Summer Scholar Report

Towards an All-Atom Model for m-Poly(phenylene ethynylene) Class of Foldamers using the CHARMM27 Force Field: Structure and Unfolding Pathway of Foldamers of Various Lengths

Morris E. Cohen, Adam Moser,† and Binyomin Abrams,* Department of Chemistry, Boston University, Boston MA

*To whom correspondence should be addressed. E-mail: abramsb@bu.edu

† Current address: Loras College, Dubuque, IA

Abstract

An all-atom model for the m-Poly(phenylene ethynylene) (mPPE) class of foldamers was developed for the CHARMM27 force field. The folding behavior of mPPEs of different residue lengths with triethylene glycol exohelical side chains in explicit water were studied by all-atom molecular dynamics simulations with the CHARMM27 force field. mPPEs were studied in both the helical and extended conformations in water to find the minimum energy structures. The unfolding pathway and free energy landscapes of the foldamers were probed using metadynamics, in which the folded state was determined to be 8.5 and 16.3 kcal · mol^{-1} more favorable in implicit solvent and explicit solvent, respectively.

Introduction

Biological macromolecules, such as proteins, lipids, and DNA self-assemble into compact solution structures.[1] Folded conformations are determined by temperature, hydrogen bonding, and van der Waals interactions. Due to the prevalence of folding in bio-macromolecules, a deep understanding of the folding process is of particular importance.[2] Consequently, significant work has been done to understand the fundamental science behind the folding process.[3] Because of the complex nature of proteins and other bio-macromolecules, compounds that mimic the way proteins behave have been sought out.[4] Nonetheless, despite α-helices being one of the most common protein structures,[5] there exist relatively few synthetic molecules that mimic an α-helix.[1] This has led to the development of m-Poly(phenylene ethynylene) or mPPEs,[6] oligomers that fold using non- covalent interactions between atoms to adopt a helical structure that is more thermodynamically favored in solution. J. Moore and others[6–10] have shown that mPPEs with ester side chains and a residue length of 12 or more adopt a specific α-helix conformation in acetonitrile and water.

mPPEs have been both synthesized[7,11,12] and modeled,[13–15] but few are water soluble, which presents a significant limitation on their application to biological systems.[9,16] Both Saven[13] and Sen[17] have independently shown that the folded α-helix conformation is a minimum energy

structure. Bruce and others1[2, 18–20] have shown that mPPEs with small exohelical side chains (COOMe, CONH2 , etc.) prefer a folded conformation over an extended conformation. Additionally, Elmer [21] has shown that mPPEs can get trapped in misfolded states during the folding process, similar to that of proteins.

In this study, we examine the effect of residue length on the folding propensity of mPPE with triethylene glycol ester side chains (COOTg) in both explicit and implicit solvent. Moore has previously shown experimentally that an mPPE with residue length of eight will not fold, and a residue length of ten will form a semi-folded state.[10] Additionally, the free energy landscape of unfolding was investigated using metadynamics, an accelerated sampling method used to study free energy surfaces and force systems out of local minima in molecular dynamics simulations.[22]

Parameterization of mPPE and Simulation Details

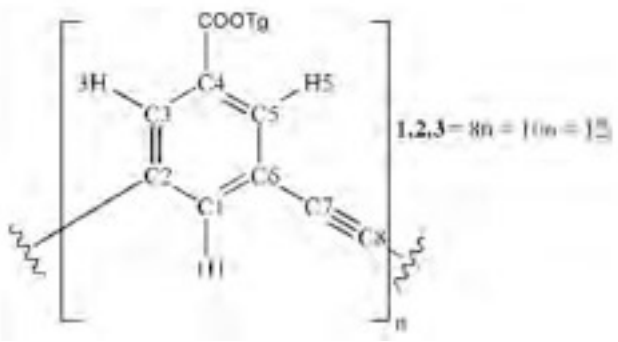

Figure 1: *Structure of mPPE monomer. C8 is connected to C2 of the following unit. 1, 2, and 3 have monomer lengths of 8, 10, and 12, respectively. Triethylene glycol (T_g) side chains were used in this study.*

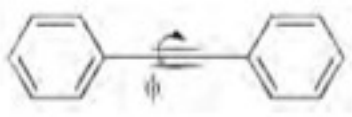

Figure 2: *Structure of tolane. The potential for the dihedral angle ϕ between the two benzene rings was calculated by Okuyama experimentally to be 0.6 kcal·mol^{-1}.[29]*

