

The Structure of a γ -Modified Peptide Nucleic Acid Duplex

Wei He,¹ Matthew J. Crawford,¹ Srinivas Rapireddy,¹ Marcela Madrid,^{2,*} Roberto R. Gil,^{1,*} Danith H. Ly,^{1,*} Catalina Achim^{1,*}

¹ Department of Chemistry, Carnegie Mellon University, 4400 5th Avenue, Pittsburgh, PA 15213, USA

² Pittsburgh Supercomputing Center, Pittsburgh, PA, 15213

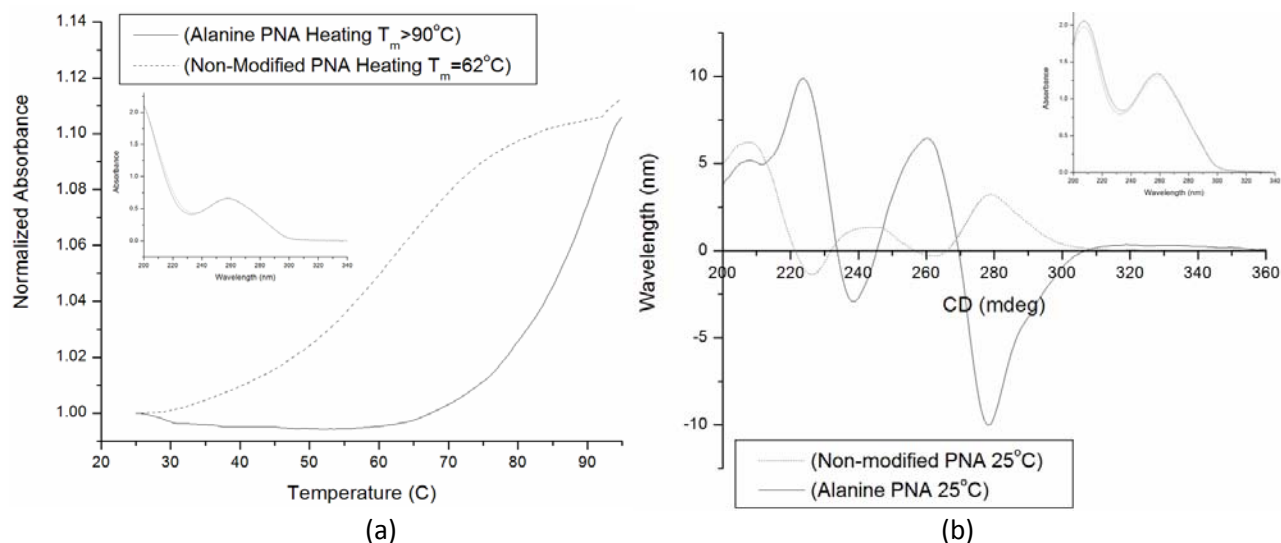


Figure S1. Melting curves (a) and CD spectra (b) of 10 μM solutions of γ -PNA (solid line) and aeg PNA (dashed line) in pH 7.0 10 mM sodium phosphate buffer.

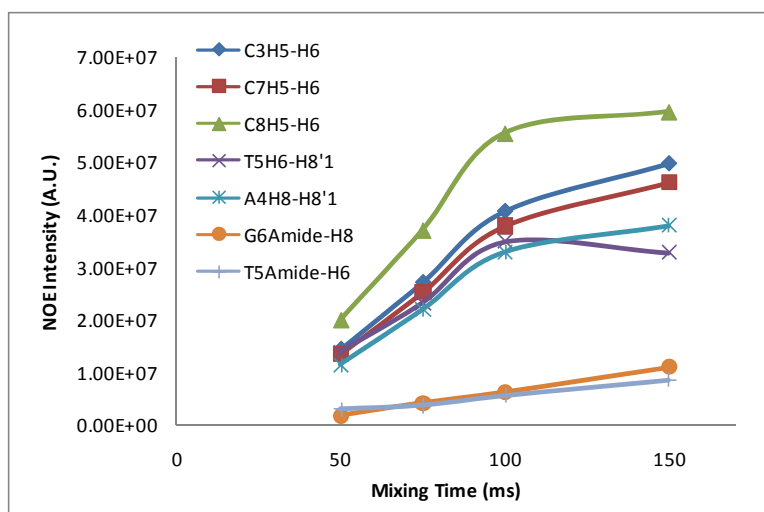


Figure S2. Experimental NOE build-up curves of the γ -PNA duplex with respect to NOESY mixing time.

