

1. Introduction

1.1 Surfactants

A surfactant molecule is composed of a hydrophilic head and a hydrophobic tail (Figure 1). The head can be an anionic, a cationic, a zwitterionic, or a nonionic group,

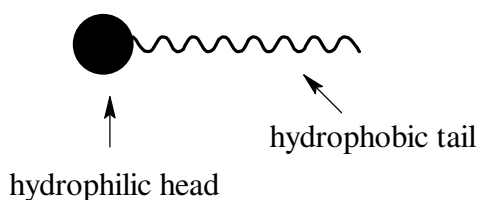


Figure 1. Schematic of a surfactant molecule.

while the tail is a nonpolar hydrocarbon chain. These molecules are surface active as a result of the combination of the hydrophobic and hydrophilic properties. They have affinity for two immiscible phases, one of which is water.

A zwitterionic surfactant (Figure 2) has both a positive and negative charge in the head

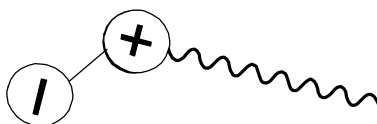


Figure 2. Schematic of a zwitterionic surfactant.

group. This charge separation causes a dipole moment to be present in the head group. Therefore, the head group is polar and soluble in water.

The zwitterionic surfactants have chemical structures similar to those surfactants synthesized by the Gandour group (Figure 3).¹ The specific zwitterionic surfactants that are

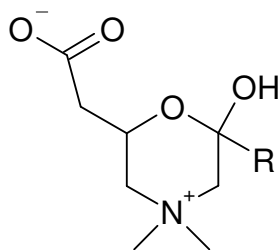
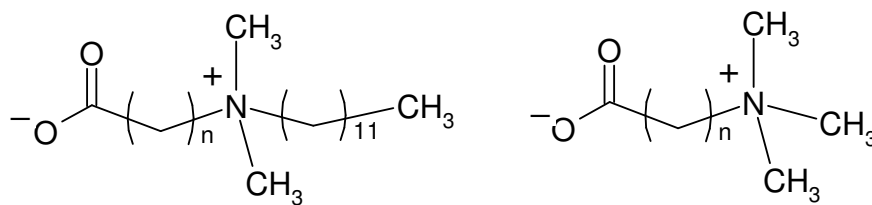


Figure 3. Z-14. A zwitterionic surfactant synthesized by the Gandour group.

being studied here are a series of molecules called trimethylammoniocarboxylates, commonly known as betaines. Betaines are a simpler zwitterionic surfactant model than the Z-14.¹ The betaines, Figure 4, have a carboxylate as a negative charge and a quaternary ammonium nitrogen as a positive charge. This characteristic yields a large dipole



4(a) n-12 Betaine surfactant.

4(b) Betaine surfactant head group.

Figure 4(a,b). Zwitterionic betaine surfactants.

moment, which makes the zwitterion soluble in solutions over a wide range of pH's, salinities, and concentrations. Betaines are also assumed to be less irritating than other surfactants.²

The Gandour group synthesizes surfactants,^{3,4} also known as amphiphiles, to be used as topical microbicides in biological testing to determine whether they will be effective against sexually transmitted diseases without being irritating to the mucosal tissue. These amphiphiles can inactivate pathogens (or spermatozoa) by various mechanisms:⁵ (1) disrupting cell walls and solubilizing membranes as micelles (see below), (2) changing fluidity and porosity of membranes, (3) altering signal pathways, (4) changing the flow of constituents and nutrients between a cell and a medium, and (5) inhibiting specific enzymes, all as monomers, and finally (6) coating non-envelope viruses to prevent infection as monomers or micelles. The mechanisms (1) and (2) are rapid and have a broad spectrum of activity, but are likely to be irritating and damage human tissue. Solubilization of the mucosal tissue is what leads to the vaginal irritation.

Therefore, an amphiphile needs to be designed such that it kills pathogens and immobilizes sperm, yet does not solubilize the vaginal membrane. If a surfactant can be designed and the properties of a surfactant calculated theoretically, the surfactant can be accepted or rejected as a possible amphiphile before a chemist attempts to synthesize it.

To test the idea that computational chemistry can aid in the design, this project calculates the structures and solvation energies of betaine molecules (Figure 4(b)) where n varies from one to six. These calculations model the hydrophilic part of the surfactant. The hydrocarbon tail is only a methyl, which is not considered a hydrocarbon chain, and so the resulting structure is the head of a zwitterionic surfactant. The solvation results are compared to experimentally determined cmc data obtained from a series of betaines (Figure 4(a)) with a long hydrocarbon chain with n also varying from one to six.⁶ The reasoning for the hydrophilic interest will be explained later.

If a relationship can be made between the solvation energy and the cmc , then there exists a method to predict cmc 's of a surfactant before entering a laboratory for synthesis. Synthesis is time consuming and possibly expensive, depending on the chemicals needed. Amphiphiles with biological activity have been found to exhibit cmc values within a certain range. If cmc 's can be predicted by calculations, then we can eliminate unnecessary synthesis time, money, and effort. Biologically active amphiphiles with computationally predicted cmc 's within the desired range can be targeted.

Before proceeding with results, a review micelles and the prediction of the critical micelle concentration (cmc) is in order. In addition, a review of experiments where end-to-end distances are measured for the betaines in Figure 4(b) is presented. These results provide another benchmark to compare with my computational results. Finally, a review of some computational background of solvation energy methods is presented.

1.2 Micelles

Surfactant molecules aggregate to form micelles. In water, the hydrophobic tails associate, leaving the hydrophilic head groups exposed to water. This structure is known as the “normal” micelle. A micelle composed of zwitterionic surfactants⁷ is shown in Figure 5.

Micelles can contain anywhere from tens to thousands of monomers. The shape and size of a micelle can vary; smaller micelles are spheres, while the larger ones are cylinders, discoidal aggregates, or infinite bilayers.⁸⁻¹⁰

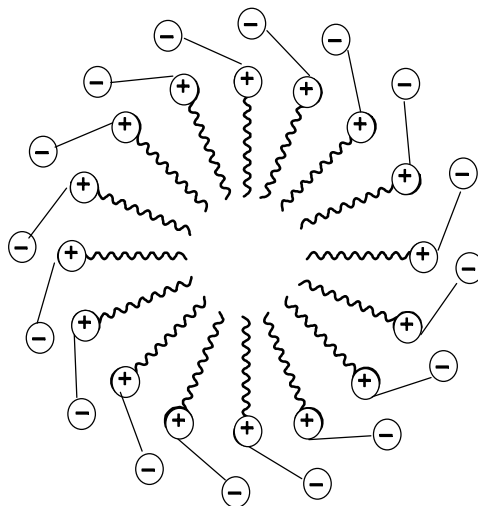


Figure 5. Micelle comprised of zwitterionic surfactants.

The thermodynamic principles behind micelle formation are relatively straightforward. A molecule with a long hydrocarbon chain attached to a polar group has a measurable solubility in water as a result of its polar head group. This is known as the hydrophobic effect. These surfactant molecules migrate to the surface of water. This hydrophobic effect causes the aggregation to occur, and the repulsion among the head groups limits the size that the micelle can achieve. Furthermore, varying solution conditions such as overall surfactant concentration, surfactant composition, temperature, pressure, ionic strength, and pH can produce morphological changes.⁹⁻¹¹ Once the micelle has formed, it goes into solution and is no longer on the surface of the water.

The concentration of surfactant at which micelles first appear in a solution is called the critical micelle concentration (*cmc*). Below the *cmc*, increasing surfactant concentration increases the monomer concentration. Above the *cmc*, increasing the total surfactant concentration does not change the concentration of monomer. The transition from pre-micellar to micellar solutions is revealed in a change in the slope of plots of surfactant concentration versus a physical property, such as viscosity, conductance, surface tension, osmotic pressure, or molar conductivity.¹²⁻¹⁵

The *cmc* is the most useful quantity in characterizing surfactants due to the sharp changes of the physical properties in the solution. As a result, the *cmc* is a central element in studying the physical properties of surfactants and in researching several industrial and biological uses of surfactants. Several possible effects influence the *cmc*; both attractive and repulsive interactions of the hydrocarbon section of the molecule with water, attractive interactions between hydrocarbon tails on different molecules, solvation of the hydrophilic head group by water, interactions between solvated head groups, and geometric and packing constraints resulting from the specific molecular structure involved.¹⁶ Therefore, the *cmc* can reveal much about the surface activity of a surfactant molecule. The surface activity is the adsorption of the surfactants at the hydrocarbon/aqueous interface in the form of a monolayer.¹² Surfactant molecules must aggregate into micelles to perform the processes required for the various functions of surfactants, such as the solubilization of otherwise insoluble organic material.

1.3 Predicting the *cmc*

1.3.1 Background. Qualitatively, both the hydrophilic head group and the length of an unbranched hydrophobic tail contribute to the determination of the *cmc*. Increasing the volume, polarity, dipole moment, and polar surface area of the hydrophilic head group, increases the *cmc*; however, increasing the hydrophobic part of the surfactant, decreases the *cmc*.

Quantitatively, measuring the *cmc* for a variety of surfactants in different solvents has often been performed within a range of a few percent uncertainties. From experimental data, empirical correlations between the structural characteristics of surfactants and the *cmc* have been developed for nonionic, cationic, and anionic surfactants.¹⁷

The *cmc* is predicted from both the hydrophilicity of the head group and the hydrophobicity of the tail

$$\log (cmc) = a(\text{hydrophilic descriptor}) + b(\text{hydrophobic descriptor}) \quad (1)$$

$$+ c(\text{hydrophilic descriptor} \times \text{hydrophobic descriptor}) + \text{constant}$$

where a, b, and c are constants and (hydrophilic descriptor x hydrophobic descriptor) is a cross-term linking the two effects. If $c(\text{hydrophilic descriptor} \times \text{hydrophobic descriptor}) = 0$, the hydrophilic and hydrophobic effects are independent of each other and can be treated separately.

However, if $c(\text{hydrophilic descriptor} \times \text{hydrophobic descriptor}) \neq 0$, the two effects are not independent and cannot be treated separately.

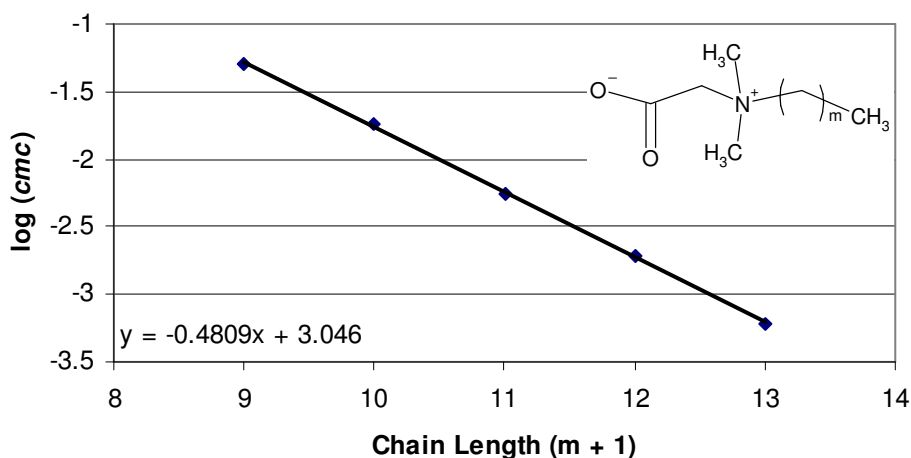


Figure 6. Graph of $\log(\text{cmc})$ vs chain length.

The hydrophobic term has already been investigated.^{4, 18-20} The relationship between the length of the tail and the cmc for betaines can be seen in Figure 6.²⁰ The $\log \text{cmc}$ vs chain length of the tail was measured for betaines with the $n = 1$ head group Figure 4. The hydrocarbon tail varies in length based on the graph.

According to the Klevens equation,^{18, 19} the cmc can generally be expressed empirically as a function of the number of carbon atoms ($n = m + 1$) in the hydrophobic chain for an ionic surfactant

$$\log(\text{cmc}) = A - Bn \quad (2)$$

where A depends on the ionic strength of the system and type of hydrophilic head group and B is determined by the number of hydrophobic chains and hydrophilic head groups. Values of A and B have been determined for some common surfactant classes. Figure 6 is an example of such a plot where $\log \text{cmc}$ is graphed versus chain length, and n is the number of methylene groups in the tail. In this example, $A = 3.046$ and $B = -0.4809$. According to this graph, the structure where $n = 1$ is the same as the first betaine in the series that was calculated for this project (Figure 4(b)). Therefore, according to the equation, the $\log \text{cmc}$ is 2.5651 when $n = 1$. The cmc is then 3.50×10^2 , or 350 M. This value can be directly related to the solvation energy that was calculated for this betaine.

This empirical relationship suggests that a relationship can be developed where the properties of the surfactants, such as the ionization degree, can be predicted from the structure of a homologous series of surfactant molecules.^{19, 21} The surfactant molecule varies in the surface area of the head group²¹ and the chain length.^{19, 21}

The hydrophilic contributions to the *cmc* have been examined.^{18, 19, 22} The *cmc*'s are not affected by the hydrophilic head group as much as the hydrophobic tail. In many cases the *cmc*'s were identical for molecules with the same tail, but differing head groups,^{18, 19} but not zwitterionic head groups.

1.3.2 An Approach to Zwitterionic Head Group Hydrophilicity. Because the tail can be accounted for with the Klevens' equation, the zwitterionic head group needs to be examined. This is why the R group is a methyl and the betaines being studied are essentially a simple model of a series of polar zwitterionic surfactant head groups.

The hydrophilicity of surfactant head groups has been previously examined. This effect can be accounted for based on conductivity,¹⁸ dipole moment,^{17, 22, 23} relative number of nitrogen and oxygen atoms,²⁴ total surfactant area,²³ polarity,²⁵ and charged partial surface areas.²⁵

The proposal for this project is to test if the measure of stability of a homologous series of betaine molecules in water (solvation energy) will mirror the trend in micelle formation (*cmc*) for a similar series of surfactants. The model of solvation energy to the monomer-micelle equilibrium is illustrated in Figure 7. It is hypothesized that the solvation energies of the monomer structures in this series differ due to conformational changes in the different homologues. However, the solvation energy of the micelle is projected to remain constant because of the similar (spherical) shape of the micelle structure. Therefore, the hypothesis is that the differences in the *cmc* of the homologous series of betaines are reflected in the solvation energy of the monomer. It is presumed that the higher the solvation energy, the more difficult it is for the micelle to form. This means the higher the solvation energy, the higher the *cmc* for each corresponding surfactant.

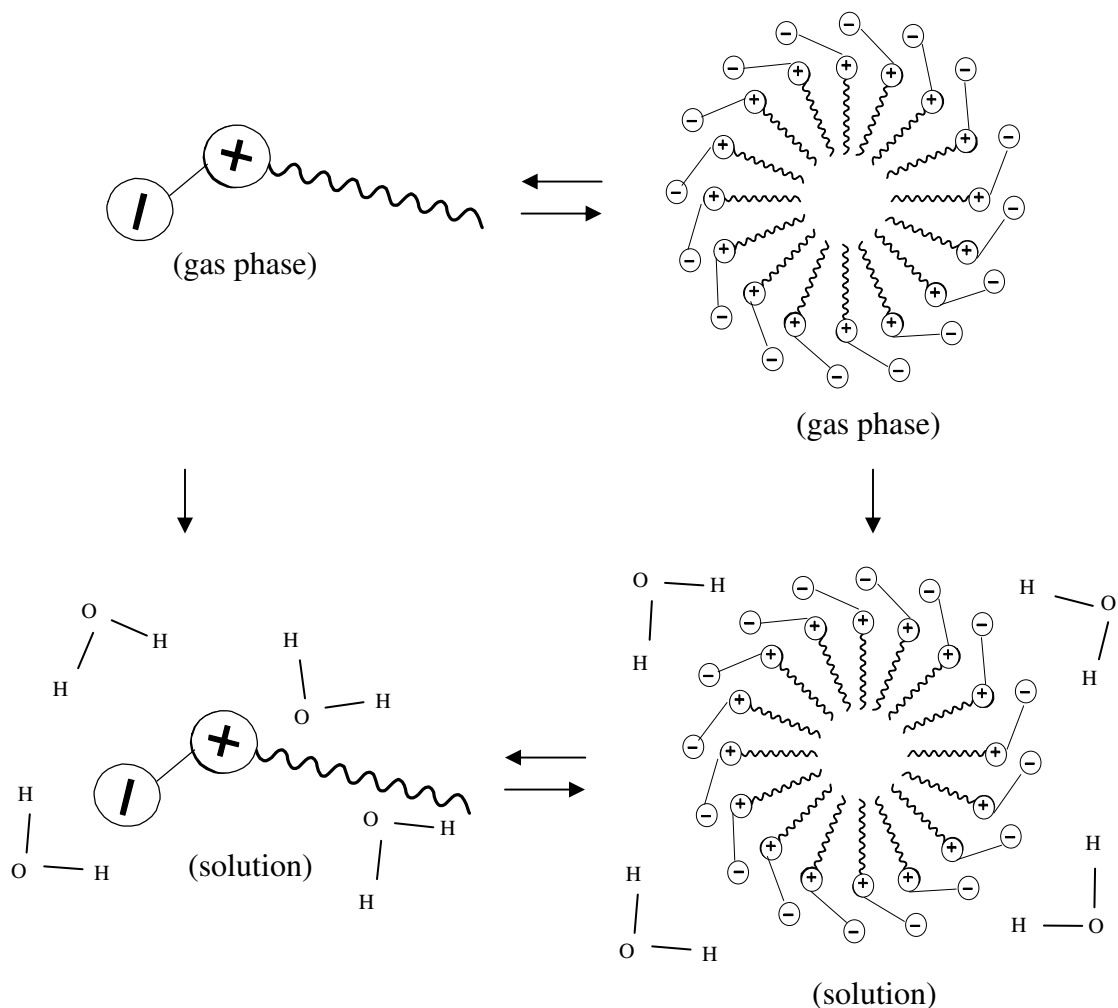


Figure 7. Model for solvation energy of monomer-micelle equilibria.

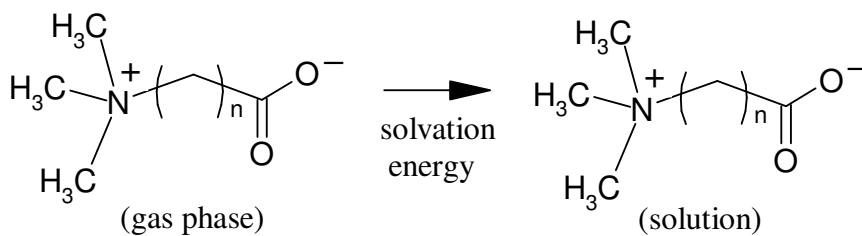


Figure 8. Illustration of project.

This research predicts the solvation energy of a series of betaine molecules in water (Figure 8); solvation energy in turn predicts the trend in hydrophilicity of these molecules. These hydrophilicity effects are a result of the properties of the head group. Therefore, if these

computational data can reproduce the trend in *cmc* values for experimental data (Figure 9),⁶ then we can say that *c* from equation 1 is approximately equal to zero. Then the solvation energy contribution from the head group can be combined with the slope (B) of the Klevens equation, which gives the hydrophobicity contribution from the tail, to predict the *cmc* of a surfactant. In other words, the solvation energy is directly proportional to A in the Kleven's equation.

The literature *cmc* values graphed in Figure 9 were determined from n-12 betaines,⁶ Figure 4(a), in water. The data are graphed such that the *cmc* is plotted versus n, the number of methylenes in the head group. In this series of molecules, the head group spans the quaternary ammonium nitrogen to the oxygens of the carboxylate, all molecules have the same hydrocarbon tail containing 12 carbons.

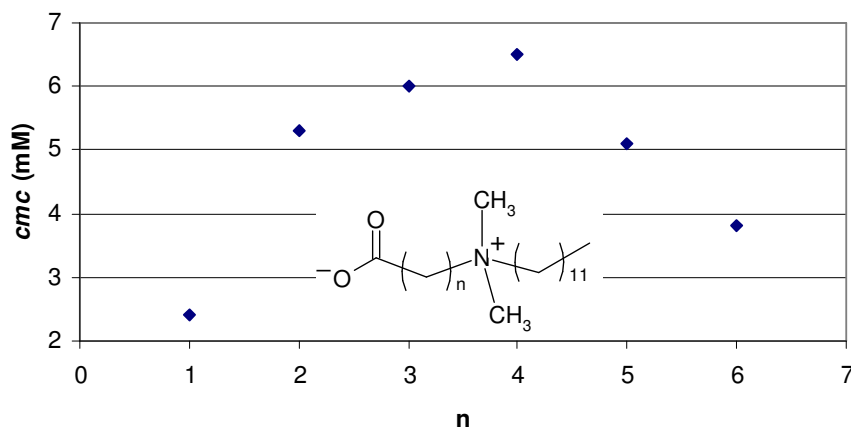


Figure 9. Graph of experimental *cmc* data versus number of methylene groups.⁶

1.4 Intercharge Distances

The hydrophilicity of the headgroup of a surfactant molecule increases as the distance between the two charged areas increase. This distance increase in turn increases the dipole moment of the molecule. However, to increase the distance, additional methylene groups must be added, making the head group more hydrophobic. The distance between the charged centers is not necessarily proportional to the number of methylene groups added. As the chain grows

longer, it is flexible and can assume many conformations. That is, as long as the strain and steric interactions are not too strong.

For a single surfactant, the electrostatic effects suggest the opposite charges would prefer to be as close as possible. This would be true for a surfactant in the gas phase and would result in a cyclic-like structure. This would minimize the charge separation and, therefore, the dipole moment also. However, for a surfactant in water, the larger the charge separation, the larger the dipole moment. When the dipole moment is increased, the molecule becomes more polar. Because water is a polar solvent, the more polar the head group, the more soluble the surfactant. So, in water the charge separation would be expected to be greater than that of a surfactant in the gas phase.

Results from two experimental methods^{26,27} for determining interchange distances were compared to the theoretically calculated interchange distances of the homologous trimethylammoniocarboxylates (Figure 4(b)). The first method²⁶ measures the interchange distance from the nitrogen to the carboxyl carbon (Figure 10(a)). The maximum distance was estimated from crystallographic data on glycine combined with the geometry of polymethylene chains. The overall interchange distance was measured with ¹³C NMR by measuring the chemical shift difference between two homologous compounds; one zwitterion with both its charges and the other only having one charge, keeping all other variables constant.

The second method²⁷ measures the interchange distance as the distance from the nitrogen to a point symmetrically between the oxygen atoms in the same plane and in line with the methylene carboxyl bond. (Figure 10(b)) The maximum distances were determined from known

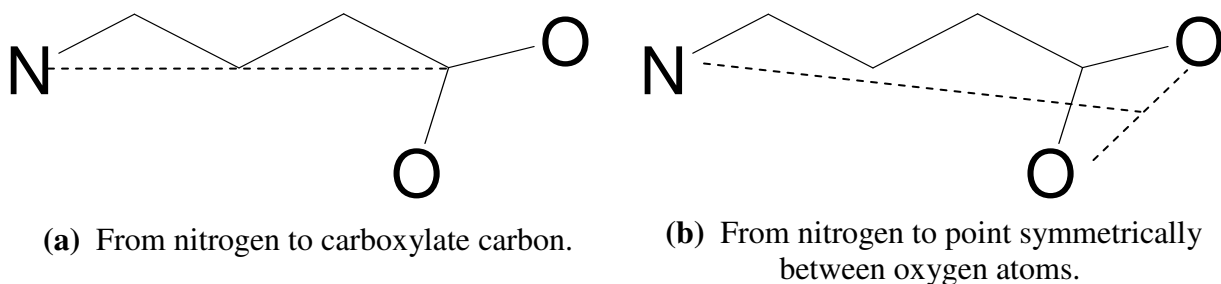
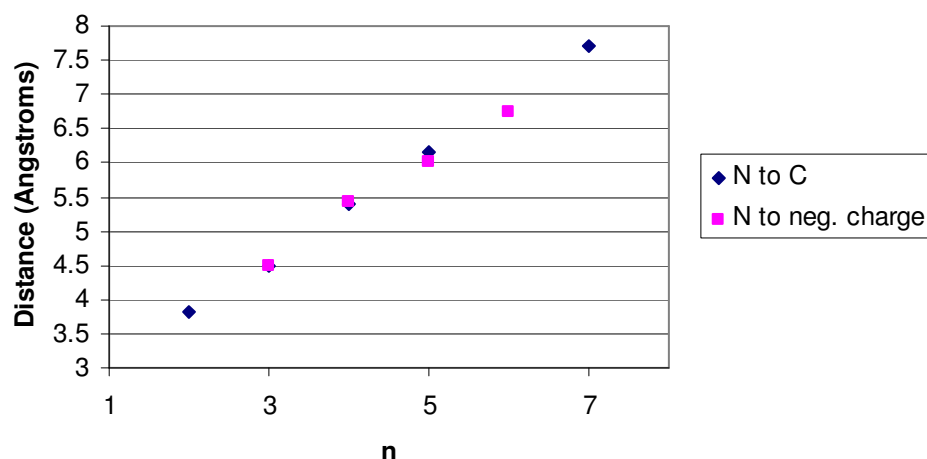


Figure 10(a,b). Interchange distance illustrations.

bond lengths and angles for amino acids and betaines. The overall distances were determined from experimentally measured densities, refractive indices, and dielectric constant values. The results for both methods are in Table 1, and the overall distances are graphed in Figure 11.

Table 1. Results for experimental interchange distances (\AA).^{26, 27}

n	Maximum Distance (N to C) ²⁶	Overall Distance (N to C) ²⁶	Maximum Distance (N to negative charge) ²⁷	Overall Distance (N to negative charge) ²⁷
1	2.48	--	--	--
2	3.81	3.81	--	--
3	4.98	4.50	5.45	4.49
4	6.30	5.40	6.93	5.43
5	7.51	6.15	7.96	6.00
6	--	--	9.38	6.74
7	10.04	7.7	--	--

**Figure 11.** Overall experimental interchange distances.

1.5 Computational Methods

The solvation energies were determined computationally by implementing a program, called HONDOPLUS.^{28, 29} The Hartree-Fock (HF) method and the basis set 6-31G(d), with the solvation model (SM), SM5.42 to calculate the solvation energies were chosen. Charge Model 2 (CM2) was used by the program to determine charges. The computational models were chosen based on the methods available through the HONDOPLUS program, which contains the most robust selection of solvation models. CM2 was the only charge model parametrized to allow optimization in solution with the solvation model. SM5.42 was the only solvation model

available here that would optimize the structure in solution. The program is only parametrized for a limited number of methods, namely HF, density functional theory B3LYP, and the hybrid density functional MPWX. Each method only has a few optional basis sets.

1.5.1 Basis Set. The 6-31G(d) basis set was chosen. Calculations were performed to compare the 6-31G(d) basis set with the 6-31+G(d,p) basis set, the values were noted as very similar. The bond lengths only vary by 0.006Å at most, and the angles vary by 0.436 degrees at most. Also comparing the HF/6-31G(d) level of theory values to the MP2/aug-cc-pVDZ+3s3p2d level of theory and experiment, it can be noted that the values either correspond well with the MP2 values or they are closer to experiment than the higher levels of theory. Other optional basis sets were either not reliable with experiment, would be too time consuming when running the larger molecules, or were not available within the limits of the HONDOPLUS program.

1.5.2 Solvation. One of the most fundamental quantities used to describe the interaction between a solute and a solvent is called the free energy of solvation ($\Delta G_{sol}(A)$). This interaction is sometimes called the free energy of transfer, where molecule *A* is transferred from the gas phase to a condensed phase.³⁰ The free energy is

$$\Delta G_{sol}(A) = \lim_{[A]_{sol} \rightarrow 0} \left\{ -RT \ln \frac{[A]_{sol}}{[A]_{gas}|_{eq}} \right\} \quad (3)$$

where ideal solution behavior is implied.

Due to the energy of interaction between a solute molecule and the solvent, the solute properties dependent on energy also depend on the solvent. These properties include geometry, vibrational frequencies, total energy, and electronic spectrum. The solvation effect is characterized by a change in the molecular behavior as a result of the solvent.