Atom	HF/631G** Charges	mPPE Charges
C2, C6	0.29	0.20
C4	0.01	-0.04
C7, C8	-0.17	-0.20

Table 1: *Comparison of HF charges to the charges used for mPPE (atoms labeled in Figure 1)*

Parameterization

The potential energy term for the CHARMM force field[23] is given by:

$$V = \sum_{bonds} K_b(b - b_0)^2 + \sum_{angles} K_\theta(\theta - \theta_0)^2 + \sum_{UB} K_{UB}(s - s_0)^2 + \sum_{dihedrals} K_\phi[1 + \cos(n\phi - \delta)] + \sum_{impropers} K_\psi(\psi - \psi_0)^2 + \sum_{nonbonded} \left[\epsilon_{ij}\left(\left(\frac{R_{min,ij}}{r_{ij}}\right)^{12} - 2\left(\frac{R_{min,ij}}{r_{ij}}\right)^{6}\right) + \frac{q_i q_j}{\epsilon_1 r_{ij}}\right] \quad (1)$$

where the force constants K_b, K_θ, K_{UB}, K_ϕ, K_ψ are the bond, angle, Urey-Bradley, dihedral, and improper dihedral force constants, respectively; b, θ , s, ϕ , ψ are the bond length, bond angle, Urey-Bradley distance, dihedral angle, and improper dihedral, respectively.

Force field parameters for 3 (mPPE 12-mer, Figure 1) were either obtained from analogous residues in the CHARMM27 force field[24] or were developed to fit experimental results and quantum mechanical calculations. The triethylene glycol side chain parameters were taken from the CHARMM ether force field. [25] The ester functional group parameters were taken from the CHARMM lipid parameters for deoxylysophosphatidyl-choline (LPPC).[26] The triple bond carbon parameters were taken from the CGenFF parameters of 2-butyne.[27] The benzene carbons and hydro- gens were matched to the benzene parameters of phenylala- nine.[24] The missing charges (C2, C4, C6-8, Figure 1) were calculated using Gaussian0928 using HF/631G** basis set and fit to ensure charge neutral groups. The parameters for the dihedral, ϕ, between the two benzene planes (Figure 2) were derived from Gaussian09 HF/631G** calculations, and using experimental data of tolane,[29] to produce the correct rotational potential energy surface. The *ab initio* torsional potential energy surface was obtained from a 90 ˚ rotational scan in 10○ increments due to the periodicity of the potential energy surface. Parameters were then created K_ϕ , n, and δ to match the potential energy surface. The best-fit parameters were determined to be a force constant K_ϕ = 0.06 kcal · mol^{-1} · rad^{-2}, periodicity n = 2, and δ = 180 ˚.

Simulation Details

Initial helical configurations of mPPE were generated using the CHARMM36 package and the new mPPE param- eters. Explicit solvent simulations had one mPPE solvated in TIP3P water in truncated octahedra (8469 TIP3P waters and side of length of 80 Å for 3). Equilibration runs (500 ps) and production runs (10-50 ns) were performed using 1 fs integration timesteps using the NAMD2 package.[30] The NPT ensemble was generated using Langevin piston method at 1 atm with a piston temperature of 300 K, a piston period of 200 fs, and a damping time constant of 100 fs. Electro- statics were treated using Particle Mesh Ewald. Non-bonded interactions were calculated using a switching function between 12 and 14 Å.

For implicit solvent simulations, Generalized Born Implicit Solvent modeling water as a dielectric continuum with a dielectric of 78.5 was used. Equilibration runs (1 ns) and production runs (100 ns) were performed using 1 fs integration timesteps using the NAMD2 package[30] with langevin temperature control. An alpha cutoff of 16 was used for calculating Born radius. Non-bonded interactions were calculated using a switching function between 16 and 18 Å.

