Computational Studies of Gas-Phase Peptide Acidities and Decomposition Mechanisms

Ashley S. McNeill*, Michele L. Stover, Carolynn Cassady, and David A. Dixon

Department of Chemistry, The University of Alabama
Box 870336, Shelby Hall, Tuscaloosa, AL 35487-0336

This thesis is submitted as part of the requirements for an ACS Certified B.S. in Chemistry from Austin Peay State University.

Chemistry 4990
Spring Semester 2011

Oral Defense: Monday April 25, 2011
Final Submission Date: Monday May 2, 2011

Thesis committee

Rebecca M. Jones, chair
G. Robert Shelton
Justin Oelgoetz
Robin B. Reed
Abstract

Computational chemistry methods were used to predict the thermodynamic properties of glycine and alanine dimers and trimers, as well as the effects of substitutions along the polymer backbone of whole-molecule properties. The predicted properties include the gas-phase acidity, aqueous \( \text{pK}_a \) and heat of formation. The anions were created by deprotonating the parent molecule at specific sites along the backbone, including both the N- and C-termini. Conformational searches of the neutrals and anions were performed with density functional theory with the B3LYP functional. The resulting optimized structures were used as starting guesses for the G3MP2 calculations. The freedom afforded by skeletal bond rotation allowed intramolecular hydrogen bonding to play a key role in stabilizing conformations. The acidities of the C-H and N-H bonds were much larger than expected due to the formation of enolate anions and show that it is thermodynamically stable for the protein backbone to be deprotonated.
Introduction

The study of gas-phase proton transfer reaction provides unique insights into the structures and energetics of peptides. Hydrogen bonding is critical to the determination of the three-dimensional structures and biological activities of polypeptides and proteins. Changing the protonation state can impact the hydrogen bonding in the molecule and alter properties such as solubility, hydrophobicity, and electrostatic interactions. The analysis of biomolecules by mass spectrometry requires an understanding of proton transfer reactions because the two most commonly used ionization techniques, electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI), involve the addition and subtraction of protons. Also, the sites of proton transfer reactions can affect the fragmentation patterns of peptide ions, which consequently impact the sequence information obtained from mass spectrometry experiments.

The sequencing of peptides and proteins by mass spectrometry has become a common tool in proteomics research. Gaining information on sequence is frequently a first step to understanding protein structure and function, which are very important in biological, biochemical, and medical studies. Biomolecule sequencing by tandem mass spectrometry (MS/MS) has most often employed fragmentation of positive ions. To date, mass spectral studies of the dissociation of negative peptide ions have been far less common than studies of positive peptide ions. Complementing positive ion studies with negative ion studies can substantially increase the amount of peptide sequence information obtained without adding significantly to either the cost or time involved. Using modern mass spectrometers, it is very easy to perform negative ion studies.

Although fragmentation of protonated peptides is what is most commonly studied,
deprotonated peptide fragmentation can also be used in sequencing. Thermodynamic properties, such as the gas-phase acidity (\(GA = \Delta G_{\text{acid}}\), where \(\Delta G\) is for the deprotonation reaction \(HA \rightarrow H^+ + A^-\), and is calculated by subtracting the reactant’s free energy value from the products’ combined free energy values, the values for which are given directly by the G3MP2 calculations) can provide valuable information to help in understanding peptide fragmentation using the less studied negative ion mode in mass spectrometry. A range of studies on positive ions of the amino acids have been performed; however, very few thermodynamic measurements have been performed for amino acids in terms of studies of anions. There are only five published reports that have involved measurements of amino acid GAs. Locke and McIver used the proton transfer equilibrium method to obtain the GAs of glycine and alanine, and Kebarle and coworkers also included glycine in their proton transfer equilibrium measurements of the GAs of ninety-six aliphatic carboxylic acids. Bowie and coworkers determined the GAs of nineteen amino acids from kinetic measurements of collision-induced dissociation (CID) on a proton bound dimer; however, they were unable to measure the GAs of aspartic acid and glutamic acid, the two most acidic amino acids of the nineteen, because these compounds were too stable to form the necessary dimer ions. Their low volatilities also prevented gas-phase equilibrium measurements from being performed on these compounds. The GA of serine has been reported. Recently, Poutsma and co-workers reported the GAs of all of the amino acids from experiments and also calculated them at the B3LYP level using the aug-cc-pVDZ basis set. Cassady, Dixon and coworkers reported the acidities of glutamic and aspartic acid from ion cyclotron resonance bracketing measurements and acidities of these two amino acids and glycine at the G3(MP2) level. This computational method was chosen on the basis of calculations of the acidities of very strong acids.
Computational Methods

