to a general expression for the maximum value of ρ, ρ_{max}, for an arbitrary segment (i-1)+n at a specific z-coordinate. Defining

$$\rho_{peak}=(n - 1)(l \sin 60^o)$$

if n is even, then z must lie between

$$\{z_{min}, z_{max}\} = \{(-3l/4)(n-4), (3l/4)(n)\}$$

and two cases define ρ_{max}:

(a.1) for $z >= 3l/2$,

$\rho_{max} = \rho_{peak} - (\tan 30^o)(z - (3l/2))$

(a.2) for $z < 3l/2$,

$\rho_{max} = \rho_{peak} + (\tan 30^o)(z - (3l/2))$

If n is odd, then z must lie in the range

$\{z_{min}, z_{max}\} = \{(-l/4)(2+3(n-5)), (l/4)(4+3(n-1))\}$

and two cases define ρ_{max}:

(b.1) for $z >= l$,

$\rho_{max} = \rho_{peak} - (\tan 30^o)(z - l)$

(b.2) for $z < l$,

$\rho_{max} = \rho_{peak} + (\tan 30^o)(z - l)$

Thus, the equations above specify the greatest distance in (z, ρ) space that any segment (i-1+n) may be placed from segments i-1 and i.

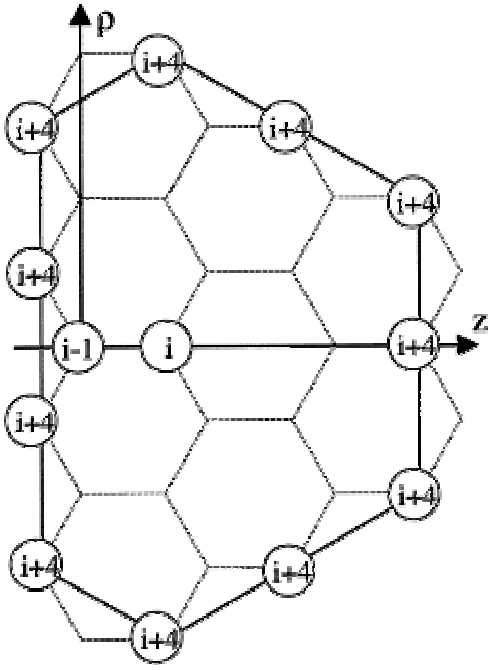


Figure 2. The allowed positions of segment *i*+4 in relation to segment *i*-1 and segment *i* when segments *i*-1, *i*, *i*+1, *i*+2, *i*+3 and *i*+4 all lie in the same plane. For the cylindrical coordinate system (*z*, *r*), the maximum value or *r* for segment *i*+4 may be expressed as a function of *z*. (From Debe, D. A.; Carlson, M. J.; Sadanobu, J.; Chan, S. I.; Goddard, W. A. *J. Phys. Chem. B* **1999**, *103*, 3001)

4. Efficiency of the CCBTX method

Our implemented code shows much better results than the commercial software. Table 1 shows the comparison results of our CCB code with commercial AMB code to generate amorphous condense polyethylene chains at 0.90g/cm^3 density. Our CCB code generates much better initial guess structures than AMB code.

Table 1. Comparing our CCB code with commercial AMB code

Sample (kcal/mol)	1	2	3	4	5
CCB	138.77	198.61	456.63	215.63	175.55
AMB	0.15E+07	0.29E+07	0.32E+07	0.27E+07	0.41E+07

5. Summary

Although computer simulation has developed as a powerful research tool to study polymer/dendrimer materials properties recently, it has been hampered by the difficulties of sampling amorphous polymer/dendrimer configurations efficiently. We develop the efficient Continuous Configurational Biased TX (CCBTX) method to generate high quality amorphous polymer and dendrimer atomistic structures directly. The method developed here can also be used for protein fold prediction.

6. Acknowledgment

This research was partially supported by grants from NSF (CHE99-85574), DOE ASCI ASAP, and ARO-MURI. The facilities of the MSC are also supported by grants from DOE, NSF (CHE 99-77872), ARO (MURI), ARO (DURIP), IBM-SUR, NIH, ChevronTexaco, 3M, Seiko-Epson, Avery-Dennison, General Motors, Kellogg's, Asahi Chemical, Beckman Institute, and Nippon Steel.

7. Reference

(1) Kotelyanskii, M. *TRENDS IN POLYMER SCIENCE* **1997**, *5*, 192.
(2) Allen, M. P. and Tildesley, D. J. *Computer simulation of liquids*; Clarendon Press, 1993.
(3) McKechnie, J. I.; Brown, D. and Clarke, J. H. R. *Macromolecules* **1992**, *25*, 1562.
(4) van der Vegt, N. F. A.; Briels, W. J.; Wessling, M. and Strathmann, H.. *J. Chem. Phys.* **1996**, *105*, 8849.
(5) Doi, M. and Edwards, S. F. *The theory of polymer dynamics*; Clarendon press, 1994.

(6) Siepmann, J. I. and Frenkel, D. *Mol. Phys.* **1992**, *75*, 59.
(7) Dodd, L. R.; Boone, T. D. and Theodorou, D. N. *Mol. Phys.* **1993**, *78*, 961.
(8) Leontidis, E.; Forrest, B. M.; Widmann, A. H. and Suter, U. W. *J. Chem. Soc., Faraday Trans.* **1995**, *91*, 2355.
(9) dePablo, J. J.; Laso, M. and Suter, U. W. *J. Chem. Phys.* **1992**, *96*, 2395.
(10) Kremer, K. and Binder, K. *Computer physics Reports* **1988**, *7*, 25.
(11) Sadanobu, J. and Goddard, W. A., III. *J. Chem. Phys.* **1997**, *106*, 6722.
(12) Siepmann, J. I. *ibid* **1990**, *70*, 1145.
(13) Siepmann, J. I. and Frenkel, D. *Mol. Phys.* **1991**, *72*, 169.
(14) Frenkel, D.; Mooji, G. C. A. M.; and Smit, B. *J. Phys. Condensed Matter* **1991**, *3*, 3035.
(15) Sadanobu, J. and Goddard, W. A., III. *Fluid phase equilibria* **1998**, *144*, 415.
(16) Siepmann, J. I.; Karaborni, S.; Smit, B. *Nature* **1993**, *365*, 330.

C h a p t e r 5

THERMODYNAMIC FUNCTIONS, CRITICAL EXPONENTS AND THETA TEMPERATURES FOR POLYMER CHAINS FROM CONTINUOUS CONFIGURATION BOLTZMANN BIASED DIRECT MONTE CARLO CALCULATIONS

ABSTRACT

We use Continuous Configuration Boltzmann Biased (CCBB) Monte Carlo sampling (no lattice) to derive the thermodynamic properties [free energy (A), entropy (S), internal energy (E), and radius gyration ($<Rg^2>$)] for isolated polymer chains from temperatures below the theta temperature (T_θ) to the high temperature limit using chain lengths from N=6 to 400. This is carried out for a force field (FF) accurately describing polyethylene (PE) and for modifications that implicitly simulate a range of solvents.

This work establishes the existence of a phase we denote as the Flory Phase, for which the partition function scales as $Z_N \propto N^{\gamma-1}\mu^N$, leading to a free energy of the form $\frac{A_N}{k_B T} = -\ln Z_N = -N\ln\mu - (\gamma-1)\ln N + C_A$. Throughout the Flory phase the polymer size scales as $R_g^2 \sim N^{2\nu}$. In the high temperature limit, the critical constants are γ=1.153, and 2ν=1.168, which are consistent with previous self-avoiding walk (SAW) and renormalization group (RG) estimates. As the temperature decreases, both γ and 2ν decrease, leading to points at which γ=1 and at which 2ν=1. For 2ν=1, the size of the polymer chain scales linearly with length, describing a Gaussian coil; this temperature is denoted as the Flory temperature (T_θ). For the temperature at which γ=1 (no free energy pre-factor), the free energy of the system behaves as if it were an ideal gas of monomers. We define this as the free energy theta temperature (denoted as $T_{\theta\gamma}$. This Flory phase exists from ~ 0.7 T_θ to T = infinite.

Throughout the Flory Phase, there is a balance between repulsive entropic effects arising from self-avoidance (Pauli principle) that favors extension and attractive enthalpic effects arising from London dispersion (van der Waals attraction) that favors collapse. We find that γ can be partitioned as $(\gamma\text{-}1) = (\gamma_S - 1) - (\gamma_E\text{-}1)$. At infinite temperature, $\gamma_E = 0$ and $\gamma_S = 1.155$. As the temperature decreases, the chain shrinks leading to an increase in both factors. At the free energy theta temperature, we find $\gamma_E = \gamma_S \sim 1.65$. At lower temperature, attractive effects dominate, leading to $\gamma_E > \gamma_S$ and $\gamma < 1$.

Throughout the Flory Phase we find that the scaled chemical potential per monomer is μ=5.649. Ignoring chain avoidance this would be 2π=6.28, indicating that the self-avoidance factor is 10%. In the high temperature limit we find that the internal energy per monomer is constant (no pre-factor) with e_0=0.0187, confirming a prediction by Des Cloizeau.

We derive a general mean field model valid throughout Flory phase. This shows that with increasing temperature, μ increases and e_0 decreases uniformly. At the high temperature limit we find that the non-bond interaction becomes temperature independent.

We find that $T_{\theta\gamma} \sim 1.1\ T_\theta$. By considering an expansion of the free energy with concentration, we find that $T_{\theta\gamma} > T_\theta$ results from three-body terms. At the point where γ=1, the free energy is linear in the number of monomers, but since both γ_E and γ_S are greater than 1, the system cannot be treated exactly as an ideal gas of monomers.

To determine how solvent affect the thermodynamics for the Flory Phase, we use a model in which the van der Waals interactions are scaled to mimic the relative importance of polymer-solvent interactions versus intra-chain interactions. We find that the solvent effects correlate linearly with the theta temperature.

By comparing to the results of molecular dynamics studies that allowed variable bonds and angles, we conclude that the assumption of constant bonds and valence angles within the polymer chain has negligible affect on the properties of the Flory phase.

1. Introduction

Many properties of polymer materials (e.g. rubber elasticity, liquid crystallinity) are well characterized in terms of single chain conformations and dynamics. Indeed the development of experimental and theoretical tools for describing the physics of long flexible chains provides remarkably simple scaling properties that dominate the behavior for many properties and allow common features of many chemically different systems to be understood.

Neutron diffraction [1] and light scattering [2] have been used to characterize the conformations and molecular weights of real polymers in dilute solutions, providing results that sometimes challenge the simple scaling laws. To obtain a theoretical understanding of how the physical properties depend on the chemical nature of the polymer mixture (blend or copolymer), many theoretical methods (Functional integrals, Feynman diagrams, many-body theory, [3] Exact summation [4] of short chains and Monte Carlo method [5] on lattice models) have been developed and applied to real polymers. There has been significant progress[6-13] over the last 40 years, but most of this work has employed an approximation in which the monomers are assumed to be located on a three-dimensional lattice. Here we report the first nonlattice study of temperature dependence of the thermodynamic functions with critical exponents / parameters over the higher temperature range. To do this we use the continuous configurational Monte Carlo method. The elimination of the lattice approximation allows us to obtain physical insight into the thermodynamics.

The calculation details are described in Section 2, and the results are discussed in Section 3 and 4. The summary is in Section 5.

2. Calculation details

The behavior of the thermodynamic quantities Helmholtz free energy (A), entropy (S), and internal energy (E) have not been previously reported for continuous polymer systems at higher temperatures (at the theta temperature and above). We will use the CCBB Monte Carlo method[14] to examine these quantities for continuous distributions of polymer chain conformations using a force field (FF) fitted to linear alkane systems. (This is the SKS FF).

To determine how sensitive the results are to the specifics of the FF, we also consider an early generation PE force field (RBFF) as described below. We consider modifications in the FF to include implicitly the effect of solvent. We will consider temperatures from 288K to 50400K to cover the range from the theta temperature to the infinite temperature. The plausibility of simulating the polymer systems at the high temperature range is discussed in section 2.5.