Results and Discussion

Solvent	Oligomer Length	R_g (Å)	Pitch (Å)
Explicit	8-mer (**1**)	6.66 ± 0.14	3.42 ± 0.31
	12-mer (**3**)	6.54 ± 0.23	3.42 ± 0.15
Implicit	8-mer (**1**)	6.33 ± 0.21	3.53 ± 0.32
	12-mer (**3**)	6.68 ± 0.13	3.42 ± 0.31

Table 2: *Structural conformation data for mPPEs of different oligomer lengths. Rg refers to radius gyration of the mPPE backbone. The pitch refers to the average distance between the planes of benzenes six residues apart*

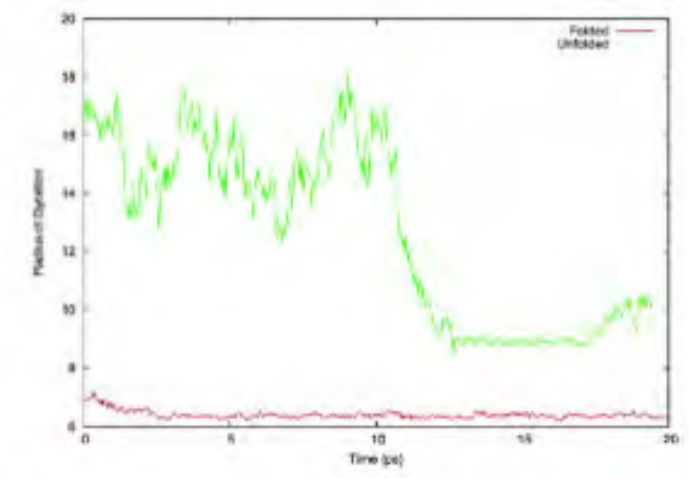

Figure 3: *Time evolution of Rg for 3 in explicit solvent beginning in the folded conformation (red) and extended-coil conformation (green). The extended-coil conformation collapses into a misfolded knotted structure that is not escaped for the rest of the simulation.*

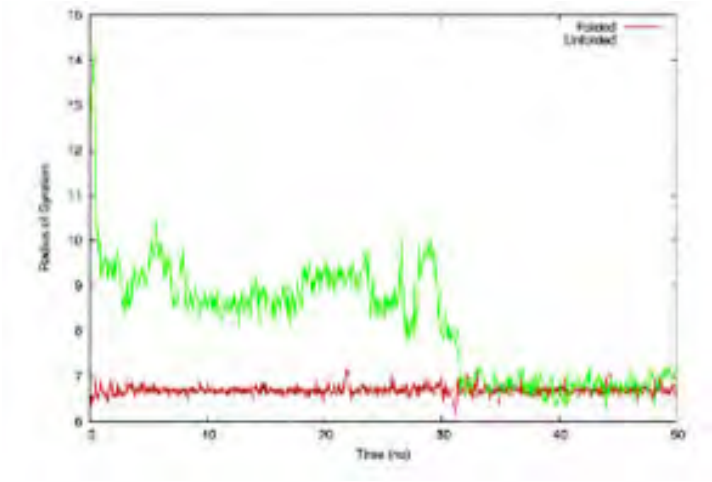

Figure 4: *Time evolution of R_g for 3 in implicit solvent beginning in the folded conformation (red) and extended-coil conformation (green). The extended-coil conformation collapses into a semi-folded structure after 2 ns and eventually collapses into a misfolded state, which it does not escape.*

Structure

In explicit solvent, mPPEs 1-3 remained in the folded conformation throughout simulations up to 50 ns, and for simulations of 100 ns in implicit solvent. The time evolution of Rg for 3 is shown in Figure 3 (explicit solvent) and Figure 4 (implicit solvent). Rg remained constant at a value of nearly 6.5 Å (Table 2) for all three foldamers, in both explicit and implicit solvents, which indicates a strong favoring of the folded structure. The helix pitch, measured as the average distance between the planes of benzenes six residues apart, was determined to be 3.42 Å, which is consistent with the experimental value of 3.5 Å.[10]

Simulations of 3 beginning in a random coil show a sharp decrease in R_g at 12 ns in explicit solvent, which corresponds to a collapse in structure from an extended-coil into a misfolded structure (Figure 3). In implicit solvent, a semi-folded structure was achieved after 2 ns (Figure 4). Furthermore, a misfolded structure, with R_g equivalent to the folded structure, occurs at 30 ns

