Conformational Analysis of Siloxane-Based Enzyme-Mimic Precursors[1]

Carol A. Parish, Martel Zeldin, Sean Hilson, Jennifer Pratt*

Department of Chemistry, Hobart and William Smith Colleges, NY 14456, USA
E-mail: parish@hws.edu

Summary: The conformational flexibility of six hybrid organodisiloxane oligomers were studied using the Low Mode-Monte Carlo conformational search method with the MM2* force field and the Generalized Born/Surface Area continuum solvent model for water., These systems have enzyme-like properties as synthetic acyltransferases and contain aminopyridine groups in various states of protonation. An ensemble of low energy structures was generated and used to investigate the dependence of molecular shape and flexibility on protonation, which plays an important role in catalyst solubility and self-association. The results as measured by the number of unique conformations, end-to-end or longest intramolecular distance and radius of gyration of the conformational point cloud indicate that the number of protonated pyridines plays a significant role in the overall molecular shape. A similar study was also carried out on various POSS-substitutive organodisiloxane oligomers.

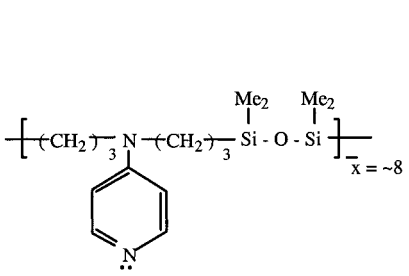
Keywords: acyltransferase; conformational analysis; 4-dialkylaminopyridine; molecular modeling; organosiloxane ladder and cage compounds; siloxane oligomers

Introduction

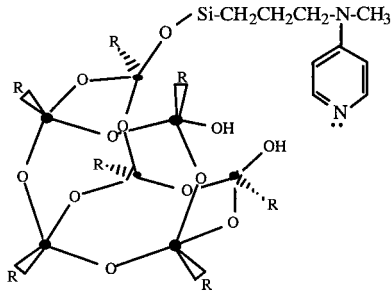
Oligomeric *bis*(trimethylene)aminopyridinedimethyldisiloxanes (**A**) are effective catalysts for the hydrolysis of activated esters of alkanoic acids in aqueous media.[2-4] The design of the oligomeric catalyst took into consideration (1) the supercatalytic behavior of the aminopyridine group, (2) the water solubility of the partially protonated oligomers at pH 8, and (3) the inclusion of the siloxane units in the hydrocarbon backbone to impart orientational flexibility and lipophilicity toward organic substrates. A remarkable feature of **A** is that it exhibits saturation kinetics in the hydrolysis of activated esters (e.g. *p*-nitrophenyl alkanoates) with a maximum effect at C_n=14. Moreover, the reaction kinetics conform to a Michaelis-Menten model for enzyme-like behavior with a turnover number comparable to known natural acyltransferases; e.g., for cholesterol and chymotrypsin, k_{cat}/K_M = 1.25 x 10^6 for n = 6 and 7.6 x 10^4 for n = 7,

 CCC 1022-1360/00/$ 17,50+.50/0

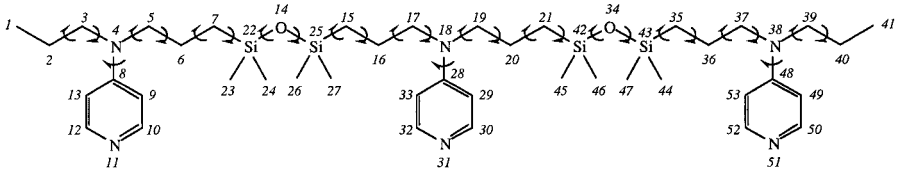
respectively; for **A**, k_{cat}/K_M = 1.1 x 10^4 for n = 14. We have launched a computational study to gain insight into the structural conformations of **A** in polar solvents by examining some model oligomers (i.e. trimers). The atom numbering sequence and the torsional degrees of freedom for the unprotonated (**B**) and all the possible protonated (**C–G**) species that were evaluated computationally are represented in Fig. 1. We have also extended the study to other analogous catalytically active organosiloxanes that have 3D-cage structures comparable to oligomeric silsesquioxanes (**H**).



A



H



B, no H^+; **C**, H^+ on N11; **D**, H^+ on N31; **E**, H^+ on N11 & N31;
F, H^+ on N11 & N 51; **G**, H^+ on N11, N31 & N51

Fig. 1. Model oligomers (trimers) B-G and the atom numbering sequence and torsions used in the conformational calculations.