2.1 Rychaert-Belleman united atom force field (RBFF)

The Rychaert-Belleman united atom FF[15] was developed for polyethylene (PE). This uses a Lennard-Jones 12-6 van der Waals (vdW) potential plus a six-term torsional potential.

The Lennard-Jones 12-6 potential between atom i and atom j in the same chain is described as

$$E_{LJ}(r_{ij}) = 4\varepsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] \quad |i-j| \geq 4 \tag{1}$$

where the well depth ε/k_B=72K and inner wall σ = 0.3923 nm. Note that we do *not* include 1-4 nonbond interactions (or 1-2 and 1-3) so that the torsional terms are fully included in the torsional potential.

The torsion potential is described as

$$\frac{E_t(\phi_i)}{k_B \cdot 10^3 K} = \sum_{n=0}^{5} a_n (\cos\phi_i)^n \tag{2}$$

where ϕ_i is ith torsion angle and

$$a_0 = 1.116,\ a_1 = 1.462,\ a_2 = -1.578,\ a_3 = -0.368,\ a_4 = 3.156,\ a_5 = -3.788.$$

This leads to that the gauche minimum is 0.699 kcal/mol higher than the trans minimum, and T->G barrier of 2.949 kcal/mol, G->G barrier of 10.698 kcal/mol. (ϕ_{trans}=0)

The bond distances (R_b=0.153 nm) and bond angles (θ = 109.47°) are fixed.

The RB-FF has been used for many studies, including the earlier CCBB MC paper.[14]

2.2 Siepmann-Karanorni-Smit force field (SKS-FF)

The Siepmann-Karanorni-Smit (SKS) force field [16-18] developed to describe thermodynamic properties of *n*-alkanes is expected to yield higher accuracy than RB-FF. It has been used to discuss thermodynamic properties of dendrimer liquid crystals.[19] SKS-FF also uses the LJ12-6 in Eq. (1) to describe the vdW interaction but uses different parameters:

- the well depth is ε_{CH2}/k_B=47.0K, ε_{CH3}/k_B=114.0K, and
- the inner wall is σ_{CH2} =σ_{CH3}= 0.393 nm.

To facilitate the Monte Carlo analysis of chain length, we approximate ε_{CH3}/k_B = ε_{CH2}/k_B = 47.0K.

The torsion potential of the SKS force field was based on the OPLS (optimized potentials for liquid simulation) force field,[20] which is described as

$$E_t(\phi_i) = \sum_n \frac{1}{2} V_n \left[1 - d_n \cos(n\phi_i)\right] \tag{3}$$

where ϕ_i is *i*th torsion angle

V_1 = 1.4109, V_2 = -0.271, V_3 = 2.787, (in kcal/mol), and

d_1 = -1, d_2 = 1, d_3 = -1.

This leads to that the gauche minimum is 0.826 kcal/mol higher than the trans minimum, and T->G barrier of 2.942 kcal/mol, G->G barrier of 4.198 kcal/mol. (ϕ_{trans}=π)

The bond distances (R_b=0.154 nm) and bond angles (θ= 114.0^o) are fixed.

2.3 Scaled vdW potentials for including solvent effects

In order to include solvent effects without the enormous cost of explicitly treating huge numbers of solvent molecules in MD simulations, we use the approach of Jang, Cagin, and

Goddard (JCG) [21] in which a *van der Waals scaling factor f* is applied to the nonbond potential, Eq. (4), to mimic solvent effects using an implicit model.

$$E_{LJ}(r_{ij}) = 4f\varepsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] \quad |i-j| \geq 4 \tag{4}$$

This approach assumes that solvent molecules between polymer segments affect both the repulsive interaction and the attractive interaction to a similar extent. Thus, the van der Waals screening arises from the solvent molecules occupying the space between polymer segments. Here solvents with different screening efficiency or different solvent power will lead to a different scaling factor, *f*. This is similar to using a dielectric constant to mimic the screening of the charge-charge interactions by a solvent. For the calculations here we consider f=0.5, and f=0.25.

To elucidate the role of torsional and attractive contributions to the energy, we also report the results of CCBB calculations using simplified FF:

- **NoTorSKSFF** excludes the torsional contributions to the energy. This could be useful in developing scaling rules that distinguish chain stiffness from non-bond interaction and torsion potential.
- **NoTorNoAttSKSFF** excludes both torsional and the attractive part of the nonbond terms in SKSFF. In this case the potential energy has only the self-avoiding terms given by Eq. (5).

$$E_{LJ}(r_{ij}) = 4\varepsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12}\right] \quad |i-j| \geq 4 \tag{5}$$

2.4 CCBB Monte Carlo method

The CCBB method [14] combines continuous configuration biased sampling (CCB) with Boltzmann-factor biased enrichment (BFB) in a highly efficient variant of Monte Carlo [14] for direct evaluation of the partition function, free energy, and other configurationally dependent physical properties for polymer chains.

2.4.1 Independent rotational sampling IRS

The CCB method uses independent rotational sampling (IRS), in which torsional degrees of the polymer chains are sampled using a weighting function based on the Boltzmann factor of the torsion energy. The normalized torsion weighting function (TWF), W_{IRS}, for IRS is defined as

$$W_{IRS}(\phi) = \frac{g_{IRS}(\phi)}{z_{IRS}} \tag{6}$$

where,

$$z_{IRS} = \int_0^{2\pi} g_{IRS}(\phi) d\phi \tag{7}$$

$$g_{IRS}(\phi) = \exp[-\beta E_t(\phi)] \tag{8}$$

Torsion angles are generated in accordance with Eq. (6). The partition function for IRS after bias correction is evaluated by

$$Z_N = N_c^{-1} (z_{IRS})^{N-3} \sum_{1}^{N_c} \exp\left[-\beta \sum_{i-5}^{N} \sum_{j=1}^{i-4} E_{LJ}(r_{ij}) \right] \tag{9}$$

2.4.2 Continuous configuration biased sampling (CCB)

Spatial overlaps between nonbonding atoms are inevitable, leading to high configurational energies. In order to exclude these overlaps, information about the spatial environment in the vicinity of the growing chain end should be introduced into the TWF.

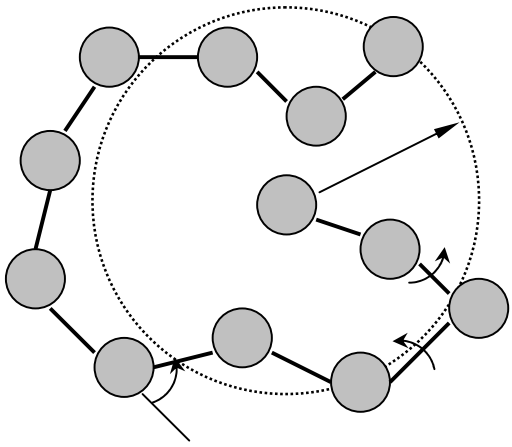


Figure 1. Illustration of quantities used in evaluating the energy for growing an additional

The CCB method includes non-bonding interactions within a cutoff length in the TWF calculation. On constructing the TWF for the *i*th torsion, we define a sphere of radius R_{cut}, centered at the (*i*-1)th atom position, as shown in Figure 1. In this paper we take the vdW cutoff to be R_{cut} = 6 Å. This length of R_{cut} is taken larger than $l+\sigma$ in order to check all possible atomic overlaps. Boltzmann factors for the nonbond energy between *i*th

atom and all other atoms inside the cutoff sphere are included in the W_{CCB} for TWF as

$$W_i^{CCB}(\phi_i;\phi_4,...,\phi_i) = \frac{g_i^{CCB}(\phi_i;\phi_4,...,\phi_{i-1})}{z_i^{CCB}(\phi_4,...,\phi_{i-1})} \tag{10}$$

where,

$$z_i^{CCB}(\phi_4,...,\phi_{i-1}) = \int_0^{2\pi} g_i^{CCB}(\phi_i;\phi_4,...,\phi_{i-1})d\phi_i \tag{11a}$$

$$g_i^{CCB}(\phi_i;\phi_4,...,\phi_{i-1}) = g_{IRS}(\phi_i)\exp\left[-\beta\sum_{j=1}^{i-4}\Theta(R_C - r_{ij})E_{LJ}(r_{ij})\right] \tag{11b}$$

and $\Theta(R)$ is the (Heavyside) step function

$\Theta(R)=0$, if $R<0$; $\Theta(R)=1$, if $R\geq 0$.

The bias-corrected partition function has the form of Eq. (12), which also includes those nonbonding energies that did not appear in the TWF calculation.

$$Z_N = N_c^{-1} z_{IRS} \sum_1^{N_c}\left\{\prod_{i=5}^{N} z_i^{CCB}(\phi_4,...,\phi_{i-1})\right\}\cdot\exp\left[-\beta\sum_{i=5}^{N}\sum_{j=1}^{i-4}\Theta(r_{ij} - R_C)E_{LJ}(r_{ij})\right] \tag{12}$$

2.4.3 Boltzmann factor biased enrichment (BFB)

In CCB sampling, we would like to exclude high-energy chains with nonbond overlaps. However, such sampling would not lead to a Boltzmann distribution since low energy samples would have too high a contribution to the partition function. To accomplish this, we extended the chain enrichment procedure to control sampling so that all collected samples make nearly equal contributions to the partition function. This does not describe a Boltzmann distribution, but by including the proper statistical factors we can correctly calculate the partition function.

We define step i in the sampling as the stage where sampling of atom i is completed. The enrichment is achieved by continuing the sampling to step i+1 in m_i different ways after completion of sampling step i. The m_i is denoted as the enrichment factor. The enrichment factor must be fixed ahead of the MC simulation in order to avoid bias. We define the sample multiplicity M_i for step I as in Eq. (13).

$$M_i = \prod_{j=1}^{i-1} m_j \tag{13}$$

The samples obtained in this enrichment process are not statistically independent. We denote as a ***cluster*** such a group of interdependent chains (all of which have the same set of atom positions at the first step).

In CCBB Monte Carlo, the multiplicity M_i is determined at every step as proportional to the ratio of the Boltzmann factor of the just sampled step i-1 to that of the running average value for the chain with the same length. The partition function is explicitly calculated as the average of weighting-bias-corrected Boltzmann factor divided by the chain multiplicity. This is denoted as CCBB method.[14]

The partition function in Eq. (12) is rewritten in terms of a sum over K clusters as:

$$Z_N(K) = K^{-1}\sum_{C=1}^{K}\zeta_N(C) \tag{14}$$

We denote C as the cluster serial number, n as the chain serial number in cluster C, i as the growing step number in chain n, and N as the chain length.

Denoting $L_n(C)$ as the total number of chains generated for cluster C by using an arbitrary choice for the enrichment factor, the partition function in Eq. (14) is calculated by

$$\zeta_N(C) = \sum_{n=1}^{L_N(C)}\frac{\zeta_N^n(C)}{M_N^n(C)} \tag{15}$$

$$M_N^n(C) = \prod_{i=1}^{N-1} m_i^n(C) \tag{16}$$

$$L_N(C) = \sum_{n=1}^{L_{N-1}(C)} m_{N-1}^n(C) \tag{17}$$

In CCBB, the chain multiplicity, $M_i^n(C)$, is determined as proportional to the ratio of $\zeta_{i-1}^n(C)$ to $Z_{i-1}(C-1)$:

$$Q_i^n(C) = \frac{p\cdot\zeta_{i-1}^n(C)}{Z_{i-1}(C-1)} \tag{18}$$

$$M_i^n(C) = INT[Q_i^n(C)] \qquad \text{if } Q_i^n(C) > 1$$

$$= 1 \quad \text{if } Q_i^n(C) \le 1 \tag{19}$$

The enrichment factor m_{i-1}^n is evaluated from the ratio of M_i^n to M_{i-1}^n. This procedure keeps the chain multiplicity to always be approximately proportional to the Boltzmann factor of the chain at the just previous step.

$$P_i^n(C) = \frac{M_i^n(C)}{M_{i-1}^n(C)} \tag{20}$$

$$m_{i-1}^n(C) = INT[P_i^n(C)] \quad \text{if } P_i^n(C) > 1$$

$$= 1 \quad \text{if } P_i^n(C) \le 1 \tag{21}$$

For $i<5$ the Boltzmann population of the chain collection is completely satisfied in CCB and we set the chain multiplicity to unity: $M_0^n = M_1^n = M_2^n = M_3^n = M_4^n = 1$.