Thermodynamic properties of molecular systems may be calculated based on several assumptions, such as the Born-Oppenheimer approximation and the treatment of the molecule as an ideal gas. The Gaussian programs (03, 09, etc.) generate the thermodynamic properties of molecules by calculation from basic statistical mechanic foundations. The partition function, the key to deriving most of the properties of the atoms within the molecule, may be calculated based on the unique identities of the atoms within the molecule and well-defined constants (i.e. atomic mass, Boltzmann’s constant, etc.). The total energy of a molecular system is a combination of several components – the zero-point energy and the translational, rotational, vibrational, and electronic motion contributions to the energy of the molecule.\textsuperscript{\textit{xxxiv}}

Calculations were performed at the DFT (density functional theory) and molecular orbital theory levels with the program Gaussian-09.\textsuperscript{\textit{xxiv}} The geometries were initially optimized with the B3LYP exchange-correlation functional\textsuperscript{\textit{xxv},\textit{xxvi}} and with the DZVP2 DFT-optimized basis set.\textsuperscript{\textit{xxvii}} Frequencies were calculated to show that all structures were energetic minima and to provide zero point and thermal corrections to the enthalpy and entropy so that free energies could be calculated for direct comparison to experiment. A range of structures were optimized for diglycine, dialanine, and trialanine, to determine the most stable structures which, in general, involve the strongest hydrogen bonding and deprotonation at the C-terminus. Protons were removed from the backbone chain and the resulting anions were geometrically optimized. The main sites of deprotonation were the N-terminus, backbone carbon, and the carboxylic alcohol group at the C-terminus.

In a recent study,\textsuperscript{\textit{xxiii}} it was shown that MP2/CBS (CBS = complete basis set) calculations with the augmented correlation-consistent basis sets\textsuperscript{\textit{xxviii}} can predict the acidities of
organic acids to better than 4 kcal/mol accuracy. The calculated values proved to be more acidic than the experimental values. Furthermore, it was shown that the G3(MP2) method\textsuperscript{xxix} improved the agreement between the experimental values and coupled cluster CCSD(T)/CBS values for the acidities to within about 1 kcal/mol. For example, the G3(MP2) value for the acidity of acetic acid, CH$_3$CO$_2$H, is 340.3 kcal/mol at 298 K, the experimental value is 341.5 ± 2.0 kcal/mol,\textsuperscript{xxx} and the MP2/CBS value is 337.2 kcal/mol. Because of these results, the G3(MP2) approach was used to calculate the gas phase acidities for the amino acids.

G3(MP2) is a modification on the Gaussian-3 (G3) technique for the calculation of molecular energies, which is a composite calculation that combines several \textit{ab initio} molecular orbital calculations in order to arrive at a total molecular energy. Typically the G3 calculations involve Moller-Plesset correlation terms up to fourth order – however, it is computationally cheaper to reduce the order of these MP terms, which is the origin of G3(MP2). G3(MP2) has a lesser accuracy than the higher-order G3 calculation, but not so much so that this calculation is not a reliable technique. The values derived by the G3(MP2) calculation are found by performing single point energy calculations at the QCISD(T)/6-31G(d) level and extrapolated to the complete basis set (G3MP2large) by combining with MP2 correlation level calculations. Additional corrections for zero point energy, spin-orbit, and empirical data are included to improve agreement with experimental values.\textsuperscript{xxxiii} G3(MP2) has an added advantage over DFT methods as the correlated molecular orbital methods used in G3(MP2) perform better in the prediction of hydrogen bond energies than do DFT methods. The G3(MP2) calculations directly generate values for the enthalpy and free energy of formation of the molecule at 298 K. The values for the parent molecule and the anion created by simple deprotonation may then be used to compute the free energy of the deprotonation reaction.
\[ \Delta G_{rxn} = G_f (\text{anion}) + G_f (\text{proton}) - G_f (\text{parent}) \]