Methods

The conformational ensembles were calculated using version 7.2 of the MacroModel[5] suite of software programs running on 800 MHz Athlon PCs under the RedHat LINUX 6.2 operating system. The MM2* force field[6] as implemented in MacroModel V7.2 was chosen since it was

the most well parameterized MacroModel force field available for the molecular system under study and it contained siloxane bond lengths, angles and torsional barriers that were in good agreement with previously published results.[7] Solvent effects were included using the Generalized Born/Surface Area (GB/SA) continuum model for water.[8] The Low Mode (LM) search method[9] was used in a 1:1 combination with the Monte Carlo (MC) search method[10] to explore the potential energy surfaces of the model systems. Torsions that were varied during the searches are shown Figure 1. Ensembles generated with each conformational search method were grouped into geometrically similar families using the Xcluster program.[11]

Results and Discussion

Trimer **B** contains three aminopyridinium moieties connected through trimethylenedisiloxane units. In addition to the unprotonated **B**, the trimer can also exist as five uniquely different protonated molecules; viz. two monoprotonated (**C**, **D**), two diprotonated (**E**, **F**) and one triprotonated (**G**) species. The conformational space for each of these trimers in water was exhaustively examined. An analysis of the enthalpic and solvation energies, along with distance parameters of the lowest energy structure of each species (**B'-G'**) is summarized in Table 1. The size and shape of the lowest energy structures are dependent on the number of pyridinium moieties and their distribution in the oligomer. For example, **B'** has one of the more compressed low energy structures as indicated by its longest molecular dimension (13.20 Å), which is measured as the N11–N51 distance. In addition, the hydrocarbon backbone bonds in **B'** are all *trans* and the orientation about the siloxane bonds Si(22)–O(14), O(14)–Si(25), Si(42)–O(34) and O(34)–Si(43) is θ=53 (gauche$^+$), θ=–128 (2*gauche$^-$), θ=163 (*trans*) and θ=–177 (*trans*), respectively. Structure **D'**, with one protonated pyridine in the middle of the oligomer backbone, is even more compressed than **B'** with a longest molecular dimension (10.71 Å) also corresponding to a distance between atoms N11–N51. It is noteworthy that one of the hydrocarbon backbone torsions in **D'** is twisted away from the *trans* orientation producing a "kink' in the chain that allows the protonated pyridinium group to extend away from the rest of the molecule. The driving force for the structure of **D'** is a desire to extend the protonated moiety as far into the solvent medium as possible. This causes the rest of the molecule to collapse on itself leading to a very compressed structure.

Table 1. Energetic and structural comparison of the lowest energy structures (B'-G') found using the LM:MC 50:50 method.

Structure	Lowest Energy (kJ/mol)	Solvation Energy of Lowest Energy Structure (kJ/mol)	Longest Molecular Distance (Å)	End-to-End Distance (C1 – C41) (Å)
B'	98.5	-11.9	13.2 *(N11-N51)*	9.7
C'	-176.3	-302.2	15.3 *(N11-N51)*	11.2
D'	-177.3	-330.2	10.7 *(N11-N51)*	10.6
E'	-454.1	-726.5	19.5 *(N11-N31)*	9.2
F'	-470.8	-730.6	20.0 *(N11-N51)*	11.1
G'	-745.5	-1266.1	19.9 *(N11-N51)*	10.7

The driving force for the structure of **B'** appears to be an attractive intramolecular π-stacking interaction between two adjacent pyridines. The only other minimum energy structure that displays a π-stacking interaction is **C'** where the siloxane bonds sample the *trans*, gauche$^+$ and gauche$^-$ orientations, and all hydrocarbon backbone atoms are in the *trans* orientation but one. The one non-*trans* hydrocarbon orientation produces a kink in the backbone that allows the chain to wrap back on itself to assume less than a 4 Å interaction between the aromatic rings. Structures **E'** and **F'** with two and **G'** with three protonated sites are relatively rigid, extended structures with longest molecular distances of 19 Å or more. In these molecules, the longest molecular dimension corresponds to the distance between positively charged nitrogen atoms. This causes the chain to fold back on itself and results in a shorter end–to–end distance (viz., 9.2-11.1 Å). The lowest energy structures of **E'-G'** adopt conformations in which the positively charged pyridine nitrogens are as far apart from each other as possible, presumably driven by both intramolecular electrostatic repulsions and solvation. This is achieved by forming kinks in the hydrocarbon backbone.