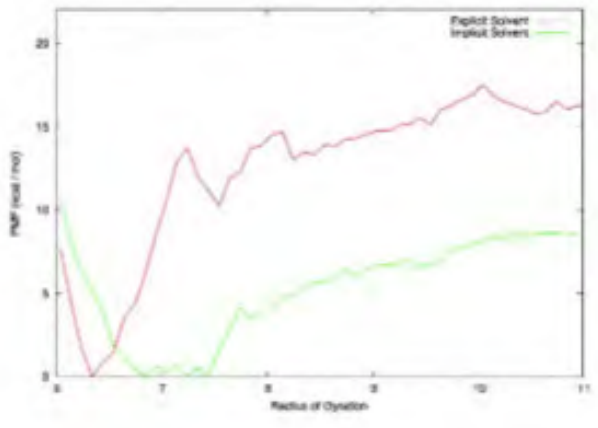

Figure 5: *Potential of mean force for 3 calculated using metadynamics on R^g collective variable, in explicit (red) and implicit (green) solvent. The folded state is found to be 8.5 kcal · mol^{-1} (implicit solvent) and 16.3 kcal · mol^{-1} (explicit solvent) more stable than the extended conformation.*

and persists for the remainder of the 100 ns simulation. While R_g of 3 in implicit solvent beginning in an extended-coil eventually reaches that of the folded structure, examination of the structure shows a compressed, misfolded structure.

Metadynamics

The free energy associated with uncoiling of the mPPEs was determined using metadynamics[31] where a history- dependent bias potential is used to fill the free energy surface (FES) over time and allow efficient sampling of the FES. The form of the bias potential on a collective variable ξ is given as:

$$V_{meta}(\xi) = \sum_{t'=\delta t, 2\delta t}^{t'<t} W \prod_{i=1}^{N_{dv}} exp\left(-\frac{(\xi_i - \xi_i(t'))^2}{2\delta_{\xi_i}^2}\right) \qquad (2)$$

where W is the gaussian hill height and δ is the hill width. R_g was chosen as the collective variable to study the unfolding pathway (W = 0.1 kcal · mol^{-1}, δ = 0.3 Å, and gaussians were added every 1000 steps).

The folded conformations were the global minima in the free energy profiles along R_g, in both explicit and implicit solvent, which is consistent with the experimental results (Figure 5). In explicit solvent, the folded conformation was found to be 16.3 kcal · mol^{-1} more stable than the extended-coil conformation. As expected, the implicit sol- vent model leads to increased sampling of the non-native states[32] and the free energy of unfolding (8.5 kcal · mol^{-1} for 3 in implicit solvent) is significantly lower than that predicted by the explicit model. Furthermore, the minimum observed in implicit solvent was located at a slightly larger radius of gyration (R_g = 6.8) than in explicit solvent (R_g = 6.4), due to the fact that the helix in implicit solvent achieved a slightly larger helix radius before unfolding. Additionally, the broadening of the implicit solvent free energy well is due to the prevalence of thermodynamically- favorable semi-folded structures that occur with R_g near 7. Nonetheless, there is good qualitative agreement between the results for the unfolding pathway generated in implicit and explicit solvent.

Conclusions

In this paper, the parameterization of the mPPE class of foldamers is discussed. The new CHARMM parameters for mPPE have been shown to correctly reproduce the experimentally determined folded structure for oligomers of 8, 10, and 12 monomers. Simulations on these mPPEs starting in the folded conformation remain folded in both implicit and explicit solvent. Additionally, mPPEs starting in the extended-coil conformations collapse into mis-folded and semi-folded conformations, which are more stable than the extended-coil conformations. Free energy sampling from metadynamics further confirms that the folded conformation is preferred over the extended state by 8.5 and 16.3 kcal·mol^{-1} in implicit and explicit solvent, respectively. As expected, implicit solvent simulation underestimate the free energy of folding.

Our results indicate that the mPPE model correctly pre- dicts the equilibrium structure of the native state for oligomers of length 8-12 monomers and provides a free energy of unfolding that is consistent with the experimental results in acetonitrile. The relative free energies of unfolding of different length oligomers, and the relative free energies of unfolding in different solvents, deserve a separate study.

Acknowledgementsx0

The authors gratefully acknowledge the support of the Boston University Undergraduate Research Opportunities Program, the Northeastern Section of the American Chemical Society, the Boston University Scientific Computing and Visualization Group, and the National Science Foundation. MC acknowledges the James Flack Norris and Theodore William Richards Undergraduate Summer Scholarship.