In order to estimate the solution phase acidities of the amino acid, the Gaussian-03\textsuperscript{xxiv} program was used to calculate the free energy of solvation by using a self-consistent reaction field (SCRF) approach\textsuperscript{xxxi} with the COSMO (Conductor-like Screening Model) formalism\textsuperscript{xxxii} using the dielectric constant for water of 78.39. These methods were chosen based on past performance modeling molecules in an aqueous environment,\textsuperscript{xxii} and so that they may, in future work, be compared to the results on other similar systems by the Dickson group. On the basis of previous studies of acids,\textsuperscript{xxiii} the acidities were calculated relative to that of acetic acid, \( \text{CH}_3\text{CO}_2\text{H} \), which is well-established as \( pK_a = 4.76 \) in aqueous solution as shown by reaction (1).

\[
\text{CH}_3\text{CO}_2\text{H} + \text{amino acid anion} \rightarrow \text{CH}_3\text{CO}_2^- + \text{amino acid}
\]

The COSMO calculations directly generated values for the energies associated with placing the molecule in an aqueous environment. These energies may be combined to get the change in free energy resulting from solvation. The free energy of the above solvation reaction (reaction (1)) could then be calculated as follows.

\[
\Delta G_{solv} = (\Delta G_{rxn} (\text{parent}) + \Delta G_{rxn} (\text{acetate})) - (\Delta G_{rxn} (\text{anion}) + \Delta G_{rxn} (\text{acetatic acid}))
\]

This value for the solvated free energy of the reaction could then be used in a simple calculation to arrive at the reaction constant for reaction (1).

\[ K_a = e^{-\frac{\Delta G_{solv}}{RT}} \]

It is trivial, then, to calculate the \( pK_a \).

\[ pK_a = -\log (K_a) \]
Results and Discussion

The diglycine and dialanine parent molecules were modeled as well as the anions resulting from deprotonation along the peptide backbone. Also, the effect of substitution of various hydrogens with methyl groups was studied in order to better understand the mechanism for energy stabilization for the peptide chain. Trialanine was also modeled with its respective anions (derived using the same approach). In the following results, hydrogen bonding distances are shown in the images for which they apply. The site of removal of the proton is indicated by a bold arrow. The sites of deprotonation are named numerically beginning with the N-terminus (site 1) and continuing along the backbone, as is displayed in Figure 1 at right. For a dimeric peptide chain, the C-terminus acid group is labeled as site 5. For a trimeric peptide chain, the C-terminus acid group would be site 7. The dominant trend for the molecules explored was that, when available, C-terminus deprotonation resulted in the most acidic anion. All values for enthalpy, Gibb's free energy, and pKₐ are given in Tables 1, 2, and 3, and in Figures 2, 3, and 4 (See Appendix A).

Diglycines. The deprotonation at the C-terminus for diglycine was the most acidic (ΔH_{gas,rxn} = 332.3 kcal/mol) and had the lowest pK_a (6.01). Figure 2.a shows the deprotonation reactions of this unaltered diglycine molecule. The substituted diglycine deprotonation reactions will be compared to the results for this unaltered diglycine reaction.