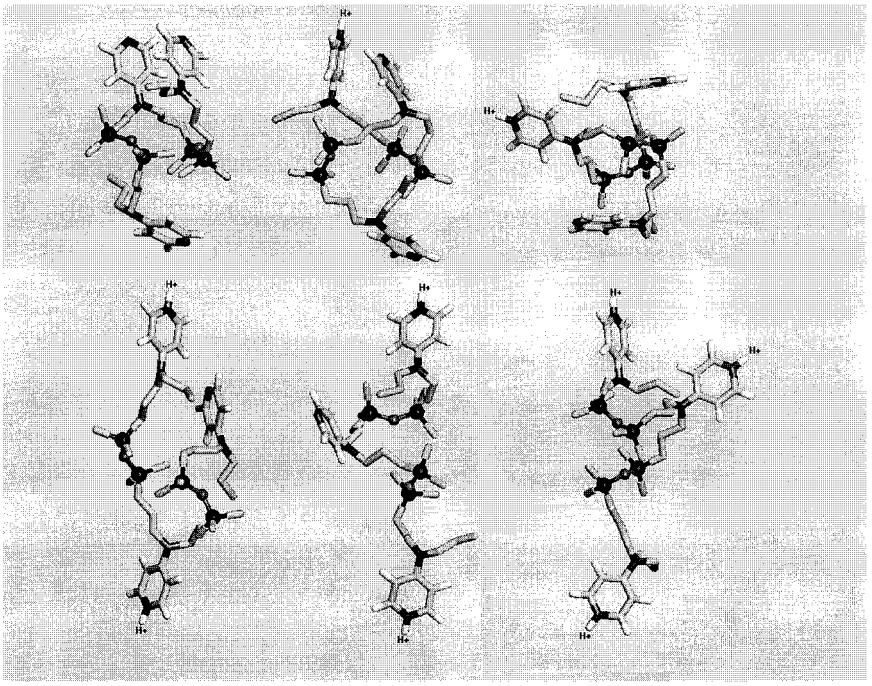


Fig. 2. Lowest energy structures, B'-G' (in sequence from upper left to lower right).

A detailed energetic comparison of isomeric pairs **C'/D'** and **E'/F'** is shown in Table 2. For example, the energies of **C'** and **D'** are similar. The stabilization of **D'** arises mainly from the more negative solvation energy (Δ=28 kJ/mol), which is determined from the polarization energy by the Generalized Born (GB) equation and the solvent accessible surface area (SASA). The SASA of **D'** is smaller than **C'** by ~28 $Å^2$ in agreement with length analysis, and indicates that **D'** is more compressed. The polarization energy for **D'** is larger than **C'** by about 26 kJ/mol, and larger than the SASA contribution. This leads to a significantly larger stabilization for **D'** by solvation. Thus, the molecule having a protonated pyridinium moiety that is sandwiched between the two-unprotonated pyridines is slightly more stable than having the pyridinium ion at a terminal position.

A comparison of the energies of the low energy structures **E'** and **F'** indicates that their isomeric stabilities differ by ~17 kJ/mol, whereas the solvation energies are comparable.

Table 2. Comparison of energetics and distances for the lowest energy isomeric structures C'-F'.

Structure	E (kJ/mol)	Solvation E (kJ/mol)	GB Polarization (kJ/mol)	SASA (Å^2)	Longest Distance (Å)	End–to–End Distance (Å)
C'	-176.3	-302.2	-332.7	876.4	15.3 *(N11 – N51)*	11.2
D'	-177.3	-330.2	-359.4	848.5	10.7 *(N11 – N51)*	10.6
E'	-454.1	-726.5	-759.7	954.2	19.5 *(N11 – N31)*	9.2
F'	-470.8	-730.6	-763.4	973.7	20.0 *(N11 – N51)*	11.1

The SASA and GB polarization energies are similar and lead us to conclude that the isomer having the pyridinium groups on either end is an enthalphically favored configuration.

It is informative to examine the ensembles of all low-lying structures that are found during the conformational search since the ensemble behavior is representative of the true nature of the molecule. The number of unique low energy minima for each system is shown in Table 3. After 55 000 steps, the number of new conformations found within 25 kJ/mol of the lowest energy structure represented less than a 10% increase in the number of structures on average.

Table 3. Comparison of B-G ensembles (55 000 search steps) using the LM-MC 50:50 method.