The choice of p in Eq. (20) is arbitrary and can be selected to maximize efficiency. Too large a value of p leads to an exploding number of samples of highly correlated configurations. Too small a value leads to too few chains per cluster. We found for the systems considered here that p=1 leads enriched chains having nearly equal contributions to the partition function.

To obtain an initial guess for the partition function, $Z_i(0)$, a short non-BFB run is performed. In this study, we sampled 200 chains prior to BFB sampling.

Prior to the chain sampling, the energy is calculated for a fixed number of grid points (in this study, we used 200 equally separated grid points from 0 to 2). W^{CCB}, the normalized TWF for CCB, is then evaluated using numerical integration for z^{CCB} as in Eq. (12).

Implementation of the CCBB method was described in Ref [14].

2.5 Simulation at high temperatures

In order to remove specific dependence on the detailed structure and the London dispersion interactions (van der Waals attraction), Flory, des Cloizeau, and others examined the ensemble behavior at high temperature. This led to general scaling arguments valid at high temperature. Indeed we find that the critical exponents / parameters depend little on the specific force fields (Section 3 and 4).

Previous studies of the critical exponents / parameters for polymer chains based on lattice always assumed infinite temperature. Our continuous study is based on Monte Carlo sampling over a range of temperatures from 288K to 50400K. Obviously no polymer (or molecule) will exist at 50000K. However, at high temperature, the ensemble description is insensitive to details in the torsion energy barrier, the self-avoidance (Pauli repulsion), the

van der Waals attraction, and the detailed chemical structure of the monomer. By performing Monte Carlo sampling at temperature as high as 50000K, we minimize the influence introduced by self-avoidance, van der Waals attraction, and chain stiffness to obtain general scaling relationships. In addition, the analysis at 50000K is necessary to compare with the previous lattice studies.

Since we want to find the limiting behavior for sampling of the polymer chain conformation ensemble at high temperature, we keep the bond length and the bond angle fixed at their most favored values (not the correct value at high temperature, which would involve broken bonds leading to a completely different ensemble of conformation). Thus anharmonicities important in the real system at high temperature are not relevant. In addition, it is well know that the local interaction doesn't change the scale properties of the polymer chain. For example, when we turn off the torsion energy (local interaction) we find the similar results for scale properties. And we find the same theta temperatures as from MD studies[21] with relaxed bond length / bond angle (Section 3).

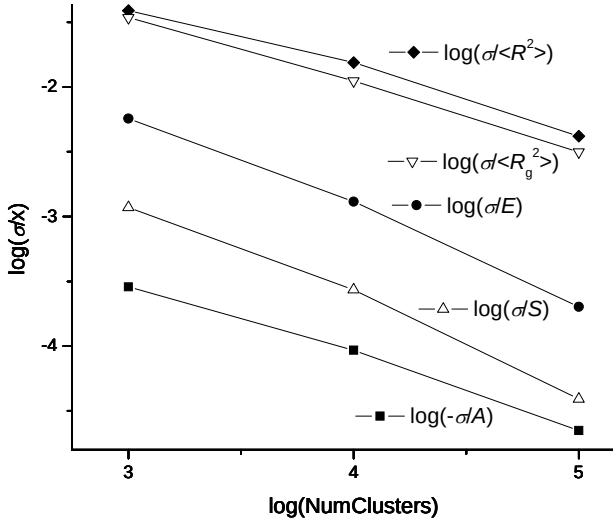


Figure 2. Standard deviation σ among 10 independent samples of CCBB calculations as a function of the number of clusters sampled. Shown here are the results for free energy A, energy E, entropy S, radius gyration $\langle R_g^2 \rangle$, and squared end to end distance $\langle R^2 \rangle$ (using Rychaert-Belleman

2.6 Convergence

To establish criteria for assessing convergence, we considered chain lengths from N= 6 to 400 and examined temperatures ranging from 288K to 50400K.

To establish the uncertainties, the CCBB MC calculations were carried out using 10 independent samples to describe the equilibrium ensemble of configurations. Each ensemble contains 10^5 clusters of chains with a total of ~ 10^6 chains. The differences among the samples arise from the choice of the initial seed for Monte Carlo sampling.

Figure 2 shows how the RMS deviation, σ, converges with sampling length. We find a deviation of 0.0022% in A, 0.0039% in S, 0.020% in E, 0.31% in $<R_g^2>$, and 0.42% in $<R^2>$ (using RBFF and T=2160K, N=400). Since the statistics of $<R_g^2>$ is slightly better than $<R^2>$ we use $<R_g^2>$ as the measure of polymer size.

3. Radius gyration, Critical exponent 2ν, and the Flory theta temperature T_θ from CCBB

Figure 2 shows that the statistical accuracy of $<R_g^2>$ is slightly better than for $<R^2>$. Consequently we base our analysis of 2ν at various temperatures on analysis of $<R_g^2>$.

Figure 3 shows the log-log plot of $<R_g^2>/(N-1)$ against $(N-1)$ at different temperatures (using CCBB with SKSFF). The results from other force fields (not shown here) are essential the same. The slope of this line gives the Flory exponent, 2ν-1, as in Eq. (22). The Flory exponent 2ν decreases as N increases and increases as T increases. The N dependence of Flory exponent is significant at low temperature and small N. The N dependence reduces at high temperature. (e. g. 50400K).

$$\left\langle R_g^{\,2} \right\rangle \propto N^{2\nu} \tag{22}$$

Figure 4 shows the Flory exponent 2ν as a function of $1/N$ from CCBB with various force fields at 50400K. Each point of 2ν at N_0 is derived from the slope of $\ln<R_g^2>$ vs $\ln(N-1)$ from (N_0-49) to (N_0+50). We see that 2ν is approximately a linear function of $1/N$ (the quality of fit R^2 ranges from 0.787 to 0.962). Using the 50400K results (the highest

temperature studied), leads to the 2ν listed in Table 1. We see here that the Flory exponent depends significantly on N near the theta point.[9,10,12,13]

The value 2ν=1.168 derived from CCBB is close to the value 2ν=1.2 by Flory. This can be compared to results from renormalization group theory of [22] 2ν=1.1764±0.0022 and [23] 2ν=1.208. Various lattice models give [24-26] 2ν=1.176±0.002. Interpreting the neutron scattering experiments on FeF_2 in terms of an Ising antiferromagnet gives [27] 2ν=1.28±0.02. The light scattering of polystyrene in toluene and benzene at around room temperature gives 2ν=1.1872[28] (for molecular weights greater than 10E4 g/mol). For linear high-density polyethylene in trichlorobenzene at 135°C from on-line light scattering detector in size exclusion chromatography, 2ν=1.106.[29]

Table 1 The Flory critical exponents 2ν (derived at infinite N) from CCBB calculations with various force fields. This is compared with other models.

	SKSFF	**SKSFF, f=0.5**	**SKSFF, f=0.25**	**RBFF**			
$2\nu(N_\infty)$	1.168	1.199	1.210	1.196			
Deviation	0.009	0.002	0.003	0.006			
	Flory	**RGT[22]**	**RGT[23]**	**Lattice model[24-26]**	**Exptl[27]**	**Exptl[28]**	**Exptl[29]**
2ν	1.2	1.1764	1.208	1.176	1.28	1.1872	1.106
Deviation	N/A	0.0022	N/A	0.002	0.02	N/A	N/A

From Fig. 3(A), $\langle R_g^2\rangle/(N-1)$ increases monotonically with increasing N for $T\geq$1368K, while it has a maximum and decreases with increasing N in the range of large N for $T\leq$1224K. Thus the theta state must lie in the range of 1224K<T<1368K. Fig. 3(B) zooms in the range of N 310~400 part of Fig. 3(A) and lists the linear regression results of $\ln(\langle R_g^2\rangle/(N-1))$ against $(N-1)$ at 1368K, 1296K, and 1224K. At 1296K, $\langle R_g^2\rangle/(N-1)$ becomes almost a constant independent of N, indicating that the theta temperature is 1296K.

Similarly, the theta temperatures derived from other force fields are listed in Table 2 and we term the temperature at which $2\nu=1$ as the *Flory theta temperature*.

Table 2 also lists the Flory theta temperatures from the JCG molecular dynamics simulations (also using the SKS force field) which allowed bond stretch and angle bend motions[21]. We obtain Flory theta temperatures the same as those from JCG.[21]

Table 2. A comparison of the Flory theta temperatures, $T_{\theta\nu}$ ($2\nu=1$ definition) from CCBB calculations with the MD results from JCG (bonds and angles not fixed). [21]

Force fields	RBFF	SKSFF	SKSFF, f=0.5	SKSFF, f=0.25
$T_{\theta\nu}$(CCBB)	2108K	1296K	648K	336K
$T_{\theta\nu}$(MD)[21]	N/A	1300K	650K	328K

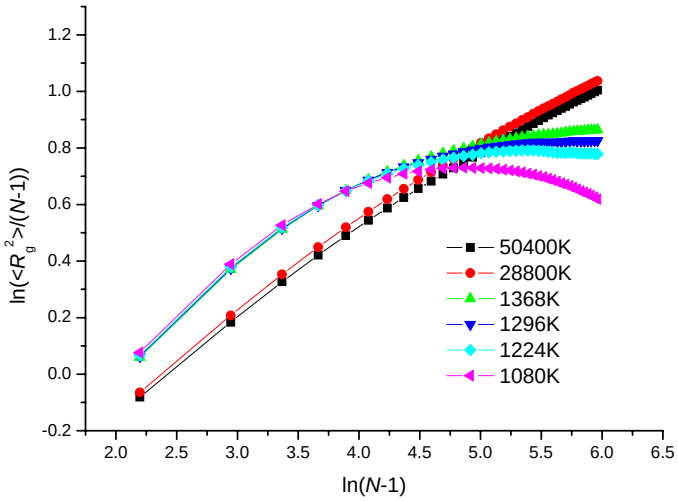


Fig. 3(A)

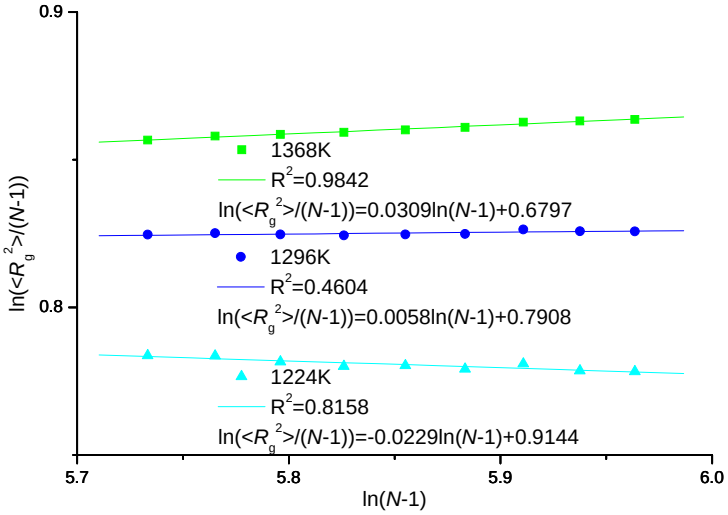


Fig. 3(B)

Figure 3. Log-log plot of $\langle R_g^2\rangle/(N\text{-}1)$ against $(N\text{-}1)$ at various temperatures (CCBB with SKSFF). Figure 3(A), shows that $\langle R_g^2\rangle/(N\text{-}1)$ increases monotonically with increasing N for $T\geq$1368K, while it has a maximum and decreases with increasing N in the range of large N for $T\leq$1224K. Thus the Flory theta state lies in the range of 1224K$<T<$1368K. Figure 3(B) considers just the range of N 310~400 from (A). Listed is the R^2 for linear regression of $\ln(\langle R_g^2\rangle/(N\text{-}1))$ against $(N\text{-}1)$ at 1368K, 1296K, and 1224K. This shows that $\langle R_g^2\rangle/(N\text{-}1)$ is nearly independent of N at 1296K leading to T_θ= 1296K.