When the solute molecule meets the first solvent shell, cavitation energy and the attractive force between the solute and solvent are noticed (Figure 12). The energy needed to disperse the solvent molecules, thereby making a cavity for the solute molecule, is called the cavitation energy. The forces attracting the solute to the solvent are van der Waals, dispersion,

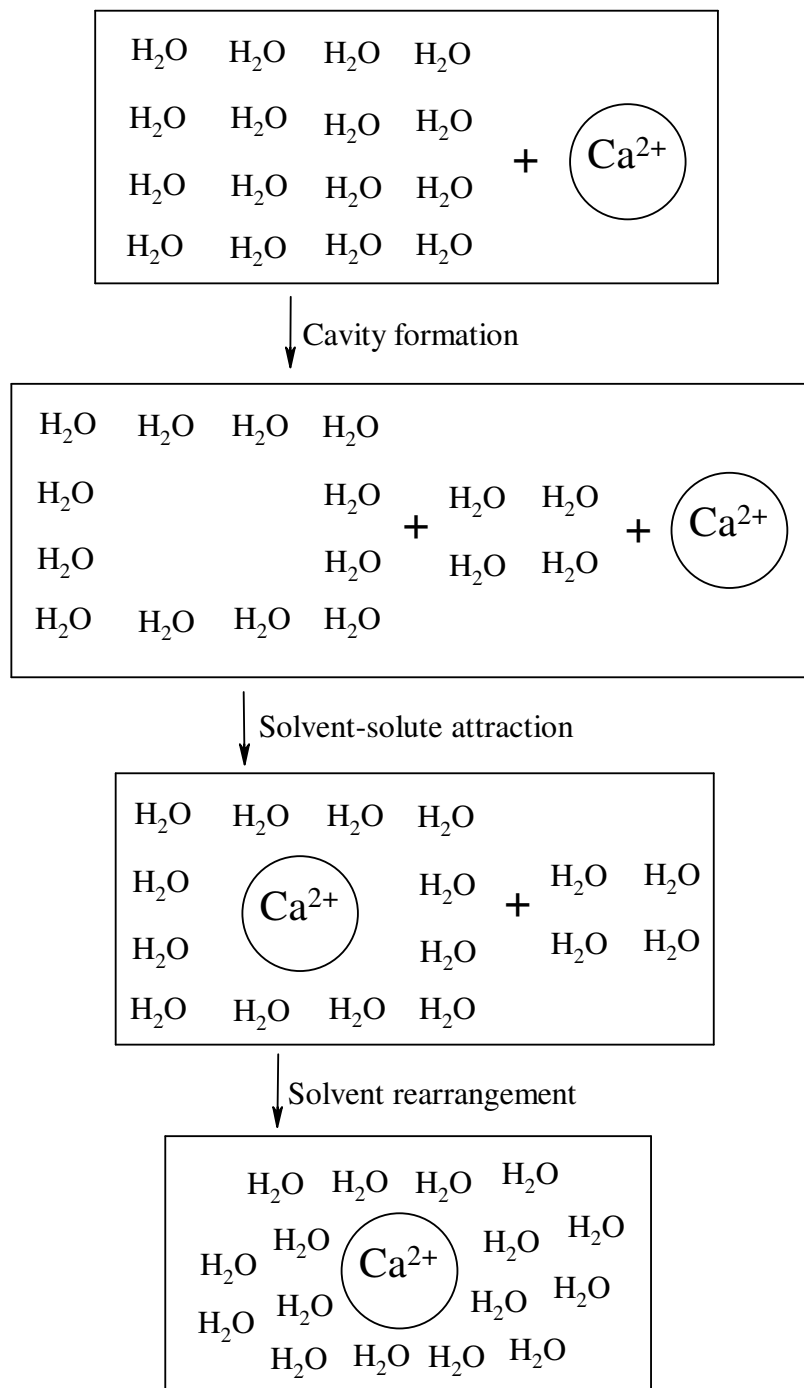


Figure 12. Illustration of first solvation shell and cavitation.

and hydrogen bonding interactions, for example, a proton donating solvent surrounding the carbonyl group of a solute. In addition to the attracting forces, solvent rearranging can occur in the first-shell molecules to maximize solute interactions. Cavitation is most important for non-

hydrogen bonding solvents because there are no hydrogen bonding interactions to aid in forming the cavity.³¹

Ben-Naim's definition³²⁻³⁴ of a solvation process can be easily adopted. Here, at constant chemical composition, pressure, and temperature, a solute M in a solvent S transfers M from a fixed position in the ideal gaseous phase to a fixed position in S . Therefore, ΔG_{sol} can be related to the work $W(M/S)$ needed to build up M in S . $W(M/S)$ can also be referred as the coupling work of M with S . Consequently, the free energy of solvation is

$$\Delta G_{sol} = W(M/S) + RT \ln \left(\frac{q_{rot,g} q_{vib,g}}{q_{rot,s} q_{vib,s}} \right) - RT \ln \left(\frac{n_{M,g} \Lambda_{M,g}^3}{n_{M,s} \Lambda_{M,s}^3} \right) \quad (4)$$

where $q_{rot,g}, q_{vib,g}, q_{rot,s}$, and $q_{vib,s}$ are microscopic partition functions of M for rotation and vibration; $n_{M,g}$ and $n_{M,s}$ are the numerical densities of M molecules; $\Lambda_{M,g}^3$ and $\Lambda_{M,s}^3$ are momentum partition functions; R is the gas constant; and T is temperature.

The cavitation energy is proportional to the surface area of the cavity according to an approach by Uhlig.³⁵ Assuming a spherical cavity

$$\Delta G_{cav} = 4\pi R^2 \gamma \quad (5)$$

where R is the radius of the cavity and γ is the surface tension of the solvent. A formula related to equation 5 was proposed by Tuñón et al.³⁶

$$\Delta G_{cav} = \gamma S_M - RT \ln(1 - V_S n_S) \quad (6)$$

where S_M is the surface area of the cavity, V_S is the volume of a solvent molecule, and n_S is the density of the solvent.

To generalize the cavitation energy and to make the equation useful for arbitrary shapes, the form is adjusted³⁷⁻⁴²

$$\Delta G_{cav} = k_S^g (V_S / V_M) \gamma \Sigma \quad (7)$$

where Σ is the cavity area and $k_S^g (V_S / V_M)$ is a correction factor that depends on the ratio of molecular volumes V of S and M as well as being characteristic of the solvent. The correction factor can be expressed as

$$k_S^g (V_S / V_M) = 1 + \left(\frac{V_S}{V_M} \right)^{2/3} (k_S^g(1) - 1) \quad (8)$$

where $k_S^g(1)$ is estimated from solubility data.

1.5.2.1 CM2 Theory.^{43,44} Atomic charges can be calculated by many methods in theoretical chemistry. The Mulliken population analysis⁴⁵ is the most widely used scheme for atomic charges in theoretical calculations. Here, the charge in a molecular orbital is partitioned such that contributions associated with atomic one-electron basis functions are centered at the nuclei. The Löwdin population analysis⁴⁶ is also a popular scheme. In this scheme, Löwdin's symmetric orthogonalization procedure is used to partition the charges into transformed one-electron basis functions.⁴³ The Mulliken analysis could possibly calculate a negative electron population on an atomic basis function due to its arbitrary equal partitioning of overlap charges. Therefore, CM2 maps to Löwdin charges. Löwdin charges are sensitive to the one-electron basis set and to the level of theory used for treating electron correlation. CM2 mapping of Löwdin charges corrects the systematic error contributed by each type of bond.

CM2 charges are atomic partial charges designed to give the best possible agreement of molecular dipole moments with experiment. CM2 charges are obtained by mapping Löwdin charges in a parametrized procedure.⁴⁴ CM2 charges correct systematic errors resulting from the population analysis of wave functions. The dipole moments calculated from the CM2 charges are more accurate than expectation values calculated from the original wave function.⁴³

Let P be the density matrix expanded in a set of nonorthogonal basis functions $\Phi = \{\phi_1, \phi_2, \dots, \phi_N\}$. S , the overlap matrix is defined as

$$S_{ij} = \langle \phi_i | \phi_j \rangle \quad (9)$$

The Löwdin charge on an atom k , denoted q_k , is

$$q_k = Z_k - \sum_{i \in k} (S^{1/2} P S^{1/2})_{ii} \quad (10)$$

where Z_k is the nuclear charge. Therefore, the modified CM2 charge on atom k , q_k , is

$$q_k = q_k^0 + \sum_{k' \neq k} T_{kk'}(B_{kk'}) \quad (11)$$

where q_k^0 is the atomic charge resulting from Löwdin population analysis, $T_{kk'}$ is the amount of charge transfer from atom k' to atom k that corrects the bond dipole, and $B_{kk'}$ is the bond order.

Here, $T_{kk'}$ is a function of the bond order $B_{kk'}$ such that

$$T_{kk'} = B_{kk'}(D_{kk'} + C_{kk'} B_{kk'}) \quad (12)$$

where $D_{kk'}$ and $C_{kk'}$ are empirically determined parameters. Because the total charge of the molecule remains constant, the charge transfer from atom k to k' should have opposite signs, yet have equal magnitudes (i.e., $B_{kk'} = B_{k'k}$ and $T_{kk'} = -T_{k'k}$). The bond order can be defined as

$$B_{kk'} = \sum_{i \in k} \sum_{j \in k'} (PS)_{im} (PS)_{mj} \quad (13)$$

If atoms k and k' do not interact, the bond order is zero and the charge mapped from k to k' should also go to zero.

CM2 is parametrized⁴³ according to a training set of 200 molecules: 186 neutral molecules containing C, H, O, and N, 13 phosphorus containing molecules, and two nonpolar molecules, benzene and ethene. Experimental dipole moments were used for most of the molecules. Whenever experimental data were not available, theoretical calculations at the B3LYP level with the cc-pVTZ basis set were used. A secondary set was also used for parametrization including 13 compounds containing N-O, O-S, and P-S bonds. The performance of the model was tested with 22 ions. Their charges obtained at the MP2/cc-pVDZ level of theory were used for comparison.

The geometries of the training set molecules were optimized at the HF/MIDI! level. The C-H parameters were fixed based on the ChElPG scheme and the remaining CM2 parameters for all other atomic pairs are determined based on a genetic algorithm. First, atomic pairs that are bonded in the primary training set are optimized. Looking at the subset containing only C, H, N, and O, six parameters were optimized. For the next step, the previous C, H, N, O parameters were fixed and the halogen parameters were optimized. Then, all previously optimized parameters were fixed and the S and Si parameters were optimized. Next, holding all previous parameters constant, the P containing bonds were optimized.

The secondary training set was then brought in to parametrize N-O, O-S, and P-S bonds. As before, all previous parameters are fixed. Four parameters were optimized: $C_{kk'}$ and $D_{kk'}$ for N-O bonds and $D_{kk'}$ for O-S and P-S bonds. For the final step, all previously obtained parameters were re-optimized to confirm a stable minimum for all of the parameter space.

1.5.2.2 SM5.42 Theory.⁴⁷ Analytic gradients, the first derivative of the energy with respect to nuclear coordinates, allow for geometry optimization in the liquid phase. The standard state free energy of solvation is defined as⁴⁷

$$G_S^0 = G_{ENP} + G_{CDS} \quad (14)$$

where G_{ENP} is the electrostatic term and G_{CDS} is the free energy of solvation due to the first solvent shell effects, including cavitation, dispersion, and solvent structural changes. The electrostatic term contains the ground state electronic energy in the presence of solvent (E_E), nuclear repulsion energy (E_N), and the polarization energy (G_P)

$$G_{ENP} = E_E + E_N + G_P \quad (15)$$

Note that E_E and E_N are internal energy terms for the distorted solute in the presence of the solvent and G_P includes both the solute-solvent interaction and the free energy cost of polarizing the solvent.

E_E is the solution Hartree-Fock electronic energy

$$E_E = \frac{1}{2} \sum_{rs}^n P_{rs}^{(1)} (h_{rs} + F_{rs}^{(0)}) \quad (16)$$

where r and s are indices of atomic basis functions, n is the number of basis functions, $P_{rs}^{(1)}$ is the density matrix optimized in the presence of solvent, h_{rs} is the one-electron Hamiltonian, and $F_{rs}^{(0)}$ is the gas-phase Fock operator. Note here that $F_{rs}^{(0)}$ is computed from the liquid-phase density matrix $P_{rs}^{(1)}$. The liquid-phase matrix representation of the gas-phase Fock operator differs from the gas-phase matrix representation of the gas-phase Fock operator due to the self-consistent field (SCF). The SCF changes when the orbitals relax in solution.⁴⁷ The one-electron Hamiltonian is applied with dielectric perturbation to compute the density matrix $P_{rs}^{(1)}$ in the presence of solvent.

E_N is the nuclear repulsion energy and is represented by

$$E_N = \sum_{a < b}^N \frac{Z_a Z_b}{|R_a - R_b|} \quad (17)$$

where Z represents nuclear charge and R_a and R_b are the atomic positions of atoms a and b , respectively.

The polarization energy (G_P) is a function of the dielectric constant (ϵ), atomic charges (q), and the Coulomb integral (γ) between atoms a and b

$$G_P = -\frac{1}{2} \left(1 - \frac{1}{\epsilon}\right) \sum_{ab}^N q_a q_b \gamma_{ab} \quad (18)$$

The second term of the standard-state free energy of solution (G_{CDS}) is due to the solvent forming a cavity around the solute molecule. This term is computed as

$$G_{CDS} = \sum_a (\sigma_a^A + \sigma^M) A_a \quad (19)$$

where σ_a^A is the contribution to the atomic surface tension of atom a and depends on the atomic number of a and possibly the solute geometry, σ^M is the contribution to the atomic surface tension that does not depend on atomic numbers or geometry, and A_a is the surface area of atom a .

2. Methods

2.1 Computational Methods

The gas-phase calculations were performed with Gaussian98.⁴⁸ A few of the smaller molecules were also run with HONDOPLUS^{28,29} to confirm HONDOPLUS was compiled correctly and that both programs yielded the same geometries and energies in the gas phase. A geometry was first optimized with AM1,⁴⁹ holding the specified dihedral angles fixed at each specific conformation. (See Figure 11) The AM1 optimized structure was then used as input for the geometry optimization at the HF/6-31G(d) level of theory, where all parameters were allowed to optimize. Then these optimized parameters were extracted from the output files, tabulated, and analyzed. Each optimized structure was subjected to vibrational analysis to check for imaginary frequencies.

The solution calculations were run using the gas-phase optimized output structures as input. The dihedral angles were re-defined to correspond with the specific conformation being run. This is important so that the calculation begins with the conformations at the exact gauche (+60 or -60 degrees) and anti (180 degrees) values. As the calculations are fully optimized, these dihedral angles could vary considerably from their starting values. Once all the solution jobs were run, the data were analyzed for the optimized conformations. The reported conformations are based on ranges from Table 2. It appeared the carboxylate rotation was having an effect on the solvation energy. Therefore, for the lowest energy conformations in each set of molecules (n = 1 to n = 6), the carboxylate was rotated 90 degrees. This will be explained further and illustrated in the n = 2 results section.

Table 2. Gauche and anti value ranges.

	Gas phase	Solution
g+	28.2 to 119.9	39.5 to 88.1
g-	-31.2 to -116.8	-54.2 to -113.1
a	122.3 to 239.9	161.3 to 215.3

2.2 Conformations

Six homologues of the betaines were analyzed, each with a different number of methylene groups in the hydrocarbon chain separating the ammonio and carboxylate groups. The number of methylene groups, n , range from $n = 1$ to $n = 6$. (See Figure 4(b)) In Figure 13, the blue arrows represent carboxyl rotations. These rotations are important for the solution calculations and are mentioned later.

For the $n = 1$ homologue (Figure 13(a)), a carboxylate rotation in the gas phase was performed to locate the lowest energy conformation of the structure. The rotation results confirmed previous calculations in the gas phase.

For $n = 2$, the N-C-C-C dihedral (Figure 13(b)) was rotated every 12 degrees and an energy curve was made for the rotation in the gas phase. Three of the conformations were energy minima: a, g+, and g-. A carboxylate rotation was performed on the three conformations to determine if the optimization located the global minimum of the structure. Once this was confirmed, no more carboxylate rotations were performed in the gas phase.

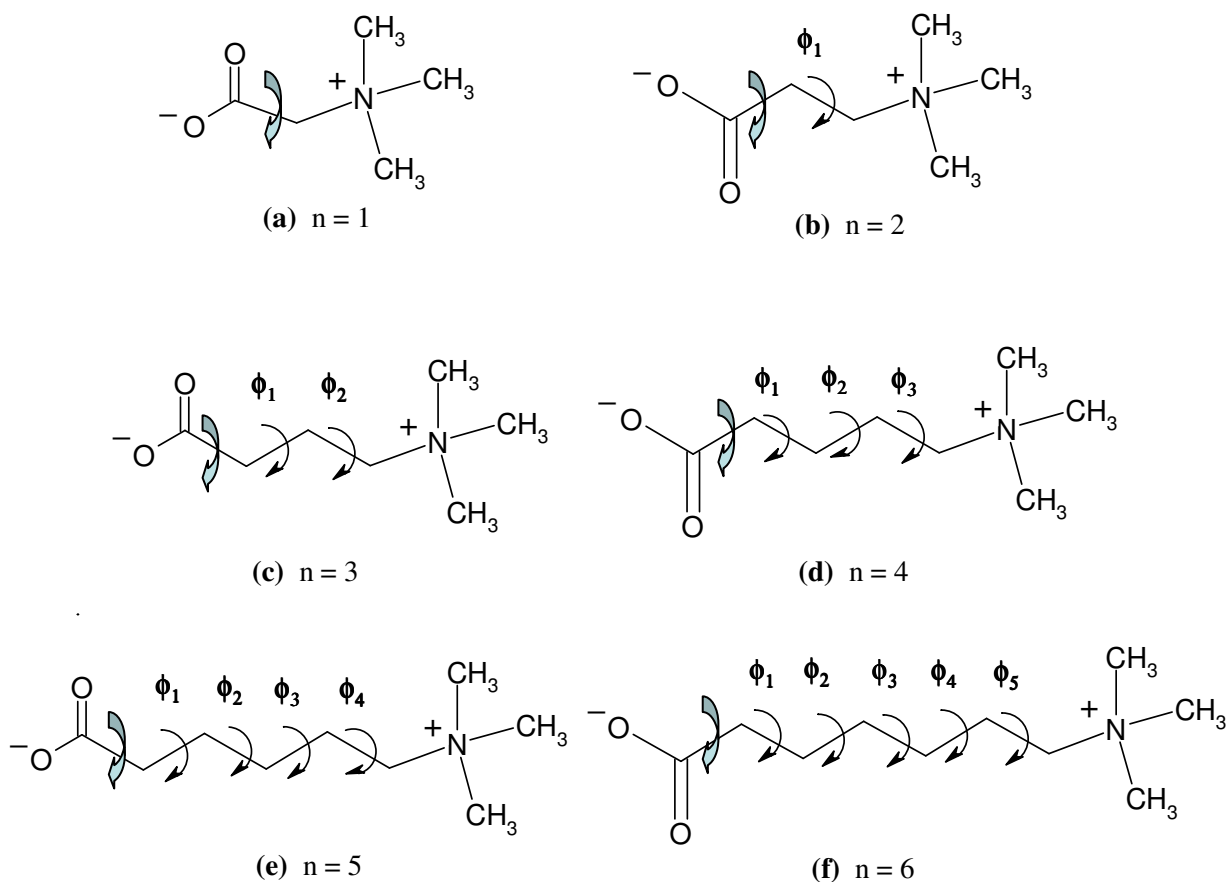


Figure 13. Schematic of dihedral angles for $n = 1$ to $n = 6$.

For $n = 3$, there are two dihedrals to rotate, C-C-C-C and C-C-C-N. (See Figure 13(c)) Therefore, nine different conformations were examined. (See Table 3) For this homologue and future homologues, symmetry was taken into account. Only one structure from a pair of conformational enantiomers was optimized. For example, ag- is equal in energy to ag+.

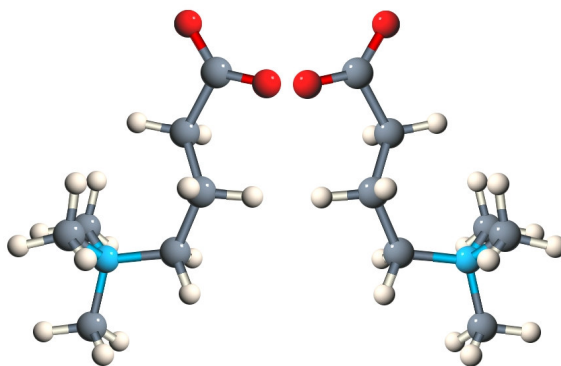


Figure 14. A set of conformational enantiomers of $n = 3$ (ag- and ag+).

(See Figure 14) These two conformations are conformational isomers. They are mirror images of each other. Also, for $n = 3$, the “hexane effect” is observed in the g-g+ and g+g- conformations (ϕ_1, ϕ_2). This effect has been previously examined with hydrocarbon chains.⁵⁰⁻⁵² (See Figure 15) The conformations orient such that a “cyclohexane ring” is formed intramolecularly. This causes a significant amount of steric hindrance when more than one atom tries to occupy the same space. These calculations do not terminate successfully. Therefore, this allows conformations where g-g+ are the last two dihedrals preceding and including the N to not be run. The g-g+ conformations would need to occur at (ϕ_1, ϕ_2) for $n = 3$, (ϕ_2, ϕ_3) for $n = 4$, (ϕ_3, ϕ_4) for $n = 5$, and (ϕ_4, ϕ_5) for $n = 6$. With all this taken into account, there are actually only four conformations that require calculation for the $n = 3$ homologue.

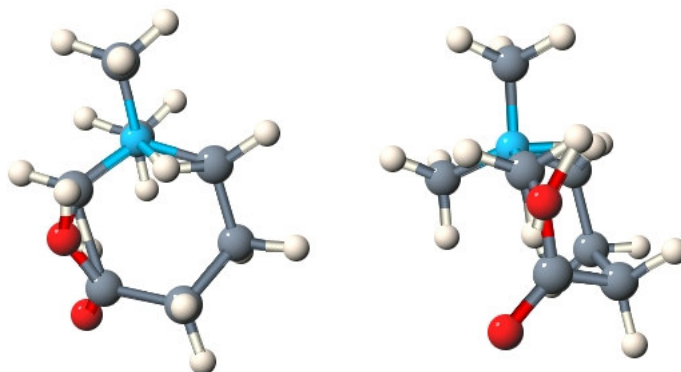


Figure 15. Illustration of the “hexane” effect.

For $n = 4$, there are three dihedral angles to rotate. (See Figure 13(d)) Therefore, of the 27 possible conformations, optimizations were performed on 11 initial conformations. (See Table 3)

For $n = 5$, there are four dihedral angles to rotate. (Figure 13(e)) Therefore, of the 81 possible conformations, optimizations were performed on 32 initial conformations. (See Table 3)

For $n = 6$, there are five dihedral angles to rotate. (Figure 13(f)) Therefore, of the 243 possible conformations, optimizations were performed on 94 initial conformations. (See Table 3)

Table 3. The conformations of betaines optimized.

n	Initial Conformations
2	N-C-C-C rotation from 0-180, increments of 12 degrees
3	aa, ag+, ag-, g+g+, g+g-, g+a, g-g-, g-g+, and g-a
4	aaa, aag+, ag+a, ag+g+, g+aa, g-ag+, g+ag+, g-g+a, g+g+a, g-g+g+, and g+g+g+
5	aaaa, g+g+g+g+, aaag+, g+aaa, ag+aa, aag+a, aag+g+, ag+g+g+, ag+g+a, ag+g-g-, ag+ag+, ag+ag-, g+aag+, g+aaag-, g+ag-a, g+ag-a, g+ag+a, g+ag+g+, g+ag-g-, g+g+g+a, g+g+g-g-, g+g+g-a, g+g+ag+, g+g+ag-, g+g+aa, g+g-g-g-, g+g-g-a, g+g-g+g+, g+g-g+a, g+g-ag-, g+g-ag+, g+g-aa.
6	aaaaa, aaaag+, aaag+a, aaag+g+, g+aaaa, g+aaaag+, g+aaaag-, g+aag+a, g+aag+g+, g+aag-a, g+aag-g-, aag+g+g+, aag+g+a, aag+g-g-, aag+g-a, aag+ag+, aag+ag-, aag+aa, g+ag+g+g+, g+ag+g+a, g+ag+g-g-, g+ag+g-a, g+ag+ag+, g+ag+ag-, g+ag+aa, g-ag+g+g+, g-ag+g+a, g-ag+g-g-, g-ag+g-a, g-ag+ag+, g-ag+ag-, g-ag+aa, ag+aaa, ag+aaag+, ag+aaag-, ag+ag+a, ag+ag+g+, ag+ag-a, ag+ag-g-, g+g+aaa, g+g+aaag+, g+g+aaag-, g+g+ag+a, g+g+ag+g+, g+g+ag-a, g+g+ag-g-, g-g+aaa, g-g+aaag+, g-g+aaag-, g-g+ag+a, g-g+ag+g+, g-g+ag-a, g-g+ag-g-, ag+g+g+g+, ag+g+g+a, ag+g+g-g-, ag+g+g-a, ag+g+ag+, ag+g+ag- ag+g+aa, g+g+g+g+g+, g+g+g+g+a, g+g+g+g-g-, g+g+g+g-a, g+g+g+ag+, g+g+g+ag-, g+g+g+aa, g-g+g+g+g+, g-g+g+g+a, g-g+g+g-g-, g-g+g+g-a, g-g+g+ag+, g-g+g+ag-, g-g+g+aa, ag+g-g-g-, ag+g-g-a, ag+g-g+g+, ag+g-g+a, ag+g-ag-, ag+g-ag+, ag+g-aa, g+g+g-g-g-, g+g+g-g-a, g+g+g-g+g+, g+g+g-g+a, g+g+g-ag-, g+g+g-ag+, g+g+g-aa, g-g+g-g-g-, g-g+g-g-a, g-g+g-g+g+, g-g+g-g+a, g-g+g-ag-, g-g+g-ag+, and g-g+g-aa.

(where a = anti, g+ = gauche plus (+60.0 degrees), and g- = gauche minus (-60.0 degrees))

Table 3 shows all the conformations optimized in the gas phase. The conformations optimized in solution are identical for n = 4, 5, and 6. However, for n = 2, only g+ and a were optimized in solution, and for n = 3, the conformations optimized in solution were aa, ag+, g+g+, and g+a.

The molecules were also optimized in water. This was accomplished via the HONDOPLUS solvation program, version 4.4.^{28, 29} Table 3 shows each conformation that was optimized in solution for the six homologues. For the most part the same conformations optimized in the gas phase were also optimized in water. In solution, four conformations were run to take into account the carboxylate rotation for n = 2. For n = 3, eight conformations were run, taking into account the carboxyl rotation. For n = 4, fifteen conformations were run in solution due to the carboxylate. From n = 4 and up, carboxylate rotations were only run on the conformations which contributed most to the Boltzmann population. Finally, 37 conformations were run for n = 5.

The input for these calculations came from the gas-phase optimized geometries. However, the dihedral angles for each initial conformation (ϕ_1 , ϕ_2 , ϕ_3 , ϕ_4 , and ϕ_5 from Figure 13) were redefined to their gas-phase input value, but all other parameters were kept the same as the

gas-phase output geometry for the solvation calculation. This geometry was then used as input for an SM5.42 solvation calculation.

2.3 Free Energies and Boltzmann Analysis.

The gas-phase free energy is computed from the equation

$$\Delta G = \Delta H - T\Delta S \quad (18)$$

The Gaussian program calculates a value for ΔH , the enthalpy, for each conformation, which needs corrected. The ΔH is then used in equation 18 along with the temperature (298.15K) and entropy to calculate the gas-phase free energy. The entropy is determined based on the vibrational frequency values with a spreadsheet that includes the following set of equations.

The partition function(q) includes a translational, rotational, vibrational, and electronic component.^{53, 54}

$$q_{tot} = q_{trans} q_{rot} q_{vib} q_{elec} \quad (19)$$

where

$$q_{trans} = \left(\frac{2\pi kT}{h^2} \right)^{\frac{3}{2}} \left(\frac{kT}{P} \right) \quad (20)$$

and $T = 273.15$ K, $P = 101325$ Pa, $k = 1.38 \times 10^{-23}$ J/K, and $h = 6.63 \times 10^{-34}$ J*s.

q_{rot} is calculated by

$$q_{rot} = \frac{\pi^{\frac{1}{2}}}{\sigma_r} \left(\frac{T^{\frac{3}{2}}}{\left(\Theta_{r,x} \Theta_{r,y} \Theta_{r,z} \right)^{\frac{1}{2}}} \right) \quad (21)$$

where σ_r is the symmetry index and Θ is $h\nu/k$. The vibrational contribution is

$$q_{vib} = \frac{e^{\frac{-h\nu}{2kT}}}{1 - e^{\frac{-h\nu}{kT}}} \quad (22)$$

q_{elec} is the electronic spin multiplicity of the molecule and in this case is equal to one.

The entropy is also composed of translational, rotational, vibrational, and electronic contributions.⁵³

$$S_{tot} = S_{trans} + S_{rot} + S_{vib} + S_{elec} \quad (23)$$

Each contribution is given by

$$S_{trans} = R \left(\ln(q_{trans}) + 1 + \frac{3}{2} \right) \quad (24)$$

$$S_{rot} = R(\ln(q_r) + 1) \quad (25)$$

$$S_{vib} = R \sum_i \left(\frac{h\nu_i}{kT} \frac{1}{e^{\frac{h\nu_i}{kT}} - 1} - \ln \left(1 - e^{-\frac{h\nu_i}{kT}} \right) \right) \quad (26)$$

$$S_{elec} = R(\ln(q_e)) \quad (27)$$

respectively.

The enthalpy (H) is calculated by the Gaussian program. In order to use that value for the determination of G, that energy is added to RT. Now the gas-phase free energy (G) can be calculated

$$G = H - TS_{tot} \quad (28)$$

Once all of the solvation energies have been computed, a Boltzmann population analysis was performed on the different optimized conformers. The gas-phase energy and free energy of solution was determined for each molecule based on the Boltzmann population of the optimized conformers for each molecule. The final solvation energy was found by subtracting the gas-phase energy from the free energy in solution for each molecule. The theoretically determined solvation energies were compared to literature values of experimentally determined *cmc*'s.

The solvation energy of each molecule was calculated by taking into account the gas-phase free energy, free energy of solvation, and the Boltzmann population for each final optimized conformation. The gas-phase free energy (E_{Gas}) for each molecule was calculated by

$$E_{Gas} = \sum_x (E_{Gas_x} * B_x) \quad (29)$$

where x represents a given conformation of one of six molecules, B is the respective Boltzmann distribution value corresponding to the given conformation, and E_{Gas_x} is the gas-phase free energy for conformation x . The free energy of solvation (E_{Solv}) is computed similarly

$$E_{Solv} = \sum_x (E_{Solv_x} * B_x) \quad (30)$$

where E_{Solv_x} is the free energy of solvation calculated for conformation x and B is again the Boltzmann value. The final solvation energy is the difference between the gas-phase free energy and the free energy of solvation

$$\Delta G_{Solv} = E_{Solv} - E_{Gas} \quad (31)$$

2.4 Intercharge Distances.

The theoretically calculated intercharge distances were measured directly from the calculation output geometries with PCMODEL.⁵⁵ For the nitrogen to carboxylate carbon intercharge distance, the maximum intercharge distances were measured from the all anti conformations for each molecule. The overall intercharge distances were determined by multiplying the nitrogen to carbon distance for each conformation by its Boltzmann value and summing over all the contributions.