Structure	Number of Conformations Found	Weighted Average End-to-End Distance (C1 – C41) (Å)	Weighted Average N11 – N51 Distance (Å)	Radius of Gyration, R_g (Å)
	8460	9.0	10.1	3.0
C	9171	9.6	12.8	2.7
D	9198	7.1	8.6	2.3
E	7540	8.1	14.9*	2.8
F	1526	10.4	19.3	1.7
G	1401	10.4	19.4	1.9

* The N11–N31 distance is used for ensemble **E**.

An analysis of this ensemble data yields information about molecular flexibility and conformationally accessible states. Structures **B-E** are much more flexible than either **F** or **G** as evidenced by the number of unique conformations found on the respective potential energy surfaces (Table 3). As more positive charge is introduced into the molecule, the rigidity of the molecule increases; e.g., **E** has the same number of positively charged nitrogens as **F**, however, **E** has two charges on adjacent pyridines. Thus, there is an uncharged "tail" on **E**, which samples the locally available conformational space and leads to greater conformational flexibility. The introduction of positive charge at the termini restricts flexibility by anchoring the charged ends into the solvent medium.

The rotational orientation of the siloxane bonds in each molecule in each ensemble was examined. The torsional bonds favor the *trans*, gauche$^+$ and gauche$^-$ orientation. The rotational orientation of the hydrocarbon backbone atoms was also examined. Unlike the siloxane bonds the hydrocarbon torsions adopt almost exclusively *trans* conformations; in those instances where the bond was other than *trans*, a kink was introduced and readily apparent in the corresponding molecular structure. The XCluster program was used to determine if the ensembles naturally form structurally related groupings and if the groupings are the same or different for each ensemble. Clustering by atomic and backbone torsional RMS after rigid body superposition of all heavy atoms did not lead to strong clustering in any of the ensembles as evidenced by distance maps, mosaics and separation ratios. This provides further evidence of conformational flexibility.

The ensembles were examined for molecular size. The C(1)–C(41) distance was measured for each structure in each ensemble to determine the change in intramolecular distance from molecule to molecule. Ensembles for **B** and **C** contain a similar number of unique structures (8460 and 9171, respectively). Their end-to-end distance histograms are also quite similar. The distances range from 3.6 Å to 17.0 Å with a weighted average of 9.0 Å and 9.6 Å, respectively. The end–to–end distance data for **C** indicates that the structure behaves more like **B** than like **D**. The molecular ensemble for **D** displays the shortest weighted average end–to–end distance of 7.07 Å and suggests that the two unprotonated terminal pyridines are very flexible and allow the molecule to adopt a more compressed conformation. **E** displays significantly different end–to–end distances than **F**. The weighted average distance for **E** is 8.14 Å, whereas the average for **F** is 10.4 Å. **F** is clearly a more rigid system as evidenced by the significantly smaller number of

unique conformations found; 7540 for **E** and 1526 for **F**. The flexibility of **E** versus **F** is also confirmed by the distribution of end–to-end distances; 3.5 Å to 14.9 Å for **E** and is 6.1 Å to 16.5 Å for **F**.

G is clearly a rigid system; i.e. 1401 unique structures with end-to-end distances ranging from 6.8 Å to 17.5 Å and a weighted average of 10.4 Å. The low energy structures for **G** are mainly extended with large end–to–end distances. It is apparent that the number of protonated pyridines controls the overall molecular shape and flexibility. To illustrate, Figure 3 superimposes the end–to–end distance data for ensembles **B** (gray tone) and **G** (black).

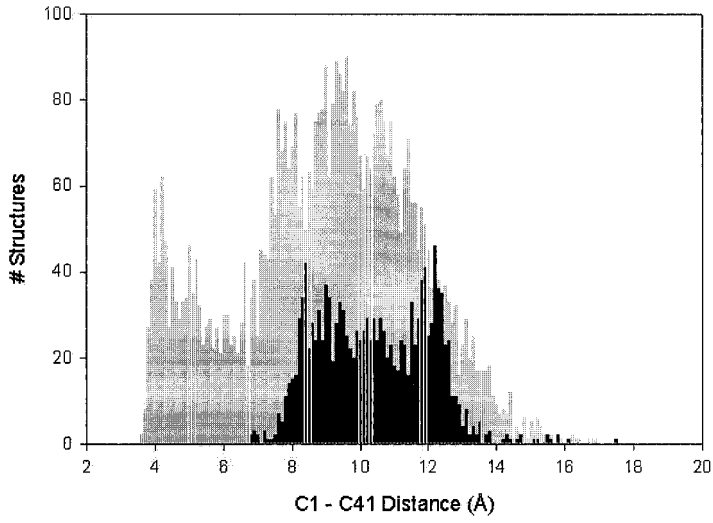


Fig. 3. Comparison of the end-to-end distance for the ensembles B (gray) and G (black).