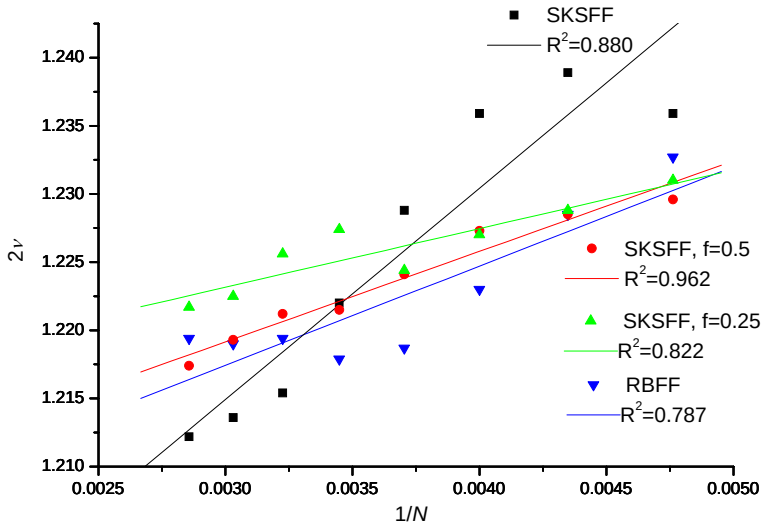


Figure 4. The Flory exponent 2ν at 50400K as a function of $1/N$ from CCBB with various force fields. Each point of 2ν at N_0 is derived from the slope of $\ln\langle R_g^2\rangle$ vs $\ln(N\text{-}1)$ from $(N_0\text{-}49)$ to $(N_0\text{+}50)$. The 2ν is approximately a linear function of $1/N$ (quality of fit R^2 ranges from 0.787 to 0.962). The derived 2ν at 50400K are listed in Table 1.

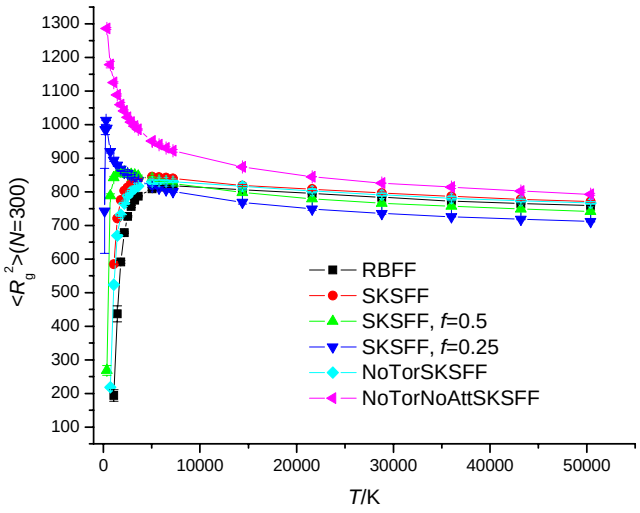


Figure 5. $\langle R_g^2\rangle$ against temperature from CCBB with various force fields, all for a chain length of N=300. The uncertainties are shown but in most cases are less than the size of the symbols. When only the repulsive van der Waals energy is present (NoTorNoAttSKSFF),

the polymer chains shrink with increasing temperature, reflecting the increasing number of available conformations. When the attractive van der Waals energy is present (the other five force fields), the polymer chains collapse at low temperature due to the favorable enthalpy.

Figure 5 illustrates the dependence of the overall chain dimensions, $<R_g^2>$, on temperature from CCBB calculations with various force fields, all for a chain length of N=300. When only the repulsive van der Waals energy is present (NoTorNoAttSKSFF) the polymer chains *shrink* with increasing temperature, reflecting the increasing number of available conformations. For the other five force fields that include attractive van der Waals terms, the polymer chains collapse at low temperature due to the favorable enthalpy.

4. Thermodynamic functions from CCBB

4.1 Thermodynamic properties as function of chain length *N* at different temperatures

The thermodynamic properties of polymer chains from CCBB calculations with SKS-FF are shown in Figure 6 (free energy, A), Figure 7 (internal energy, E), and Figure 8 (entropy, S). CCBB calculations with Rychaert-Belleman FF give results similar to the Siepmann-Karanorni-Smit force field in Figure 6~8 which are not shown here.

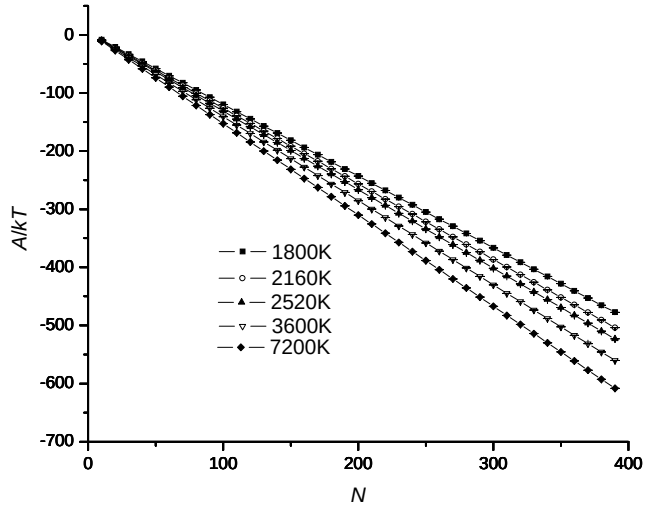


Figure 6. Free energy A/kT for polymer chains of length N at different temperatures from CCBB calculations with Siepmann-Karanorni-Smit (SKS) force field. The uncertainties are shown, but are less than the size of the symbols. We see that A/kT is proportional to N.

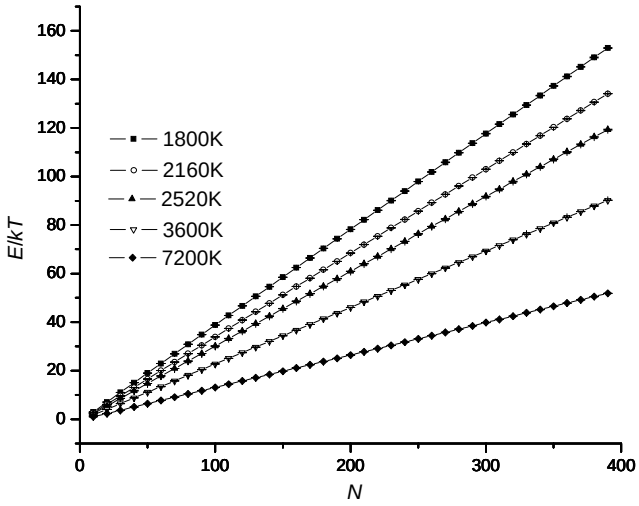


Figure 7. Internal energy E/kT for polymer chains of length N at different temperatures from CCBB calculations with Siepmann-Karanorni-Smit (SKS) force field. The uncertainties are shown, but are less than the size of the symbols. We find that E/kT is proportional to N.

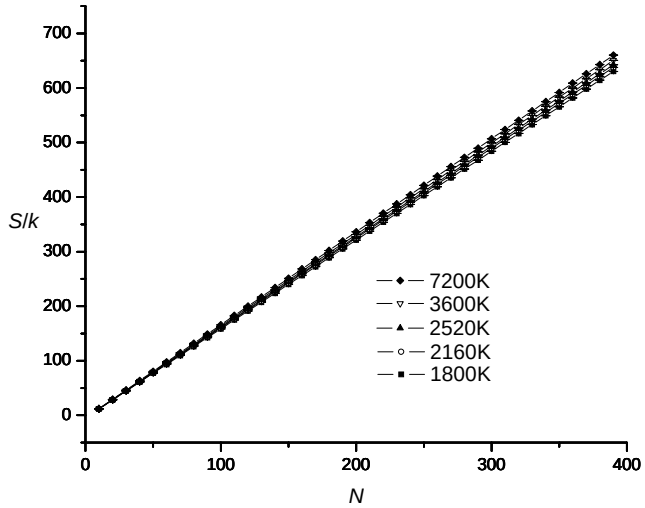


Figure 8. Entropy S/k for polymer chains of length N at different temperatures from CCBB calculations with SKS-FF. The uncertainties are shown, but are less than the size of the symbols. We find that S/k is proportional to N.

Figure 6~8 shows that *A*, *S*, and *E* are linear with chain length *N*, in which the partition function can be written as a product of local partition functions. Such an expression would be expected if there were only short-range interactions.

4.2 The correction to the *f*(*N*) term in the free energy

The dominant term in the *N* dependence of the free energy is proportional to *N* (the volume term) however here we analyze the form of the dominant corrections to this volume term, which is shown as in Eq. (23).

$$\frac{A_N}{k_B T} = -a * N + b * f(N) + c \tag{23}$$

We will consider two functional forms for (23) as described in the following two sections.

4.2.1 SAW lattice function at infinite temperature

The results on self-avoiding walks (SAW)[30] on a lattice in the high temperature limit (T_{high}), suggest that an ensemble of polymer chains of length *N* would lead to a partition function (Z_N) of the form

$$Z_N \propto N^{\gamma-1} \mu^N \tag{24}$$

Here μ is denoted as the ***critical attrition*** and γ is the ***free energy critical exponent***, which is related to the number of different polymer conformations. This leads to a free energy *A* of the form with ln*N* term

$$\frac{A_N}{k_B T} = -\ln Z_N = -N \ln \mu - (\gamma - 1) \ln N + C_A = -aN + b \ln N + c \tag{25}$$

These results were derived for the infinite temperature case and have not previously been suggested to be valid for finite temperature. Rationalizations for the forms in (24) and (25) have not previously been proposed. Our studies find that these forms are valid to temperatures down to ~ 0.7 T_θ.

4.2.2 Surface energy corrections at low temperatures

For compact systems one expects the free energy to have a volume term (N) and a surface term ($N^{2/3}$), leading to

$$\frac{A_N}{k_B T} = -aN + bN^{2/3} + c \tag{26}$$

At sufficiently low temperature, we expect collapse of the polymer into a complex with $R \sim N^{1/3}$ and a free energy as in (26).

4.2.3. Fitting numerical results

To compare these two forms with correction terms, ln$N$ and $N^{2/3}$, we fit Eq. (25) and (26) to our Monte Carlo results using multiple linear regressions (using the Origin 7.0 software). The results are listed in Table 3 and Figure 9, 10.

Table 3. Comparison of the multiple linear regressions fitting results with ln$N$ (Eq. 25) and $N^{2/3}$ (Eq. 26) in the range of N=101~400 (300 data points) from SKSFF results.