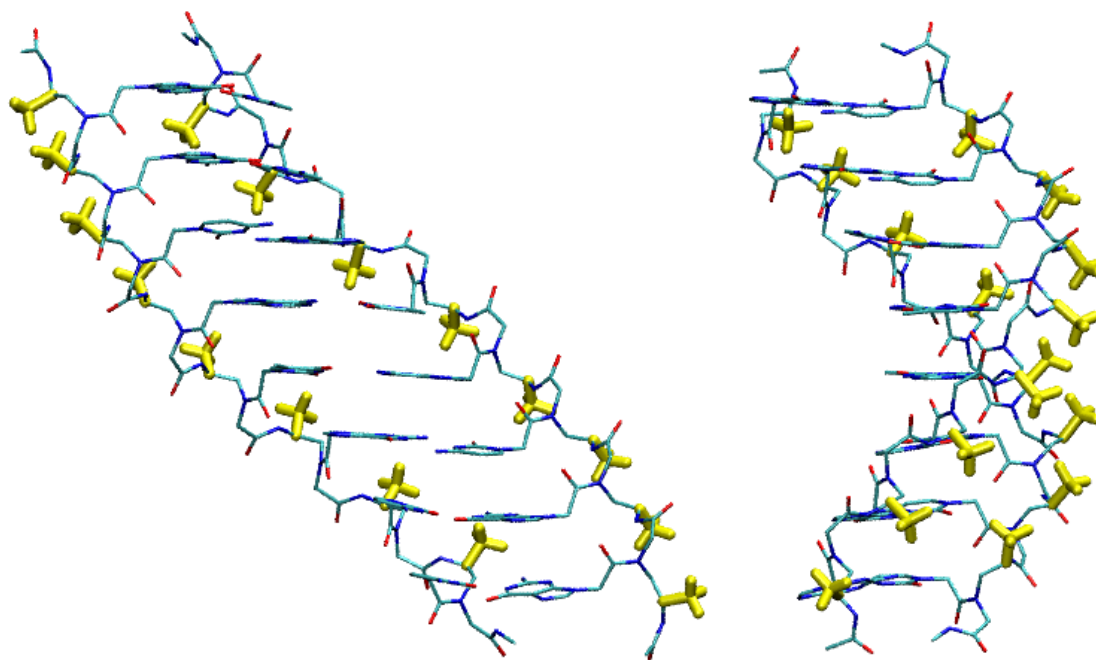


Figure S3. Two views of the initial model for γ -methylated PNA duplex based on the solution structure of the non-modified aeg PNA. The γ -methyl is labeled in yellow.

Table S1. NOE Violations for the ten γ -PNA models after 2 ns MD simulation.

#MD	Number of violations > 0.1 Å and <0.2 Å	Number of violations > 0.2 Å	Largest violation (Å) (Å)
1	4	5	0.46
2	9	4	0.29
3	7	0	0.19
4	8	2	0.26
5	5	3	0.32
6	7	4	0.32
7	4	6	0.38
8	5	4	0.25
9	5	7	0.44
10	9	4	0.28