The longest molecular distance in each lowest energy structure was also determined (Table 1). The longest molecular distance in all structures is between N11 and N51 except for **E'** where the longest distance is between N11 and N31. It is interesting to note that these lengths correspond to the space between pyridines or pyridinium nitrogens at either end of the trimeric

chain except for **E'** where the distance corresponds to the length between the two charged pyridinium nitrogens, one in the middle and one at the end of the chain. The longest dimension in **E'-G'** is noticeably larger than in **B'-D'**. The difference between **E'-G'** and **B'-D'** is attributed to a repulsive interaction between positively charged pyridinium ions. If we examine the ensemble behavior of these distances, it is observed that **F'** and **G'**, with two pyridinium ions anchored at either end or three pyridinium ions, display markedly different length data than **B'-E'**. For example, the N11–N51 distances in **B'-D'** have a range between ~3 Å to ~17 Å with weighted averages of 10.1 Å, 12.8 Å and 8.6 Å, respectively. The N11–N31 distance in **E'** ranged from 6.7 Å to 19.9 Å with a weighted average distance of 14.9 Å. Yet the longest distances in **F'** and **G'** ranged from 12.3 Å to 20.9 Å and from 10.6 Å to 21.0 Å with averages of 19.3 Å and 19.4 Å, respectively.

There is no guarantee that either the end-to-end or longest molecular distances in the lowest energy structures are representative of the shape of the structures in the corresponding ensembles. Therefore, we used the XCluster program to compute the radius of gyration, R_g, of the conformational point cloud for all structures in each ensemble. The structures in each ensemble are subjected to an optimal atomic RMS superimposition and from this an ensemble centroid is determined for use in calculating the radius of gyration. Used in this sense, R_g is a practical measure of the conformational diversity of each ensemble; i.e., ensembles with smaller R_g values contain conformations that are very similar whereas ensembles with larger R_g values contain structurally more diverse conformations. This provides an approximate measure of the molecular shapes involved in each ensemble without biasing the measurement to any one intramolecular distance criterion. The overall ensemble radius of gyration for **B'-G'** is summarized in Table 3. The results indicate that the distance criteria are representative of the overall ensemble behavior.

Acknowledgment

This project was partially supported by NSF grant CHE-0116435 as part of the MERCURY supercomputer consortium, the Donors of the Petroleum Research Fund, administered by the American Chemical Society, and by a grant from the Merck Foundation administered by the AAAS.

[1] This study has been published in part: C. A. Parish, M. Zeldin and J. Pratt, *J. Inorg.Organometal. Polym.,* **12**, 31 (2002).
[2] S. Rubinsztajn; M. Zeldin; W. K. Fife, *Macromolecules* **23,** 4026 (1990).
[3] W. K. Fife; S. Rubinsztajn; M. Zeldin, *J. Am. Chem. Soc.* **113,** 8535 (1991).
[4] S. Rubinsztajn; M. Zeldin; W. K. Fife, *Macromolecules* **24,** 2682 (1991) and refs.therein.
[5] F. Mohamadi; N. G. J. Richards; W. C. Guida; R. Liskamp; M. Lipton; C. Caufield; G. Chang; T. Hendrickson; W. C. Still, *J. Comput. Chem.* **11,** 440 (1990).
[6] N. L. Allinger, *J. Am. Chem. Soc.* **99,** 8127 (1977).
[7] I. Bahar; I. Zuniga; R. Dodge; W. L. Mattice, *Macromolecules* **24,** 2986 (1991), S. Grigoras; T. H. Lane In *Advances in chemistry series 224: Silicon-based polymer science a comprehensive resource*; J. M. Zeigler, F. W. G. Fearon, Eds.; American Chemical Society: Washington, D. C., 1990; pp 127, S. Grigoras; T. H. Lane, *J. Comput. Chem.* **9,** 25 (1988).
[8] J. Weiser; P. S. Shenkin; W. C. Still, *J. Comput. Chem.* **20,** 586 (1999).
[9] I. Kolossvary; W. C. Guida, *J. Comput. Chem.* **20,** 1671 (1999).
[10] G. Chang; W. C. Guida; W. C. Still, *J. Am. Chem. Soc.* **111,** 4379 (1989).
[11] P. S. Shenkin; D. Q. Mcdonald, *J. Comput. Chem.* **15,** 899 (1994).