T/K		Chi^2/DoF	a	error	b	error	c	error	
50400	**LnN**	**4.66E-06**	**1.7087**	**8.55E-07**	**-0.15212**	**1.97E-03**	**5.20621**	**0.00859**	
50400	**N^(2/3)**	**4.11E-06**	**1.70728**	**3.00E-05**	**-0.01938**	**0.00024**	**4.77789**	**0.00286**	
3600	**LnN**	**3.64E-06**	**1.44618**	**8.31E-03**	**-0.12018**	**0.00174**	**4.47254**	**0.0076**	
3600	**N^(2/3)**	**2.56E-06**	**1.44504**	**2.00E-05**	**-0.0154**	**0.00019**	**4.13526**	**0.00226**	**theta/K**
1440	**LnN**	**5.63E-06**	**1.13825**	**7.38E-07**	**-0.01961**	**0.00217**	**3.5418**	**0.00945**	**1399**
1440	**N^(2/3)**	**5.26E-06**	**1.13804**	**3.00E-05**	**-0.00277**	**0.00027**	**3.48984**	**0.00323**	**1394**
1368	**LnN**	**6.08E-06**	**1.11457**	**0.00001**	**0.01474**	**0.00225**	**3.37896**	**0.00981**	
1368	**N^(2/3)**	**6.35E-06**	**1.11467**	**8.44E-07**	**0.00157**	**0.00029**	**3.42419**	**0.00355**	
1080	**LnN**	**9.50E-06**	**0.99709**	**0.00001**	**0.40513**	**0.00282**	**1.74419**	**0.01227**	
1080	**N^(2/3)**	**2.59E-06**	**1.00091**	**0.00002**	**0.05173**	**0.00019**	**2.88357**	**0.00227**	

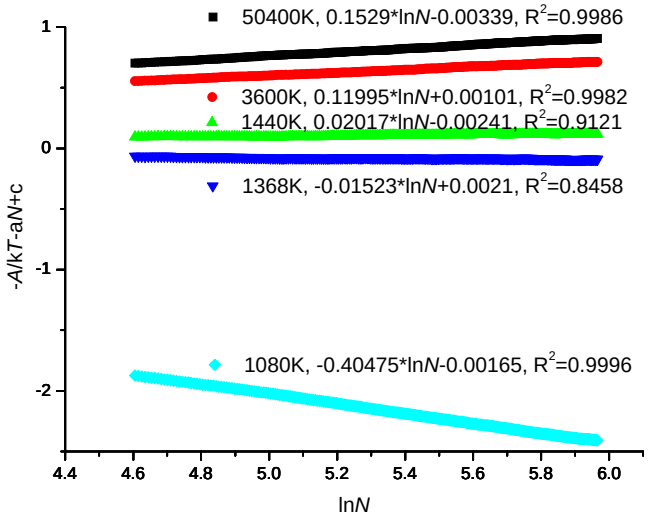


Figure 9. Plotted is (-A/kT-aN+c) as a function of ln$N$ at different temperatures from CCBB with SKSFF, where a and c come from the multiple linear regressions of A/kT vs N in the form of Eq. (25). The linear relationship here validates Eq. (25) and the slope of each line represents b=(γ-1).

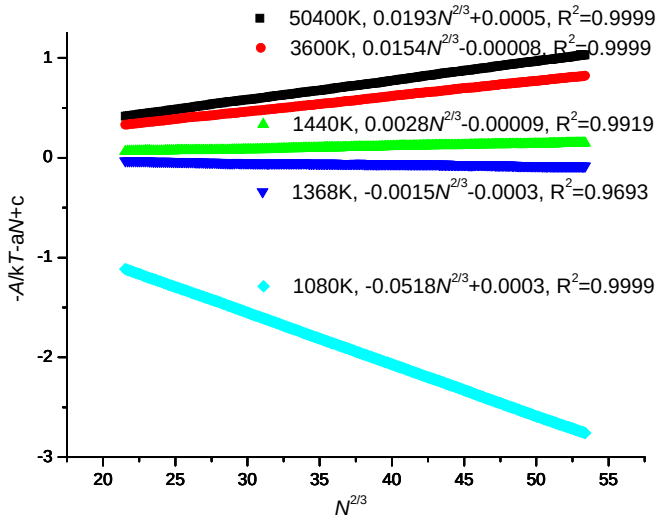


Figure 10. Plot of (-A/kT-aN+c) as a function of $N^{2/3}$ at different temperatures from CCBB with SKSFF, where a and c come from the multiple linear regressions of A/kT vs N

in the form of Eq. (26). The multiple linear regression also gives high correlation with term $N^{2/3}$.

These results show that our Monte Carlo data can be fit accurate by either Eq. (25) or (26). This indicates that the $b*f(N)$ term is negligible compared with the $a*N$ term. With either description we find that:

- b is *negative* at high temperature, and the free energy increment *increases* with chain length (it is extended)
- b is *positive* at low temperature, the free energy increment *decreases* with chain length (it collapses).
- b passes through a point in between for which b=0. In model (25) this corresponds to the free energy theta temperature, where the free energy behaves as if the monomers were independent of each other (a "free monomer gas"). For model (26) this corresponds to a state in which the surface energy is zero.

In the following section we derive the form $f(N) = \ln N$ from simple physical arguments. We argue that even though the free energy curve can be fitted with $f(N)=N^{2/3}$ the negative b at high temperature corresponds to a negative surface energy, which we consider to be nonphysical.

4.3. Physics of ln*N* term and γ

Consider one chain with length 2*N*, as two coils with length *N,* but connected as shown in Figure 11.

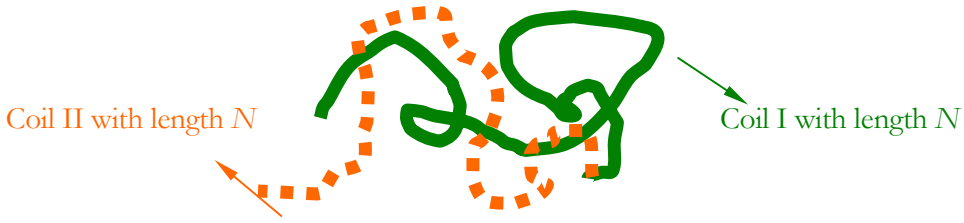


Figure 11. One chain with length 2*N*, considered as connecting two coils with length *N*

We write the free energy form as in Eq. (23)

$$\frac{A_{2N}}{k_B T} = -a * 2N + b * f(2N) + c \tag{27}$$

If the two coils, Coil I and Coil II, do not interact with other, then the total free energy would be

$$\frac{A'_{2N}}{k_B T} = (\frac{A_N}{k_B T} - c) + (\frac{A_N}{k_B T} - c) + c = -2a * N + 2b * f(N) + c \tag{28}$$

The difference between (27) and (28) gives the interaction between Coil I and Coil II as

$$\frac{\Delta A}{k_B T} = \frac{A_{2N}}{k_B T} - \frac{A'_{2N}}{k_B T} = b * f(2N) - 2b * f(N) \tag{29}$$

The interaction between Coil I and Coil II includes two parts: repulsive (self avoiding entropy) and attractive (enthalpy).

At infinite temperature, there is no attractive energy (athermal) leaving only the repulsive part (self-avoiding entropy) between these two coils. Thus $\Delta A > 0$. Because $f(N) < N$ and $f(2N) < 2f(N)$, b must < 0. The volume occupied by the other coil scales as N and the reduced entropy scales as $\ln(\Delta V) \sim \ln N$. This suggests $f(N)$ has the $\ln N$ form.

As the temperature decreases from infinity, the attractive energy between two coils increases. For sufficiently low temperature, the attractive energy exactly balances with the repulsive part between these two coils leading to $\Delta A = 0$. At this temperature, the correction term for the attractive energy will scale as ln$N$ to cancel the same correction term from the reduced entropy. We denote this point as the *free energy theta temperature,* $T_{\theta F}$. De Gennes[31] discusses such an analysis (see p306 of Ref.[31]) and points out that the renormalization group analysis shows that this temperature coincides with the "bare theta temperature".

For temperatures lower than $T_{\theta F}$, the attractive energy dominates and $\Delta A < 0$. This leads to $b > 0$.

At sufficiently low temperature that the chain collapses due to the dominance of the enthalpic attractive forces, we expect $V \sim N$. In this range the correction term is expected to have the form of a surface energy, leading to an $N^{2/3}$ dependence as described in section 4.2.2. This would eventually dominate over the ln$N$ term.

Summarizing, we find that over the temperature range from infinite temperature to far below theta temperature, the free energy can be accurately described using a ln$N$ correction term arising from two parts: self-avoiding entropy and attractive enthalpy. In the chain-collapsed phase, one expects that the $N^{2/3}$ correction term arising from surface energies will dominate, but we have not examined such low temperatures.

4.4 critical exponent γ and Free energy theta temperature

In Figure 9, we plot ($-A/kT - aN + c$) as a function of ln$N$ at different temperatures from CCBB with SKSFF, where a and c come from the multiple linear regressions of A/kT vs N in the form of Eq. (25). The linear relationship here validates Eq. (25) and the slope of each line represents $b = (\gamma - 1)$. The results from other force fields are similar, which are not shown here. γ increases with the temperature, reflecting the increased number of available conformations and partition function. Indeed, the detailed analysis of its components (γ_E and γ_S) in the next section, 4.5, gives much more information of the behavior of γ over temperature.

Comparing the N-dependence of ν in Fig. 3 of section 3, we find that γ is nearly independent of N. Fig. 12 shows N-dependence of (γ-1) derived from CCBB calculations at 50400K with SKSFF. Each point of (γ-1) at N_0 is derived from the slope of ($-A_N/k_BT-N\ln\mu$) vs ln(N-1) from (N_0-49) to (N_0+50). Thus (γ-1) fluctuates in the range of 0.14~0.18 as N increases, an N-dependence far weaker than for 2ν (see Fig 3). Thus we evaluate (γ-1) from the slope of ($-A_N/k_BT-N\ln\mu$) vs ln(N-1) from N=100 to N=400, the N range studied here. This leads to γ = 1.153 with a quality of fit of R^2 =0.9986. The results from other force fields are listed in Table 4.

In the high temperature limit, RGT with high order calculations[22] gives γ=1.1596 ± 0.0020. Another RGT analysis gives γ=1.187.[23]

Exact enumeration with data extrapolation from the SAW lattice model gives γ=1.1595 ± 0.0012[24] and γ=1.16193 ± 0.0001[25]. Monte Carlo on the lattice model gives γ=1.1575 ± 0.0006.[32] Neutron scattering experiments analyzed in terms of an Ising antiferromagnet γ=1.25 ± 0.02 for FeF_2,[27] γ=1.233 ± 0.01[33] for N_2, and γ=1.25 ± 0.01[33] for Ne.

Fig. 9 also shows that ($-A_N/k_BT-N\ln\mu$) increases monotonically with increasing N for $T \geq 1440$K, while it decreases monotonically with increasing N for $T \leq 1368$K. Thus the free energy theta temperature is in the range of 1368K<T<1440K. With linear interception, we obtain $T_{\theta FE}(\gamma=1.0)$ = 1399 K. Similarly, we can derive the $T_{\theta FE}$ for other force fields, leading to the results listed in Table 5.

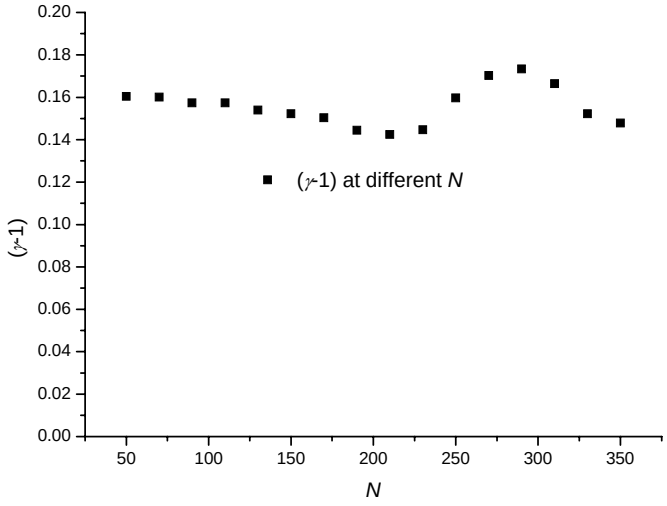


Figure 12. N-dependence of (γ-1) derived from CCBB calculations at 50400K with SKSFF. Each point of (γ-1) at N_0 is derived from the slope of ($-A_N/k_BT-N\ln\mu$) vs ln(N-1) from (N_0-49) to (N_0+50). (γ-1) depends weakly on N in contrast to 2ν (see Fig 3) and it fluctuates in the range of 0.14~0.18 as N increases. Thus we evaluate (γ-1) from the slope of ($-A_N/k_BT-N\ln\mu$) vs ln(N-1) in the range of N=100 to N=400 (the maximum N studied here).