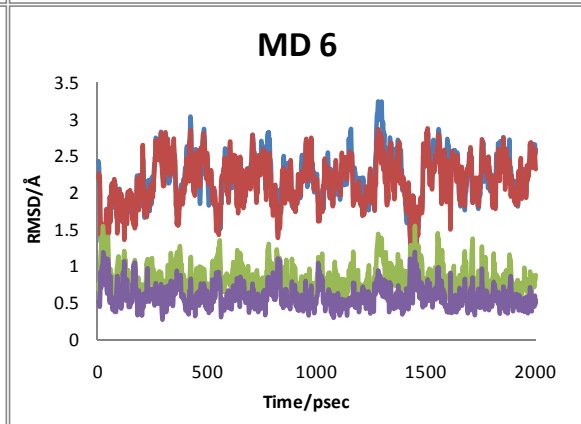
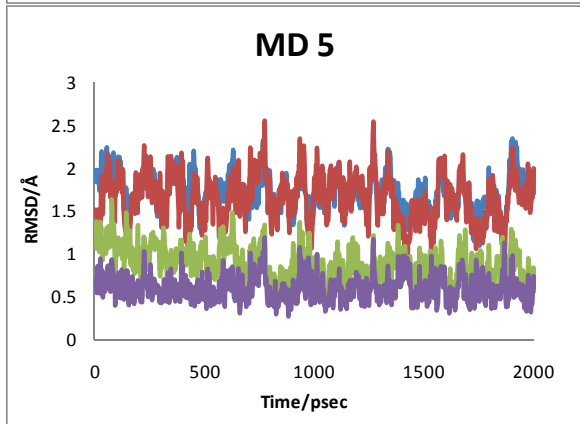
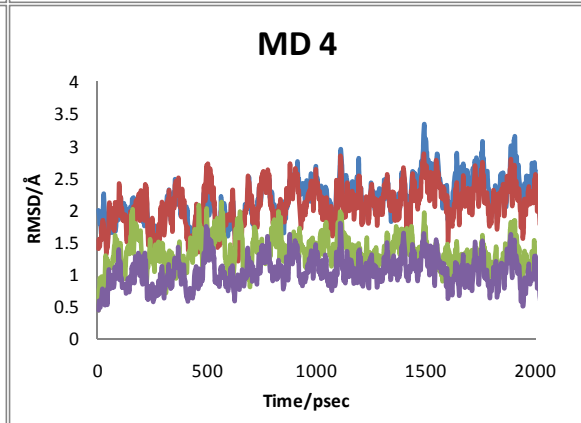
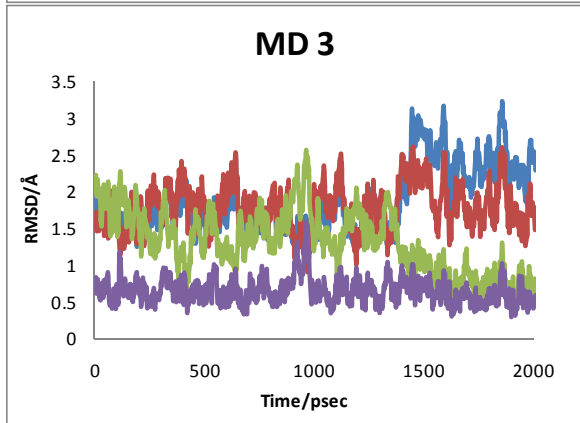
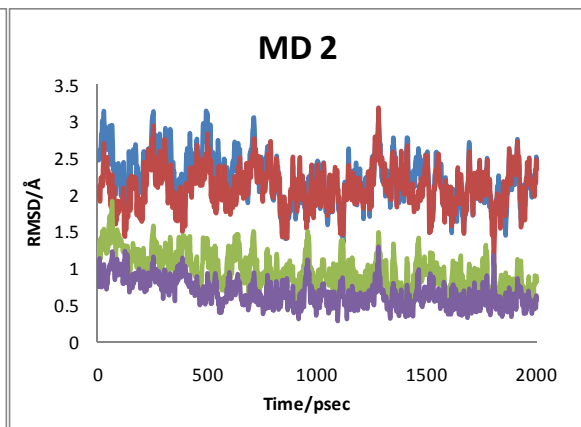
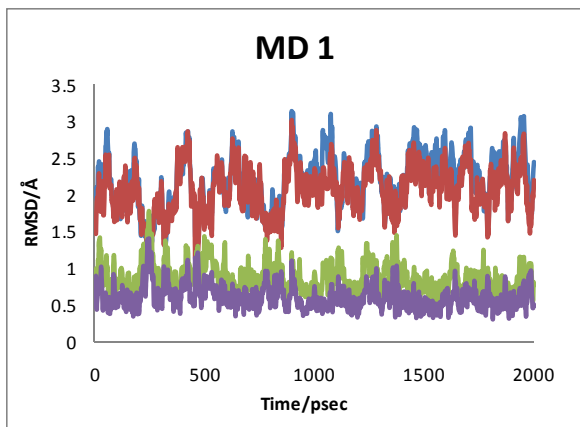
Molecular Modelling