For the nitrogen to the point symmetrically between the oxygens, the theoretically calculated maximum intercharge distances were again taken from the all anti conformations. The overall distances were also calculated by multiplying the nitrogen to point between oxygen atoms distance for each conformation by its Boltzmann value and summing all the contributions. These distances were determined by applying trigonometry to the known geometrical information from the output geometries. The Law of Cosines was applied to get final distances.

3. Results

3.1 n = 1 Results

Calculations performed on betaine (where $n = 1$ in Figure 4(b)) are presented in the following tables. Table 4 shows dipole moments and Table 5 shows calculated energies. Table 5 shows calculated and experimental geometrical parameters. These literature data are compared to the ab initio data that is calculated in the gas phase and in solution. These lower levels of theory are in good agreement with the higher levels of theory.

Table 4. Previously calculated dipole moments for $n = 1$.^{56, 57}

Method	Basis Set	Dipole Moment (Debye)	Reference
SCF	aug-cc-pVDZ+3s3p3d	12.24	54
MP2	aug-cc-pVDZ+3s3p3d	11.53	54
B3LYP	aug-cc-pVDZ+3s3p3d	11.42	54
HF	6-31G(d)	11.97	55
MP2	6-31G(d)	11.18	55
B3LYP	6-31G(d)	10.85	55
HF	6-31G(d)	11.98	This Work
SM5.42 (HF)	6-31G(d)	15.58	This Work

Table 5. Previously calculated energies for $n = 1$.^{58, 59}

Method	Basis Set	Energy (a.u.)	Reference
B3LYP	6-31+G(d,p)	-402.3505	56
HF	6-31+G(d,p)	-399.9009	57
HF	6-31G(d)	-399.9115	This Work
SM5.42 (HF)	6-31G(d)	-399.9182	This Work

Table 6. Previously calculated geometrical parameters for $n = 1$ (distances in Å and angles in degrees).^{56, 60, This Work (TW)}

Parameter	MP2 ⁵⁶	B3LYP ⁵⁶	Experimental ⁶⁰	HF ^{TW}	SM5.42 ^{TW}
	aug-cc-pVDZ+3s3p2d	aug-cc-pVDZ+3s3p2d		6-31G(d)	
C ₁ C ₂	1.575	1.577	1.534	1.573	1.531
C ₂ N ₃	1.525	1.537	1.505	1.519	1.488
N ₃ C ₅ , N ₃ C ₄	1.502	1.505	1.498	1.494	1.487
N ₃ C ₆	1.493	1.495	1.505	1.483	1.494
C ₁ O ₇	1.267	1.257	1.242	1.231	1.231
C ₁ O ₈	1.250	1.239	1.238	1.213	1.236
C ₁ C ₂ N ₃	116.5	116.9	118.4	117.2	119.4
C ₂ N ₃ C ₆	108.9	109.1	107.5	108.9	107.6
C ₄ N ₃ C ₅	109.7	109.5	109.9	109.6	110.2
O ₇ C ₁ C ₂	117.1	117.3	120.1	116.8	119.8
O ₈ C ₁ C ₂	110.5	110.5	111.9	110.1	111.6

Table 7 displays the root-mean-square deviations comparing the bond lengths and angles of Table 6 determined by different methods. The experimental values seem to be correlated as well with the gas phase as they are with solution; however, the bond lengths are about two times better in solution.

Table 7. Root-mean-square deviations for bond lengths and angles.

	Bond lengths	Angles
RMSD (exp. vs HF)	0.022	1.9
RMSD (exp. vs SM5.42)	0.011	0.5
RMSD (exp. vs MP2)	0.022	1.8
RMSD (exp. vs B3LYP)	0.023	1.7
RMSD (HF vs MP2)	0.022	0.4
RMSD (HF vs B3LYP)	0.018	0.3

The numbers corresponding to particular atoms in Table 5 are shown Figure 16.

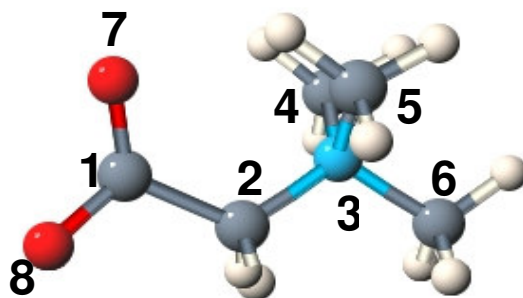


Figure 16. Atom labels.

Table 5 shows the structure is lower in energy in water compared to the gas phase. Also the dipole moment is higher in water than in the gas phase (Table 4). This is due to polar interactions that occur between the water molecules and the $n = 1$ betaine. These interactions increase the charge separation in the molecule, therefore increasing the dipole moment. The charge separation increase is illustrated in Figure 17. The distance between the oxygen and the nearest methyl hydrogens in the gas phase is smaller than the corresponding distance when the molecule is in water.

Table 6 includes the optimized parameters for $n = 1$ betaine in water. These can be compared to the gas-phase values. The bond lengths tend to be shorter in solution and the angles vary slightly in either direction. It is worth noting that the optimized structure in solution shows better agreement with the experimental (crystal) structure than the calculated structure in the gas phase.

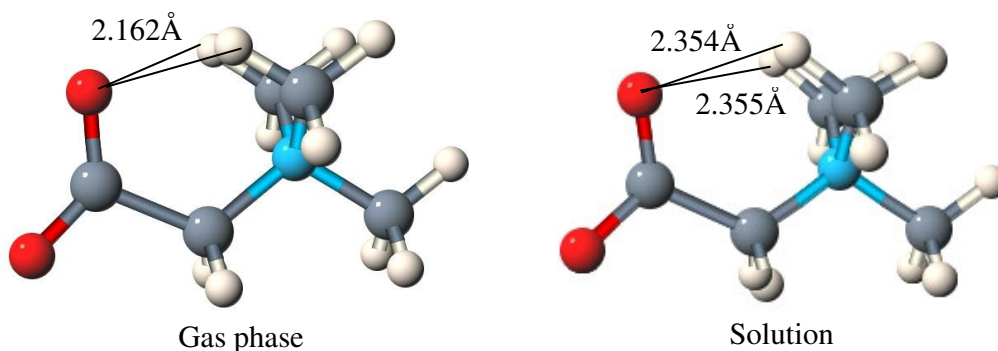


Figure 17. Comparison of the oxygen to hydrogen distances in the gas-phase and in solution.

Figure 17 illustrates the oxygen to hydrogen distances for the gas phase and solution structures. It can be noted that in the gas phase the distance is shorter than in solution. This is

due to polar interactions between the water and the molecule. When the distance is increased the charge separation increases. Therefore, the molecule is more polar. Because it is more polar, the solvation energy is higher, yielding a more stable structure in solution.

Figure 18 shows the partial charges on the oxygens, nearby methyl hydrogens, and the nitrogen. In this molecule, the nitrogen has a negative charge because the positive charge is distributed among the methyl hydrogens.

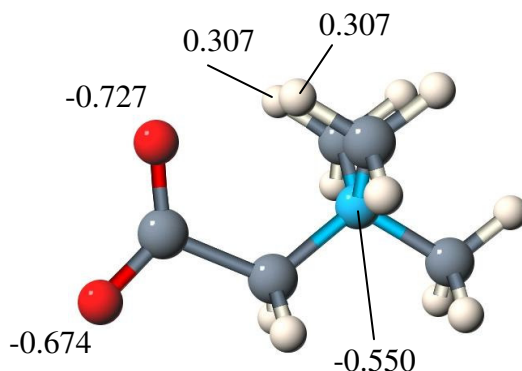


Figure 18. Some partial charges for $n = 1$ gas phase.

The final solvation energy is 4.2 kcal/mol for the $n = 1$ molecule.

3.2. $n = 2$ Results

For $n = 2$ ϕ_1 was rotated 180 degrees with 12 degree increments. This rotation was performed to confirm the presence of only three minima. All ϕ_1 values except 168.0 and 180.0 optimized to the 60 degree conformation. The 168.0 and 180.0 conformations optimized to the 180.0 degrees. Therefore, due to symmetry, this rotation confirmed the presence of only three minima: 60, -60, and 180.

Table 8 contains data in the gas phase and in water for the $n = 2$ betaine. For the anti conformation the molecule is 7.3 kcal/mol more stable in solution. However, for the $g+$ and $g-$ conformations, the structures are 1.6 kcal/mol more stable in the gas phase. This makes sense due to the interchange distances being less for the $g+$ and $g-$ conformations causing a lower dipole moment, therefore being less stable in water. For the anti conformation, the interchange

distance is the greatest, yielding a higher dipole moment making the molecule more polar and more stable in water.

Table 8. $n = 2$ Data in the gas phase and in solution.

Dihedral Angles	Energy (a.u.)	Dipole Moment (Debye)	Relative Energy (kcal/mol)	P(n)
Gas phase				
g+ [g-]	-438.9345	12.58	0	1
a	-438.9228	17.91	7.3	0
Solution				
a (180)	-438.9529	22.10	0	0.7665
a(90)	-438.9508	21.89	1.3	0.0890
g+ [g-] (90)	-438.9503	18.06	1.6	0.1027 ^a
g+ [g-] (180)	-438.9495	19.27	2.1	0.0418 ^a

^apopulation for two conformations.

The Boltzmann populations (P(n)) for $n = 2$ are in Table 8. The (180) and (90) indicate the position of the carbonyl oxygens. The (180) indicates dihedral angles of 0 and 180 degrees are formed (Figure 19(a)) and (90) indicates dihedral angles of -90 and 90 degrees are formed with respect to the oxygens (Figure 19(b)). This notation will be used throughout the

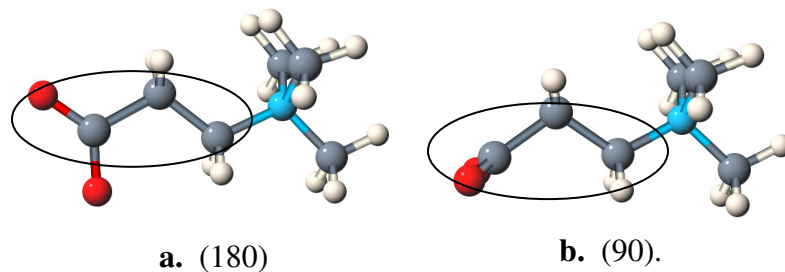


Figure 19(a,b). Illustration of (180) and (90) with $n = 2$ anti conformation.

remainder of the paper. Based on the data from this table, the anti conformation contributes most to the final structure in solution and the gas-phase structure is only g+ and g- conformations. In solution the dominate conformation is a(180). The gauche does, however, contribute with a Boltzmann population of 0.1027 for the g+(90). The structures that contribute most to the Boltzmann population are illustrated in Figure 20.

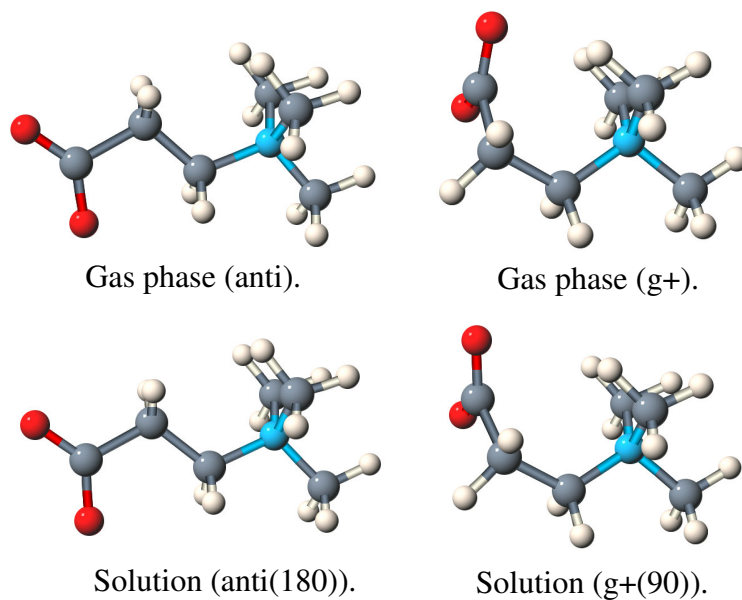


Figure 20. $n = 2$ Optimized structures.

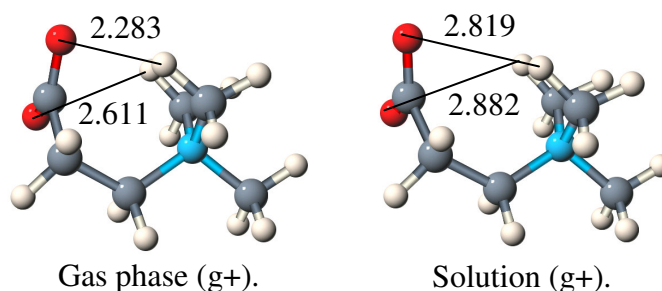


Figure 21. Oxygen to hydrogen distances.

Figure 21 illustrates the oxygen to nearest methyl hydrogen distances for the g+ conformation in the gas phase and in solution for $n = 2$. As was seen with the $n = 1$ molecule, the distances in solution are longer than those in the gas phase.

Figure 22 reveals some partial charges for the g+ conformation in the gas phase. Here again, the nitrogen has a negative charge and the methyl hydrogens have a strong positive charge. These strong positive charges from the methyl hydrogens are stabilizing the negative charges from the oxygens.

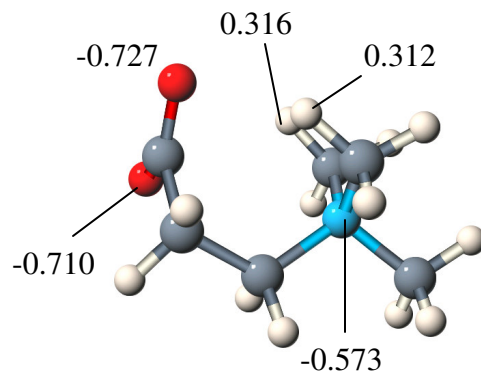


Figure 22. Some partial charges for $n = 2$ (g+).

The $n = 2$ conformations have been previously studied with NMR spectroscopy.⁶¹ The results indicate that the fraction of gauche are 0.62 and 0.69, calculated from $^3J_{\text{HH}}$ coupling constants. This varies from my calculated Boltzmann population of 0.1445 for the gauche conformation. The authors of this previous work were surprised by the outcome of their experiment. They were expecting the dominant conformation to be anti, which is the result that my calculations show.

There have also been calculations performed on the $n = 2$ conformations using PM3 for the gas phase and the self-consistent reaction field (SCRF) method in solution.⁶² The relative energies are tabulated in Table 9. These relative energies follow the same trend as my calculations with the gauche conformation lowest in energy in the gas phase and the anti conformation lowest in energy in solution.

Table 9. $n = 2$ Literature data in the gas phase and in solution.⁶²

Dihedral Angles	Relative Energy (kcal/mol)
Gas phase	
a	16.5
g+ [g-]	0
Solution	
a	0
g+ [g-]	20.4

The final solvation energy for this work is 11.2 kcal/mol for the $n = 2$ molecule.

3.3 n = 3 Results

Tables 10 and 11 show the optimization results of the n = 3 betaine in the gas phase and in solution, respectively. Following the trend of the n = 2 betaine, the gas-phase dipole moment is less than the solution dipole and hence is less polar. The gas-phase final structure is composed of the g+a and g+g+ conformations. These conformations have the smallest dipole moments and the smallest intercharge distances. However, in solution, the final structure is comprised primarily of the aa and g+a conformations. These conformations have the highest dipole moments and are the most polar. Therefore, they are the most stabilized by water.

In Table 10 the Boltzmann populations for g+a and g+g+ also take into account the symmetrically equivalent conformations of g-a and g-g-. This means, for example, that the P(n) value of 0.704 for g+a is actually 0.3520 for g+a and 0.3520 for g-a. The same holds true for the g+g+ conformation as tabulated. The P(n) for g+g+ is actually 0.1480 and for g-g- is also 0.1480. The doubling of the P(n) value to take into account symmetric conformations will be carried through for the remainder of the paper. Please note accordingly.

Table 10. n = 3 Optimization results in the gas phase.

Dihedral Angles (ϕ_1, ϕ_2)	Energy (a.u.)	Dipole moment (Debye)	Relative Energy (kcal/mol)	P(n)
g+a [g-a]	-477.9626	16.48	0	0.7040
g+g+ [g-g-]	-477.9618	13.32	0.5	0.2960
aa	-477.9419	22.88	13.0	0
ag+ [ag-]	-477.9377	21.84	15.6	0

Table 11. n = 3 Optimization results in water.

Dihedral Angles (ϕ_1, ϕ_2)	Energy (a.u.)	Dipole moment (Debye)	Relative Energy (kcal/mol)	P(n)
aa (180)	-477.9879	27.94	0	0.3153
aa (90)	-477.9863	27.76	1.0	0.0579
g+a [g-a] (180)	-477.9878	23.50	0.1	0.5673
g+a [g-a] (90)	-477.9856	23.32	1.4	0.0552
g+g+ [g-g-] (90)	-477.9823	21.93	3.5	0.0017
g+g+ [g-g-] (180)	-477.9822	22.92	3.6	0.0016
ag+ [ag-] (180)	-477.9815	26.60	4.0	0.0007
ag+ [ag-] (90)	-477.9809	26.51	4.4	0.0004

The structures that contribute most to the Boltzmann population in the gas phase and in solution are shown in Figure 23. The lower energy structures in the gas phase have a smaller interchange distance, and the structures lower in energy in solution have a larger interchange separation.

Figure 24 includes partial charges for the optimized gas-phase conformations for $n = 3$. This figure again shows the negative charge on the nitrogen and the relatively large positive and negative partial charges on the methyl hydrogen and oxygens, respectively.

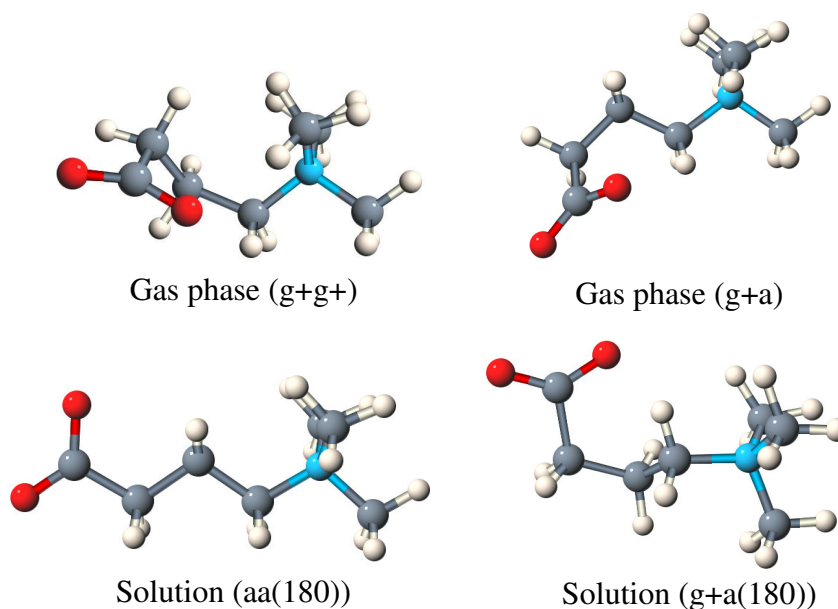


Figure 23. $n = 3$ Optimized structures.

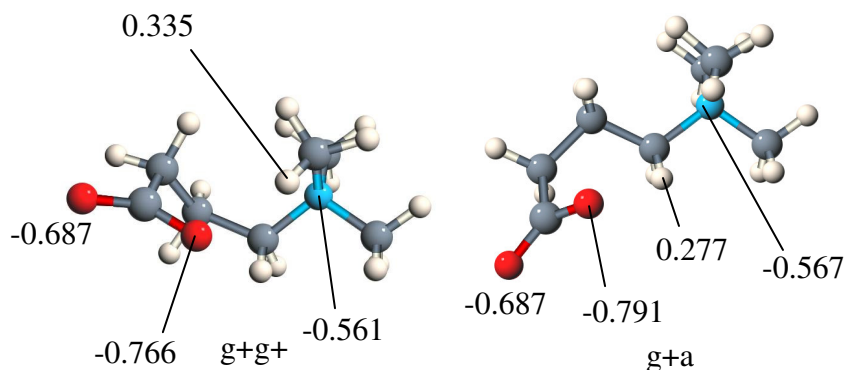


Figure 24. Some partial charges for $n = 3$ optimized structures.

Previous calculations have been performed on the $n = 3$ conformations.⁶² The gas-phase calculations were run using PM3 and the solution calculations were run using the SCRF method. Table 12 includes the results of the study. For the gas phase, the results are similar to my calculations; however, the aa and ag+ results are reversed based on my calculations, but both conformations have very high relative energies. Comparing the solution results, my determinations are quite different from the literature values. The aa conformations of both works have relative energies equal to 0 kcal/mol and the g+g+ conformations have one of the highest relative energies. The other values do not show any similar trend.

Table 12. $n = 3$ Literature data in the gas phase and in solution.⁶²

Dihedral Angles (ϕ_1, ϕ_2)	Relative Energy (kcal/mol)
Gas phase	
g+g+	0
g+a	1.8
ag+	14.6
aa	17.2
Solution	
aa	0
ag+	9.5
g+a	14.3
g+g+	23.9

The final calculated solvation energy for the $n = 3$ molecule is 15.8 kcal/mol.

3.4 $n = 4$ Results

The $n = 4$ betaine results for the gas phase are revealed in Table 13. Table 13 shows that the anti conformation has the highest energy in the gas phase. This makes sense because the dipole moment is the highest when the dihedral angles are all anti. Therefore, the final structure is composed primarily of the g-g+a, g-g+g+(180), g-g+g+(90), and g+g+a conformations. These conformations are illustrated in Figure 25.

Table 13. $n = 4$ Optimization results in the gas phase.

Optimized Dihedral Angles (ϕ_1, ϕ_2, ϕ_3)	Energy (a.u.)	Dipole moment (Debye)	Relative Energy (kcal/mol)	P(n)
g-g+g+ [g+g-g-] (90)	-517.0046	12.01	0	0.6682
g-g+g+ [g+ag-] (180)	-517.0026	14.37	1.2	0.0813
g-g+a [g+g-a]	-517.0030	11.74	1.0	0.1343
g+g+a [g-g-a]	-517.0029	12.79	1.1	0.1131
g+ag+ [g-ag-]	-516.9995	15.73	3.2	0.0031
ag+a [ag-a]	-516.9662	27.53	24.0	0
aaa	-516.9642	29.49	25.3	0

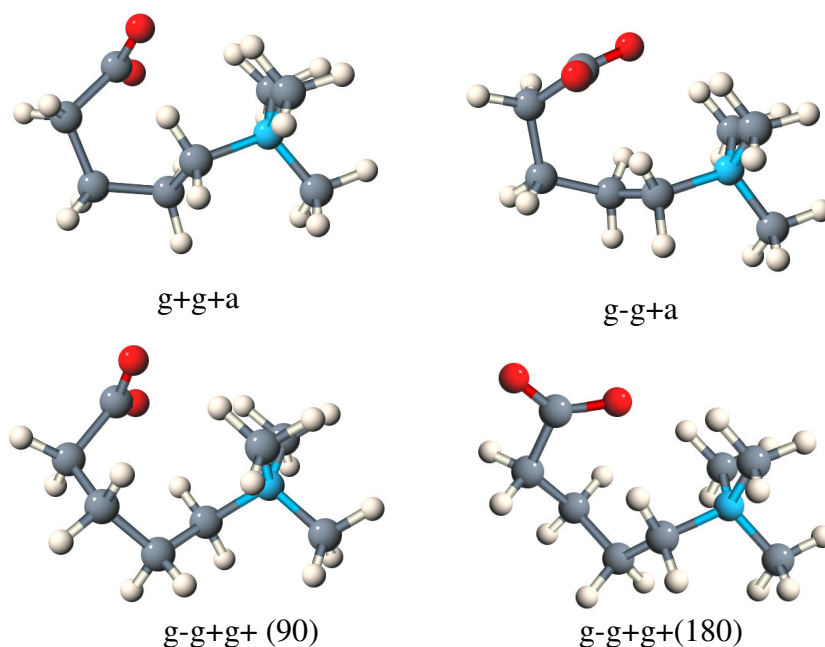
**Figure 25.** $n = 4$ Gas-phase optimized structures.

Figure 26 shows the partial charges for the $n = 4$ gas-phase conformations g+g+a, g-g+a, g-g+g+(90), and g-g+g+(180). This figure confirms the partial charge trend that has been detected to this point. The nitrogen and oxygens have negative partial charges and the methyl hydrogens have positive partial charges. The positive charges of the hydrogens are balancing the negative charges of the oxygens. They orient themselves such that the hydrogens and oxygens line up for the opposite charge attraction.

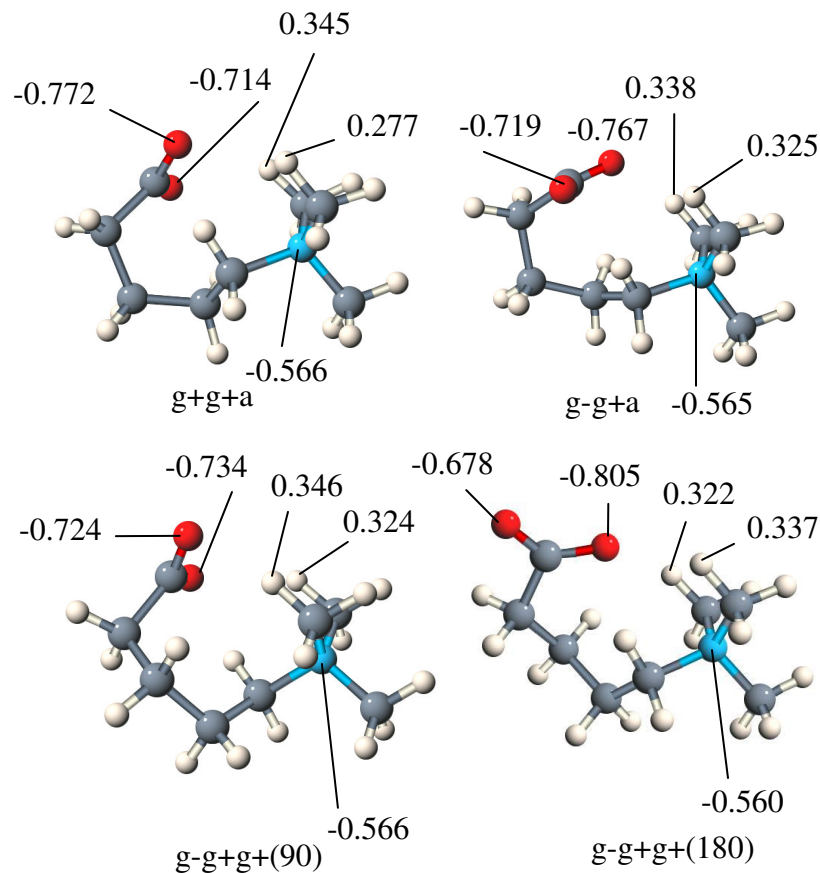
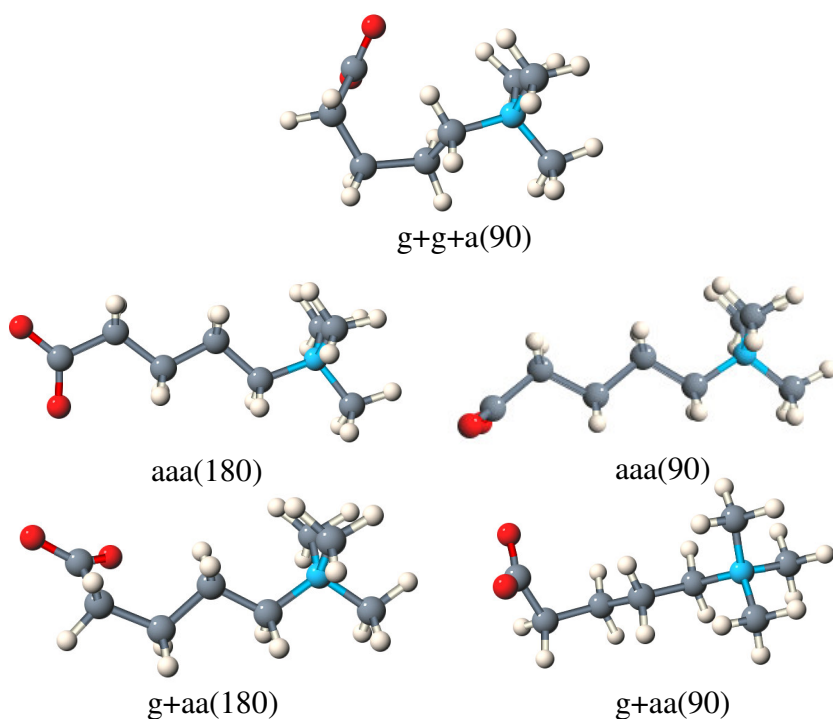


Figure 26. Some partial charges for $n = 4$ optimized gas-phase structures.

The $n = 4$ betaine solution optimization data is tabulated in Table 14. Surprisingly, the lowest energy conformation is g+aa. This conformation is not the most polar, and so does not have the largest dipole moment. Therefore, the dipole moment is not the only contribution to the solvation energy and may not even be the strongest. This notion will be discussed further later in the paper. The conformations that contribute most to the final structure are aaa, ag+a, and g+g+a. Figure 27 illustrates these conformations. The final calculated solvation energy is 11.2 kcal/mol for the $n = 4$ molecule.

Table 14. $n = 4$ Optimization results in water.