Table 4. The Free Energy critical exponent γ from CCBB calculations using various force fields. These results are compared with other results from the literature.

	RBFF	**SKSFF**	**SKSFF, *f*=0.5**	**SKSFF, *f*=0.25**			
γ	1.158	1.153	1.165	1.152			
Deviation	0.003	0.002	0.003	0.004			
	RGT[22]	**RGT[23]**	**Enum-lattice[24]**	**Enum-lattice[25]**	**MC-lattice[32]**	**Exptl[27]**	**Exptl[33]**
γ	1.1596	1.187	1.1595	1.16193	1.1575	1.25	1.233/1.25
Deviation	0.002	N/A	0.0012	0.0001	0.0006	0.01	0.01

Table 5. Free energy Theta temperatures ($T_{\theta FE}$) predicted from CCBB calculations using various force fields

Force fields	RBFF	SKSFF	SKSFF, f=0.5	SKSFF, f=0.25
$T_{\theta FE}$(CCBB)	2247K	1399K	727K	379K

4.5 Critical exponent γ's components: γ_E and γ_S

In Figure 9, we plot (-A/kT-aN+c) as a function of ln$N$ at different temperatures from CCBB with SKSFF, where a and c come from the multiple linear regressions of A/kT vs N in the form of Eq. (25). Similar analysis of internal energy E and entropy S based on Eq. (30) and (31) are shown in Figure 13 and Figure 14. The slopes in Figure 13 and Figure 14 are (γ_E-1) and (γ_S-1) respectively.

$$\frac{E_N}{k_B T} = -a_E N + b_E \ln N + c_E = N e_0 + (\gamma_E - 1)\ln N + C_E \tag{30}$$

$$\frac{S_N}{k_B} = -a_S N + b_S \ln N + c_S = N s_0 + (\gamma_S - 1)\ln N + C_S \tag{31}$$

The physics of this analysis is described in section 4.3. Briefly, the self-avoiding entropy between two connected coils scales as the reduced space, ln*N*. At the theta temperature, the attractive enthalpy exactly balances the self-avoiding entropy, which should also scale as ln*N* term. From Eq. (30) and (31), we can see that the magnitude of the self-avoiding entropy and the attractive energy are indexed by γ_E and γ_S.

With Eq. (25), we have the following relationships

$$b = b_E - b_S \tag{32}$$

$$(\gamma - 1) = (\gamma_S - 1) - (\gamma_E - 1) \tag{33}$$

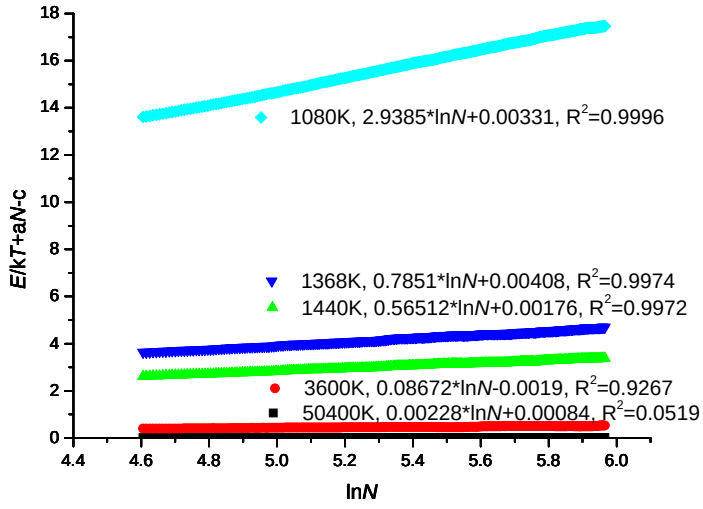


Figure 13. Plotted is (E/kT+aN-c) as a function of ln$N$ at different temperatures from CCBB with SKSFF. This is based on the same analysis as the free energy in Figure 9 with Eq. (30).

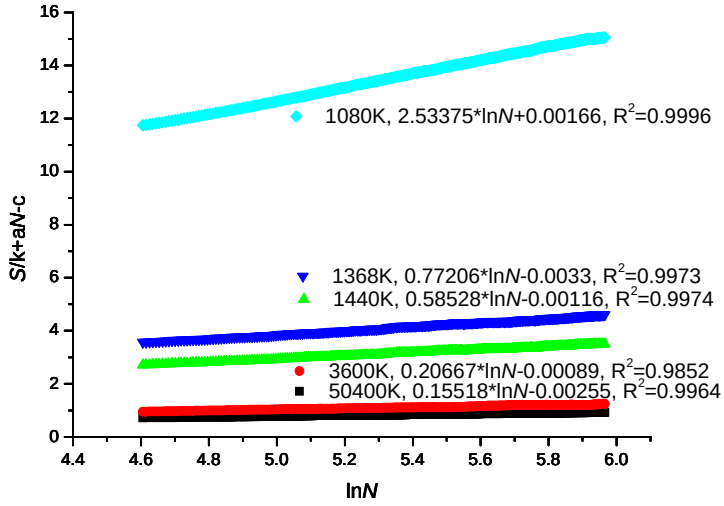


Figure 14. Plotted is (S/k+aN-c) as a function of ln$N$ at different temperatures from CCBB with SKSFF. This is based on the same analysis as the free energy in Figure 9 with Eq. (31).

From Figure 13 and 14, we can see that both γ_E and γ_S increase as the temperature decreases. This can be easily understood from the physics described in section 4.3. As the temperature decreases, the chain with chain length 2N, which connected 2 coils, will shrink normally. This wills strength the two interactions between the 2 coils: self-avoiding entropy and attractive energy.

Figure 13 and 14 show that at the infinite temperature, $\gamma_E = 0$ and $\gamma_S = 1.155$. Both of them increase as the temperature decreases. At the theta temperature, 1399 K, $\gamma_E = \gamma_S \sim$ 1.65 and the self-avoiding entropy exactly balances off the attractive energy part. At the lower temperature, when attractive energy dominates, $\gamma_E > \gamma_S$ and $\gamma < 1$.

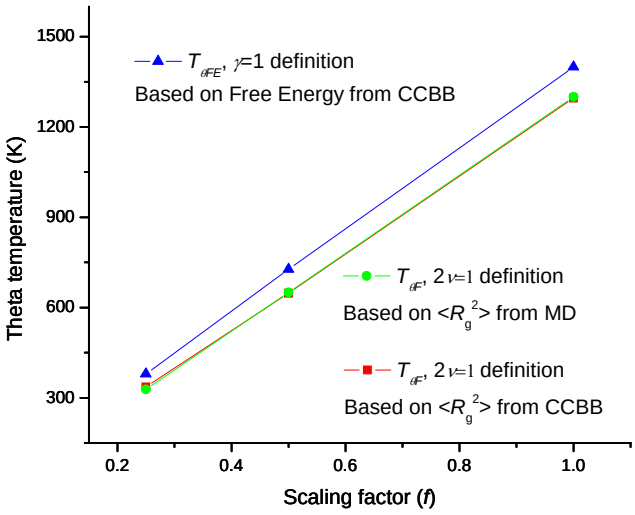


Figure 15. Dependence of the theta temperature on the solvent scale factor *f* from CCBB calculations using SKS-FF. We find that $T_{\theta F}$ is the same for CCBB and for MD (JCG). [21] However $T_{\theta FE}$ is ~10% higher than $T_{\theta F}$ for all solvent scale factors.

4.6 The shift of the theta point

As illustrated in Figure 15, the free energy theta temperature $T_{\theta FE}$ (Table 5) is about 10% larger than the Flory theta temperature $T_{\theta F}$ (Table 2).

In lattice-based theories, the free energy can be expressed in terms of Φ, the fraction of sites occupied by monomers as (34)

$$\frac{F}{T}|_{site} = \frac{\Phi}{N}\ln\Phi + \frac{1}{2}(1-2\chi)\Phi^2 + \frac{1}{6}\Phi^3 + \ldots \qquad (34)$$

(Eq. IV. 30 of Ref[31]).At the temperature for which χ=1/2, we see that there is no quadratic term in Φ, leading to μ=1-2χ = 0. de Gennes refers to this as "the bare theta temperature" (See the discussion in section 4.3 and p306 of Ref.[31]). At this temperature there is an exact cancellation between steric repulsion and van der Waals attraction between monomers. Indeed this point corresponds to the free energy theta temperature, $T_{\theta FE.}$

The shift of $T_{\theta FE}$ to a value 10% higher than T_{θ}, the Flory theta point arises from three-body and higher order terms in the concentration expansion in (34). A discussion of this

topic can be found in p113~p116 of Ref[31]. When χ =1/2, the Φ^2 term disappears, leaving higher order terms dominated by the Φ^3 term. This 3-body term causes a repulsive interaction leading to swelling the chain from the Gaussian distribution. (ν>0.5) This is confirmed by our results as shown in Figure 15. Our results show provide a quantitative measure of this effect, $T_{\theta FE}$ = 1.1 $T_{\theta.}$

The Flory theta temperature, $T_{\theta F}$, is the temperature at which the chain describes a Gaussian coil (that is, $<R_g^2>$ scales linearly with N, leading to 2ν=1). The free energy theta temperature, $T_{\theta FE}$, is defined as the point at which the self-avoiding entropy balances off the attractive energy (γ=1), leading to an ideal solution. In the case that the free energy depends quadratically on concentration, we would obtain $T_{\theta FE}$, = $T_{\theta F}$. In fact for alkanes (PE) we find that $T_{\theta FE}$, = 1.1 $T_{\theta F}$,, reflecting the higher order effects arising from Pauli exclusion. When γ=1, we find that both γ_E and γ_S are greater than 1 (section 4.3 and 4.5) indicating that the chain cannot be treated exactly as ideal gas of monomers.

We represent solvation by including a van der Waals scaling factor in the Lennard-Jones potential. This leads to a linear relationship between the solvent factor and the theta temperature as shown in Figure 15. This allows us to determine the vdW scaling factor to mimic the effects of a specific solvent environment, by comparing with the observed theta temperature in the specific solvent. This permits the use of the general thermodynamic relations described here for all solvent conditions.

The most common solvents for linear polyethylene lead to experimental theta temperatures between 360K and 460K.[34,35] From MD studies [21] this corresponds to solvent scale factors ranging from f=0.28 (θ=360K) to f=0.35 (θ=460K).

4.7 Critical attrition μ and self-avoidance factor (SAF)

Figure 16 shows the critical attrition μ as a function of 1/T from calculations with various force fields. We find that μ increases with temperature (except for NoTorSKSFF), reflecting the increased number of available conformations and partition function. The

relationship between μ and T is given by Eq. (42) in section 4.9. NoTorSKSFF includes only repulsive and attractive van der Waals energies so that the attractive energy dominates at low temperature. Thus NoTorSKSFF leads to a minimum critical attrition around 3600K.

To derive μ at infinite temperature, we fit the CCBB value for μ to a second order function of $1/T$ in the range of T= 5040K ~ 50400K. For PE (SKS) this leads to the critical attrition μ at infinite temperature of 5.6489 (see Table 6). This can be compared with the limiting value of 2π=6.2832 when intra-chain and torsion terms are ignored, indicating a 10% decrease.

The critical attrition μ from lattice studies [30] are also given in Table 6. The maximum partition functions for Simple Cubic, BCC, FCC, and CCBB are 6, 8, 12, 2π respectively. Thus the lattice models are not adequate to determine μ.

The calculated partition functions listed in Table 6 are always less than the maximum partition functions due to self-avoidance of the monomers in the same polymer chain. We define the *self-avoidance factor, SAF* as the ratio of the difference between the maximum and calculated partition functions to the maximum partition function. The results are listed in Table 6. For PE (SKS), CCBB leads to SAF = 0.10 while lattice studies give 0.16 to 0.22. Thus Lattice models underestimate the partition function because they build in too much self-avoidance.