The molecular dynamics simulations were performed using AMBER10.¹ The force field ff99SB² was complemented with the electrostatic charges of the γ -PNA monomer atoms. The γ -methyls were added to the non-modified PNA monomers to fulfill the *S*-configuration at backbone C2' position using the program Gaussview.³ The monomer backbones were capped with an acetyl group (-COCH₃) on the N-terminus and an N-methylamide (-NHCH₃) group on the C-terminus. The atomic partial charges were then computed separately for the four monomers following the Cornell *et al.* protocol.⁴ The geometries were optimized using the HF/6-31G* basis set.³ The ESP (Electrostatic Potential) charges on each atom were obtained by running a single point calculation using the MP2/6-31G* basis set, using Gaussian 03, Rev. E.01,³ and then refined by the restrained electrostatic potential (RESP) method implemented in the Antechamber module of AMBER.¹ Using these initial charges, RESP was ran twice more: first constraining the charges of the acetyl and N-methylamide groups to the values defined in the Cornell *et al.* force field,⁴ and then restraining the charges of chemically identical hydrogens.

The simulations were conducted using a starting structure based on the aeg PNA duplex (PDB ID: 2K4G).⁵ The H2'2 proton was first converted to carbon and then three methyl hydrogens were added to the mirror image of the 8-bp aeg PNA using the xleap module of AMBER.¹ The γ -PNA structure obtained by this procedure was subjected to ten independent 250 ps runs of NMR-restrained simulated annealing in vacuum. Each of the annealing runs was started with a different initial random velocity. This protocol was employed in order to increase the sampling of the conformational space. A time step of 1 fs was employed. The temperature was increased to 600 K during the first 80 ps, and then slowly decreased to 300 K during the rest of the simulation. The simulated annealing runs employed a total of 138 distance restraints obtained from the NMR experiments and 36 Watson-Crick base pairing restraints. The potential function for the restraints has a flat square bottom between the minimum and maximum inter-proton distances estimated from the NOESY cross-peaks and parabolic sides for 0.5 Å beyond that, and grows linearly thereafter. The force constant for the restraints was 20 kcal/mol. The weight of this force constant was gradually increased from 0.1 to 1 during the first 80 ps, and then kept at 1 for the remaining of the simulation.

The ten γ -PNA structures resulting from the simulated annealing runs were solvated in a truncated octahedral TIP3P water box, such that the distance between the walls of the box and the closest PNA atom was 9 Å. After energy minimization, the solvated structures were subjected to 2 ns of MD using

the same distance restraints as in the simulated annealing step. The simulations employed an NPT ensemble ($T = 300$ K, $P = 1$ atm) in periodic boundary conditions. The Particle Mesh Ewald method was used to compute full electrostatic interactions.^{1,6} Average structures for each of the ten MD simulations were generated from the last 400 ps of the trajectories. The convergence of the simulations was inferred using RMSDs from either the initial structure or the time-averaged structures. The RMSDs were computed after the alignment of either the backbone atoms N1', C2', C3', N4' and C5' or the base pair atoms N1, N3, C5, N7 and N9. Helical parameters were calculated with the program Curves,⁷ using only the 6 central base pairs of the γ -PNA duplex. The terminal nucleobase pairs have been excluded from the calculation because these nucleobase pairs are subject to fraying in solution. They were significantly more tilted towards the major groove than the central base pairs, which led to a substantial fluctuation in the calculated helical parameters when these base pairs were included in the calculation of the helical parameters. A linear global helical axis was chosen for the duplex and standard nucleotide geometry was applied in the calculation (using command settings MINI=true, LINE=true, FIT=true).⁷