The deprotonation at the carboxylic acid site was the most acidic for diglycine with methylated nitrogens on the

Figure 1
Deprotonation naming scheme for polypeptides.
backbone. The enthalpy value for the gas phase acidity forming this anion was 335.9 kcal/mol, which was 3.6 kcal/mol higher than that of the parent unaltered diglycine anion. Also, the GA was 328.2 kcal/mol, which was 3.9 kcal/mol higher than that of the parent. The pK\textsubscript{a} of this reaction is 1.88, which indicated that this deprotonation is actually more favorable than the unaltered diglycine deprotonation in solution. The lowest-energy conformation for this anion is comparable to the parent diglycine molecule, which is displayed in Figure 2.b. The primary difference in energy is likely related to the fact that the substituted methyl group on the interior nitrogen prevents additional hydrogen bonding from occurring that would further stabilize the anion. This lack of ability to hydrogen bond may make the carboxylic acid region of the molecule more susceptible to deprotonation.

Deprotonation of the interior nitrogen was the most acidic for diglycine with the carboxylic acid hydrogen replaced by a methyl group to form an ester at the C-terminus. However, this final conformation gave the greatest acidity when the original site of deprotonation was the backbone carbon closest to the ester group. Proton transfer then occurred from the neighboring nitrogen to the deprotonated carbon. The deprotonation enthalpy for this anion (\(\Delta H_{\text{gas,rxn}} = 353.5\) kcal/mol) is 20.4 kcal/mol higher than the most stable anion for the parent diglycine and 15.0 kcal/mol higher than the deprotonation enthalpy of the anion formed by the interior nitrogen deprotonation in the unaltered diglycine. The difference in the GA values for the ester and the parent diglycine is 21.8 kcal/mol, and the difference between the ester GA and that for interior nitrogen deprotonation of the parent diglycine is 15.7 kcal/mol. The pK\textsubscript{a} of this deprotonation is 17.93, which is 11.92 higher than the pK\textsubscript{a} for the unaltered species, indicating that this deprotonation is significantly less spontaneous in solution. The geometry for this anion is comparable to the geometry for the interior nitrogen deprotonation of
diglycine, and may be found in Figure 2.c. The steric hindrance of the added methyl group together with the reduction in the hydrogen bonding potential of the molecule most likely resulted in making the ester less acidic in the gas phase.

Deprotonation at the interior nitrogen also led to the most stable anion for the diglycine with the –OH of the carboxylic acid group replaced with an NH₂ group to form the terminal amide. The greatest acidity occurred when the original site of deprotonation was either of the backbone nitrogens, which both simplified to the structure with the final deprotonation site as the interior nitrogen. The deprotonation enthalpy for this anion (ΔH_{gas,rxn} = 346.2 kcal/mol) is 13.9 kcal/mol higher than for the most stable anion for the unaltered parent and 8.5 kcal/mol higher than that of the interior nitrogen anion formed from the unaltered parent. The GA for the interior nitrogen for the terminal amide is 15.2 kcal/mol higher than the GA for diglycine and 9.1 kcal/mol higher than the GA for the interior nitrogen anion for the parent. The pKₐ for this deprotonation is 19.99, which is 13.98 greater than that of the most favorable unaltered diglycine deprotonation. The geometric conformation of this anion does not closely resemble the geometry of the anion generated by deprotonation at the interior nitrogen for the parent diglycine, as the terminal amide anion structure has twisted to increase the hydrogen bonding between backbone atoms. The GA for the structure of the anion generated by loss of a proton from the interior nitrogen in the terminal amide is smaller than the GA for the ester because the terminal NH₂ group can still participate in stabilizing hydrogen bonding. The deprotonation pathways for this molecule may be found in Figure 2.d.

**Dialanines.** Deprotonation at the C-terminus for dialanine was the most acidic (ΔH_{gas,rxn} = 332 kcal/mol) and had the lowest pKₐ (3.56). The deprotonation enthalpy for dialanine is only 0.3 kcal/mol lower than that of diglycine, and the GA of dialanine is 0.2 kcal/mol higher than that of
diglycine. The dialanine deprotonation pKa is 2.45 lower than that of the diglycine deprotonation. These structures have comparable gas phase acidities due to the fact that the added methyl groups to the carbons of the backbone do not sterically hinder the geometric conformation that was most stable for the diglycine anion. In fact, the hydrogen bonding distances for the two most stable anions are comparable for diglycine and dialanine. An image of the deprotonation pathways for unaltered dialanine may be found in Figure 3.a. The following dialanine deprotonation reactions will be compared to the results for this unaltered dialanine reaction.