Table 6. Critical attrition μ at infinite temperature from CCBB calculations. This is derived from the second order fit of μ vs. $1/T$ over the range of T= 5040K to 50400K. Comparisons are made to values from lattice models. [30] We define the self-avoidance factor, SAF, as the difference between the limiting value (2π for CCBB) and the calculated value divided by the limiting value (2π).

Lattice	CCBB	Unbiased Pade [30]	Biased Pade [30]	Ratio [30]	SAF [a]
SC		4.6838±0.001	4.6834±0.001	4.6835±0.0005	0.22[b]
BCC		6.5295±0.002	6.5295±0.002	6.5295±0.0005	0.18[c]
FCC		10.035±0.005	10.0346±0.001	10.0355±0.001	0.16[d]
RBFF	5.3648±0.0003				0.14[e]
SKSFF	5.6489±0.0004				0.10
SKSFF(f=0.5)	5.7204±0.0004				0.09
SKSFF(f=0.25)	5.7837±0.0004				0.08[f]
NoTorSKSFF	5.6622±0.0004				0.10
NoTorNoAttSKSFF	5.6432±0.0005				0.10

[a] Self avoidance factor (SAF) is defined as the difference between the limiting value (2π for CCBB) and the calculated value divided by the limiting value (2π).

[b]. $0.22 = (6 - 4.68)/6$

[c]. $0.18 = (8 - 6.53)/8$

[d] $0.16 = (12 - 10.04)/12$

[e] $0.14 = (2\pi - 5.36)/2\pi$

[f] $0.08 = (2\pi - 5.78)/2\pi$

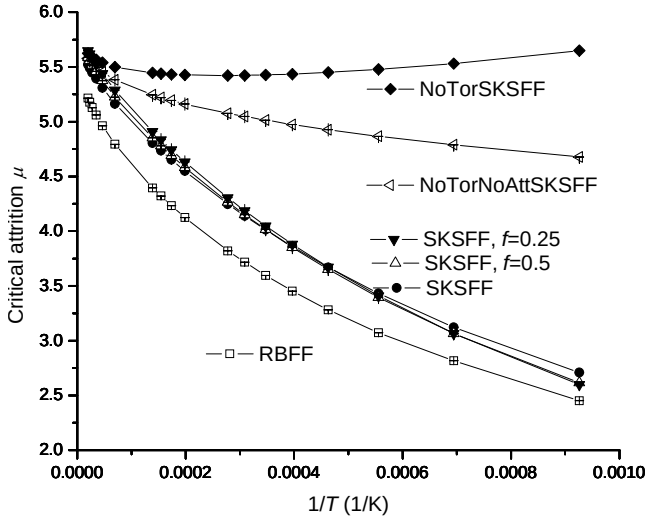


Figure 16. Critical attrition μ as a function of $1/T$ from CCBB with various force fields. In each case the uncertainties are shown less than the size of the symbols. We find that μ increases with temperature (except for NoTorSKSFF) reflecting the increasing number of available conformations and partition function. The relationship between μ and T is given in Eq. (42) in section 4.9.

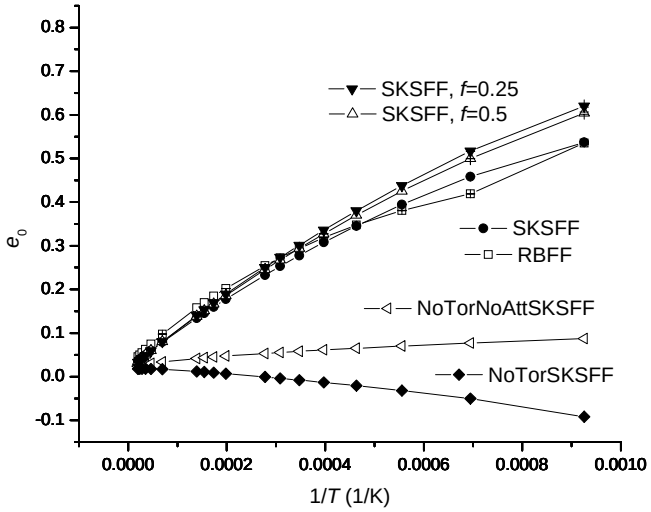


Figure 17. The critical energy increment, e_0 (see Eq. 30), as a function of $1/T$ from CCBB with various force fields. The uncertainties are less than the size of the symbols. We find

that e_0 decreases as temperature increases, except for NoTorSKSFF. This indicates that the average energy $<E>$ increases slower than kT, except for NoTorSKSFF. The relationship between e_0 and T is given in Eq. (45) of section 4.9.

Table 7. The critical energy increment, e_0, (see Eqn. 30) derived from CCBB at infinite temperature using various force fields. Here e_0 is obtain from fitting $e_0(T)$ to a quadratic expansion in $1/T$ over the range of 5040K ~ 50400K.

e_0	RBFF	SKSFF	SKSFF(f=0.5)	SKSFF(f=0.25)	NoTorSKSFF	NoTorNoAtt-SKSFF
CCBB	0.0255±0.0001	0.01871±0.0003	0.01673±0.0003	0.01436±0.0004	0.0186±0.0001	0.0209±0.0001

4.8 critical energy increment e_0

e_0 from the fitting data to Eq. (30) is shown in Figure 17, which decreases as the temperature increases (except for NoTorSKSFF). [Eq. (45) in section 4.9 shows the relationship between e_0 and T.]. Figure 17 shows the relationship between e_0 and $1/T$ from various force fields. As the temperature increases, the average energy $<E>$ increases slower than kT, except for NoTorSKSFF.

For NoTorSKSFF which has only repulsive and attractive van der Waals energies, the attractive energy dominates over the repulsive part at low temperature. This leads to an average energy $<E>$ that increases faster than kT.

To derive e_0 at infinite temperature, we fit e_0 to a second order function of $1/T$ in the range of 5040K to 50400K. Table 7 lists the values of e_0 at infinite temperature derived from this analysis. CCBB calculations with various force fields give e_0 ranging from 0.014~0.025.

Figure 17 shows that the e_0 from various force fields are closer at high temperature than low temperature, due to the decreased effect of well depth ε/k_B at high temperature.

4.9 The mean-field model for the Flory phase

The results in section 4.7 and 4.8 can be used to develop a mean field model valid throughout the Flory phase that *includes* the non-bond interactions between monomer units averaged over the ensemble (including both the repulsive and attractive parts). Assuming pair-wise interactions, the non-bond interaction for each monomer is:

$$E_n = k_B T\lambda \frac{N}{R_g^3} = C_n k_B T = (C_R - C_A)k_B T = E_R - E_A \tag{35}$$

where the λ factor represents both excluded volume repulsive effects and attractive effects. Here

$$C_n = \lambda \frac{N}{R_g^{\ 3}} = C_R - C_A \tag{36}$$

where $C_R(E_R)$ and $C_A(E_A)$ represent the Repulsive and Attractive parts, respectively.

Simplifying the torsion energy of each monomer as:

$$E_T(\phi) = \frac{k_\phi}{2}(1+\cos 3\phi) \tag{37}$$

the total energy for each monomer becomes

$$E = E_T + E_n = \frac{k_\phi}{2}(1+\cos 3\phi) + C_n k_B T \tag{38}$$

This leads to a partition function for each monomer of the form

$$\mu = \int_0^{2\pi} e^{-E/k_B T} d\phi = \int_0^{2\pi} e^{-\frac{k_\phi}{2k_B T}(1+\cos 3\phi)-C_n} d\phi = c_2 \int_0^{2\pi} e^{-c_1 \cos 3\phi} d\phi \tag{39}$$

where

$$c_1 = \frac{k_\phi}{2k_B T} \tag{40}$$

$$c_2 = e^{-C_n - c_1} \tag{41}$$

and

$$\mu = c_2 \int_0^{2\pi} e^{-c_1 \cos 3\phi} d\phi = c_2 \cdot 2\pi \cdot BesselI[0, c_1] \tag{42}$$

When the chain stiffness is zero ($c_1 = 0$), this (42) becomes

$$\mu = 2\pi \cdot e^{-C_n} \tag{43}$$

The average energy is

$$\begin{aligned} < E >= & \int_0^{2\pi} E \cdot e^{-E/k_B T} d\phi = C_n k_B T \mu + \frac{k_\phi}{2} \mu + \frac{k_\phi}{2} \cdot c_2 \cdot \int_0^{2\pi} \cos 3\phi \cdot e^{-c_1 \cos 3\phi} d\phi \\ = & (C_n k_B T + \frac{k_\phi}{2}) c_2 \cdot 2\pi \cdot BesselI[0, c_1] - \frac{k_\phi}{2} \cdot c_2 \cdot 2\pi \cdot BesselI[1, c_1] \\ = & k_B T \cdot C_n \cdot c_2 \cdot 2\pi \cdot BesselI[0, c_1] + \frac{k_\phi}{2} \cdot c_2 \cdot 2\pi \cdot \{BesselI[0, c_1] - BesselI[1, c_1]\} \end{aligned} \tag{44}$$

And we have

$$e_0 = \frac{< E >}{k_B T} = C_n \cdot c_2 \cdot 2\pi \cdot BesselI[0, c_1] + \frac{k_\phi}{2k_B T} \cdot c_2 \cdot 2\pi \cdot \{BesselI[0, c_1] - BesselI[1, c_1]\} \tag{45}$$

Setting $c_1 = 0$ (torsions negligible compared to *kT*), leads to

$$e_0 = C_n \cdot e^{-C_n} \cdot 2\pi + \frac{k_\phi}{2k_B T} \cdot e^{-C_n} \cdot 2\pi \tag{46}$$

while turning off torsions entirely (k_ϕ=0), leads to

$$e_0 = C_n \cdot e^{-C_n} \cdot 2\pi \tag{47}$$

Turning off both torsions (k_ϕ=0) and the attractive nonbonds, C_A=0, leads to

$$e_0 = C_R \cdot e^{-C_R} \cdot 2\pi \tag{48}$$

Using Eq. (42), (46), and (47), we can deduce C_n, C_R from both μ and e_0. We choose e_0 instead of μ because C_n, C_R are in the exponent in Eq. (45), making it more sensitive to the values. We can derive C_R directly from NoTorNoAttSKSFF, and we can get C_n. from NoTorSKSFF.

Figure 18 shows C_R, C_n, C_A derived from e_0 in Eq. (46), (47) at different temperatures from CCBB with NoTorNoAttSKSFF and NoTorSKSFF. Here, C_R, C_n, C_A are found to be linear with $1/T$, which indicates E_R, E_n, E_A are independent of temperature.

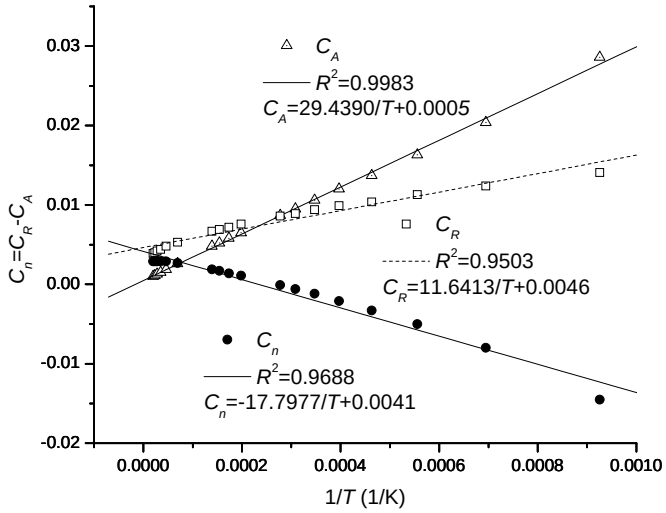


Figure 18. The generalized repulsive (C_R), torsion (C_n), and attractive (C_A) energy components, derived from e_0 at various temperatures from CCBB using NoTorNoAttSKSFF and NoTorSKSFF. We find that C_R, C_n, and C_A are linear with $1/T$, indicating that the energy components, E_R, E_n, E_A, are independent of temperature over this temperature range.