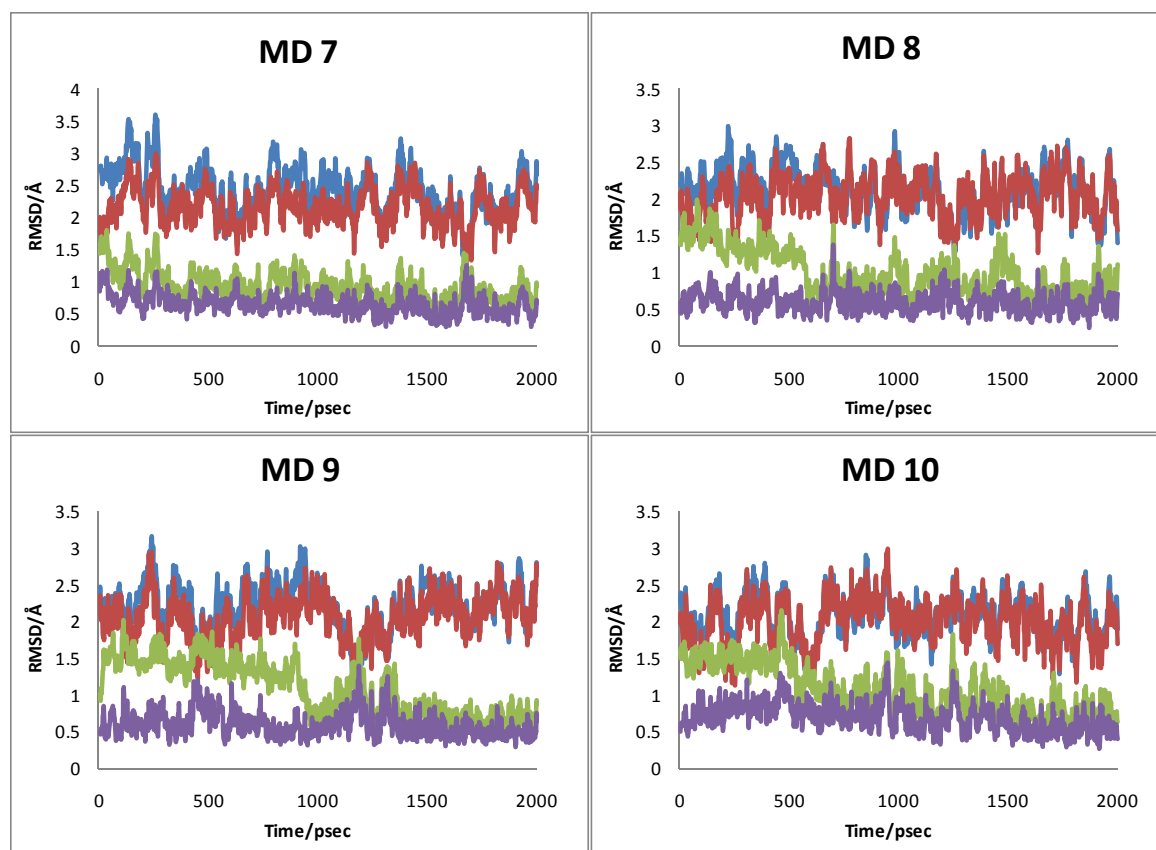


Figure S4. Root mean square deviation (RMSD) of the backbone and base pair regions of γ -PNA with respect to the starting PNA structure (backbone in blue and base pair in red) and with respect to the structure time-averaged over the last 400 ps of each trajectory (backbone in green and base pair in purple).