The most stable deprotonation site for anion formation for the dialanine molecule with methylated backbone carbons was the carboxylic acid site (pathways in Figure 3.b). The deprotonation enthalpy for this reaction is only 0.5 kcal/mol higher than that of the parent dialanine, and the GA is only 0.8 kcal/mol larger. The pKa (6.66) is only 3.1 higher than that of the unaltered dialanine anion. This small difference is likely due to a slight increase of steric hindrance in the dialanine, which is cancelled by the additional hydrogen bond found in the methylated species. The overall geometric conformation for both the parent and anion here are virtually identical to that of the parent dialanine and the corresponding anion.

The most acidic anion for dialanine with methylated backbone nitrogen was derived from deprotonation at the carboxylic acid site. The deprotonation enthalpy for this anion ($\Delta H_{\text{gas,rxn}} = 338.6$ kcal/mol) was 6.1 kcal/mol higher than that of the parent dialanine, and the GA was 6.0 kcal/mol higher. The pK$_a$ for this deprotonation is only 0.07 higher than that of the unaltered dialanine deprotonation, which indicates that this alteration to the backbone of dialanine has virtually no effect on the acidity of the molecule. This anion no longer has the ability to hydrogen bond as in the parent dialanine anion, resulting in a less sable anion in the substituted
species and a higher GA. An image of the deprotonation pathways for this molecule may be found in Figure 3.c.

The most energetically favorable deprotonation site for dialanine with an ester replacing the carboxylic acid group was at the interior nitrogen (pathway in Figure 3.d). The original site of deprotonation that was most energetically favorable was at the carbon closest to the C-terminus of the anion, though a proton transfer occurred between the interior nitrogen and the C-terminus carbon to arrive at the final optimized structure. The deprotonation enthalpy for this anion ($\Delta H_{\text{gas,rxn}} = 343.3$ kcal/mol) is 11.1 kcal/mol higher than the most stable anion for the parent dialanine and 6.0 kcal/mol higher than the anion for site 3 deprotonation in the parent. The difference in GA values between the ester and unaltered dialanine is 11.7 kcal/mol, and between the ester and deprotonation at the interior nitrogen of the parent dialanine is 5.7 kcal/mol. The pKa for this deprotonation is 16.61, which is 13.05 higher than that of the parent dialanine, indicating that this reaction requires a larger amount of energy to proceed forward.

The geometry of the interior nitrogen ester anion is comparable to that of the anion formed by deprotonation at the interior nitrogen of the parent dialanine, although the presence of the additional methyl group prevents hydrogen bonding between the backbone oxygen and the carboxylic acid group. The difference between most stable ester anion and the corresponding anion formed from the parent dialanine is less severe than the difference noted earlier between the analogous diglycine molecules because the additional methyl group doesn’t significantly alter the steric hindrance of the dialanine molecule, whereas it notably limits the flexibility of the unaltered diglycine molecule.

Deprotonation at the interior nitrogen led to the most stable anion for the dialanine with the –OH of the carboxylic acid group replaced with an NH$_2$. The original site of deprotonation
that leads to the lowest energy conformation is the N-terminus, which then undergoes a proton transfer from the interior nitrogen to give the final product. The deprotonation enthalpy for this anion ($\Delta H_{\text{gas,rxn}} = 345.8$ kcal/mol) is 13.8 kcal/mol higher than the most stable anion for the parent dialanine and 7.1 kcal/mol higher than that of the interior nitrogen anion for the parent dialanine. The GA for the terminal amide is 14.1 kcal/mol higher than that of the parent dialanine and 7.1 kcal/mol higher than the GA for anion formed by deprotonation of the interior nitrogen in the parent dialanine. The $pK_a$ for this deprotonation is 19.71, which is 16.15 higher than that of the parent dialanine. The conformation of the anion of the terminal amide does not closely resemble the geometry of the anion formed by deprotonation at the interior nitrogen for the parent dialanine, as the former structure has twisted to increase the hydrogen bonding between backbone atoms. The structure of the most stable anion for the terminal amide leads to a more acidic structure than the structure from the ester due to the fact that the NH$_2$ group may still participate in stabilizing hydrogen bonding.