Optimized Dihedral Angles (ϕ_1, ϕ_2, ϕ_3)	Energy (a.u.)	Dipole moment (Debye)	Relative Energy (kcal/mol)	P(n)
g+aa [g-aa] (180)	-517.0226	29.01	0	0.5302
g+aa [g-aa] (90)	-517.0210	27.97	1.0	0.1004
aaa (180)	-517.0220	34.25	0.4	0.1416
aaa (90)	-517.0215	34.10	0.7	0.0807
g+g+a [g-g-a] (90)	-517.0206	22.86	1.3	0.0620
g+g+a [g-g-a] (180)	-517.0197	25.42	1.9	0.0229
ag+a [ag-a] (180)	-517.0197	32.31	1.8	0.0238
ag+a [ag-a] (90)	-517.0196	32.35	1.9	0.0209
g+ag+ [g-ag-]	-517.0186	24.39	2.5	0.0079
g-g+a [g+g-a]	-517.0183	22.69	2.7	0.0057
g-ag+ [g+ag-]	-517.0174	30.33	3.3	0.0020
aag+ [aag-]	-517.0169	30.96	3.6	0.0013
ag+g+ [ag-g-]	-517.0159	22.35	4.2	0.0004

**Figure 27.** $n = 4$ Solution optimized structures.

3.5 n = 5 Results

The gas-phase results for the n = 5 betaine is included in Table 15. As is common with all of the betaines, the aaaa conformation has the highest dipole moment and so has the highest energy for the gas-phase results. The lowest in energy is the g+g-g-a conformation. The conformations that contribute most to the final structure are the g+ag+g+, g+g-g+a, and g+g-g-a conformations (Figure 28).

Table 15. n = 5 Optimization results in the gas phase.

Optimized Dihedral Angles ($\phi_1, \phi_2, \phi_3, \phi_4$)	Energy (a.u.)	Dipole Moment (Debye)	Relative Energy (kcal/mol)	P(n)
g-g+g+a [g+g-g-a]	-556.0484	11.49	0	0.6436
g+g-g+a [g-g+g-a]	-556.0478	11.30	0.4	0.3471
g+ag+g+ [g-ag-g-]	-556.0444	11.49	2.5	0.0093

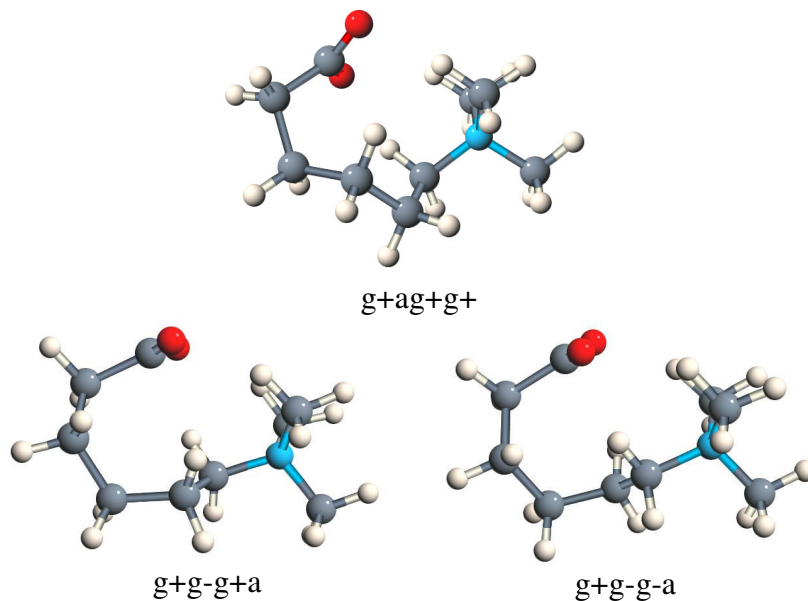


Figure 28. n = 5 Gas-phase optimized structures.

Figure 29 shows some partial charges for the n = 5 gas-phase optimized conformations. The partial charges shown are consistent with those previously determined for other molecules.

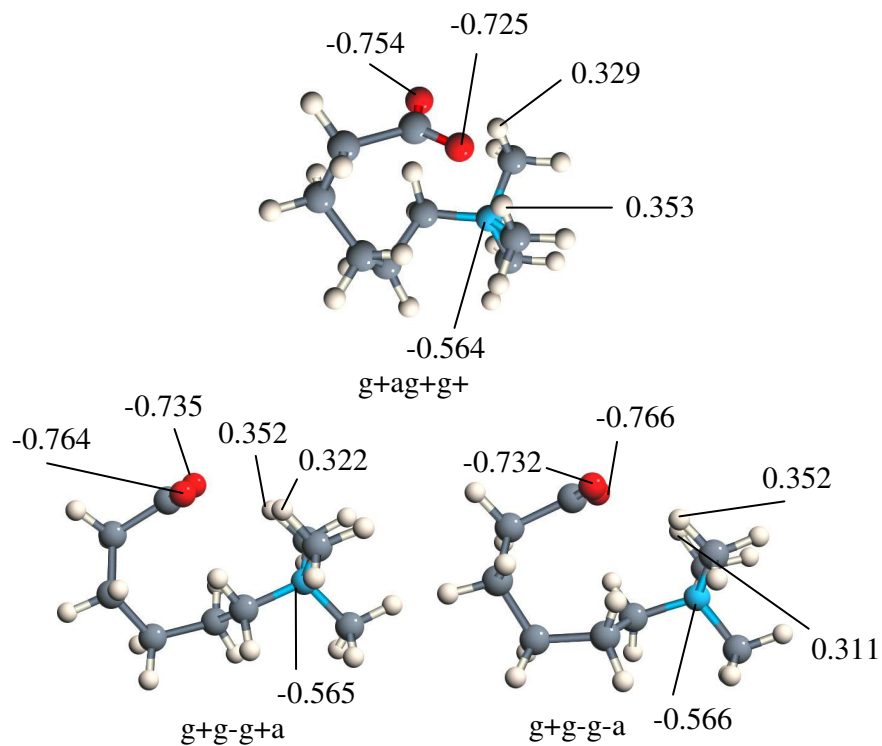


Figure 29. Partial charges for $n = 5$ gas-phase optimized structures.

The solution results for the $n = 5$ betaine are in Table 16. The conformation lowest in energy is the all anti (aaaa). The conformations that contribute most to the final structure are aaaa, g+aaa, g+ag-a, ag+aa, and aag+a (Figure 30).

Table 16. n = 5 Optimization results in water.

Optimized Conformations ($\phi_1, \phi_2, \phi_3, \phi_4$)	Dipole moment (Debye)	Energy (Hartrees)	Relative Energy (kcal/mol)	P(n)
aaaa (180)	40.05	-556.0567	0	0.1890
aaaa (90)	39.87	-556.0564	0.2	0.1461
g+aaa [g-aaa] (180)	34.37	-556.0564	0.2	0.2736
g+aaa [g-aaa] (90)	34.24	-556.0553	0.9	0.0890
g+ag-a [g-ag+a] (180)	35.59	-556.0554	0.8	0.0946
g+ag-a [g-ag+a] (90)	35.52	-556.0541	1.6	0.0253
ag+aa [ag-aa] (90)	37.31	-556.0547	1.3	0.0441
ag+aa [ag-aa] (180)	37.26	-556.0545	1.4	0.0375
aag+a [aag-a] (180)	35.92	-556.0545	1.4	0.0371
aag+a [aag-a] (90)	35.46	-556.0541	1.6	0.0252
g+ag+a [g-ag-a]	28.44	-556.0533	2.2	0.0100
g+g-g-a [g-g+g+a]	19.59	-556.0532	2.2	0.0097
g+g+aa [g-g-aa]	25.41	-556.0527	2.5	0.0057
g+g+ag+ [g-g-ag-]	23.82	-556.0526	2.6	0.0049
g+g+g+a [g-g-g-a]	27.20	-556.0525	2.6	0.0044
ag+g+a [ag-g-a]	24.28	-556.0510	3.5	0.0010
aaag+ [aaag-]	37.96	-556.0507	3.7	0.0007
g+g-aa [g-g+aa]	20.33	-556.0502	4.1	0.0004
ag+ag- [ag-ag+]	38.13	-556.0500	4.2	0.0003
g+aag+ [g-aag-]	29.92	-556.0496	4.4	0.0002

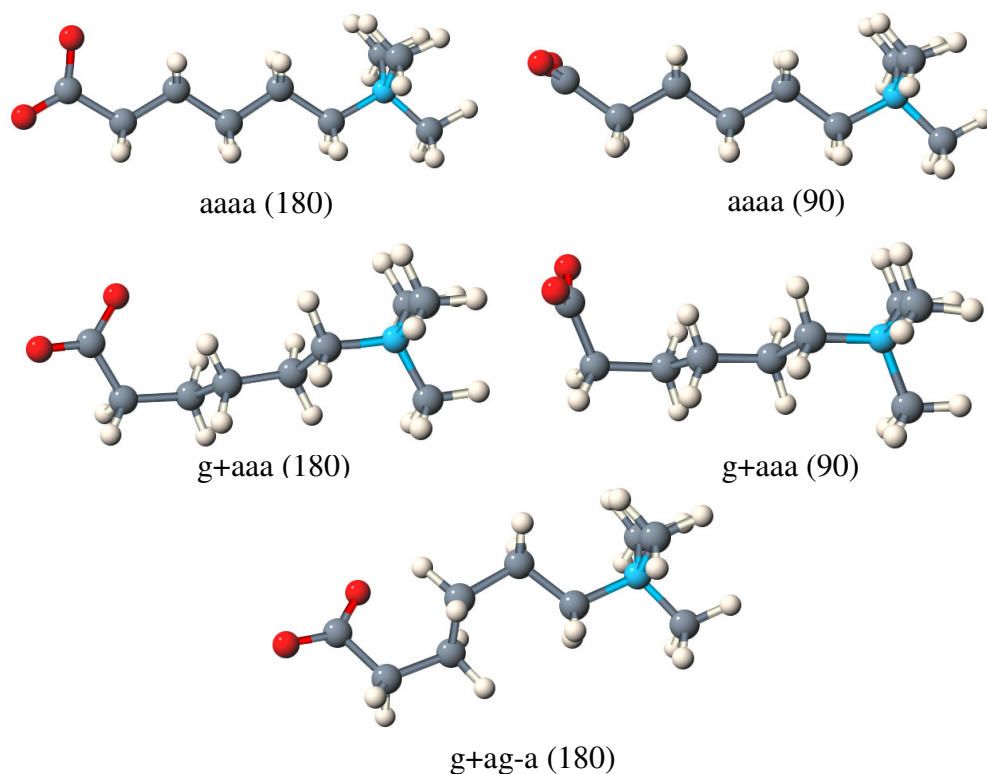


Figure 30. $n = 5$ Solution optimized structures.

The final calculated solvation energy for the $n = 5$ molecule is 4.8 kcal/mol.

3.6 $n = 6$ Results

The $n = 6$ betaine optimization results in the gas phase are included in Table 17. The most polar conformation again being the all anti. Ninety-four possible conformations were optimized. Of all ninety-four only seven conformations were found in the final geometries to contribute to the Boltzmann population: g+ag+g+a, g-g+ag+g+, g-g+g-g+a, g-g+g+g-g-, g-g+g+ag+, g-g+g+aa, and g+g-g+aa. The lowest in energy is g+ag+g+a. Figure 31 illustrates the g-g+g-g+a, g-g+ag+g+, and g+ag+g+a conformations.

Table 17. n = 6 optimization results in the gas phase.

Optimized Conformations ($\phi_1, \phi_2, \phi_3, \phi_4, \phi_5$)	Energy (a.u.)	Dipole Moment (Debye)	Relative Energy (kcal/mol)	P(n)
g+ag+g+a [g-ag-g-a]	-595.0883	10.78	0	0.8907
g-g+g-g+a [g+g-g-g-a]	-595.0854	10.70	1.8	0.0420
g-g+ag+g+ [g+g-ag-g-]	-595.0852	11.12	2.0	0.0327
g-g+g+aa [g+g-g-aa]	-595.0846	12.28	2.3	0.0177
g-g+g+ag+ [g+g-g-ag-]	-595.0837	11.95	2.9	0.0069
g-g+g+g-g- [g+g-g-g+g+]	-595.0836	11.10	2.9	0.0064
g+g-g+aa [g-g+g-aa]	-595.0831	11.02	3.3	0.0036

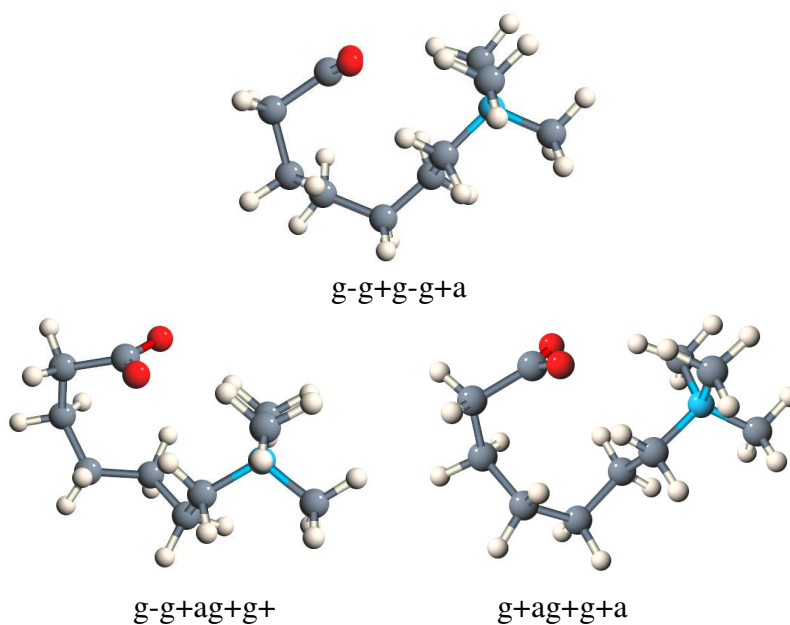
**Figure 31.** n = 6 Gas-phase optimized structures.

Figure 32 illustrates some partial atomic charges for the n = 6 optimized conformations. Yet again, the charges are distributed according to the pattern set from the previous sets of molecules.

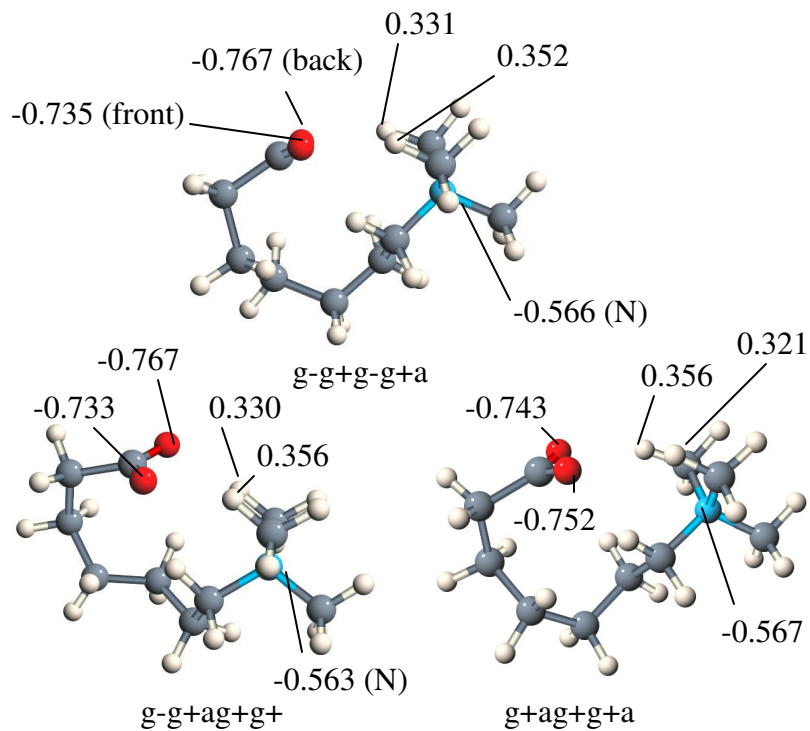


Figure 32. Partial charges for $n = 6$ gas-phase optimized structures.

Preliminary solution results for $n = 6$ are tabulated in Table 18. Solvation energy calculations are still being performed. Vibrational frequency analysis has not been run on any of these geometries yet.

Table 18. n = 6 optimization results in solution.

Optimized Conformations ($\phi_1, \phi_2, \phi_3, \phi_4, \phi_5$)	Energy (a.u.)	Dipole Moment (Debye)	Relative Energy (kcal/mol)	P(n)
aaaaa	-595.0910	46.33	0	0.3551
g+aaaa	-595.0905	38.84	0.3	0.2131
g+ag+aa	-595.0899	36.56	0.7	0.1138
g-ag+aa	-595.0898	41.89	0.7	0.1024
ag+aaa	-595.0893	43.63	1.1	0.0557
aag+g+a	-595.0888	24.27	1.4	0.0336
aag+aa	-595.0887	38.59	1.4	0.0308
aaag+a	-595.0885	43.24	1.5	0.0261
g+aaag+a	-595.0884	36.23	1.6	0.0235
g-ag+g+a	-595.0874	38.20	2.3	0.0078
ag+ag-a	-595.0872	43.77	2.4	0.0065
g+g+ag+a	-595.0871	31.85	2.5	0.0056
g-g+aaa	-595.0871	30.82	2.5	0.0054
ag+ag+a	-595.0870	41.49	2.5	0.0051
ag+g+aa	-595.0870	35.86	2.5	0.0050
g+aaag-a	-595.0862	24.17	3.0	0.0021
g-g+ag+a	-595.0861	32.75	3.1	0.0020
g-g+ag-a	-595.0859	35.14	3.2	0.0017
g+aaag-	-595.0858	39.61	3.2	0.0015
g+g+aaag-	-595.0855	29.70	3.5	0.0010
g+g+aaag+	-595.0850	36.00	3.8	0.0006
g-ag+ag-	-595.0846	40.77	4.0	0.0004
g-g+g+g+a	-595.0845	26.21	4.1	0.0003
aaaag+	-595.0845	42.21	4.1	0.0003
ag+g+g+a	-595.0843	28.04	4.2	0.0003

4. Conclusions

4.1 Solvation Energies and Dipole Moments

4.1.1 Solvation Energies. The solvation energies of the homologous betaine molecules are tabulated in Table 19. The data is plotted in Figure 33. The solvation energy shows a similar trend compared to the *cmc* data.⁶ Both curves are parabolic. However, the *cmc* curve peaks at $n = 4$ and the solvation energy curve peaks at $n = 3$.

Table 19. Absolute Solvation Energies for $n = 1$ to $n = 5$.

n	Solvation Energy (kcal/mol)
1	4.2
2	11.2
3	15.8
4	11.2
5	4.8

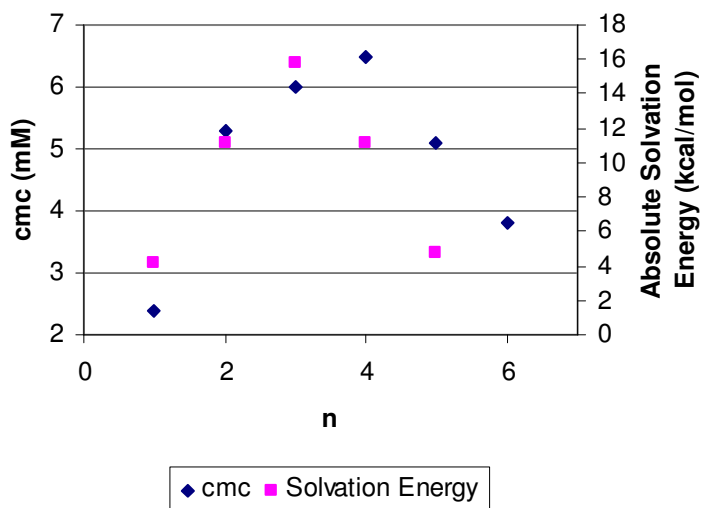


Figure 33. Solvation Energy Plot for $n = 1$ to $n = 5$.

Even though the curvatures of the two data sets are consistent, the curves are not exactly the same. Therefore, the cross-term $c(\text{hydrophilic descriptor} \times \text{hydrophobic descriptor})$ in equation 1 is not quite equal to zero,

$$\log (cmc) = a(\text{hydrophilic descriptor}) + b(\text{hydrophobic descriptor}) \quad (1) \\ + c(\text{hydrophilic descriptor} \times \text{hydrophobic descriptor}) + \text{constant}$$

and the hydrophilicity and hydrophobicity terms cannot be treated separately. The hydrophilicity of the head group can be predicted with solvation energies and the hydrophobicity of the tail can be accounted for with previously known data according to the Kleven's equation and Figure 6. The two sets of data in Figure 33 can be correlated and a relationship found. This can lead to an approximation for the cross-term.

Being able to reproduce the cmc curve with solvation data determined from the head group of a surfactant is very useful because calculation expense increases considerably when increasing the number of atoms. Without having to account for the tail, the number of atoms and computational expense is significantly decreased.

Figure 34 gives a more specific look at the model in Figure 7. The initial proposal was that the solvation energies of the monomers would predict the cmc of micelles because the solvation energy of the micelle was thought to be constant. The graph in Figure 33 shows that the solvation energy curve does not exactly model the cmc curve. The solvation energy peaks at $n = 3$ and the cmc peaks at $n = 4$. Therefore, the solvation energy of the micelle must differ slightly depending on the monomers that comprise it.

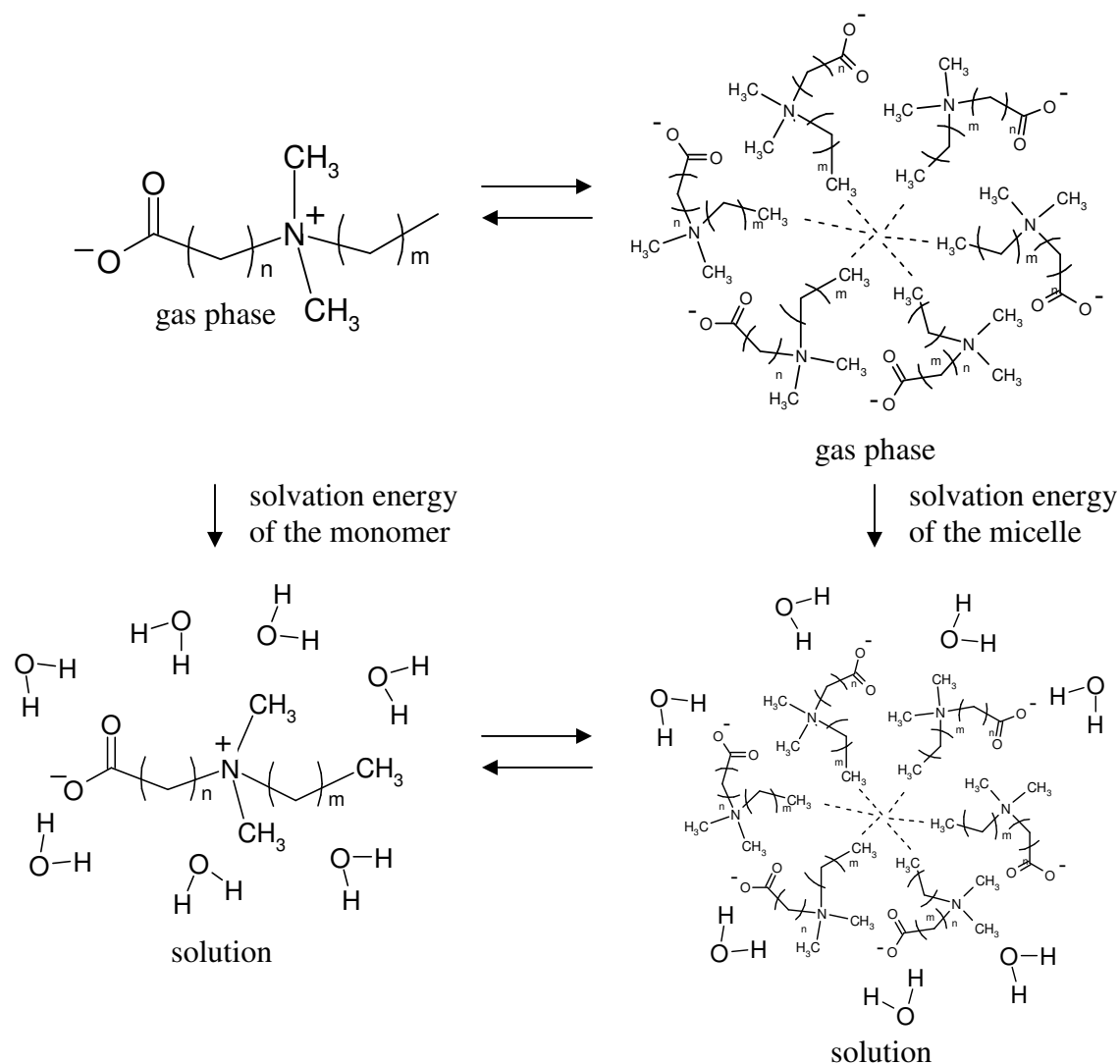


Figure 34. Betaine series model for solvation energy of monomer-micelle equilibria.

The calculated gas-phase free energy and free energy of solution were graphed versus n . A linear trend line was determined to correlate the data for each. Then the predicted gas-phase and solution energies were determined based on the trend line equations. The predicted gas-phase energies were subtracted from the calculated gas-phase energies, and the same was done for the solution energies. The energy differences for the gas phase and solution are plotted in Figures 35 and 36, respectively.

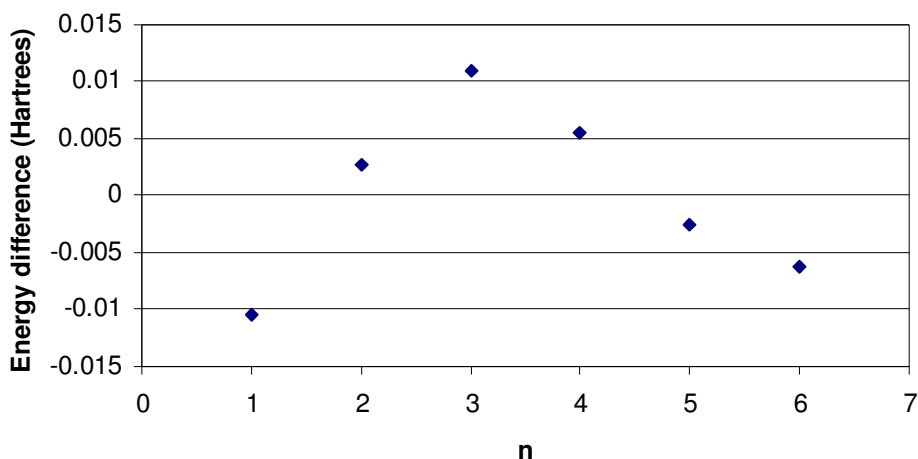


Figure 35. Gas-phase energy differences.

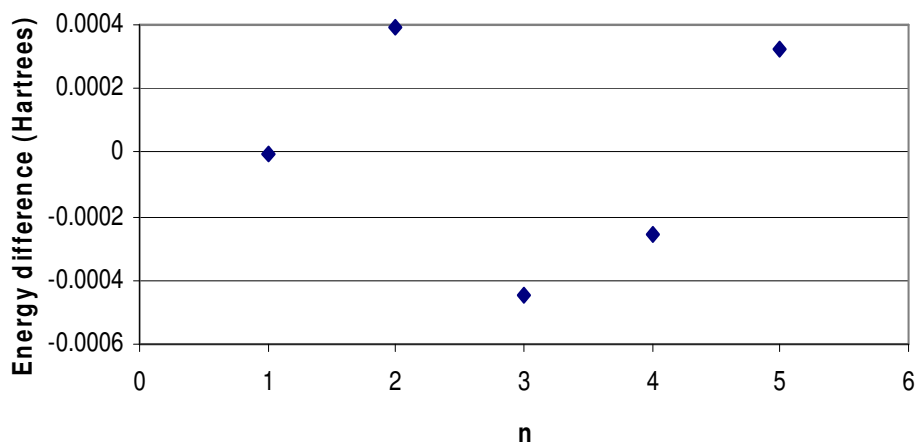


Figure 36. Solution energy differences.

The gas-phase energy difference graph shows a maximum at $n = 3$, which is similar to the solvation energy curve. The solution energy difference graph does not show anything interesting. There is no specific pattern that relates the solution energies to the solvation energies. This indicates that the gas-phase free energies, hence the stability of the gas-phase structures, have more of an effect on the solvation energy than do the free energies in solution, and so the structures in solution.

4.1.2 Dipole Moments. Figures 37-39 illustrate the effect of varying n on the dipole moments. The dipole moments that are plotted are expectation values based on the Boltzmann populations of each conformation. In Figure 37 the gas-phase dipole moments show a similar trend compared to the final solvation energy graphed in Figure 33, both have a maximum at $n = 3$. This curve indicates that as n increases, the dipole moment does not continually increase or level off. As the molecule increases in length, the intramolecular charges become more stabilized in the gas phase. The molecule is able to twist around and decrease the charge distribution between atoms. This effect is illustrated in the gas-phase optimization figures for $n = 2$ to $n = 6$, Figures 20, 23, 25, 28, and 31, respectively. Once these charges become more stabilized, the dipole moments decrease as is seen in Figure 37. This result indicates that the gas-phase energy is a major contributor to the final solvation energy. This is very interesting because it seems more likely the solvation energy would be affected by the free energy of solvation more than the gas-phase energy. The stabilization of the intramolecular charges in the gas-phase structures appears to be stronger than the stabilization of the molecules in water.