References

(1) Guichard, G.; Huc, I. Chem. Commun. 2011, 47, 5933–5941.

(2) Vendruscolo, M.; Zurdo, J.; MacPhee, C. E.; Dobson, C. M. Phil. Trans. R. Soc. Lond. A. 2003, 361, 1205–1222.

(3) Dobson, C. M. Nature 2003, 426, 884–890.

(4) Hecht, S., Huc, I., Eds. Foldamers: Structure, Properties, and Applications; Wiley-VCH: Federal Republic of Germany, 2007.

(5) de Oliveira Santos, J. S.; Voisin- Chiret, A. S.; Burzicki, G.; Sebaoun, L.; Sebban, M.; Lohier, J.-F.; Legay, R.; Oulyadi, H.; Bureau, R.; Rault, S. J. Chem. Inf. Model 2012, 52, 429– 439.

(6) Nelson, J. C.; Saven, J. G.; Moore, J. S.; Wolynes, P. G. Science 1997, 277, 1793–1796.

(7) Prince, R. B.; Barnes, S. A.; Moore, J. S. J. Am. Chem. Soc. 2000, 122, 2758–2762.

(8) Heemstra, J. M.; Moore, J. S. Org. Lett. 2004, 6, 659–662.

(9) Stone, M. T.; Moore, J. S. Org. Lett. 2004, 6, 469–472.

(10) Lahiri, S.; Thompson, J. L.; Moore, J. S. J. Am. Chem. Soc. 2000, 122, 11315–11319.

(11) Zhao, D.; Moore, J. S. J. Am. Chem. Soc. 2002, 124, 9996–9997. 7

(12) Nguyen, H. H.; McAliley, J. H.; Bruce, D. A. Macromolecules 2011, 44, 60–67.

(13) Lee, O.-S.; Saven, J. G. J. Phys. Chem. B 2004, 108, 11988–11994.

(14) Blatchly, R. A.; Tew, G. N. J. Org. Chem. 2003, 68, 8780–8785.

(15) Tie, C.; Gallucci, J. C.; Parquette, J. R. J. Am. Chem. Soc. 2006, 128, 1162–1171.

(16) Arnt, L.; Tew, G. N. Macromolecules 2004, 37, 1283–1288.

(17) Sen, S. J. Phys. Chem. B 2002, 106, 11343–11350.

(18) Adisa, B.; Bruce, D. A. J. Phys. Chem. B 2005, 109, 7548–7556.

(19) Adisa, B.; Bruce, D. A. J. Phys. Chem. B 2005, 109, 19952–19959.

(20) Nguyen, H. H.; McAliley, J. H.; III, W. R. B.; Bruce, D. A. Macromolecules 2010, 43, 5932– 5942.

(21) Elmer, S.; Pande, V. S. J. Phys. Chem. B 2001, 105, 482–485

(22) Laio, A.; Gervasio, F. L. Rep. Prog. Phys. 2008, 71, 12660.

(23) Brooks, B. R. et al. J. Comput. Chem. 2009, 30, 1545–1614.

(24) MacKerell, A. D. et al. J. Phys. Chem. B 1998, 102, 3586–3616.

(25) Vorobyov, I.; Anisimov, V. M.; Greene, S.; Venable, R. M.; Moser, A.; Pastor, R. W.; Alexander D. MacKerell, J. J. Chem. Theory Com- put. 2007, 3, 1120–1133.

(26) Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; Alexander D. MacKerell, J.; Pastor, R. W. J. Phys. Chem. B 2010, 114, 7830–7843.

(27) Vanommeslaeghe, K.; Hatcher, E.; Acharay, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; I. Vorobyov, A. D. M. J. J. Comput. Chem. 2010, 31, 671–690.

(28) Gaussian09, Frisch, M. J. et. al. Gaussian, Inc., Wallingford CT, 2009.

(29) Okuyama, K.; Hasegawa, T.; Ito, M.; Mikami, N. J. Phys. Chem. 1984, 88, 1711–1716.

(30) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kaleé, L.; Schulten, K. J. Comput. Chem. 2005, 26, 1781–1802.

(31) Laio, A.; Parrinello, M. Proc. Natl. Acad. Sci. USA 2002, 99, 12562–12566.

(32) Zhou, R. Proteins 2003, 53, 148–161.