**Trialanine.** Deprotonation at the C-terminus for trialanine resulted in the greatest acidity. As compared to dialanine, the deprotonation enthalpy and GA for trialanine are lower by –6.8 and -4.9 kcal/mol lower respectively. Trialanine is a longer peptide chain, so it has greater flexibility and a greater number of potential hydrogen bonding sites. Deprotonation at the N-terminus also gave a reasonably small $pK_a$, which was only 4.05 higher than the $pK_a$ for the C-terminus deprotonation. An image of the possible deprotonation pathways may be found in Figure 4.

**Conclusions**

A series of amino acid dimers, with and without various modifications to the backbone, and trialanine were geometrically optimized and the amino acids resulting from backbone deprotonation were explored. Molecules for which the carboxylic acid site was a viable option
for deprotonation were always the most acidic sites (lowest $\Delta H_{\text{gas}}$ and $pK_a$). Molecules for which the carboxylic acid group was either methylated or the OH was replaced with an NH$_2$ group always resulted in deprotonation at the interior nitrogen being the most acidic. Overall, increased hydrogen bonding within the molecule led to a more stable structure, thus somewhat lowering the energy of the molecule. Also, longer chains of amino acids provide greater flexibility and more sites of potential hydrogen bonding that serve to reduce the overall enthalpy for anion formation.
Table 1 Gas phase acidities and deprotonation enthalpies in kcal/mol and pKₐ’s for substituted diglycines.

<table>
<thead>
<tr>
<th>Substitution Pattern</th>
<th>Thermo Property</th>
<th>Site 1 (N-terminus)</th>
<th>Site 2</th>
<th>Site 3</th>
<th>Site 4</th>
<th>Site 5 (C-terminus)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent</td>
<td>ΔGᵢ</td>
<td>331.1</td>
<td>347.2</td>
<td>331.1</td>
<td>345.1</td>
<td>324.3</td>
</tr>
<tr>
<td></td>
<td>ΔHᵢ</td>
<td>337.7</td>
<td>353.6</td>
<td>337.7</td>
<td>353.0</td>
<td>332.3</td>
</tr>
<tr>
<td>N-Me</td>
<td>ΔGᵢrxn</td>
<td>364.3</td>
<td>346.4</td>
<td>a</td>
<td>345.9</td>
<td>328.2</td>
</tr>
<tr>
<td></td>
<td>ΔHᵢrxn</td>
<td>370.8</td>
<td>353.5</td>
<td>a</td>
<td>353.4</td>
<td>335.9</td>
</tr>
<tr>
<td></td>
<td>pKₐ</td>
<td>38.06</td>
<td>21.34</td>
<td>a</td>
<td>21.06</td>
<td>1.88</td>
</tr>
<tr>
<td>C(O)OCH₃</td>
<td>ΔGᵢrxn</td>
<td>346.1</td>
<td>355.4</td>
<td>346.1</td>
<td>346.0</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>ΔHᵢrxn</td>
<td>352.7</td>
<td>362.9</td>
<td>352.7</td>
<td>353.5</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>pKₐ</td>
<td>21.29</td>
<td>22.21</td>
<td>21.29</td>
<td>17.93</td>
<td>a</td>
</tr>
<tr>
<td>C(O)NH₂</td>
<td>ΔGᵢrxn</td>
<td>339.5</td>
<td>359.4</td>
<td>339.5</td>
<td>356.4</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>ΔHᵢrxn</td>
<td>346.2</td>
<td>365.7</td>
<td>346.2</td>
<td>363.8</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>pKₐ</td>
<td>19.99</td>
<td>30.44</td>
<td>19.99</td>
<td>29.28</td>
<td>a</td>
</tr>
</tbody>
</table>

a: Not possible for this substitution pattern
Table 2 Gas phase acidities and deprotonation enthalpies in kcal/mol and pKₐ's for substituted dialalanines.