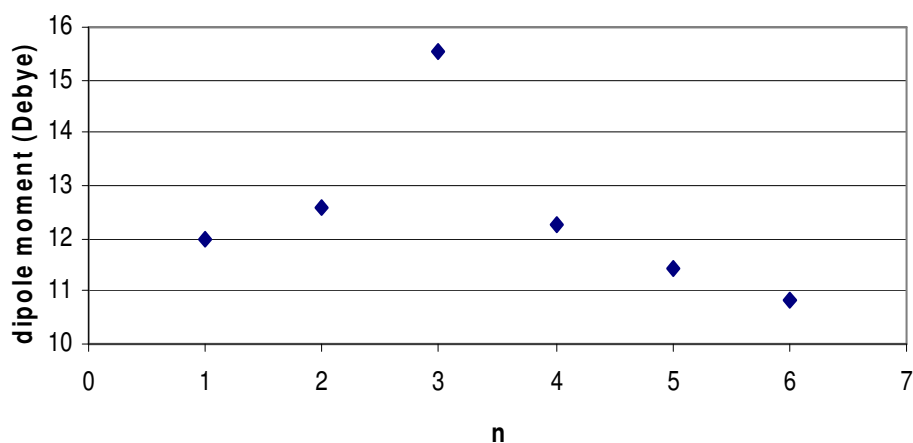


Figure 37. Gas-phase dipole moments vs n .

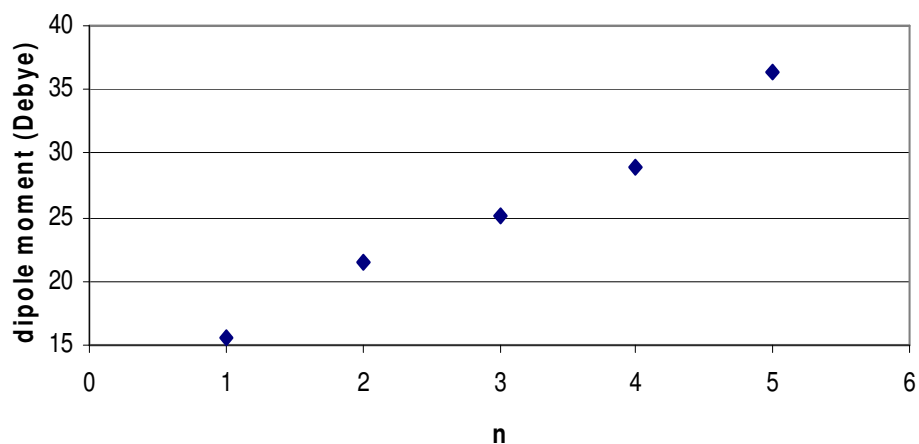


Figure 38. Solution dipole moments vs n.

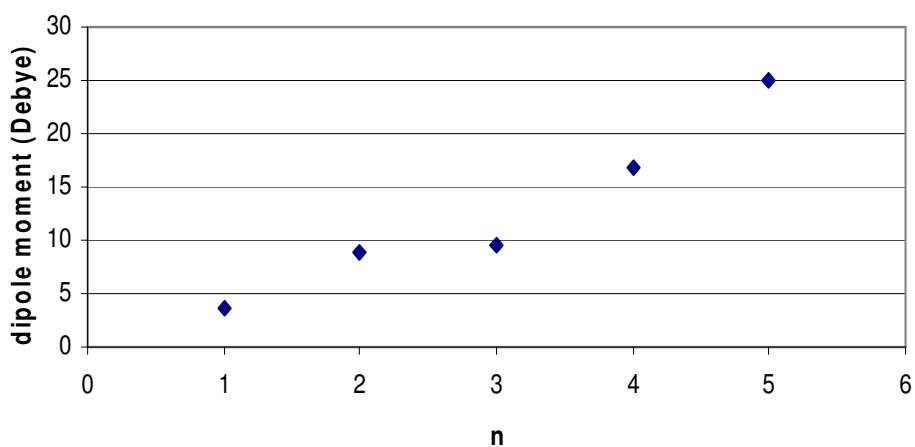


Figure 39. Δ Dipole moments vs n.

There is a linear trend in Figures 38 and 39 for the solution dipole moments and the plot of Δ dipole moments (Δ dipole moment = solution dipole moment – gas-phase dipole moment). Because the solution dipole moment graph is linear and the solvation energy curve is parabolic, it can be presumed that the dipole moment is not the major factor that contributes to the solvation energy.

4.2 Distances

4.2.1 Intercharge Distances. The nitrogen to carboxyl carbon intercharge distances are in Table 20. The experimental values can be compared to the calculated values by their ratios.

Table 20. Nitrogen to carbon intercharge distances, experimental²⁶ and calculated (distances in Å) in water.

n	Maximum Distance (exp.)	Overall Distance (exp.)	Ratio (exp.)	Maximum Distance (calc'd)	Overall Distance (calc'd)	Ratio (calc'd)
1	2.48	--	--	2.607	--	--
2	3.81	3.81	1	3.892	3.802	0.98
3	4.98	4.50	0.90	5.108	4.621	0.90
4	6.30	5.40	0.86	6.406	5.671	0.89
5	7.51	6.15	0.82	7.622	7.021	0.92

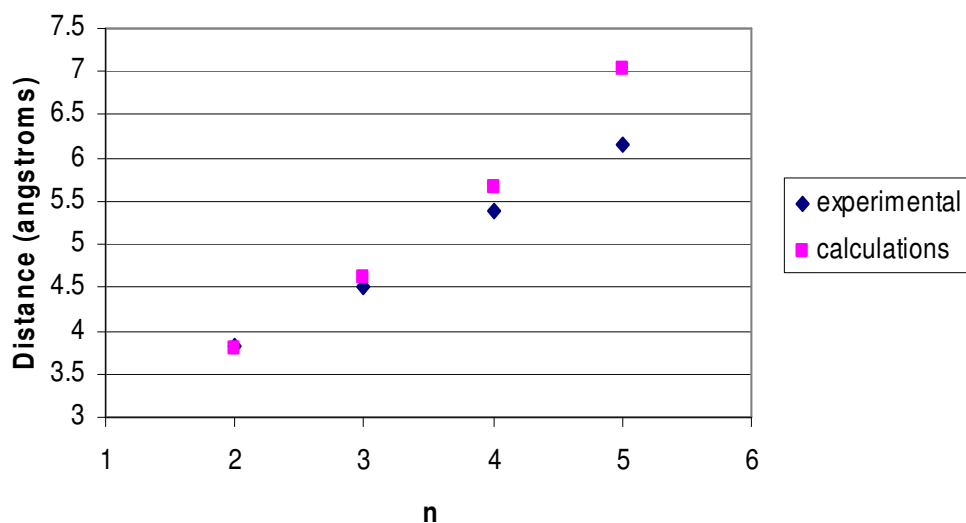


Figure 40. Overall Nitrogen to Carboxylate Carbon Distances.

The intercharge distances measured from the nitrogen to a point symmetrically between the oxygen atoms are tabulated in Table 21. The experimental values can again be compared to the calculated values by observing the ratios between the maximum and overall distances.

Table 21. Intercharge distances in Å (from nitrogen to point symmetrically between oxygen atoms) including experimental²⁷ and theoretical values in water.

n	Maximum Distance (exp.)	Overall Distance (exp.)	Ratio (exp.)	Maximum Distance (calc'd)	Overall Distance (calc'd)	Ratio (calc'd)
1	--	--	--	3.098	--	--
2	--	--	--	4.410	4.321	0.98
3	5.45	4.49	0.82	5.611	5.123	0.91
4	6.93	5.43	0.78	6.932	6.174	0.89
5	7.96	6.00	0.75	8.071	7.500	0.93

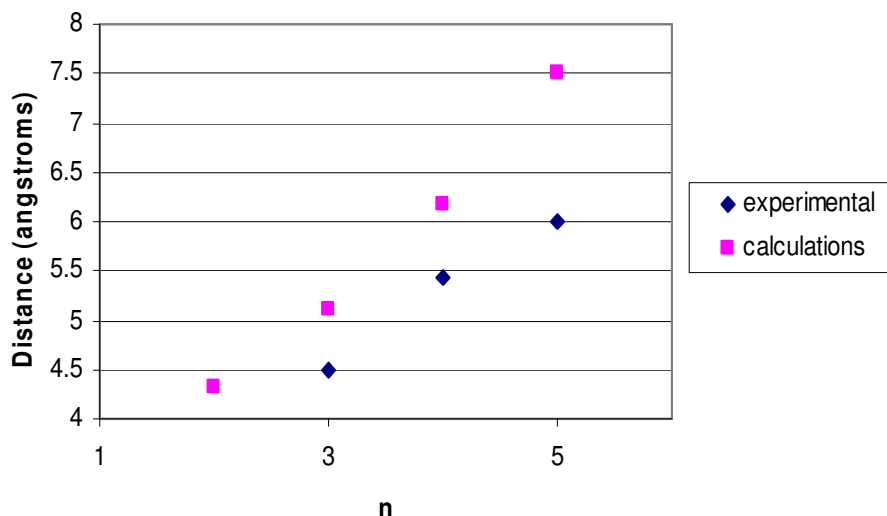


Figure 41. Overall Nitrogen to Negative Charge Distances.

The calculated intercharge distances did not correlate well with experiment. For $n = 2$ and $n = 3$ in Figure 40, the distances were predicted well. However, as n increases, the correlation between the two methods begins to drop off. Even though, they still both show a linear trend. The correlation decrease is especially seen in Figure 41. This does not prove the calculation method is poor. The experimentally measured distances were predicted to be an underestimation of the intercharge distances due to assumptions made in the models used.^{26, 27}

4.2.2 Nitrogen to Oxygens Distances. The nitrogen to oxygens distances are in Tables 22 and 23. Figure 42 illustrates these distances for $n = 2$ anti in the gas phase. The calculated distances are compared to the literature values.⁶²

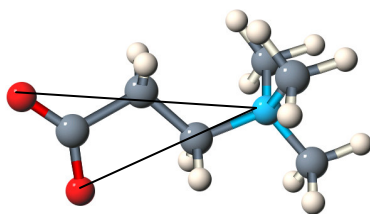


Figure 42. N – O distances for n = 2 anti (gas phase).

Table 22. n = 2 Nitrogen to oxygens distances (Å).

	Literature ⁶²		Calculated	
	Gas phase			
a	4.431	4.647	4.107	4.921
g+	2.948	4.247	3.365	3.549
Solution				
a	4.623	5.130	4.218	4.842
g+	3.509	4.773	3.671	3.930

Table 23. n = 3 Nitrogen to oxygens distances (Å).

	Literature ⁶²		Calculated	
	Gas phase			
aa	5.182	5.972	5.019	6.166
ag+	4.866	5.488	5.262	5.549
g+g+	3.370	4.879	3.196	4.309
g+a	3.876	5.591	3.421	5.325
Solution				
aa	5.182	5.972	5.289	6.106
ag+	5.277	5.829	5.267	5.532
g+g+	3.504	5.055	4.026	4.932
g+a	4.086	5.760	3.856	5.398

4.3 Summary

Therefore, solvation energy can be used as a molecular descriptor to calculate the *cmc* of a surfactant. The hydrophobic effects are accounted for with the slope in the Kleven's equation and the hydrophilic effects are accounted for by calculated solvation energies. Because the solvation energies did not follow the *cmc* curve exactly, either the solvation energy of the micelles needs to be calculated or the surfactant head group model used in this project was insufficient. For example, changing the tail from a methyl to an ethyl group could give a better

basis for the head group hydrophilicity and structure. This extra methyl would add more conformations to examine.

Also, the solvation energy appears to be affected more from the gas-phase energies than the solution energies. The gas-phase structures are influenced more by intramolecular interactions, which significantly lower the energy and dipole moment of the structures. The calculations appear to reproduce the trend of the intercharge distances of the betaine series fairly well. The experimentally determined distances from the literature were based on fixed bond lengths and angles. The calculated distances in this work were determined from bond lengths and angles that were allowed to optimize in solution.

4.4 Future Work

Based on the calculated results for the betaines, a relationship can be found between the *cmc*'s and the solvation energies. Both curves show a parabolic trend. Mathematical correlations can be determined to link the two properties. These correlations are beyond the scope of this project.

Once all the solvation energy calculations are completed through $n = 6$, the next step will apply a Quantitative Structure Property Relationship (QSPR) approach to determine the *cmc*. Solvation energy can be the molecular descriptor to describe the hydrophilic effects and chain length (n) can be the molecular descriptor to describe the hydrophobic effects. Thus equation 1 becomes

$$\log(\text{cmc}) = a(\text{solvation energy}) + b(\text{chain length}) + c(\text{cross-term}) + \text{constant} \quad (32)$$

The cross-term can be estimated based on correlations between the experimentally determined *cmc* data and the solvation energy.

4.5 Final Notes

Calculations give both an energetic and geometrical perspective on problems. In this work, the energies were calculated in the gas phase and in solution. These energies were used to

learn the lowest energy conformations of each molecule, as well as to determine the Boltzmann population for each conformation. These data in turn lead to the numerical final solvation energy for each molecule. The geometrical side of computations aided in understanding why each conformation had the energy that was calculated for it. The geometries explained different effects such as the “hexane” effect and illustrated intramolecular interactions that lead to charge stabilization within the molecule. It is the geometrical perspective that is most interesting because the structures reveal explanations for the energy values that are given only by the computational program. Numbers mean nothing if they are not understood. Both the energetic and geometrical perspectives are required to fully understand, interpret, and solve any problem.

5. Cartesian Coordinates

5.1. Gas-phase Cartesian Coordinates (Å)

n = 1

N	0.977815	-0.059250	0.000005
C	2.124988	-0.999589	-0.000032
H	3.049801	-0.439430	0.000138
C	1.050102	0.798431	1.220811
H	1.984568	1.342400	1.206516
C	1.050244	0.798656	-1.220628
H	1.984685	1.342667	-1.206104
C	-0.313575	-0.858540	-0.000146
C	-1.660461	-0.046399	-0.000023
O	-1.567921	1.180879	-0.000023
O	-2.605820	-0.805929	0.000022
H	-0.285849	-1.489959	-0.876977
H	-0.285887	-1.490278	0.876455
H	2.076501	-1.621151	-0.881703
H	2.076329	-1.621395	0.881458
H	0.203080	1.461177	1.208409
H	1.015364	0.157013	2.089897
H	1.015651	0.157396	-2.089834
H	0.203192	1.461358	-1.208181

n = 2 (anti)

N	1.550419	0.003396	-0.000003
C	0.083315	0.470792	-0.001238
C	-0.967425	-0.619343	-0.000008
C	-2.392656	0.064839	0.000058
O	-2.348275	1.295540	0.000090
H	-0.023997	1.098276	0.868297
H	-0.023230	1.095855	-0.872640
C	1.846838	-0.803777	-1.213010
H	1.586048	-0.227916	-2.088290
C	2.412748	1.214769	-0.002646
H	3.452772	0.919964	-0.001200
C	1.846610	-0.798274	1.216735
H	2.900448	-1.038723	1.235125
H	2.900858	-1.043491	-1.230571
H	1.264941	-1.709459	-1.191424
H	2.196053	1.802534	0.876228
H	2.197209	1.797883	-0.884899
H	1.264053	-1.703617	1.199637
H	1.586538	-0.218084	2.089361
O	-3.315686	-0.728104	-0.000079
H	-0.910112	-1.255952	-0.875732
H	-0.909398	-1.254572	0.876698

n = 2 (g+)

N	1.187795	-0.033939	0.056232
C	0.338423	1.194242	-0.288147
C	-1.010622	0.908776	-0.932454
C	-1.844122	-0.020846	-0.011813
O	-2.178151	0.496522	1.047491
H	0.192122	1.717623	0.644110
H	0.963468	1.791890	-0.938875
C	1.192590	-1.014104	-1.071638
H	1.518003	-0.509605	-1.970423
C	2.572751	0.441775	0.303860
H	3.194936	-0.395385	0.583663
C	0.706685	-0.734903	1.299063
H	1.465494	-1.452023	1.578920
H	1.885277	-1.807725	-0.828534
H	0.193269	-1.404452	-1.187173
H	2.558851	1.167080	1.104339
H	2.962755	0.897903	-0.593805
H	-0.223126	-1.233477	1.092533
H	0.570211	-0.005127	2.080394
O	-1.938902	-1.190456	-0.398720
H	-1.498209	1.874224	-1.014413
H	-0.895418	0.498478	-1.927759

n = 3 (aa)

C	-1.730431	-0.823467	-0.131695
C	-2.999119	0.095289	-0.026525
O	-2.756969	1.304237	-0.091131
C	-0.470746	-0.042029	0.236734
C	0.790948	-0.791236	-0.153099
N	2.090478	0.007678	-0.008905
C	3.230592	-0.887985	-0.341991
H	3.104100	-1.261237	-1.347045
C	2.253028	0.503155	1.387402
H	2.183272	-0.333296	2.066743
C	2.097890	1.167194	-0.948341
H	1.961948	0.802040	-1.955126
O	-4.052711	-0.512718	0.082679
H	-1.865345	-1.699437	0.493828
H	-1.668702	-1.172885	-1.162284
H	-0.557967	0.908973	-0.267088
H	-0.481613	0.157284	1.302780
H	0.931478	-1.678392	0.450610
H	0.758930	-1.089412	-1.192348
H	4.156146	-0.335077	-0.271401
H	3.245644	-1.714827	0.351866
H	3.221652	0.972958	1.480398
H	1.476674	1.215849	1.607619
H	3.047879	1.675344	-0.866276
H	1.297028	1.840695	-0.697235

n = 3 (ag+)

C	-1.310234	0.347972	-0.252049
C	-2.844632	0.099110	-0.013101
O	-3.178886	-1.084036	-0.047142
C	-0.481992	-0.875845	0.122871
C	0.935683	-0.999399	-0.421552
N	1.956222	0.062561	0.009032
C	3.318443	-0.465849	-0.274776
H	3.381890	-0.732967	-1.319118
C	1.843320	0.350770	1.467276
H	1.919445	-0.575858	2.015961
C	1.778586	1.330185	-0.762098
H	1.899459	1.116634	-1.813314
O	-3.488218	1.125579	0.153718
H	-1.186356	0.580146	-1.310065
H	-1.040203	1.237554	0.309787
H	-0.480594	-1.021057	1.198432
H	-1.016126	-1.731311	-0.273104
H	1.368153	-1.938349	-0.101534
H	0.943899	-0.973566	-1.503102
H	4.052619	0.293221	-0.047006
H	3.496441	-1.339018	0.334207
H	2.647366	1.013980	1.752290
H	0.893248	0.815969	1.667130
H	2.532434	2.035679	-0.443139
H	0.796562	1.727007	-0.582682

n = 3 (g+g+)

C	1.670846	0.964021	0.699369
C	2.022865	-0.408723	0.077931
O	2.752824	-0.380355	-0.899924
C	0.550456	1.609586	-0.150871
C	-0.383095	0.635611	-0.874603
N	-1.463471	-0.093200	-0.059106
C	-1.558821	-1.490917	-0.584319
H	-0.610584	-1.971389	-0.394410
C	-2.761281	0.603304	-0.241897
H	-3.037806	0.583599	-1.285573
C	-1.176877	-0.180033	1.410299
H	-0.242992	-0.706445	1.533972
O	1.397508	-1.374884	0.562006
H	1.364060	0.850572	1.731559
H	2.531416	1.620756	0.674352
H	-0.011259	2.353156	0.409750
H	1.021392	2.143989	-0.969011
H	-0.930869	1.161436	-1.644068
H	0.212033	-0.122304	-1.351063
H	-2.365455	-2.000996	-0.076391
H	-1.763610	-1.448860	-1.644662
H	-3.520863	0.105401	0.342920
H	-2.662098	1.627073	0.088670
H	-1.996166	-0.716522	1.868523
H	-1.130116	0.817749	1.817050

n =3 (g+a)

C	1.881531	1.136745	0.142619
C	2.395413	-0.328799	-0.003994
O	1.516377	-1.187989	-0.231316
C	0.417307	1.325309	-0.269378
C	-0.493597	0.539765	0.677070
N	-1.717607	-0.103786	0.036819
C	-2.355428	-0.960184	1.073121
H	-2.563771	-0.364145	1.949329
C	-1.337208	-0.981723	-1.125651
H	-0.413834	-1.488359	-0.886085
C	-2.680144	0.932857	-0.408501
H	-3.012857	1.502973	0.446983
O	3.593492	-0.469758	0.141766
H	2.032044	1.448915	1.174608
H	2.515002	1.780263	-0.457261
H	0.140421	2.376819	-0.265424
H	0.317018	0.969777	-1.285204
H	0.062420	-0.278456	1.093732
H	-0.886187	1.168761	1.466956
H	-3.275650	-1.370308	0.682628
H	-1.674702	-1.758402	1.326176
H	-2.157950	-1.660423	-1.309590
H	-1.185339	-0.364748	-1.995394
H	-3.526415	0.454263	-0.880987
H	-2.193145	1.587730	-1.113518

n = 4 (aaa)

C	-2.376804	-0.677814	0.000444
C	-3.741330	0.085661	0.000080
O	-3.646070	1.315401	0.000497
C	-1.182454	0.266535	-0.000805
C	0.167746	-0.466629	0.000304
C	1.283368	0.568037	-0.000837
N	2.715488	0.035517	-0.000006
C	2.975532	-0.783557	1.218435
H	2.742265	-0.195436	2.093271
C	3.641363	1.203797	-0.001366
H	3.458909	1.799800	0.879824
C	2.975778	-0.786591	-1.216343
H	2.358526	-1.668485	-1.194030
O	-4.724277	-0.646195	-0.000502
H	-2.352273	-1.329790	0.872017
H	-2.352737	-1.331721	-0.869670
H	-1.249398	0.917397	0.864009
H	-1.249157	0.915014	-0.867456
H	0.229046	-1.102923	0.877382
H	0.229323	-1.105179	-0.875115
H	1.210989	1.199332	0.873926
H	1.211203	1.197193	-0.877159
H	4.017850	-1.067150	1.233214
H	2.358175	-1.665427	1.198256
H	3.459176	1.797524	-0.884146
H	4.662586	0.851739	-0.000757
H	2.742559	-0.200695	-2.092685
H	4.018133	-1.070086	-1.230259

n = 4 (ag+a)

C	1.933353	0.311074	0.231541
C	3.480107	0.269548	-0.012865
O	3.953542	-0.858502	-0.143897
C	1.228334	-0.982949	-0.161344
C	-0.247697	-1.059686	0.271894
C	-1.045739	0.023105	-0.441422
N	-2.510492	0.172413	-0.034624
C	-3.258960	-1.097693	-0.246129
H	-3.134759	-1.412099	-1.271617
C	-3.113780	1.237247	-0.886406
H	-3.049343	0.939843	-1.921929
C	-2.620703	0.585159	1.394780
H	-2.247587	-0.201664	2.027702
O	4.007764	1.376519	-0.023220
H	1.777253	0.510594	1.293308
H	1.550159	1.175311	-0.305714
H	1.765589	-1.805684	0.287853
H	1.298931	-1.136671	-1.235556
H	-0.649642	-2.040912	0.035877
H	-0.298065	-0.930581	1.348347
H	-1.059105	-0.162731	-1.507028
H	-0.605335	0.992450	-0.272512
H	-4.305429	-0.927128	-0.038526
H	-2.875363	-1.855892	0.415446
H	-2.568427	2.156976	-0.739468
H	-4.148403	1.375019	-0.608104
H	-2.037943	1.480948	1.547191
H	-3.659022	0.776369	1.623733

n = 4 (g+g+a)

C	-2.423599	0.378045	-0.679918
C	-1.884267	-0.881526	0.033472
O	-1.200453	-1.651696	-0.672070
C	-1.752716	1.668904	-0.145550
C	-0.323343	1.474511	0.400132
C	0.583976	0.769716	-0.610966
N	1.731180	-0.040596	-0.015199
C	2.743589	0.859017	0.586209
H	3.122255	1.530790	-0.170958
C	2.337400	-0.810776	-1.136922
H	2.637089	-0.126654	-1.917707
C	1.267965	-1.038511	1.015773
H	0.951048	-0.517786	1.902048
O	-2.077094	-0.944905	1.243974
H	-2.261895	0.279981	-1.747705
H	-3.493024	0.457112	-0.516622
H	-2.335973	2.066739	0.676385
H	-1.754821	2.429854	-0.923080
H	0.094876	2.437824	0.680478
H	-0.425061	0.882592	1.295498
H	1.056334	1.478165	-1.280739
H	0.022416	0.053409	-1.181548
H	3.554198	0.267043	0.986619
H	2.285754	1.427107	1.380755
H	1.595045	-1.500497	-1.508535
H	3.200190	-1.350153	-0.773299
H	0.444411	-1.606762	0.606190
H	2.115248	-1.668059	1.250007

n = 4 (g-g+a)

C	-2.583392	0.391124	0.322984
C	-1.809006	-0.917114	0.083547
O	-1.861090	-1.386152	-1.052898
C	-1.811879	1.582814	-0.279728
C	-0.333172	1.633322	0.164684
C	0.510048	0.684402	-0.703101
N	1.702522	-0.033409	-0.045118
C	2.964616	0.647903	-0.422043
H	3.087717	0.608972	-1.494131
C	1.706461	-1.439121	-0.567854
H	1.751157	-1.404829	-1.646512
C	1.632420	-0.109052	1.449106
H	1.739916	0.883595	1.858390
O	-1.080340	-1.296776	1.019453
H	-3.559359	0.336518	-0.146192
H	-2.714916	0.549886	1.388218
H	-2.290804	2.510316	0.019692
H	-1.869190	1.540259	-1.363782
H	0.054655	2.644970	0.083135
H	-0.298707	1.358356	1.206630
H	0.933442	1.225680	-1.538203
H	-0.109517	-0.094123	-1.105215
H	3.797628	0.152276	0.055567
H	2.921012	1.677870	-0.098109
H	0.790919	-1.912263	-0.243408
H	2.576334	-1.950300	-0.179239
H	0.687998	-0.553587	1.720534
H	2.458928	-0.721970	1.780430

n = 4 (g-g+g+ (180))

C	2.348610	0.618844	-0.249061
C	1.871401	-0.836190	-0.038861
O	1.773564	-1.189821	1.141263
C	1.380511	1.621961	0.402032
C	0.058071	1.857507	-0.361118
C	-0.683297	0.637900	-0.919181
N	-1.596325	-0.150453	0.026489
C	-1.117240	-0.205638	1.445963
H	-0.130849	-0.646872	1.485371
C	-1.663162	-1.552927	-0.496869
H	-0.666611	-1.971307	-0.484712
C	-2.952327	0.451850	0.016821
H	-2.887347	1.480794	0.338384
O	1.531809	-1.470050	-1.047363
H	3.314441	0.731339	0.232885
H	2.474457	0.828463	-1.307180
H	1.857765	2.595946	0.474478
H	1.199863	1.297194	1.417816
H	-0.605607	2.492332	0.222897
H	0.296049	2.433883	-1.251401
H	0.022430	-0.071749	-1.316213
H	-1.340928	0.950345	-1.719071
H	-1.822207	-0.808664	2.000795
H	-1.110068	0.794618	1.849076
H	-2.034739	-1.521900	-1.511231
H	-2.339656	-2.122482	0.124487
H	-3.358645	0.408317	-0.982650
H	-3.592458	-0.097957	0.691290

n = 4 (g-g+g+(90))

C	2.293734	0.869540	-0.058699
C	2.215492	-0.678589	0.027081
O	3.181633	-1.274561	-0.404550
C	1.109765	1.668809	0.497366
C	-0.126736	1.761542	-0.418611
C	-0.734346	0.455923	-0.933916
N	-1.743815	-0.255705	-0.030056
C	-1.426708	-0.164448	1.431697
H	-0.429452	-0.546944	1.573879
C	-1.728649	-1.702810	-0.410775
H	-0.747181	-2.088726	-0.179247
C	-3.096834	0.308068	-0.258165
H	-3.083604	1.366785	-0.046356
O	1.159086	-1.156936	0.510783
H	3.193259	1.160942	0.474807
H	2.481175	1.134387	-1.097941
H	1.426264	2.697110	0.660826
H	0.847152	1.277161	1.470568
H	-0.887760	2.390233	0.040073
H	0.175184	2.291727	-1.318484
H	0.041518	-0.258970	-1.117367
H	-1.283285	0.636908	-1.848808
H	-2.161192	-0.756005	1.960560
H	-1.506597	0.864978	1.743069
H	-1.932818	-1.782913	-1.469217
H	-2.497380	-2.218094	0.148092
H	-3.385808	0.147241	-1.286218
H	-3.802832	-0.182123	0.396417

n = 4 (g+ag+)

C	-2.421093	0.923549	0.201620
C	-2.393577	-0.615437	0.015817
O	-3.438570	-1.131406	-0.324174
C	-1.053458	1.548495	0.472217
C	-0.088187	1.360271	-0.714896
C	1.396021	1.182499	-0.388774
N	1.887171	-0.199434	0.015188
C	1.388584	-0.611942	1.369616
H	1.640275	0.165365	2.077215
C	3.371380	-0.137391	0.056233
H	3.678801	0.624350	0.758009
C	1.469668	-1.239922	-0.982330
H	0.407103	-1.408854	-0.881072
O	-1.285066	-1.168829	0.230892
H	-2.876249	1.360831	-0.683553
H	-3.097528	1.146430	1.022277
H	-0.648572	1.111851	1.374879
H	-1.157455	2.613465	0.671510
H	-0.422408	0.547467	-1.339714
H	-0.120738	2.250321	-1.338621
H	1.701747	1.845910	0.410442
H	1.982976	1.426785	-1.264492
H	1.895837	-1.527182	1.640769
H	0.323875	-0.782228	1.301429
H	3.750455	0.098494	-0.927254
H	3.757446	-1.094991	0.371296
H	1.735311	-0.893186	-1.971418
H	2.011992	-2.147631	-0.758781

n = 5 (g+ag+g+)