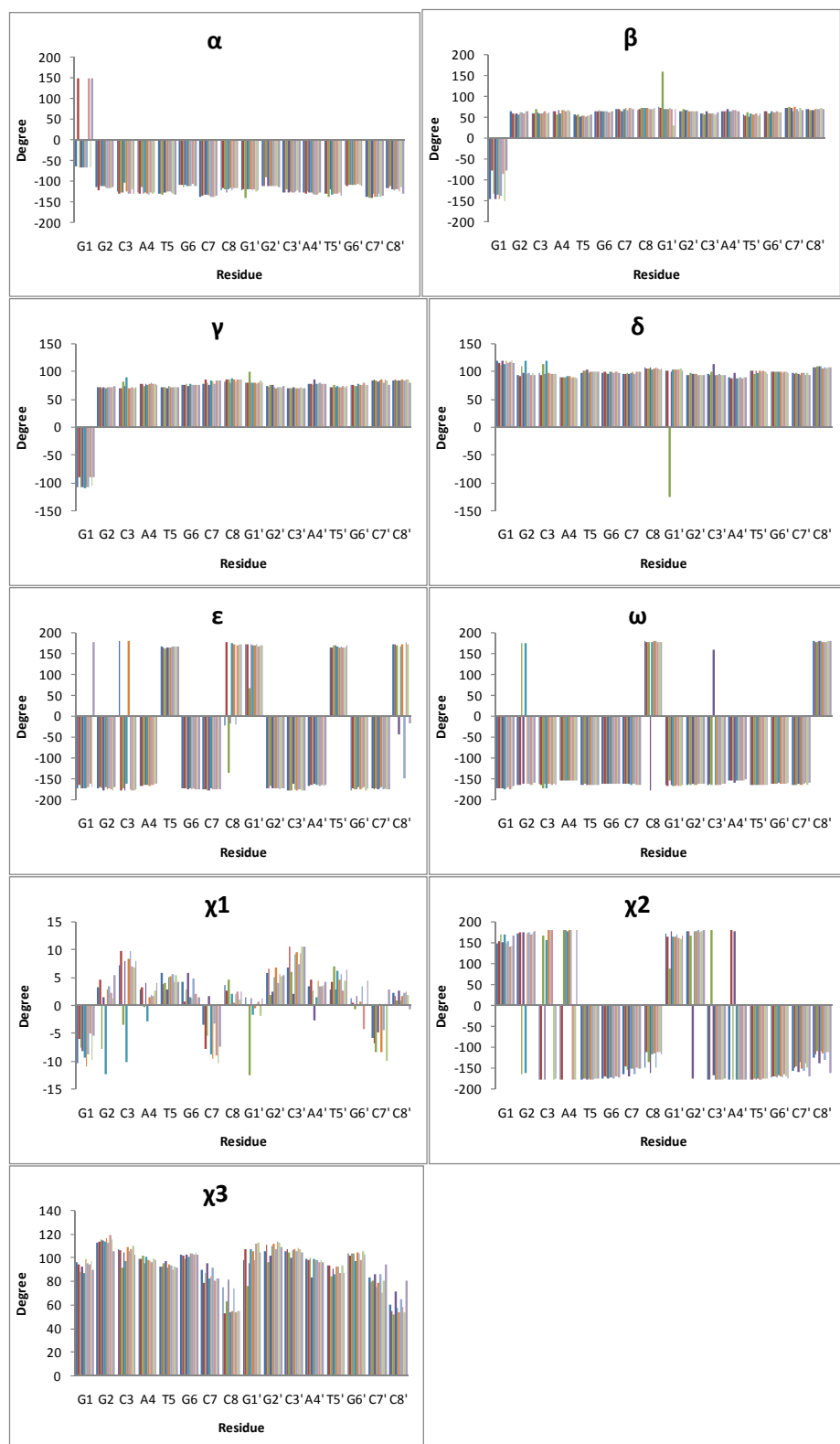


Figure S5. Torsion angles distributions of the ten time-averaged γ - PNA structures.

References

- (1) Case, D. A.; T. A. Darden; Cheatham, T. E. III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; B. Wang; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; K.F. Wong; Paesani, F.; Vanicek, J.; X. Wu; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. **2008**.
- (2) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Structure, Function and Bioinformatics* **2006**, *65*, 712-725.
- (3) Gaussian 03, Revision E.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, **2004**.
- (4) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Jr., K. M. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc* **1995**, *117*, 5179-5197.
- (5) He, W.; Hatcher, E.; Balaeff, A.; Beratan, D. N.; Gil, R. R.; Madrid, M.; Achim, C. *J. Am. Chem. Soc* **2008**, *130*, 13264-13273.
- (6) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089-10092.
- (7) Lavery, R.; Sklenar, H. *Curves 5.1, Computer Program; Institut de Biologie Physico-Chimique: CNRS*, **1996**.