The final results for SKSFF are

$$C_R, = 11.6413\text{K}/T+0.0046 \tag{49}$$

$$C_n, = -17.7977\text{K}/T+0.0041 \tag{50}$$

$$C_A = 29.4390\text{K}/T+0.0005 \tag{51}$$

$$E_R, = k_\text{B} (11.6413\text{K} +0.0046*T) \tag{52}$$

$$E_n, = k_\text{B} (-17.7977\text{K} +0.0041*T) \tag{53}$$

$$E_A = k_\text{B} (29.4390\text{K} +0.0005*T) \tag{54}$$

From Eq. (52), (53), and (54) we see that E_R, E_n, E_A are approximately independent of temperature.

Eq. (42) and (45) show that as the temperature increases, the critical attrition μ increases and e_0 decreases regularly with torsion energy.

Using these relationships we can predict the properties for other polymer chains by changing the vdW attraction term, C_A, the vdW repulsive term C_R, and the torsions potential term k_ϕ in these equations. With these factors removed, equations (49)-(54) provide a generic description for all polymers.

4.10 The Three phases for isolated polymer in solution

An isolated polymer in solution is usually argued to be in one of three states[36-38] depending on the strength of the inter-monomer interactions which are mediated by the solvent molecules and can be controlled via the temperature T.

At high temperatures and in so-called "good solvents" a polymer chain is expected to be in a swollen coil phase. Here the behavior is dominated by excluded volume interactions and is well describe by the non-interacting SAW model. Therefore at infinite temperature the partition function and average root-mean-square end-to-end distance are expected to scale as

$$Z_N \sim A\mu^N N^{\gamma_+ - 1} \tag{55}$$

$$\sqrt{< R_e^2 >_N} \sim BN^{\nu_+} \tag{56}$$

And $\gamma_+ > 1$, $\nu_+ > 1/2$.

For low temperatures, it is accepted that the partition sum is dominated by configurations that are internally dense and the collapsed polymer should have a well defined surface (and associated surface free energy). This leads to the globular or collapsed state describe with functions such as [38]

$$Z_N \sim A\mu^N \mu_S^{N^{2/3}} N^{\gamma_- - 1} \tag{57}$$

$$\sqrt{< R_e^2 >_N} \sim BN^{\nu_-} \tag{58}$$

With μ_S<1, γ_-<1, and ν_-<1/2.

Between these coil and globule states, is expected to be the transition state, refered to as the θ phase. This transition is expected to be tri-critical in nature, so that it should conform to a crossover scaling theory. This leads to

$$Z_N \sim A\mu^N N^{\gamma_t - 1} \tag{59}$$

$$\sqrt{< R_e^2 >_N} \sim BN^{\nu_t} R(By) \tag{60}$$

with γ_t=1 and ν_t=1/2. The function $R(By)$ must have asymptotic properties for large arguments (positive and negative) that match it to the high and low temperature behaviors.[38]

Using 2D SAW for walks of length up to 6000, Owcarek et al.[36] showed that Eq. (57) is valid for the collapsed phase with γ_-=1/4 (2D).

In this thesis, we use continuous conformation modeling (rather than a lattice) to obtain the first description of the temperature dependence of γ and ν from infinite temperature to below theta temperature. This describes the temperature range from swollen coil state to the theta state. These methods could also be used to describe lower temperatures and longer chain lengths, which would allow us to validate of Eq. (57). However, such calculations become more computationally demanding, requiring much additional effort and we have not yet started such studies.

We find that γ and ν decrease uniformly with the temperature, reflecting the reduction in excluded volume effects and increases in monomer-monomer attractions. We also provide the physics underlying the behavior of γ (section 4.3) and its components (section 4.5), and we illustrate its temperature dependence.

Zhen-Gang Wang has questioned whether the chain length range in current study is sufficient to establish the validity of the ln$N$ dependence on which gamma is derived. He suggests that for sufficiently large N, the results should always be fittable by a surface term

as in (26). We argue that the central feature of the free energy theta temperature is the linearity of the free energy on N. Thus in a surface type analysis this corresponds to the temperature at which the surface energy is zero *and independent of N*. This is the dominant aspect of the Flory phase. At sufficiently lower temperatures to obtain the collapsed phase, the physics demands the surface analysis as in (26). These are very interesting points and we plan to modify our programs to make the study of much longer chain lengths and lower temperatures practical.

5. Summary

We use Continuous Configuration Boltzmann Biased (CCBB) Monte Carlo sampling (no lattice) to derive the thermodynamic properties [free energy (A), entropy (S), internal energy (E), and radius gyration ($<Rg^2>$)] for isolated polymer chains from the theta temperature (T_θ) to the high temperature limit with chain length from N=6 to 400. This is carried out for a force field (FF) accurately describing polyethylene (PE) and other FFs simulating an implicit solvent.

Throughout this range of temperature we find that $R_g^2 \sim N^{2\nu}$ with 2ν=1.168 and $Z_N \propto N^{\gamma-1}\mu^N$ with γ=1.153 and μ=5.649, indicating that the self-avoidance factor is 10%.

We establish a model to illustrate the physics of the pre-factor $N^{\gamma-1}$ and its components γ_S and γ_E, which we find to arise from the self-avoiding entropy and attractive energy, respectively. At infinite temperature we find that γ_E = 0 and γ_S = 1.155. Both quantities increase as the temperature decreases due to shrinkage in the chain size. At the theta temperature we find $\gamma_E = \gamma_S \sim 1.65$. At lower temperature, where attractive energy dominates, we find $\gamma_E > \gamma_S$ and $\gamma<1$. Our results confirm the prediction by Des Cloizeau that the internal energy increment is uniform (no pre-factor) at high temperatures with e_0=0.0187.

Using a $1/T$ expansion of the temperature dependence of the thermodynamic properties, we derive a general mean field model valid throughout the Flory phase. This shows that the

non-bond interaction becomes temperature independent for high temperatures. It also explains why μ increases and e_0 decreases regularly with increasing temperature.

Within the Flory phase, there is always as temperature for which γ=1 (no pre-factor), leading to a Partition function $Z=\mu^N$ corresponding to a gas of free monomers. We denote this as the free energy theta temperature, $T_{\theta F}$. We find that $T_{\theta F}$ is ~ 10% higher than the temperature $T_{\theta \nu}$ at which the chain describes a Gaussian coil (2ν=1). This shift in the theta point arises from three-body terms in the expansion of free energy in terms of concentration. When γ=1, both γ_E and γ_S are greater than 1 so that the chain cannot be treated exactly as ideal gas of monomers.

To mimic solvent effects we use a van der Waals scaling factor (f) in the Lennard-Jones potential. We find that this solvent scaling factor correlates linearly with the theta temperature. Our CCBB MC calculations assumed fixed bond lengths and angles, however the calculated theta temperature is nearly identical with the value obtained in molecular dynamics studies that allowed variable bonds and angles. This indicates that the fixed bond length and angle assumption has little effect on the thermodynamics of the Flory phase.

6. Acknowledgments

We thank Professor Zhen-Gang Wang for helpful and stimulating discussions.

This research was partially supported by grants from NSF (CHE99-85574) and ARO-MURI (DAAG55-97-1-0126). In addition Teijin Corporation supported JS while he spent a sabbatical at the MSC. The computational facilities were provided by DURIP grants from ARO and ONR, and MRI grant from NSF, and a SUR grant from IBM. The facilities of the Materials and Process Simulation Center are also supported by NIH, DOE (ASCI ASAP), NSF (CHE), MURI-ARO, MURI-ONR, General Motors, Chevron Texaco, Seiko-Epson, Beckman Institute, Asahi Kasei, and Toray Corp.

7. Reference

(1) Marshall, W.; Lovesey, S. *Theory of neutron scattering*; Oxford University Press: London, 1971.
(2) Benedek, G. *Polarisation, Matiere, et Rayonnement*; Presses Universitaries de France: Paris, 1969.
(3) Edwards, S. F. *Fluides Moleculaires*; Gordon & Breach: New York, 1976.
(4) Domb, C. *Adv. Chem. Phys.* **1969**, *15*, 229.
(5) Wall, F. T.; Erpenbeck, J. *J. Chem. Phys.* **1959**, *30*, 634.
(6) Fixman, M. *J. Chem. Phys.* **1973**, *58*, 1553.
(7) Fixman, M. *J. Chem. Phys.* **1973**, *58*, 1559.
(8) Fixman, M. *J. Chem. Phys.* **1973**, *58*, 1564.
(9) Milchev, A.; Paul, W.; Binder, K. *J. Chem. Phys.* **1993**, *99*, 4786.
(10) Wittkop, M.; Kreitmeier, S.; Goritz, D. *J. Chem. Phys.* **1996**, *104*, 3373.
(11) Olaj, O. F.; Petrik, T.; Zifferer, G. *J. Chem. Phys.* **1997**, *107*, 10214.
(12) Grassberger, P. *Phys. Rev. E* **1997**, *56*, 3682.
(13) Yamakawa, H.; Yoshizaki, T. *J. Chem. Phys.* **2003**, *118*, 2911.
(14) Sadanobu, J.; III, W. A. G. *J. Chem. Phys.* **1997**, *106*, 6722.
(15) Ryckaert, J. P.; Bellemans, A. *Faraday Discuss. Chem. Soc.* **1978**, *66*, 95.
(16) Smit, B.; Karaborni, S.; Siepmann, J. I. *J. Chem. Phys.* **1995**, *102*, 2126.
(17) Martin, M. G.; Siepmann, J. I. *J. Am. Chem. Soc.* **1997**, *119*, 8921.
(18) Siepmann, J. I.; Karaborni, S.; Smit, B. *nature* **1993**, *365*, 330.
(19) Li, Y.; Lin, S.-T.; III, W. A. G. *J. Am. Chem. Soc.* **2004**, *126*, 1872.
(20) Jorgensen, W. L.; Madura, F. D.; Swenson, C. J. *J. Am. Chem. Soc.* **1984**, *106*, 6638.
(21) Jang, S. S.; Cagin, T.; III, W. A. G. *J. Chem. Phys.* **2003**, *119*, 1843.
(22) Guida, R.; Zinn-Justin, J. *J. Phys. A: Math. Gen.* **1998**, *31*, 8103.
(23) Allegra, G.; Colombo, E. *J. Chem. Phys.* **1994**, *101*, 4268.
(24) Butera, P.; Comi, M. *Phys. Rev. B* **1997**, *56*, 8212.
(25) MacDonald, D.; Hunter, D. L.; Kelly, K.; Jan, N. *J. Phys. A: Math. Gen.* **1992**, *25*, 1429.
(26) Li, B.; Madras, N.; Sokal, A. D. *J. Stat.Phys.* **1995**, *80*, 661.
(27) Belanger, D. P.; Yoshizawa, H. *Phys. Rev. B* **1987**, *35*, 4823.
(28) Des Cloizeau, J.; Jannink, G. *Polymers in solution: Their Modelling and Structure*; Clarendon press: Oxford, 1990.
(29) Trainoff, S.; Wyatt, P. J., Eds. *International GPC Symposium's 98*, 1999.
(30) Watts, M. G. *J. Phys. A: Math. Gen.* **1975**, *8*, 61.
(31) Gennes, P. G. d. *scaling concepts in polymer physics*; cornell university press: London, 1979.
(32) Caracciolo, S.; Causo, M. S.; Pelissetto, A. *Phys. Rev. E* **1998**, *57*, 1215.
(33) Pestak, M. W.; Chan, H. W. *Phys. Rev. B* **1984**, *30*, 274.
(34) Brandup, J.; Immergut, E. H., Eds. *Polymer Handbook*; Wiley: New York, 1989.
(35) Mark, J. E., Ed. *Physical Properties of Polymer Handbook*; AIP: Woodbury, 1996.
(36) Owczarek, A. L.; Prellberg, T.; Brak, R. *Phys. Rev. Lett.* **1993**, *70*, 951.
(37) Grassberger, P.; R, H. *J. Chem. Phys.* **1995**, *102*, 6881.
(38) Prellberg, T.; Owczarek, A. L. *Phys. Rev. E* **1995**, *51*, 2142.