C	2.840153	-0.095351	-0.088278
C	1.760003	-1.181256	0.009960
O	1.166026	-1.453931	-1.050696
C	2.195802	1.262043	-0.423166
C	1.089750	1.692725	0.577025
C	-0.288194	1.941044	-0.067354
H	-0.171855	2.718052	-0.817656
O	1.491636	-1.606397	1.136995
H	3.375868	-0.033863	0.852397
H	3.545973	-0.348400	-0.873477
H	1.792178	1.205113	-1.429157
H	2.967152	2.026160	-0.454482
H	1.012508	0.954646	1.364491
H	1.377019	2.620649	1.063407
H	-0.975467	2.354138	0.667045
C	-0.907967	0.746161	-0.805844
N	-1.860004	-0.147258	-0.013669
C	-1.368304	-0.483423	1.361371
H	-1.284949	0.425649	1.935632
C	-1.986994	-1.429441	-0.778942
H	-1.005369	-1.878866	-0.841423
C	-3.188929	0.500740	0.096997
H	-3.091690	1.437906	0.623497
H	-1.505274	1.088794	-1.640463
H	-0.139245	0.087125	-1.176356
H	-0.414609	-0.992187	1.302774
H	-2.106699	-1.126474	1.819875
H	-2.687396	-2.070907	-0.262885
H	-2.358998	-1.202810	-1.768270
H	-3.857655	-0.149816	0.641663
H	-3.584688	0.679059	-0.891940

n = 5 (g+g-g+a)

C	-2.601214	-0.426428	-0.520004
C	-1.493184	-1.310099	0.082185
O	-0.776077	-1.932059	-0.720238
C	-2.653299	0.989950	0.069377
C	-1.490880	1.945001	-0.269495
C	-0.117732	1.590239	0.350637
H	-0.295556	1.056838	1.271610
O	-1.363473	-1.265318	1.309727
H	-2.494060	-0.396897	-1.600331
H	-3.553417	-0.907638	-0.308973
H	-2.735457	0.906008	1.146870
H	-3.562850	1.474869	-0.279411
H	-1.395726	2.043176	-1.350539
H	-1.766877	2.931447	0.091585
H	0.419868	2.505249	0.586662
C	0.702310	0.739921	-0.618075
N	1.921551	0.029657	-0.038759
C	1.647311	-0.620039	1.287806
H	0.730853	-1.191315	1.238720
C	3.050229	0.978077	0.109457
H	2.757021	1.783584	0.765213
C	2.288094	-1.050588	-1.006299
H	3.201059	-1.522804	-0.671739
H	0.103357	-0.047455	-1.034867
H	1.087121	1.349397	-1.426955
H	1.560643	0.147463	2.040093
H	2.494785	-1.250650	1.517419
H	3.898465	0.459224	0.532720
H	3.315163	1.375152	-0.859711
H	2.442638	-0.606431	-1.979531
H	1.468704	-1.754001	-1.036973

n = 5 (g+g-g-a)

C	-2.751144	-0.262448	0.016870
C	-1.588801	-1.272130	-0.057841
O	-0.975251	-1.477758	1.006974
C	-2.355501	1.078958	0.654037
C	-1.442317	1.994252	-0.187920
C	-0.128727	1.379022	-0.709609
H	0.402589	2.128936	-1.292099
O	-1.316559	-1.737032	-1.166582
H	-3.160654	-0.113126	-0.976669
H	-3.531522	-0.702589	0.632367
H	-1.904644	0.875845	1.618918
H	-3.256512	1.652115	0.859978
H	-1.995236	2.333474	-1.059605
H	-1.218273	2.887874	0.393100
H	-0.382091	0.564675	-1.370596
C	0.727919	0.871448	0.446105
N	1.930043	0.006711	0.067189
C	3.095404	0.862847	-0.256832
H	3.364698	1.446946	0.610983
C	1.662055	-0.910477	-1.091608
H	1.584766	-0.321248	-1.991347
C	2.235534	-0.845286	1.256140
H	2.389345	-0.204522	2.112855
H	1.135990	1.691151	1.025041
H	0.129206	0.241921	1.076433
H	2.836727	1.521191	-1.072370
H	3.928718	0.238956	-0.546697
H	2.508956	-1.577430	-1.173784
H	0.745192	-1.461695	-0.933713
H	3.133160	-1.414503	1.059489
H	1.387229	-1.493739	1.418204

n = 6 (g-g+g+aa)

C	-2.818504	-0.846748	0.275414
C	-1.410672	-1.464103	0.151295
O	-1.077865	-1.802205	-0.990099
C	-2.997920	0.386052	-0.630580
C	-2.491403	1.724983	-0.057731
C	-1.122051	1.699595	0.632469
C	0.055639	1.279895	-0.268184
C	1.188630	0.792378	0.630484
N	2.325908	0.052382	-0.056004
C	1.832590	-1.214081	-0.705944
H	1.221414	-1.754739	-0.001296
O	-0.710217	-1.526200	1.178313
C	3.303597	-0.324600	0.998981
H	2.809856	-0.967335	1.711437
C	2.994292	0.909156	-1.063787
H	2.294344	1.149835	-1.846924
H	-3.523265	-1.606676	-0.050085
H	-3.041789	-0.613529	1.311905
H	-4.053747	0.522300	-0.853233
H	-2.518445	0.168156	-1.577515
H	-2.491256	2.469050	-0.853407
H	-3.206717	2.078673	0.681761
H	-1.179594	1.011763	1.466077
H	-0.920124	2.684938	1.048999
H	-0.259935	0.483138	-0.923436
H	0.368128	2.119522	-0.884329
H	0.781258	0.084065	1.332377
H	1.654616	1.618677	1.154926
H	2.701657	-1.775755	-1.020296
H	1.213255	-0.973922	-1.549763
H	3.662062	0.566936	1.492809
H	4.132282	-0.849160	0.545984
H	3.341110	1.816976	-0.590246
H	3.832987	0.372482	-1.483922

n = 6 (g+g-g+aa)

C	-2.510558	-1.040371	0.598244
C	-1.218242	-1.463936	-0.119033
O	-0.291589	-1.881053	0.602190
C	-3.207154	0.190759	0.005592
C	-2.562959	1.569216	0.252196
C	-1.224949	1.846589	-0.461252
C	0.024624	1.382044	0.326617
C	1.002340	0.701846	-0.623817
N	2.258393	0.088962	0.012230
C	3.393930	1.033522	-0.107546
H	3.600042	1.216414	-1.151714
O	-1.181882	-1.325802	-1.344249
C	2.556145	-1.171026	-0.738827
H	2.658960	-0.931021	-1.787257
C	2.084075	-0.272603	1.455333
H	1.941926	0.629636	2.029548
H	-3.196658	-1.882108	0.524059
H	-2.304919	-0.906352	1.656251
H	-4.212920	0.242062	0.418839
H	-3.314835	0.037929	-1.062089
H	-3.276563	2.315058	-0.089842
H	-2.448434	1.738853	1.322529
H	-1.137270	2.913442	-0.647790
H	-1.250257	1.362262	-1.430206
H	0.479272	2.230389	0.833705
H	-0.272975	0.675645	1.084546
H	1.368696	1.396366	-1.369030
H	0.487414	-0.102646	-1.120113
H	4.268021	0.608091	0.364497
H	3.134856	1.962539	0.379339
H	1.724257	-1.842364	-0.582777
H	3.480080	-1.590543	-0.364725
H	1.235747	-0.933703	1.538683
H	2.991068	-0.764081	1.779347

n = 6 (g-g+g+ag+)

C	-2.929747	-0.769048	0.101849
C	-1.545412	-1.430055	0.024200
O	-1.020762	-1.787048	1.086326
C	-2.946392	0.639219	-0.511181
C	-2.248056	1.736855	0.311893
C	-0.766252	1.521335	0.660121
C	0.134805	1.405619	-0.576236
C	1.641877	1.327948	-0.311874
N	2.266489	-0.021219	-0.000883
C	1.786346	-0.595205	1.300449
H	1.850242	0.173803	2.057216
O	-1.041330	-1.495770	-1.110557
C	3.733182	0.199200	0.097822
H	3.938012	0.898010	0.895922
C	1.999631	-1.013882	-1.095616
H	0.947566	-1.276819	-1.116844
H	-3.622834	-1.388873	-0.461090
H	-3.265739	-0.747419	1.133409
H	-3.979288	0.955876	-0.640558
H	-2.520227	0.580388	-1.506282
H	-2.350346	2.681719	-0.221400
H	-2.785887	1.861089	1.249196
H	-0.682086	0.644344	1.283978
H	-0.434170	2.370540	1.258054
H	-0.183359	0.583108	-1.198376
H	0.014651	2.311214	-1.167916
H	1.923157	1.973839	0.510632
H	2.170199	1.666336	-1.193241
H	2.445175	-1.411932	1.559413
H	0.778320	-0.978673	1.201233
H	4.103045	0.593841	-0.837039
H	4.218643	-0.741428	0.308675
H	2.313679	-0.576004	-2.033405
H	2.592659	-1.893795	-0.890103

n = 6 (g-g+g+g-g-)

C	-2.788242	-0.802497	-0.163075
C	-1.460379	-1.491047	0.177801
O	-0.741337	-1.774686	-0.801313
C	-2.591412	0.553355	-0.864303
C	-2.187382	1.727599	0.044754
C	-0.839939	1.681513	0.789681
C	0.421167	1.912718	-0.072314
C	1.009027	0.702162	-0.810533
N	2.092287	-0.083033	-0.075686
C	1.698500	-0.453639	1.322031
H	1.576220	0.448224	1.901098
O	-1.173408	-1.645426	1.367167
C	3.359048	0.685892	-0.039049
H	4.117681	0.093042	0.450998
C	2.307877	-1.352779	-0.842145
H	2.595080	-1.099058	-1.852992
H	-3.331618	-1.453935	-0.841389
H	-3.382205	-0.680999	0.736640
H	-3.533067	0.839814	-1.327476
H	-1.888820	0.426172	-1.681888
H	-2.222889	2.645153	-0.542494
H	-2.959352	1.832060	0.803452
H	-0.772243	0.762896	1.354581
H	-0.863566	2.487225	1.519688
H	0.181950	2.642903	-0.840611
H	1.185978	2.389384	0.534442
H	1.492007	1.021485	-1.725220
H	0.253761	-0.026429	-1.052502
H	2.504929	-1.038923	1.741950
H	0.780279	-1.027757	1.312172
H	3.216198	1.604260	0.508033
H	3.671585	0.908870	-1.048764
H	3.103179	-1.909021	-0.366104
H	1.377281	-1.902894	-0.837727

n = 6 (g-g+g-g+a)

C	2.744913	-1.059810	-0.195117
C	1.361965	-1.515483	0.276258
O	1.050362	-1.267813	1.447304
C	2.655136	0.345473	-0.820681
C	2.322287	1.478400	0.185126
C	1.037667	2.298434	-0.053685
C	-0.280771	1.672364	0.452379
C	-0.880617	0.703837	-0.560561
N	-2.128223	-0.042008	-0.104151
C	-1.957730	-0.644300	1.260436
H	-1.005875	-1.157453	1.312431
O	0.620033	-1.996216	-0.600624
C	-2.345177	-1.160093	-1.075547
H	-1.492216	-1.820276	-1.010692
C	-3.308868	0.852371	-0.105892
H	-3.471784	1.228909	-1.105471
H	3.438082	-1.044058	0.639369
H	3.113009	-1.752985	-0.943616
H	1.930913	0.318308	-1.630745
H	3.606976	0.570727	-1.293516
H	2.285208	1.068907	1.187696
H	3.143672	2.189561	0.177632
H	0.949777	2.564211	-1.106341
H	1.152075	3.239085	0.477346
H	-0.073496	1.157846	1.379394
H	-0.984257	2.475219	0.660365
H	-0.185963	-0.071926	-0.823047
H	-1.177700	1.229067	-1.460829
H	-2.785733	-1.319962	1.423494
H	-1.991901	0.144876	1.995126
H	-2.427780	-0.742192	-2.069139
H	-3.263487	-1.668520	-0.816943
H	-4.176279	0.294014	0.216216
H	-3.141669	1.675769	0.570590

n = 6 (g-g+ag+g+)

C	2.939539	-0.797980	-0.334123
C	1.570678	-1.381009	0.037062
O	1.333323	-1.531353	1.239714
C	3.062445	0.685644	0.062121
C	1.901838	1.567564	-0.444552
C	0.750146	1.700364	0.574709
C	-0.626272	1.952435	-0.063492
C	-1.207268	0.743066	-0.811768
N	-2.154459	-0.156461	-0.024681
C	-1.620248	-0.541344	1.322631
H	-0.660482	-1.033051	1.221536
O	0.775423	-1.577807	-0.902907
C	-2.324009	-1.413793	-0.820999
H	-1.354845	-1.886134	-0.896312
C	-3.471314	0.502951	0.142061
H	-3.353040	1.419399	0.698718
H	3.715451	-1.359380	0.177189
H	3.089659	-0.901559	-1.403343
H	4.000761	1.066094	-0.333024
H	3.129270	0.763259	1.143207
H	2.267011	2.564356	-0.680731
H	1.537884	1.154086	-1.380646
H	0.970469	2.531493	1.240350
H	0.719836	0.817873	1.195686
H	-1.323704	2.326347	0.681228
H	-0.534830	2.750957	-0.794792
H	-0.420535	0.089112	-1.151074
H	-1.794022	1.063756	-1.662821
H	-2.339548	-1.209738	1.775585
H	-1.531245	0.345827	1.929586
H	-2.700707	-1.153364	-1.800280
H	-3.036842	-2.048341	-0.313300
H	-3.888895	0.719796	-0.830358
H	-4.133610	-0.159676	0.680012

n = 6 (g+ag+g+a)

C	-2.824671	-0.978895	-0.088637
C	-1.394418	-1.516425	0.036530
O	-0.776856	-1.760605	-1.013650
C	-2.878242	0.509191	0.299216
C	-2.096077	1.440760	-0.662781
C	-0.990240	2.294764	-0.015249
C	0.205056	1.530296	0.584047
C	0.941956	0.753186	-0.501480
N	2.162420	-0.041910	-0.050163
C	3.350636	0.837754	0.041551
H	3.562755	1.257746	-0.930846
O	-0.941898	-1.574106	1.189046
C	2.391302	-1.099637	-1.083295
H	2.530031	-0.620388	-2.042256
C	1.950043	-0.725829	1.268285
H	1.921690	0.021770	2.045804
H	-3.186585	-1.118369	-1.101739
H	-3.466355	-1.536329	0.586790
H	-2.503126	0.604376	1.312517
H	-3.916481	0.827919	0.330183
H	-1.683215	0.855374	-1.477287
H	-2.789252	2.134963	-1.130051
H	-1.427714	2.889822	0.782064
H	-0.620780	3.006054	-0.752800
H	0.862815	2.249034	1.068758
H	-0.154060	0.842016	1.335038
H	1.306585	1.416984	-1.276715
H	0.286390	0.023603	-0.941133
H	4.198066	0.257808	0.378186
H	3.155432	1.631916	0.745913
H	1.516027	-1.733436	-1.108228
H	3.280010	-1.656310	-0.820923
H	1.014723	-1.273282	1.242471
H	2.794077	-1.381203	1.433105

5.2. Solution Cartesian Coordinates (Å)

n = 1

H	-0.288400	1.533277	1.240240
H	-1.019142	0.164181	2.089318
H	-1.018401	0.165412	-2.089172
H	-0.287841	1.533999	-1.239007
H	-2.034111	-1.626036	-0.884151
H	-2.035059	-1.625898	0.883388
H	0.291195	-1.467385	-0.870847
H	0.291204	-1.466949	0.871533
O	2.618967	-0.882893	0.000033
O	1.687750	1.146700	-0.000346
C	1.664540	-0.089429	-0.000069
C	0.316110	-0.827331	0.000092
H	-2.043989	1.289817	1.200808
C	-1.078394	0.807766	1.224359
H	-2.043464	1.290435	-1.200309
C	-1.078155	0.808448	-1.223872
H	-3.017876	-0.454635	-0.001100
C	-2.093173	-1.010568	-0.000617
N	-0.952233	-0.037136	0.000075

n = 2 (a(180))

N	0.005498	-0.000031	0.021173
C	-0.042600	0.000223	1.523742
C	1.306013	0.000087	2.232266
C	1.171245	0.000206	3.757571
O	0.021802	-0.000419	4.228679
H	-0.610646	0.870633	1.811394
H	-0.609165	-0.870538	1.812156
C	0.683563	-1.214637	-0.517979
H	0.210255	-2.093215	-0.108783
C	-1.408976	0.000098	-0.468048
H	-1.402968	-0.001498	-1.546976
C	0.683218	1.215047	-0.517792
H	0.579103	1.210062	-1.591991
H	0.580190	-1.208907	-1.592369
H	1.727320	-1.197881	-0.255855
H	-1.906589	0.884002	-0.102423
H	-1.907846	-0.882274	-0.099894
H	1.727605	1.197514	-0.256814
H	0.210519	2.093135	-0.107466
O	2.248261	0.000781	4.382836
H	1.899525	-0.869160	1.971957
H	1.902802	0.866862	1.971812

n = 2 (a (90))

N	0.820049	-0.766317	-0.793223
C	-0.213207	-0.020792	-0.001387
C	0.308054	1.047207	0.957504
C	-0.904482	1.679591	1.645165
O	-1.280388	1.140589	2.701840
H	-0.876032	0.436153	-0.720593
H	-0.763955	-0.764547	0.554805
C	1.763154	-1.516173	0.087795
H	1.197546	-2.155386	0.747085
C	0.092831	-1.753937	-1.652242
H	0.813749	-2.305341	-2.234177
C	1.602141	0.141997	-1.682785
H	2.246950	-0.466027	-2.298437
H	2.407054	-2.111779	-0.540783
H	2.355993	-0.823421	0.659616
H	-0.579917	-1.221121	-2.305360
H	-0.462963	-2.426749	-1.018484
H	2.197336	0.811904	-1.086923
H	0.919998	0.701367	-2.303350
O	-1.429831	2.639439	1.052368
H	0.949222	0.613257	1.712892
H	0.842930	1.821171	0.423772

n = 2 (g+(90))

N	-0.007830	-0.002668	-0.026658
C	0.022249	0.022456	1.483281
C	1.391596	0.052341	2.155144
C	2.217257	1.315538	1.909855
O	1.721227	2.377492	2.328589
H	-0.563167	0.876265	1.785928
H	-0.491264	-0.875784	1.795047
C	0.954302	-0.994908	-0.588069
H	0.761931	-1.958407	-0.141753
C	-1.392239	-0.412171	-0.430199
H	-1.461037	-0.393578	-1.506175
C	0.265656	1.335743	-0.623565
H	0.097682	1.264746	-1.687091
H	0.794261	-1.050303	-1.654054
H	1.960344	-0.675733	-0.381561
H	-2.100434	0.278802	-0.001006
H	-1.580314	-1.410943	-0.068666
H	1.282633	1.622657	-0.430478
H	-0.411709	2.056435	-0.193486
O	3.293235	1.170898	1.298827
H	1.176205	0.005581	3.218337
H	1.968733	-0.827056	1.904361

n = 2 (g+(180))

N	-0.656956	0.498270	-0.649160
C	-0.389086	-0.791915	0.090462
C	1.021540	-1.044848	0.606953
C	1.486096	-0.319109	1.871641
O	0.607326	0.111315	2.636190
H	-1.097505	-0.830289	0.902622
H	-0.631075	-1.569910	-0.618979
C	0.410990	0.810441	-1.640933
H	0.566637	-0.049163	-2.274374
C	-1.942092	0.310429	-1.399403
H	-2.199016	1.236787	-1.888510
C	-0.832670	1.663619	0.263211
H	-1.092892	2.518072	-0.342603
H	0.082925	1.647530	-2.236966
H	1.322474	1.065394	-1.127110
H	-2.715978	0.033487	-0.701592
H	-1.808297	-0.466975	-2.135127
H	0.073982	1.844144	0.808111
H	-1.626564	1.446676	0.958695
O	2.721105	-0.273489	2.033796
H	1.039795	-2.102866	0.860860
H	1.768732	-0.930858	-0.166016

n = 3 (aa(180))

C	0.042442	0.000113	0.021736
C	-0.071964	-0.000013	1.545676
O	0.982526	0.000120	2.203684
C	1.469055	-0.000526	-0.521752
C	1.461292	-0.000577	-2.048060
N	2.819767	-0.001235	-2.697772
C	2.618039	-0.001201	-4.180289
H	2.068609	0.883296	-4.460426
C	3.601279	-1.218930	-2.336541
H	3.016475	-2.092647	-2.577696
C	3.602111	1.216086	-2.336672
H	3.018088	2.090102	-2.578037
O	-1.236628	0.000169	1.994631
H	-0.511670	-0.864749	-0.337671
H	-0.510839	0.865090	-0.337555
H	1.983320	0.867413	-0.132933
H	1.982684	-0.869141	-0.132889
H	0.950144	-0.875531	-2.424654
H	0.950864	0.874609	-2.424660
H	3.584125	-0.001615	-4.659920
H	2.067946	-0.885399	-4.460291
H	4.513167	-1.218456	-2.913047
H	3.835977	-1.206430	-1.286777
H	4.514145	1.214749	-2.913203
H	3.836824	1.203513	-1.286950

n = 3 (aa(90))

C	-2.004726	0.282018	0.951001
C	-2.522915	0.860217	2.262758
O	-2.770746	2.081767	2.271389
C	-0.473633	0.361406	0.867890
C	-0.008484	-0.238664	-0.458954
N	1.476360	-0.240918	-0.702314
C	1.717879	-0.865198	-2.042765
H	1.214297	-0.283255	-2.798326
C	2.198807	-1.047282	0.323559
H	1.772779	-2.038005	0.353903
C	2.030130	1.143804	-0.727391
H	1.482701	1.726519	-1.452064
O	-2.613486	0.070004	3.222807
H	-2.320101	-0.753599	0.877670
H	-2.444095	0.838240	0.129279
H	-0.178246	1.400361	0.953798
H	-0.054843	-0.176114	1.710232
H	-0.325687	-1.269354	-0.528801
H	-0.446404	0.305833	-1.283687
H	2.779823	-0.874500	-2.233104
H	1.336759	-1.873910	-2.035771
H	3.238207	-1.104106	0.039802
H	2.111314	-0.576675	1.287417
H	3.069253	1.088602	-1.011694
H	1.944105	1.589769	0.248297

n = 3 (ag+(180))

C	0.031380	0.032223	0.056020
C	0.004439	-0.180113	1.570504
O	1.089055	-0.349173	2.154570
C	1.410538	0.324082	-0.534865
C	1.569311	0.304425	-2.056696
N	0.832844	1.339784	-2.876213
C	1.542437	1.460875	-4.191410
H	1.610551	0.483776	-4.642604
C	0.838397	2.681120	-2.229574
H	1.855042	2.953282	-1.994530
C	-0.570621	0.920737	-3.167399
H	-0.544725	0.001899	-3.731775
O	-1.131964	-0.176802	2.086345
H	-0.396986	-0.865446	-0.388565
H	-0.680512	0.824648	-0.143903
H	1.789737	1.255439	-0.132263
H	2.092071	-0.436848	-0.177872
H	2.617335	0.461310	-2.270071
H	1.287292	-0.654875	-2.467108
H	0.979107	2.123975	-4.828744
H	2.528152	1.862280	-4.022129
H	0.422788	3.392425	-2.926423
H	0.244579	2.658768	-1.332664
H	-1.042248	1.693220	-3.753576
H	-1.110804	0.773597	-2.250572

n = 3 (ag+(90))

C	-1.624253	0.881348	0.762725
C	-2.211643	2.281803	0.881986
O	-2.190680	2.979314	-0.151041
C	-0.117198	0.904570	1.055235
C	0.590749	-0.438802	1.240184
N	0.668175	-1.376889	0.059527
C	1.795580	-2.329704	0.318273
H	1.637335	-2.805406	1.273362
C	0.955626	-0.660597	-1.215112
H	1.837125	-0.052996	-1.084636
C	-0.572914	-2.193774	-0.085741
H	-0.703526	-2.782561	0.808596
O	-2.621908	2.620971	2.009620
H	-2.128424	0.226982	1.466913
H	-1.817842	0.531348	-0.243565
H	0.388592	1.501646	0.306432
H	0.030397	1.416432	1.999738
H	1.615141	-0.232556	1.515722
H	0.137279	-1.003316	2.042303
H	1.807922	-3.070171	-0.465891
H	2.724527	-1.782597	0.323662
H	1.132388	-1.399083	-1.981308
H	0.117131	-0.043718	-1.487810
H	-0.449430	-2.846468	-0.935731
H	-1.424081	-1.556108	-0.234358

n = 3 (g+g+(90))

C	0.232256	-0.002471	-0.021736
C	0.333336	0.120819	1.496288
O	1.460531	-0.064955	1.997768
C	0.144929	-1.471523	-0.481263
C	-0.928501	-2.343682	0.181128
N	-2.305287	-2.339578	-0.431939
C	-3.230604	-3.004521	0.536290
H	-3.306424	-2.397300	1.423559
C	-2.319116	-3.129790	-1.701039
H	-2.055091	-4.151746	-1.478777
C	-2.822537	-0.970709	-0.710128
H	-2.731060	-0.369249	0.177250
O	-0.727665	0.347404	2.110522
H	-0.617424	0.572464	-0.362176
H	1.123519	0.426103	-0.466729
H	0.063389	-1.504453	-1.562417
H	1.091421	-1.941298	-0.242665
H	-0.616930	-3.377975	0.148092
H	-1.051669	-2.058647	1.211775
H	-4.198806	-3.109919	0.072185
H	-2.836201	-3.977323	0.785077
H	-3.313424	-3.090102	-2.117892
H	-1.615046	-2.708181	-2.398442
H	-3.857608	-1.060380	-1.003040
H	-2.261118	-0.535888	-1.519993

n = 3 (g+g+(180))

C	-0.688444	-1.204860	2.038279
C	-2.193491	-1.420102	1.860097
O	-2.555189	-2.206600	0.965276
C	0.179511	-1.464402	0.802433
C	-0.162816	-0.659745	-0.456327
N	0.495516	0.687020	-0.621353
C	-0.173580	1.378853	-1.765661
H	-1.202719	1.571944	-1.507827
C	1.943858	0.532327	-0.960463
H	2.028036	-0.006875	-1.890963
C	0.373514	1.549149	0.585792
H	-0.660867	1.609854	0.880388
O	-2.928987	-0.799041	2.652594
H	-0.543527	-0.216806	2.453269
H	-0.373980	-1.893593	2.820981
H	1.224630	-1.345683	1.067972
H	0.057317	-2.502599	0.525209
H	0.134505	-1.218371	-1.331911
H	-1.226533	-0.501457	-0.511116
H	0.341158	2.306875	-1.958874
H	-0.124412	0.743790	-2.635696
H	2.377369	1.514741	-1.065241
H	2.445992	-0.004703	-0.174420
H	0.746012	2.530584	0.335558
H	0.965714	1.135152	1.384344

n = 3 (g+a(180))

C	0.030999	-0.038210	0.083432
C	0.092745	0.423024	1.544668
O	1.157738	0.912290	1.955949
C	1.329783	0.081764	-0.714172
C	2.343384	-0.969169	-0.267340
N	3.761017	-0.737396	-0.709023
C	4.599806	-1.828322	-0.123866
H	4.209191	-2.781424	-0.444405
C	4.288745	0.571767	-0.223236
H	4.042900	0.689335	0.819349
C	3.886244	-0.794324	-2.195113
H	3.559607	-1.763838	-2.537982
O	-0.959788	0.269847	2.196857
H	-0.344756	-1.058752	0.076489
H	-0.745389	0.551969	-0.395738
H	1.105852	-0.043761	-1.768138
H	1.716818	1.079668	-0.571981
H	2.368512	-1.013984	0.807652
H	2.068115	-1.941443	-0.653371
H	5.613244	-1.711746	-0.473378
H	4.567639	-1.757034	0.951421
H	5.358350	0.573859	-0.365632
H	3.851695	1.372678	-0.793356
H	4.922989	-0.642247	-2.455262
H	3.283681	-0.019264	-2.638037

n = 3 (g+a(90))

C	-2.408319	0.585713	-0.377909
C	-2.361555	1.927206	0.352297
O	-2.283608	2.941551	-0.369773
C	-1.030966	0.200005	-0.942210
C	-0.053459	-0.128595	0.183603
N	1.363073	-0.396125	-0.242467
C	2.182509	-0.570319	0.996346
H	1.758807	-1.366855	1.587348
C	1.935225	0.735571	-1.028402
H	1.754974	1.659593	-0.502614
C	1.464082	-1.651499	-1.045447
H	1.100921	-2.475238	-0.451241
O	-2.355639	1.894430	1.597976
H	-2.771971	-0.182277	0.295792
H	-3.102670	0.667605	-1.206791
H	-1.153201	-0.655609	-1.596313
H	-0.677261	1.027909	-1.541689
H	-0.014127	0.682420	0.893285
H	-0.377609	-1.015261	0.710227
H	3.191607	-0.822330	0.710268
H	2.174069	0.350330	1.557079
H	2.995801	0.569130	-1.133861
H	1.479383	0.769924	-2.002977
H	2.500956	-1.810367	-1.298446
H	0.882526	-1.555705	-1.946510

n = 4 (aaa(180))

C	0.039255	0.000621	0.021803
C	-0.035573	-0.001899	1.548390
O	1.034619	-0.001765	2.180833
C	1.440295	0.001887	-0.578397
C	1.439202	0.002206	-2.111501
C	2.885334	0.003750	-2.602713
N	3.094967	0.003215	-4.092443
C	2.516507	-1.215642	-4.728398
H	2.918259	-2.089648	-4.239995
C	4.571656	0.005249	-4.335333
H	5.001857	-0.877692	-3.889804
C	2.513235	1.218931	-4.730052
H	1.440815	1.200268	-4.639558
O	-1.187403	-0.004981	2.029718
H	-0.526906	-0.863406	-0.322077
H	-0.527642	0.865070	-0.320544
H	1.980863	-0.863489	-0.208206
H	1.979201	0.868581	-0.207874
H	0.907165	-0.872319	-2.469021
H	0.905468	0.875916	-2.468398
H	3.399746	-0.869650	-2.229887
H	3.397130	0.879823	-2.231388
H	2.793447	-1.218208	-5.771010
H	1.444136	-1.199182	-4.638975
H	4.999493	0.889035	-3.890160
H	4.749237	0.005074	-5.398988
H	2.912756	2.094858	-4.243794
H	2.789060	1.219935	-5.773479

n = 4 (aaa(90))

C	-2.287449	1.233405	-0.333994
C	-2.707839	2.695873	-0.269339
O	-2.630066	3.346739	-1.330138
C	-0.789156	1.064246	-0.056345
C	-0.351448	-0.403696	-0.106607
C	1.145400	-0.508514	0.176693
N	1.730332	-1.895416	0.164879
C	1.116058	-2.765054	1.209672
H	1.196280	-2.274419	2.166777
C	3.190694	-1.764138	0.462189
H	3.311096	-1.307511	1.431580
C	1.580991	-2.548196	-1.168291
H	0.540006	-2.735022	-1.368215
O	-3.051094	3.127924	0.848952
H	-2.858488	0.672852	0.400193
H	-2.518946	0.845250	-1.321308
H	-0.566287	1.487103	0.919426
H	-0.231661	1.645539	-0.785986
H	-0.923747	-0.962362	0.624905
H	-0.589401	-0.804406	-1.085226
H	1.366842	-0.101134	1.152486
H	1.701451	0.054454	-0.558669
H	1.654483	-3.699884	1.230341
H	0.082971	-2.949037	0.970103
H	3.648500	-1.149222	-0.296106
H	3.636228	-2.745766	0.459161
H	1.996995	-1.901403	-1.924470
H	2.119070	-3.483148	-1.148749

n = 4 (g+aa(90))

C	-0.004627	-0.033819	0.007167
C	0.002425	-0.008066	1.534848
O	1.125459	0.035368	2.077120
C	0.617016	-1.318358	-0.551728
C	-0.128103	-2.586701	-0.123407
C	0.553600	-3.807654	-0.730978
N	-0.053996	-5.141069	-0.388100
C	0.080381	-5.455592	1.063341
H	1.118866	-5.380639	1.344713
C	0.689393	-6.183355	-1.164121
H	1.730483	-6.150694	-0.886586
C	-1.492270	-5.209610	-0.776397
H	-2.064276	-4.518650	-0.182266
O	-1.099383	-0.065892	2.111799
H	-1.023756	0.077371	-0.349628
H	0.568880	0.816704	-0.350646
H	1.652633	-1.368971	-0.233628
H	0.619394	-1.257860	-1.637416
H	-0.148394	-2.643587	0.957661
H	-1.155128	-2.515649	-0.461308
H	1.585641	-3.858790	-0.412139
H	0.535612	-3.741171	-1.809455
H	-0.268806	-6.464086	1.220606
H	-0.512284	-4.770595	1.645486
H	0.578887	-5.980247	-2.217130
H	0.274035	-7.151419	-0.931835
H	-1.583938	-4.968607	-1.823835
H	-1.842880	-6.214677	-0.600254

n = 4 (g+aa(180))

C	0.207472	2.552028	-1.771032
C	0.183125	3.603959	-0.656667
O	0.815430	3.366019	0.386109
C	0.827936	1.198011	-1.434098
C	0.014657	0.375204	-0.431231
C	0.722374	-0.946901	-0.147576
N	0.098465	-1.804386	0.921555
C	0.176045	-1.155518	2.261227
H	1.206424	-0.923589	2.479362
C	0.863865	-3.088889	0.982060
H	1.895088	-2.873700	1.211246
C	-1.323868	-2.123126	0.606662
H	-1.918440	-1.227861	0.660235
O	-0.467931	4.637796	-0.912580
H	-0.811843	2.433811	-2.131225
H	0.749173	3.003837	-2.600012
H	1.828248	1.358815	-1.046480
H	0.932030	0.632601	-2.357972
H	-0.102761	0.958999	0.469526
H	-0.971063	0.191475	-0.844232
H	1.739267	-0.767411	0.171269
H	0.750345	-1.556356	-1.039799
H	-0.206881	-1.848143	2.994219
H	-0.412775	-0.255030	2.269589
H	0.790707	-3.586421	0.027648
H	0.436619	-3.710019	1.753446
H	-1.375010	-2.547166	-0.384512
H	-1.679524	-2.839498	1.330246

n = 4 (ag+a)

C	0.029552	0.001628	0.062418
C	0.028134	-0.068125	1.587732
O	1.128257	-0.071410	2.169481
C	1.328192	0.464056	-0.597503
C	1.293305	0.436353	-2.133772
C	0.289170	1.442600	-2.689674
N	0.209597	1.542124	-4.189362
C	1.502996	1.984889	-4.787570
H	1.818174	2.897524	-4.306871
C	-0.831688	2.565716	-4.513069
H	-0.535959	3.512509	-4.089448
C	-0.198459	0.246106	-4.805671
H	0.578796	-0.484587	-4.663088
O	-1.098441	-0.137333	2.119705
H	-0.228459	-0.997492	-0.292205
H	-0.811760	0.629691	-0.211787
H	2.135186	-0.171913	-0.256251
H	1.574036	1.464214	-0.250217
H	2.290721	0.664558	-2.492603
H	1.054324	-0.569127	-2.461027
H	0.530319	2.436042	-2.338648
H	-0.710270	1.203346	-2.362256
H	1.345823	2.160042	-5.840555
H	2.246923	1.217977	-4.657354
H	-1.774715	2.252874	-4.095429
H	-0.913992	2.651780	-5.584891
H	-1.114747	-0.089131	-4.345976
H	-0.353315	0.407063	-5.861002

n = 4 (g+g+a(90))

C	-2.535249	-0.956897	-0.794942
C	-2.772051	0.531486	-1.054838
O	-2.702826	1.296568	-0.074080
C	-1.202703	-1.458181	-1.382764
C	0.027702	-0.602749	-1.050611
C	0.246990	-0.474469	0.453468
N	1.340305	0.472134	0.869594
C	2.684680	-0.025478	0.452731
H	2.850807	-0.994675	0.896224
C	1.316412	0.556920	2.362588
H	1.434890	-0.434560	2.770998
C	1.128494	1.841951	0.317991
H	1.354774	1.847388	-0.734394
O	-2.980558	0.859228	-2.241061
H	-2.580424	-1.140406	0.272075
H	-3.341307	-1.520556	-1.254534
H	-1.294203	-1.495777	-2.461999
H	-1.039974	-2.479352	-1.048331
H	0.889395	-1.064373	-1.520455
H	-0.106423	0.371065	-1.502152
H	0.504882	-1.433378	0.882373
H	-0.648450	-0.108128	0.924266
H	3.429356	0.671828	0.804269
H	2.730456	-0.092894	-0.620696
H	0.374688	0.978136	2.675366
H	2.130597	1.185536	2.686845
H	0.103010	2.133310	0.473787
H	1.798083	2.515360	0.829928

n = 4 (g+g+a(180))

C	-2.169711	-0.157540	-1.921977
C	-2.687717	1.249971	-1.612179
O	-3.578461	1.332746	-0.744370
C	-0.681962	-0.259017	-2.272003
C	0.291602	0.260483	-1.205835
C	0.158529	-0.500790	0.109462
N	1.118956	-0.093417	1.194405
C	2.535090	-0.378270	0.812153
H	2.621532	-1.419756	0.545123
C	0.797125	-0.900899	2.411870
H	0.905391	-1.947722	2.176102
C	0.985303	1.350070	1.540445
H	1.310448	1.955658	0.712214
O	-2.195706	2.190229	-2.266295
H	-2.434929	-0.805267	-1.093934
H	-2.746947	-0.520101	-2.771519
H	-0.501873	0.299708	-3.181645
H	-0.450486	-1.298749	-2.491148
H	1.294419	0.148339	-1.601857
H	0.109345	1.316250	-1.063433
H	0.322540	-1.557964	-0.048926
H	-0.828606	-0.366468	0.523941
H	3.164341	-0.163894	1.661809
H	2.823094	0.245436	-0.015976
H	-0.215416	-0.691631	2.716497
H	1.483301	-0.629113	3.198368
H	-0.045418	1.564744	1.772326
H	1.607674	1.548703	2.398853

n = 4 (g-g+a)

C	0.004982	0.062784	-0.173857
C	-0.004395	-0.492679	1.255256
O	1.024517	-0.300411	1.931868
C	0.492753	1.510081	-0.318956
C	-0.503452	2.592703	0.121991
C	-0.864693	2.491423	1.601157
N	-1.677411	3.631881	2.146616
C	-0.865839	4.881348	2.254194
H	-0.018123	4.695163	2.894093
C	-2.129343	3.252875	3.519470
H	-1.266426	3.014559	4.120631
C	-2.883444	3.907524	1.311268
H	-2.579165	4.322617	0.365370
O	-1.032941	-1.095050	1.619631
H	0.671384	-0.580443	-0.744875
H	-0.986009	-0.055669	-0.600960
H	0.709235	1.692984	-1.367047
H	1.427688	1.619392	0.217933
H	-0.056625	3.555512	-0.100889
H	-1.397762	2.508987	-0.485511
H	0.024716	2.435187	2.209557
H	-1.441490	1.598421	1.779715
H	-1.487775	5.652241	2.681669
H	-0.535606	5.186443	1.276013
H	-2.776855	2.393605	3.452655
H	-2.660685	4.086831	3.949712
H	-3.426332	2.987916	1.160332
H	-3.501135	4.620132	1.835113

n = 4 (ag+g+)

C	-0.831055	0.190325	0.029363
C	0.589656	0.470199	0.537021
O	1.351656	-0.495424	0.714600
C	-1.125253	-1.177202	-0.596177
C	-0.728224	-1.337567	-2.073646
C	0.765581	-1.363621	-2.414814
N	1.422616	-2.721622	-2.460981
C	2.897003	-2.506425	-2.568094
H	3.098165	-1.891610	-3.431520
C	0.978007	-3.483400	-3.667698
H	1.232912	-2.919287	-4.551093
C	1.153358	-3.539732	-1.245197
H	1.418555	-2.971157	-0.371271
O	0.847970	1.669148	0.766240
H	-1.470562	0.320445	0.901205
H	-1.108036	0.995542	-0.645799
H	-2.200227	-1.331365	-0.555338
H	-0.682196	-1.945295	0.022315
H	-1.224700	-2.214571	-2.473797
H	-1.149998	-0.502639	-2.623690
H	1.324229	-0.783261	-1.704302
H	0.923138	-0.951138	-3.400766
H	3.378763	-3.464826	-2.684182
H	3.246506	-2.018034	-1.673572
H	-0.085858	-3.644723	-3.625920
H	1.488119	-4.433914	-3.679476
H	0.110232	-3.807235	-1.219364
H	1.748523	-4.437735	-1.311230

n = 4 (g+ag+)

C	0.032257	0.092927	-0.010008
C	-0.283371	-0.635273	1.302967
O	0.651252	-0.689915	2.128566
C	-1.169991	0.434582	-0.886934
C	-2.053012	1.531165	-0.276359
C	-3.444148	1.703482	-0.889631
N	-4.561831	0.832769	-0.358906
C	-4.401373	-0.602865	-0.726594
H	-4.277013	-0.675179	-1.796063
C	-5.831201	1.330506	-0.980445
H	-5.739957	1.279546	-2.054327
C	-4.696490	0.946402	1.121567
H	-3.864612	0.464444	1.602492
O	-1.414822	-1.131667	1.437035
H	0.604015	0.984040	0.236537
H	0.721162	-0.548775	-0.556447
H	-1.726486	-0.473663	-1.062928
H	-0.816191	0.775249	-1.857150
H	-2.122724	1.398296	0.793284
H	-1.554863	2.484038	-0.425069
H	-3.419280	1.541942	-1.958499
H	-3.781256	2.715024	-0.715627
H	-5.298260	-1.125753	-0.431599
H	-3.555572	-1.017034	-0.210670
H	-5.999199	2.349277	-0.669420
H	-6.646658	0.707307	-0.650404
H	-4.739458	1.991282	1.388778
H	-5.614252	0.459962	1.414081

n = 4 (g-ag+)

C	0.041733	-0.031558	0.249339
C	-0.204670	-0.054099	1.761293
O	0.799724	-0.291563	2.464027
C	-1.189829	0.095078	-0.645470
C	-2.061316	-1.168814	-0.669183
C	-3.488430	-0.961734	-1.195946
N	-3.727709	-1.190596	-2.666309
C	-2.794866	-0.407294	-3.522823
H	-2.833331	0.629873	-3.231220
C	-5.126581	-0.754385	-2.969795
H	-5.210650	0.305880	-2.793484
C	-3.612883	-2.639074	-3.007619
H	-2.620730	-2.990034	-2.782411
O	-1.354819	0.182562	2.168664
H	0.714555	0.804618	0.068902
H	0.620747	-0.917708	-0.001421
H	-0.851346	0.339498	-1.647029
H	-1.779957	0.934516	-0.295522
H	-1.552288	-1.961878	-1.206477
H	-2.168061	-1.513541	0.348723
H	-3.815401	0.046118	-0.988260
H	-4.164453	-1.638423	-0.695255
H	-3.110344	-0.512673	-4.549342
H	-1.795125	-0.789590	-3.410482
H	-5.808519	-1.291404	-2.329617
H	-5.342525	-0.975078	-4.003142
H	-4.340351	-3.191245	-2.433875
H	-3.810546	-2.758313	-4.061509

n = 4 (aag+)

C	0.030123	-0.110347	-0.000926
C	0.144076	-0.010386	1.520537
O	1.215625	-0.427156	2.007942
C	-1.316554	0.288171	-0.594346
C	-1.310931	0.251049	-2.126576
C	-2.659272	0.309289	-2.847104
N	-3.507850	1.552277	-2.698023
C	-4.263974	1.568889	-1.412438
H	-4.875239	0.681790	-1.359112
C	-4.515808	1.535087	-3.806945
H	-5.067249	0.609430	-3.761802
C	-2.697797	2.795947	-2.825599
H	-2.018875	2.875898	-1.994733
O	-0.814465	0.477636	2.142851
H	0.826048	0.507796	-0.414438
H	0.291500	-1.129473	-0.277160
H	-1.552566	1.277974	-0.225415
H	-2.085624	-0.371938	-0.201587
H	-0.648064	1.016925	-2.513406
H	-0.885514	-0.693416	-2.452891
H	-3.289467	-0.516377	-2.550140
H	-2.468213	0.219213	-3.906888
H	-4.893206	2.445014	-1.401785
H	-3.587443	1.598561	-0.579385
H	-4.002023	1.615460	-4.751350
H	-5.185847	2.370439	-3.680553
H	-2.150023	2.762642	-3.754328
H	-3.369717	3.639949	-2.828519

n = 5 (aaaa(180))

C	0.034069	-0.002116	0.017453
C	-0.027502	0.002899	1.545501
O	1.047561	-0.000338	2.169518
C	1.428352	0.001942	-0.596284
C	1.435516	-0.002204	-2.125582
C	2.869878	-0.009698	-2.664836
H	3.376177	-0.886677	-2.280332
O	-1.174500	0.008823	2.037766
H	-0.532654	-0.869900	-0.317927
H	-0.541265	0.857301	-0.324166
H	1.974810	-0.860580	-0.226317
H	1.968956	0.870364	-0.230720
H	0.895267	-0.872429	-2.492104
H	0.904053	0.870932	-2.497210
H	3.385654	0.859460	-2.275494
C	2.881375	-0.006216	-4.189233
N	4.247578	-0.009105	-4.825859
C	5.025201	-1.226437	-4.453777
H	4.444240	-2.102305	-4.696436
C	5.027294	1.207403	-4.455242
H	5.259410	1.186903	-3.404493
C	4.054926	-0.009184	-6.308845
H	3.515902	-0.899245	-6.592435
H	2.371635	-0.879109	-4.570617
H	2.378074	0.871745	-4.567458
H	5.249745	-1.210026	-3.401571
H	5.944136	-1.228399	-5.018550
H	5.942304	1.211771	-5.026367
H	4.443705	2.083327	-4.690527
H	5.022008	0.002322	-6.785303
H	3.495089	0.868493	-6.589680

n = 5 (aaaa(90))

C	-2.396288	2.291280	0.167570
C	-2.397343	3.809567	0.288192
O	-2.382405	4.454421	-0.779130
C	-0.971468	1.726815	0.120767
C	-0.941180	0.201342	0.007796
C	0.490634	-0.342000	-0.026131
H	0.998020	-0.028912	0.877826
O	-2.368134	4.278520	1.443380
H	-2.922961	1.870780	1.019265
H	-2.926740	2.011380	-0.737833
H	-0.440904	2.041682	1.016408
H	-0.447468	2.170795	-0.722664
H	-1.481382	-0.228161	0.848053
H	-1.474715	-0.099341	-0.890478
H	1.008956	0.106732	-0.864377
C	0.492898	-1.863840	-0.144784
N	1.856711	-2.504599	-0.191973
C	2.635232	-2.227817	1.049783
H	2.051585	-2.532041	1.904460
C	2.636984	-2.043201	-1.376567
H	2.870213	-0.997857	-1.275280
C	1.661641	-3.983484	-0.304183
H	1.119174	-4.332643	0.559991
H	-0.017137	-2.309130	0.697235
H	-0.012991	-2.172595	-1.048021
H	2.864254	-1.178462	1.110744
H	3.551758	-2.795160	1.010685
H	3.551394	-2.613199	-1.425255
H	2.052800	-2.209946	-2.267873
H	2.628297	-4.459204	-0.348203
H	1.105142	-4.197508	-1.202647

n = 5 (aaag+)

C	0.022395	-0.048261	0.046085
C	-0.046377	-0.000484	1.570952
O	1.032769	0.034673	2.190221
C	1.275047	-0.683911	-0.550361
C	1.301159	-0.611557	-2.078692
C	2.597605	-1.189764	-2.654867
H	2.703694	-2.211019	-2.301265
O	-1.189894	0.035582	2.067565
H	-0.875912	-0.541481	-0.314581
H	-0.055623	0.987100	-0.291106
H	1.343049	-1.720641	-0.226254
H	2.146252	-0.185344	-0.139808
H	0.442918	-1.141079	-2.487174
H	1.184432	0.430461	-2.357273
H	3.447907	-0.653067	-2.252025
C	2.727638	-1.309553	-4.174659
N	2.764622	-0.035618	-4.994465
C	3.542739	1.045551	-4.328055
H	4.523124	0.672106	-4.079252
C	1.390836	0.452954	-5.315253
H	0.851170	0.662847	-4.410431
C	3.437944	-0.356549	-6.293436
H	4.465505	-0.619129	-6.101173
H	3.660670	-1.812714	-4.381324
H	1.928490	-1.905967	-4.589786
H	3.032602	1.370115	-3.438128
H	3.631343	1.871471	-5.016760
H	1.477681	1.350609	-5.906684
H	0.877874	-0.307817	-5.881931
H	3.393136	0.511447	-6.931521
H	2.924101	-1.183368	-6.758154

n = 5 (g+aaa(180))

C	-1.787322	3.390395	0.453546
C	-1.717709	3.413002	1.984996
O	-0.644122	3.080541	2.518231
C	-0.794322	2.467840	-0.250194
C	-1.047474	0.984831	0.016873
C	0.033555	0.077169	-0.579852
H	1.007928	0.469625	-0.316703
O	-2.748174	3.801200	2.571126
H	-2.809390	3.154919	0.170151
H	-1.622067	4.417188	0.129714
H	0.207281	2.727580	0.074329
H	-0.840956	2.654973	-1.321372
H	-1.098184	0.837469	1.090052
H	-2.017709	0.703984	-0.387149
H	-0.043423	0.090078	-1.661474
C	-0.134730	-1.333857	-0.022358
N	0.770533	-2.389344	-0.594305
C	2.207094	-1.997183	-0.510343
H	2.438593	-1.717109	0.505071
C	0.425902	-2.690337	-2.015019
H	0.590390	-1.817204	-2.622702
C	0.571607	-3.633987	0.209504
H	0.876084	-3.447551	1.226834
H	0.051061	-1.321688	1.041113
H	-1.143084	-1.687731	-0.183465
H	2.398145	-1.172984	-1.175903
H	2.805313	-2.845405	-0.803915
H	1.060610	-3.493367	-2.354893
H	-0.608155	-2.991858	-2.067430
H	1.171526	-4.422406	-0.216692
H	-0.471246	-3.906638	0.180492

n = 5 (g+aaa(90))

C	-0.586377	3.157043	-2.200688
C	-0.773788	4.119372	-1.029362
O	0.208928	4.832161	-0.739394
C	0.165616	1.881657	-1.788439
C	-0.622506	0.977520	-0.841027
C	0.161395	-0.272242	-0.425837
H	1.147645	0.023245	-0.087748
O	-1.868521	4.088980	-0.435652
H	-1.558233	2.895746	-2.608249
H	-0.018754	3.665982	-2.973459
H	1.107481	2.172023	-1.329901
H	0.417349	1.322817	-2.686803
H	-0.912730	1.546615	0.036013
H	-1.548819	0.679141	-1.325909
H	0.292644	-0.915076	-1.288814
C	-0.590050	-0.987660	0.694271
N	-0.007532	-2.294940	1.157908
C	1.430153	-2.161093	1.530009
H	1.539197	-1.348074	2.230245
C	-0.149420	-3.353600	0.115669
H	0.417675	-3.083413	-0.757792
C	-0.772094	-2.724334	2.368765
H	-0.630726	-1.995181	3.150085
H	-0.627762	-0.349847	1.565060
H	-1.604863	-1.203567	0.392804
H	2.017379	-1.973966	0.647650
H	1.748405	-3.085170	1.987087
H	0.227309	-4.280387	0.520134
H	-1.193058	-3.460230	-0.134465
H	-0.402151	-3.685128	2.690303
H	-1.817684	-2.797138	2.115980

n = 5 (ag+aa(180))

C	-0.007143	0.085153	-0.067615
C	-0.039698	0.012101	1.457629
O	1.057187	0.047782	2.048521
C	1.005250	1.076699	-0.645600
C	1.106974	1.053620	-2.174828
C	-0.177409	1.489977	-2.886706
H	-0.416326	2.503149	-2.584701
O	-1.163364	-0.103976	1.984143
H	0.230600	-0.918507	-0.423829
H	-1.015323	0.300944	-0.402050
H	1.975002	0.860648	-0.216152
H	0.745203	2.081778	-0.315917
H	1.920595	1.710127	-2.472334
H	1.384172	0.052527	-2.497170
H	-0.993003	0.856295	-2.568963
C	-0.004113	1.391986	-4.398510
N	-1.191889	1.837879	-5.212041
C	-1.372165	3.316200	-5.137044
H	-0.481662	3.796671	-5.512001
C	-2.449936	1.165974	-4.776165
H	-2.739298	1.527161	-3.804183
C	-0.932829	1.468727	-6.638034
H	0.001128	1.906990	-6.949219
H	0.830407	1.991642	-4.734063
H	0.185317	0.365651	-4.678722
H	-1.553556	3.609774	-4.118355
H	-2.218581	3.587419	-5.749345
H	-3.221416	1.399346	-5.492389
H	-2.288836	0.100241	-4.742847
H	-1.738712	1.846511	-7.246402
H	-0.884365	0.393666	-6.716517

n = 5 (ag+aa(90))

C	-0.656335	-2.499793	1.458968
C	-1.622618	-3.256244	2.359674
O	-1.141966	-4.175628	3.051920
C	-0.031742	-1.299233	2.186968
C	1.028315	-0.556856	1.365746
C	0.471843	0.120825	0.109811
H	-0.270003	0.852761	0.408267
O	-2.805857	-2.863482	2.358519
H	0.131208	-3.176503	1.136948
H	-1.205454	-2.165371	0.586067
H	0.419260	-1.655402	3.108138
H	-0.825813	-0.613723	2.474697
H	1.493581	0.192678	2.001031
H	1.813202	-1.257316	1.089879
H	-0.028151	-0.616511	-0.501728
C	1.602713	0.770556	-0.680633
N	1.181902	1.516636	-1.920912
C	0.432838	2.759100	-1.574360
H	1.060778	3.385226	-0.960904
C	0.346824	0.669312	-2.818954
H	-0.613146	0.493002	-2.365752
C	2.423480	1.906142	-2.658468
H	3.054870	2.483002	-2.001610
H	2.139747	1.483809	-0.071852
H	2.300414	0.013813	-1.009516
H	-0.469398	2.505367	-1.045468
H	0.183250	3.274557	-2.488668
H	0.214320	1.192143	-3.753044
H	0.853838	-0.267008	-2.991338
H	2.147908	2.496111	-3.517644
H	2.937221	1.012062	-2.974873

n = 5 (aag+a(180))

C	-1.861711	2.938490	-0.996836
C	-1.900773	4.083078	0.014821
O	-0.874064	4.281388	0.689390
C	-0.921491	1.788839	-0.647370
C	-0.985678	0.636091	-1.655166
C	0.079592	-0.448838	-1.438223
H	1.059330	-0.005474	-1.567621
O	-2.959052	4.742541	0.059076
H	-2.877529	2.581181	-1.135853
H	-1.568180	3.375270	-1.952773
H	-1.176626	1.436890	0.348824
H	0.091615	2.172761	-0.578285
H	-1.975852	0.187237	-1.631966
H	-0.856278	1.031655	-2.657920
H	-0.043321	-1.196150	-2.214072
C	-0.036767	-1.083754	-0.055031
N	0.865257	-2.261740	0.201318
C	2.295646	-1.923257	-0.053366
H	2.548374	-1.027650	0.491632
C	0.481897	-3.439249	-0.631944
H	0.626283	-3.211721	-1.673811
C	0.716985	-2.639055	1.640033
H	1.037484	-1.813812	2.255302
H	0.201764	-0.361601	0.709565
H	-1.044016	-1.436868	0.116175
H	2.449686	-1.773141	-1.108147
H	2.904815	-2.745872	0.287140
H	1.110066	-4.270970	-0.354062
H	-0.552118	-3.680669	-0.442100
H	1.328685	-3.504722	1.837377
H	-0.318263	-2.866774	1.838409

n = 5 (aag+a(90))

C	-2.922268	1.460254	-0.629933
C	-4.042897	1.121490	0.344681
O	-4.182239	1.880438	1.324367
C	-1.607032	0.788359	-0.219505
C	-0.478438	1.033196	-1.226696
C	0.902708	0.546973	-0.766350
H	1.191831	1.100524	0.119179
O	-4.693095	0.084760	0.106924
H	-3.205910	1.120224	-1.621713
H	-2.787445	2.537145	-0.659782
H	-1.799405	-0.275265	-0.111626
H	-1.318236	1.158346	0.762643
H	-0.736596	0.563183	-2.172330
H	-0.407950	2.098482	-1.424792
H	1.615943	0.788107	-1.546222
C	0.901739	-0.951977	-0.480278
N	2.243798	-1.558967	-0.169970
C	2.917971	-0.851049	0.956623
H	2.238781	-0.792101	1.792616
C	3.138624	-1.548311	-1.364832
H	3.358678	-0.533363	-1.646284
C	2.012939	-2.980310	0.231082
H	1.407487	-3.000488	1.122782
H	0.278153	-1.169599	0.372116
H	0.516190	-1.498715	-1.329165
H	3.206435	0.136972	0.641931
H	3.795887	-1.412296	1.234633
H	4.053968	-2.058451	-1.109529
H	2.645774	-2.061898	-2.175303
H	2.965763	-3.447655	0.422164
H	1.505493	-3.493039	-0.570300

n = 5 (ag+g+a)

C	0.386925	-0.173170	-0.128254
C	-0.802585	0.133159	0.787189
O	-0.872011	-0.512758	1.851609
C	1.104120	1.059599	-0.727230
C	1.010803	1.213993	-2.253893
C	-0.389749	1.431777	-2.840300
H	-1.023443	0.600408	-2.564900
O	-1.620265	0.980617	0.389212
H	1.083697	-0.789221	0.426103
H	-0.000252	-0.804625	-0.928312
H	2.159639	0.994135	-0.481924
H	0.743502	1.956175	-0.235973
H	1.653959	2.038248	-2.553233
H	1.430971	0.325310	-2.716163
H	-0.298259	1.431004	-3.921464
C	-1.003286	2.736904	-2.348830
N	-2.358852	3.072308	-2.918919
C	-3.282244	1.901951	-2.906762
H	-3.271283	1.452505	-1.928209
C	-2.239153	3.586226	-4.315583
H	-1.766615	2.842769	-4.934866
C	-2.950853	4.149604	-2.071525
H	-3.122817	3.761686	-1.081157
H	-1.133615	2.687199	-1.284264
H	-0.367750	3.576588	-2.594203
H	-2.966785	1.190138	-3.651181
H	-4.272238	2.255285	-3.150792
H	-3.228406	3.797562	-4.690299
H	-1.650577	4.490079	-4.305171
H	-3.880964	4.469153	-2.515790
H	-2.261950	4.978827	-2.030666

n = 5 (ag+ag-)

C	2.830805	-1.259214	-0.310963
C	4.315443	-1.241344	0.048507
O	4.695230	-2.090127	0.878851
C	2.204234	0.115703	-0.553114
C	0.748083	0.074369	-1.032233
C	-0.242003	-0.457226	0.012570
H	-0.194312	0.148817	0.909904
O	5.029443	-0.403603	-0.535126
H	2.739339	-1.856769	-1.219987
H	2.314395	-1.807920	0.468410
H	2.808312	0.635878	-1.285595
H	2.270810	0.703672	0.361991
H	0.476100	1.079851	-1.331931
H	0.687611	-0.536548	-1.930839
H	0.071730	-1.447939	0.315282
C	-1.687073	-0.623085	-0.477510
N	-2.624972	0.548424	-0.318399
C	-2.125502	1.780623	-0.989008
H	-1.883182	1.552099	-2.014317
C	-2.882673	0.844321	1.120922
H	-1.969284	1.148310	1.602044
C	-3.925877	0.163902	-0.950592
H	-3.767680	-0.011852	-2.002226
H	-1.698776	-0.877049	-1.527756
H	-2.160864	-1.428513	0.063139
H	-1.256553	2.147033	-0.470997
H	-2.906118	2.524357	-0.954138
H	-3.604565	1.643246	1.179690
H	-3.275821	-0.040678	1.595644
H	-4.630119	0.969266	-0.813209
H	-4.295367	-0.731320	-0.476640

n = 5 (g+aag+)

C	-0.416125	0.069761	-0.116650
C	-0.001403	-0.445820	1.260107
O	1.012579	-1.165227	1.315154
C	-0.189039	1.579811	-0.269651
C	1.249444	2.012431	0.021805
C	1.520467	3.452636	-0.427659
H	0.662189	4.066490	-0.170572
O	-0.701144	-0.069820	2.223344
H	0.144722	-0.465958	-0.875719
H	-1.470661	-0.146954	-0.263975
H	-0.866624	2.110845	0.391910
H	-0.455569	1.865131	-1.284926
H	1.435002	1.901315	1.085996
H	1.921125	1.323379	-0.479962
H	1.594054	3.492907	-1.508557
C	2.696494	4.192385	0.209539
N	4.093417	3.653963	-0.001820
C	4.335331	3.286520	-1.425535
H	4.096117	4.132868	-2.050507
C	4.392999	2.487679	0.877398
H	3.760160	1.656583	0.629680
C	5.046889	4.749127	0.367047
H	4.911482	5.576112	-0.311030
H	2.712934	5.193803	-0.195551
H	2.565860	4.265812	1.279452
H	3.725195	2.440965	-1.692072
H	5.377466	3.031954	-1.538602
H	5.425742	2.212866	0.731847
H	4.234845	2.775965	1.904401
H	6.054429	4.372887	0.291606
H	4.844664	5.061638	1.379324

n = 5 (g+aag-)

C	-0.047422	0.311840	0.301687
C	-0.598208	1.108367	1.485018
O	0.192248	1.373281	2.411201
C	0.168074	1.166964	-0.958802
C	1.126839	2.357793	-0.774990
C	0.407083	3.666590	-0.423293
H	-0.288454	3.520266	0.388653
O	-1.792032	1.460927	1.408381
H	0.891042	-0.149898	0.590385
H	-0.752241	-0.481098	0.066259
H	-0.796002	1.518692	-1.313815
H	0.557531	0.513514	-1.732803
H	1.868136	2.100826	-0.027325
H	1.671419	2.517693	-1.701034
H	-0.201004	3.938203	-1.279232
C	1.314871	4.874948	-0.174152
N	1.751807	5.152054	1.245613
C	2.656758	4.097412	1.781548
H	3.487165	3.978362	1.102239
C	0.584683	5.315945	2.158242
H	0.107129	4.366842	2.321611
C	2.506963	6.445525	1.219819
H	3.344262	6.349957	0.546627
H	2.222831	4.801267	-0.756719
H	0.800209	5.773644	-0.480237
H	2.117584	3.175883	1.898440
H	3.021489	4.425163	2.742674
H	0.944928	5.704909	3.097794
H	-0.104935	6.019033	1.716190
H	2.858499	6.663412	2.215544
H	1.847849	7.230148	0.883324

n = 5 (g+ag-a(180))

C	-0.028699	-0.091304	0.148062
C	0.222892	-0.092783	1.659143
O	1.202837	-0.728859	2.085176
C	0.734605	-1.131094	-0.667774
C	0.255726	-2.569184	-0.441724
C	1.227547	-3.638766	-0.955632
H	0.817393	-4.609937	-0.705117
O	-0.577007	0.583126	2.338788
H	-1.100911	-0.180622	-0.008701
H	0.228238	0.909279	-0.196826
H	1.789073	-1.053458	-0.421687
H	0.630818	-0.873493	-1.719735
H	0.115158	-2.726990	0.619494
H	-0.716037	-2.704655	-0.912090
H	2.164741	-3.536214	-0.421607
C	1.454595	-3.519919	-2.460808
N	2.209254	-4.658215	-3.095255
C	1.406166	-5.915883	-3.087968
H	0.465873	-5.732075	-3.583030
C	3.514931	-4.903045	-2.417691
H	3.343882	-5.273055	-1.421709
C	2.478098	-4.284721	-4.517831
H	1.540150	-4.084001	-5.010664
H	0.510309	-3.463784	-2.983666
H	2.019622	-2.628467	-2.686647
H	1.236932	-6.236932	-2.074994
H	1.958330	-6.675114	-3.619360
H	4.059746	-5.638001	-2.989243
H	4.070829	-3.979695	-2.381100
H	2.982157	-5.104953	-5.003766
H	3.102042	-3.405561	-4.538000

n = 5 (g+ag-a(90))

C	-1.527008	2.851501	-1.451429
C	-1.514137	4.031317	-0.483352
O	-0.495429	4.751902	-0.508782
C	-0.655345	1.682778	-0.967371
C	-1.175962	0.993164	0.298027
C	-0.224502	-0.065792	0.872190
H	-0.674767	-0.456674	1.777404
O	-2.495913	4.160272	0.273350
H	-2.549519	2.515546	-1.593601
H	-1.146318	3.194898	-2.408677
H	0.354830	2.050581	-0.803514
H	-0.598633	0.962102	-1.779070
H	-1.360515	1.738245	1.062203
H	-2.140275	0.537979	0.084085
H	0.704118	0.413926	1.157974
C	0.049215	-1.189831	-0.126091
N	0.754135	-2.398261	0.434541
C	-0.141116	-3.183088	1.334136
H	-1.011076	-3.493531	0.776639
C	1.991790	-2.021286	1.176078
H	1.728348	-1.501244	2.080893
C	1.139532	-3.272894	-0.715562
H	0.255626	-3.511723	-1.285895
H	-0.873879	-1.556365	-0.551743
H	0.673990	-0.831261	-0.929361
H	-0.434831	-2.580568	2.175823
H	0.399540	-4.049471	1.681079
H	2.528952	-2.924467	1.420279
H	2.600742	-1.390650	0.547596
H	1.584922	-4.176717	-0.331179
H	1.848880	-2.747631	-1.334460

n = 5 (g+ag+a)

C	-0.008785	-0.235362	0.186087
C	0.792913	-0.072011	1.484562
O	2.014352	-0.288649	1.437239
C	0.661254	-1.093216	-0.883842
C	0.739894	-2.579490	-0.503302
C	1.874862	-3.356257	-1.182989
H	1.814817	-3.225352	-2.257999
O	0.143631	0.299711	2.483549
H	-0.991281	-0.620262	0.444690
H	-0.179395	0.770395	-0.196818
H	1.648365	-0.684501	-1.058790
H	0.113938	-0.997596	-1.819329
H	0.849039	-2.670030	0.572307
H	-0.200987	-3.057960	-0.757814
H	1.718391	-4.409317	-0.982736
C	3.242086	-2.902972	-0.666997
N	4.385015	-3.834883	-0.959530
C	4.491042	-4.128965	-2.418749
H	4.546534	-3.197642	-2.959708
C	4.256867	-5.115723	-0.204641
H	3.390278	-5.656583	-0.543180
C	5.644484	-3.161988	-0.520509
H	5.779057	-2.258526	-1.093669
H	3.516511	-1.954575	-1.100880
H	3.221677	-2.771935	0.403290
H	3.631648	-4.694223	-2.737009
H	5.387460	-4.705896	-2.585406
H	5.140744	-5.706733	-0.385907
H	4.170093	-4.894956	0.847528
H	6.472268	-3.832164	-0.690926
H	5.569707	-2.922803	0.528052

n = 5 (g+g+g+a)

C	-0.234281	-0.112795	-0.220508
C	-0.631968	-0.165844	1.253353
O	0.302665	-0.148520	2.081719
C	0.767447	-1.188432	-0.666138
C	0.379248	-2.641467	-0.346297
C	-1.074808	-3.036493	-0.636062
H	-1.724984	-2.408363	-0.044984
O	-1.849550	-0.216531	1.509530
H	-1.137746	-0.153605	-0.817422
H	0.211279	0.863170	-0.402120
H	1.726431	-0.988336	-0.201632
H	0.923119	-1.073734	-1.735784
H	1.052152	-3.301731	-0.887834
H	0.558869	-2.820821	0.707759
H	-1.214347	-4.053796	-0.288125
C	-1.423898	-2.963274	-2.119899
N	-2.898203	-3.041849	-2.423891
C	-3.597902	-1.782634	-2.039089
H	-3.186159	-0.967270	-2.613001
C	-3.542049	-4.198229	-1.736928
H	-3.552898	-4.031698	-0.673557
C	-3.051896	-3.229500	-3.900430
H	-2.535388	-2.432222	-4.411092
H	-1.082234	-2.041402	-2.565585
H	-0.961248	-3.783762	-2.650581
H	-3.471853	-1.589509	-0.988367
H	-4.646436	-1.894991	-2.268517
H	-4.553740	-4.287752	-2.101350
H	-2.988995	-5.094920	-1.970685
H	-4.101481	-3.206694	-4.147012
H	-2.629407	-4.182394	-4.176723

n = 5 (g+g+ag+)

C	-0.158662	0.029120	0.087106
C	-0.425006	0.302271	1.566560
O	0.552534	0.608461	2.275281
C	-0.693700	1.150077	-0.821724
C	-0.232843	2.564951	-0.453858
C	1.288116	2.746785	-0.470420
H	1.641438	2.532302	-1.474027
O	-1.614327	0.229386	1.936475
H	0.905839	-0.110982	-0.061890
H	-0.651335	-0.899703	-0.186424
H	-1.777842	1.123642	-0.794767
H	-0.401567	0.929972	-1.845824
H	-0.673887	3.257716	-1.167402
H	-0.647393	2.818135	0.514042
H	1.758154	2.022526	0.177357
C	1.779860	4.162843	-0.154546
N	2.058024	4.501656	1.292319
C	3.135541	3.637934	1.857250
H	3.995659	3.689553	1.207457
C	0.845759	4.399434	2.149348
H	0.555853	3.369991	2.249509
C	2.536061	5.921023	1.322853
H	3.428816	6.003814	0.723656
H	2.717021	4.340409	-0.661273
H	1.071202	4.900777	-0.503269
H	2.786716	2.625231	1.948516
H	3.394964	4.016897	2.833392
H	1.086060	4.802310	3.121266
H	0.054853	4.982310	1.703275
H	2.752074	6.193410	2.343297
H	1.762707	6.560052	0.927280

n = 5 (g+g+aa)

C	0.048641	-3.911758	-1.288886
C	-1.067568	-3.099852	-1.946688
O	-1.096306	-3.108913	-3.195612
C	1.423898	-3.226481	-1.414905
C	1.436299	-1.738713	-1.045421
C	0.989641	-1.439750	0.393301
H	1.816594	-1.617863	1.070849
O	-1.827735	-2.467288	-1.193727
H	-0.202467	-4.075485	-0.247158
H	0.103646	-4.885473	-1.766452
H	1.763780	-3.324287	-2.440436
H	2.137577	-3.767363	-0.798290
H	2.439185	-1.351045	-1.201791
H	0.793619	-1.215678	-1.746712
H	0.189950	-2.111811	0.659751
C	0.504382	0.002527	0.500360
N	-0.198943	0.358267	1.781758
C	0.665510	0.089140	2.967963
H	1.615337	0.580450	2.825550
C	-1.487044	-0.378080	1.931082
H	-1.302549	-1.434032	2.010213
C	-0.502272	1.822083	1.739816
H	0.425051	2.371561	1.705265
H	1.331995	0.692526	0.411780
H	-0.201741	0.206656	-0.289472
H	0.806659	-0.972249	3.078922
H	0.174988	0.484486	3.843666
H	-1.969871	-0.024437	2.828703
H	-2.106430	-0.188744	1.070457
H	-1.053843	2.087994	2.627695
H	-1.092605	2.032832	0.862516

n = 5 (g+g-g-a)

C	0.781139	2.797681	-1.420972
C	-0.707867	2.433661	-1.421751
O	-1.261317	2.381343	-0.304988
C	1.665895	1.925838	-0.516738
C	2.008677	0.523681	-1.047736
C	0.847883	-0.455809	-1.281430
H	1.270884	-1.366727	-1.691300
O	-1.249251	2.227291	-2.523592
H	1.152831	2.791946	-2.440933
H	0.842000	3.823173	-1.062614
H	1.205322	1.871268	0.462891
H	2.610194	2.442934	-0.372967
H	2.539589	0.631759	-1.989146
H	2.716127	0.071028	-0.357391
H	0.177334	-0.043867	-2.021622
C	0.108261	-0.758396	0.016974
N	-1.116686	-1.626550	-0.106600
C	-0.757720	-3.024300	-0.489473
H	-0.065340	-3.421294	0.236533
C	-2.097243	-1.088029	-1.091565
H	-1.701740	-1.186585	-2.087783
C	-1.770531	-1.656179	1.236904
H	-1.070655	-2.049729	1.957535
H	0.761908	-1.271933	0.709249
H	-0.231630	0.155779	0.468810
H	-0.314340	-3.027974	-1.470974
H	-1.659104	-3.617178	-0.498330
H	-3.002495	-1.669068	-1.006748
H	-2.288240	-0.050794	-0.879035
H	-2.640489	-2.291926	1.186778
H	-2.059363	-0.652714	1.504750

n = 5 (g+g-g+a)

C	-0.429065	-0.183309	-0.177535
C	-0.743036	-0.635224	1.253993
O	0.182616	-1.189339	1.879014
C	-1.070135	1.129954	-0.648012
C	-0.396984	2.446529	-0.225001
C	-0.532788	2.866754	1.251684
H	-1.496319	2.539891	1.615833
O	-1.898080	-0.432358	1.672399
H	0.648929	-0.160374	-0.308266
H	-0.789805	-0.986278	-0.819683
H	-2.110029	1.134805	-0.340442
H	-1.070263	1.121395	-1.734678
H	0.652060	2.426491	-0.514055
H	-0.848992	3.230310	-0.824423
H	-0.514413	3.950055	1.297538
C	0.596955	2.302999	2.108363
N	0.473046	2.530247	3.594257
C	-0.873149	2.161828	4.119935
H	-1.154959	1.191475	3.749235
C	0.750931	3.956575	3.942133
H	0.052257	4.595836	3.428032
C	1.498284	1.675591	4.265069
H	1.480320	1.882127	5.323952
H	0.683332	1.240146	1.981028
H	1.537197	2.755914	1.823465
H	-1.591260	2.902523	3.810034
H	-0.816982	2.153808	5.197822
H	0.636418	4.076608	5.008588
H	1.761211	4.198154	3.651781
H	2.469481	1.920728	3.863418
H	1.271154	0.639130	4.077614

n = 5 (g+g-aa)

C	-3.225301	0.445843	2.226939
C	-2.610500	-0.954692	2.181277
O	-2.535340	-1.501283	1.066209
C	-2.183551	1.574274	2.346987
C	-1.120960	1.581253	1.235336
C	0.074599	0.649729	1.527113
H	-0.226283	-0.104509	2.236630
O	-2.199143	-1.417817	3.265059
H	-3.814452	0.598058	1.327901
H	-3.894321	0.500105	3.081209
H	-1.702024	1.513327	3.319160
H	-2.720678	2.517734	2.330799
H	-1.599870	1.305866	0.300647
H	-0.754139	2.593847	1.101201
H	0.859203	1.231311	1.997039
C	0.602727	-0.002462	0.253949
N	1.566256	-1.139839	0.472335
C	0.893029	-2.327915	1.069013
H	0.086482	-2.637069	0.426464
C	2.708050	-0.731390	1.341063
H	2.351214	-0.525233	2.335563
C	2.108286	-1.537225	-0.864065
H	2.762213	-2.385289	-0.735442
H	-0.213821	-0.413955	-0.317597
H	1.127378	0.720572	-0.355397
H	0.490296	-2.086912	2.035896
H	1.628078	-3.112277	1.159347
H	3.418263	-1.542888	1.372461
H	3.172958	0.146169	0.918606
H	2.659354	-0.708496	-1.279784
H	1.286289	-1.799958	-1.510590

6. References

- [1] Savle, P.; Doncel, G.; Bryant, S.; Hubieki, M.; Robinette, R.; Gandour, R., Acylcarnitine Analogues as Topical, Microbicidal Spermicides. *Bioorganic Medicinal Chemistry Letters* **1999**, 9, 2545-2548.
- [2] Nicander, I.; Rantanen, I.; Rozell, B.; Soderling, E.; Ollmar, S., The Ability of Betaine to Reduce the Irritating Effects of Detergents Assessed Visually, Histology and by Bioengineering Methods. *Skin Research and Technology* **2003**, 9, 50-58.
- [3] Wong, Y.; Hubieki, M.; Curfman, C.; Doncel, G.; al., e., A Structure-Activity Study of Spermicidal and Anti-HIV Properties of Hydroxylated Cationic Surfactants. *Bioorganic Medicinal Chemistry* **2002**, 10, 3599-3608.
- [4] Wong, Y.; Curfman, C.; Doncel, G.; Hubieki, M.; Dudding, T.; Savle, P.; Gandour, R., Spermicidal, Anti-HIV, and Micellar Properties of Di- and Trihydroxylated Cationic Surfactants. *Tetrahedron* **2002**, 58, 45-54.
- [5] Watanabe, S.; Regen, S., Quaternary Ammonium-Based Surfactants that can Recognize Cholesterol-Rich Membranes and Proton-Ionizable Analogs that Cannot. *Journal of the American Chemical Society* **1994**, 116, 5762-5765.
- [6] Weers, J.; Rathman, J.; Axe, F.; Crichlow, C.; Foland, L.; Scheuing, D.; Weirsema, R.; Zielske, A., Effect of the Intramolecular Charge Separation Distance on the Solution Properties of Betaines and Sulfobetaines. *Langmuir* **1991**, 7, 854-867.
- [7] Beckett, A.; woodward, R., Surface-Active Betaines: N-Alkyl-NN-Dimethylglycines and Their Critical Micelle Concentrations. *Journal of Pharmacy and Pharmacology* **1963**, 15, (7), 422-431.
- [8] Mittal, K., *Micellization, Sollubilization, and Microemulsions*. ed.; Plenum: New York, NY, 1977; Vol. 1
- [9] Mittal, K., *Micellization, Sollubilization, and Microemulsions*. ed.; Plenum: New York, NY, 1977; Vol. 2
- [10] Tanford, C., *The Hydrophobic Effect: Formation of Micelles and Biological Membranes*. 2 ed.; Wiley: New York, NY, 1980.

- [11] Thurston, G.; Blankschtein, D.; Fisch, M.; Benedek, G., Theory of Thermodynamic Properties and Phase Separation of Micellar Solutions with Lower Consolute Points. *Journal of Chemical Physics* **1986**, 84, (8), 4558-4562.
- [12] Shaw, D., *Introduction to Colloid and Surface Chemistry*. 2 ed.; Butterworth Inc: Woburn, MA, 1978.
- [13] Rosen, M., *Surfactants and Interfacial Phenomenon*. ed.; Wiley: New York, 1987.
- [14] Elimelech, M.; Gregory, J.; Jis, X.; Williams, R., *Particle Deposition and Aggregation, Measurement, Modeling, and Simulation*. ed.; Butterworth: Oxford, 1995.
- [15] Heimenz, P., *Principles of Colloid and Surface Chemistry*. ed.; Marcel Dekker: New York, 1977.
- [16] Myers, D., *Surfaces, Interfaces, and Colloids*. ed.; VCH Publishers, Inc: New York, 1991.
- [17] Wang, Z.; Li, G.; Zhang, X.; Wang, R.; Lou, A., A Quantitative Structure-Property Relationship Study for the Prediction of Critical Micelle Concentration of Nonionic Surfactants. *Colloids and Surfaces A: Physicochemical and Engineering Aspects* **2002**, 197, 37-45.
- [18] Klevens, H., Structure and Aggregation in Dilute Solutions of Surface Active Agents. *Journal of the American Oil Chemists Society* **1953**, 30, 74-90.
- [19] Klevens, H., Critical Micelle Concentrations as Determined by Refraction. *Journal of Physical Colloid Chemistry* **1948**, 52, 130-148.
- [20] Kriwanek, J.; Miller, R., UV/Vis Spectroscopic Investigations of Micellisation of Homologous N-alkyl Betaines Using the Dye Indicator E_T(30). *Colloids and Surfaces A: Physicochemical and Engineering Aspects* **1995**, 105, 233-242.
- [21] Zana, R., Ionization of Cationic Micelles: Effect of the Detergent Structure. *Journal of Colloid Interface Science* **1980**, 78, (2), 330-336.
- [22] Huibers, P.; Lobanov, V.; Katritzky, A.; Shah, D.; Karelson, M., Prediction of Critical Micelle Concentration Using a Quantitative Structure-Property Relationship Approach. 2. Anionic Surfactants. *Journal of Colloid and Interface Science* **1997**, 187, 113-120.
- [23] Yuan, S.; Cai, Z.; Xu, G.; Jiang, Y., Quantitative Structure-Property Relationship of Surfactants: Critical Micelle Concentration of Anionic Surfactants. *Journal of Dispersion Science and Technology* **2002**, 23, (4), 465-472.
- [24] Huibers, P.; Lobanov, V.; Katritzky, A.; Shah, D.; Karelson, M., Prediction of Critical Micelle Concentration Using a Quantitative Structure-Property Relationship Approach. I. Nonionic Surfactants. *Langmuir* **1996**, 12, 1462-1470.

- [25] Katritzky, A.; Lobanov, V., The Correlation and Quantitative Prediction of Chemical and Physical Properties from Structure. *Chemical Society Reviews* **1995**, 279-287.
- [26] Chevalier, Y.; Le Perchec, P., Intercharge Distance of Flexible Zwitterionic Molecules in Solution. *Journal of Physical Chemistry* **1990**, 94, 1768-1774.
- [27] Kirchnerova, J.; Farrell, P.; Edward, J., Dielectric Increments and the Conformations of Amino Acids and Betaines in Water. *Journal of Physical Chemistry* **1976**, 80, (18), 1974-1980.
- [28] Dupuis, M.; Marquez, A.; Davidson, E. *HONDO 99.6*, Indiana University, Bloomington, IN 47405, 1999.
- [29] Nakamura, H.; Xidos, J.; Thompson, J.; Li, J.; Hawkins, G.; Zhu, T.; Lynch, B.; Volobuev, Y.; Rinaldi, D.; Liotard, D.; Cramer, C.; Truhlar, D. *HONDOPLUS*, 4.5; University of Minnesota, Minneapolis, 2004.
- [30] Cramer, C., *Essentials of Computational Chemistry. Theories and Models*. ed.; Wiley Sons: England, 2003.
- [31] Cramer, C.; Truhlar, D., Implicit Solvation Models. *Chemical Reviews* **1999**, 99, (8), 2161-2200.
- [32] Ben-Naim, A., *Water and Aqueous Solutions*. ed.; Plenum Press: New York, 1974.
- [33] Ben-Naim, A., Standard Thermodynamics of Transfer. Uses and Misuses. *Journal of Physical Chemistry* **1978**, 82, 792-803.
- [34] Ben-Naim, A., *Solvation Thermodynamics*. ed.; Plenum Press: New York, 1987.
- [35] Uhlig, H., The Solubilities of Gases and Surface Tension. *Journal of Physical Chemistry* **1937**, 41, 1215-1226.
- [36] Tunon, I.; Silla, E.; Pascual-Ahuir, J., Continuum-Uniform Approach Calculations of the Solubility of Hydrocarbons in Water. *Chemical Physics Letters* **1993**, 203, 289-294.
- [37] Tolman, R., Consideration of the Gibbs Theory of Surface Tension. *Journal of Chemical Physics* **1948**, 16, 758-774.
- [38] Tolman, R., The Superficial Density of Matter at a Liquid-Vapor Boundary. *Journal of Chemical Physics* **1949**, 17, 118-127.
- [39] Sinanoglu, O., An Intermolecular Potential for use in Liquids. *Chemical Physics Letters* **1967**, 1, 283.
- [40] Sinanoglu, O., The C-Potential Surface for Predicting Conformations of Molecules in Solution. *Theoretica Chimica Acta* **1974**, 33, 279-284.

- [41] Sinanoglu, O., Microscopic Surface Tension Down to Molecular Dimensions and Microthermodynamic Surface Areas of Molecules or Clusters. *Journal of Chemical Physics* **1981**, 75, 463-468.
- [42] Sinanoglu, O., Molecular Interactions Within Liquids, the Solvophobic Force and Molecular Surface Areas. In *Molecular Interactions*, ed.; Ratajczak, H.; Orville-Thomas, W., Ed. Wiley: New York, 1982; pp 283-342.
- [43] Li, J.; Zhu, T.; Cramer, C.; Truhlar, D., New Class IV Charge Model for Extracting Accurate Partial Charges from Wave Functions. *Journal of Physical Chemistry A* **1998**, 102, 1820-1831.
- [44] Nakamura, H.; Xidos, J.; Thompson, J.; Li, J.; Hawkins, G.; Zhu, T.; Lynch, B.; Volobuev, Y.; Rinaldi, D.; Liotard, D.; Cramer, C.; Truhlar, D., *HONDOPLUS Users Manual, version 4.3*. ed.; 2003.
- [45] Mulliken, R., Electronic Population Analysis on LCAO-MO (Linear Combination of Atomic Orbital - Molecular Orbital) Molecular Wave Functions. I. *Journal of Chemical Physics* **1950**, 23, 1833-1840.
- [46] Lowdin, P., The Nonorthogonality Problem Connected with the USE of Atomic Wave Functions in the Theory of Molecules and Crystals. *Journal of Chemical Physics* **1950**, 18, 365-375.
- [47] Li, J.; Zhu, T.; Liotard, D.; Cramer, C.; Truhlar, D., Analytical Energy Gradients of a Self-Consistent Reaction Field Solvation Model Based on CM2 Atomic Charges. *Journal of Chemical Physics* **1999**, 110, (12), 5503-5513.
- [48] Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Zakrzewski, V.; Montgomery, J., J.; Stratmann, R.; Burant, J.; Dapprich, S.; Millam, J.; Daniels, A.; Kudin, N.; Strain, M.; Farkas, O.; Tomasi, J.; Barone, A.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G.; Ayala, P.; Cui, Q.; Morokuma, K.; Malick, A.; Rabuck, A.; Raghavachari, K.; Foresman, J.; Cioslowski, J.; Ortiz, J.; Baboul, A.; Stefanov, B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R.; Fox, D.; Keith, T.; Al-Laham, M.; Peng, C.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P.; Johnson, B.; Chen, W.; Wong, W.; Andres, J.; Head-Gordon, M.; Replogle, E.; Pople, J. *Gaussian 98, Revision A.7*, Gaussian, Inc: Pittsburg, PA, 1998.
- [49] Dewar, M.; Zoebisch, E.; Healy, E.; Stewart, J., AM1: A New General Purpose Quantum Mechanical Molecular Model. *Journal of the American Chemical Society* **1985**, 107, 3902-3909.

- [50] Scott, R.; Scheraga, H., Conformational Analysis of Macromolecules. II. The Rotational Isomeric States of the Normal Hydrocarbons. *Journal of Chemical Physics* **1966**, 44, (8), 3054-3069.
- [51] Mirkin, N.; Krimm, S., Ab initio Studies of the Conformation Dependence of the Spectra of Stable Conformers of n-Pentane and n-Hexane. *Journal of Physical Chemistry* **1993**, 97, 13887-13895.
- [52] Tasi, G.; Mizukami, F.; Palinko, I.; Csontos, J.; Gyorffy, W.; Nair, P.; Maeda, K.; Toba, M.; Niwa, S.; Kiyozumi, Y.; Kiricsi, I., Enumeration of the Conformers of Unbranched Aliphatic Alkanes. *Journal of Physical Chemistry A* **1998**, 102, 7698-7703.
- [53] Jensen, F., *Introduction to Computational Chemistry*. ed.; John Wiley & Sons Ltd: 1999.
- [54] McQuarrie, D.; Simon, J., *Physical Chemistry: A Molecular Approach*. ed.; University Science Books: Sausalito, CA, 1997.
- [55] *PCMODEL*, 5.13; Serena Software: Box 3076, Bloomington, IN 47402-3076.
- [56] Rak, J.; Skurski, P.; Gutowski, M., An ab initio Study of the Betaine Anion-Dipole-Bound Anionic State of a Model Zwitterion System. *Journal of Chemical Physics* **2001**, 114, 10673-10681.
- [57] Lever, M.; Blunt, J.; Maclagan, R., Some Ways of Looking at Compensatory Kosmotropes and Different Water Environments. *Comparative Biochemistry and Physiology Part A* **2001**, 130, 471-486.
- [58] Zhu, C.; Lifshitz, C., The Protonated Betaine/Ammonia Complex: How Stable is the Gas-Phase Zwitterion Structure? *Chemical Physics Letters* **2000**, 320, 513-517.
- [59] Nyrogen, T.; Suontamo, R.; Pitkanen, I., Why Betaine Crystallizes in High Local Cs Symmetry. An ab initio MO and DFT Study of Anhydrous Betaine and Betaine Monohydrate. *Theoretical Chemistry Accounts* **1999**, 101, (209-214).
- [60] Viertorriine, M.; Valkonen, J.; Pitkanen, I., Crystal and Molecular Structure of Anhydrous Betaine, (CH₃)₃NH₂CO₂. *Journal of Molecular Structure* **1999**, 477, 23-29.
- [61] Gregoire, F.; Wei, S.; Streed, E.; Brameld, K.; Fort, D.; Hanely, L.; Walls, J.; Goddard, W.; Roberts, J., Conformational Equilibria of β-Alanine and Related Compounds as Studied by NMR Spectroscopy. *Journal of the American Chemical Society* **1998**, 120, 7537-7543.
- [62] Szafran, M.; Dega-Szafran, Z.; Nowak-Wydra, B.; Pietrzak, M., Differences in Proton-proton Coupling Constants of N⁺-CH₂-CH₂ Protons of Some Betaines, N⁺-(CH₂)₂₋₃-COO⁻, and Their Complexes in Aqueous Solution. *Journal of Molecular Structure* **2001**, 563-564, 555-564.