<table>
<thead>
<tr>
<th>Substitution Pattern</th>
<th>Property</th>
<th>Site 1 (N-terminus)</th>
<th>Site 2</th>
<th>Site 3</th>
<th>Site 4</th>
<th>Site 5 (C-terminus)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unaltered</td>
<td>ΔGᵣₐ</td>
<td>330.5</td>
<td>347.4</td>
<td>331.5</td>
<td>343.4</td>
<td>324.5</td>
</tr>
<tr>
<td></td>
<td>ΔHᵣₐ</td>
<td>336.9</td>
<td>353.9</td>
<td>338.7</td>
<td>351.4</td>
<td>332.0</td>
</tr>
<tr>
<td></td>
<td>pKₐ</td>
<td>13.24</td>
<td>19.53</td>
<td>15.88</td>
<td>20.11</td>
<td>3.90</td>
</tr>
<tr>
<td>C-Me</td>
<td>ΔGᵣₓₙ</td>
<td>330.5</td>
<td>a</td>
<td>330.5</td>
<td>a</td>
<td>323.7</td>
</tr>
<tr>
<td></td>
<td>ΔHᵣₓₙ</td>
<td>338.1</td>
<td>a</td>
<td>338.1</td>
<td>a</td>
<td>332.5</td>
</tr>
<tr>
<td></td>
<td>pKₐ</td>
<td>16.93</td>
<td>a</td>
<td>16.93</td>
<td>a</td>
<td>6.66</td>
</tr>
<tr>
<td>N-Me</td>
<td>ΔGᵣₓₙ</td>
<td>348.6</td>
<td>339.7</td>
<td>a</td>
<td>353.9</td>
<td>331.4</td>
</tr>
<tr>
<td></td>
<td>ΔHᵣₓₙ</td>
<td>353.3</td>
<td>346.9</td>
<td>a</td>
<td>362.4</td>
<td>338.6</td>
</tr>
<tr>
<td></td>
<td>pKₐ</td>
<td>15.15</td>
<td>17.61</td>
<td>a</td>
<td>25.11</td>
<td>3.63</td>
</tr>
<tr>
<td>C(O)OCH₃</td>
<td>ΔGᵣₓₙ</td>
<td>336.2</td>
<td>351.1</td>
<td>344.1</td>
<td>335.7</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>ΔHᵣₓₙ</td>
<td>343.1</td>
<td>358.6</td>
<td>350.9</td>
<td>343.3</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>pKₐ</td>
<td>18.75</td>
<td>26.59</td>
<td>24.51</td>
<td>16.61</td>
<td>a</td>
</tr>
<tr>
<td>C(O)NH₂</td>
<td>ΔGᵣₓₙ</td>
<td>338.6</td>
<td>360.0</td>
<td>339.6</td>
<td>340.4</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>ΔHᵣₓₙ</td>
<td>345.8</td>
<td>367.0</td>
<td>347.6</td>
<td>348.3</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>pKₐ</td>
<td>19.71</td>
<td>32.27</td>
<td>20.14</td>
<td>19.85</td>
<td>a</td>
</tr>
</tbody>
</table>

a: Not possible for this substitution pattern
Table 3 Gas phase acidities and deprotonation enthalpies in kcal/mol and pK\(_a\)'s for trialalanines.

<table>
<thead>
<tr>
<th>Thermo Property</th>
<th>Site 1 (N-terminus)</th>
<th>Site 2</th>
<th>Site 3</th>
<th>Site 4</th>
<th>Site 5</th>
<th>Site 6</th>
<th>Site 7 (C-terminus)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\Delta G_{\text{rxn}})</td>
<td>321.4</td>
<td>351.4</td>
<td>335.2</td>
<td>352.3</td>
<td>338.9</td>
<td>343.1</td>
<td>319.6</td>
</tr>
<tr>
<td>(\Delta H_{\text{rxn}})</td>
<td>326.8</td>
<td>354.9</td>
<td>342.7</td>
<td>359.0</td>
<td>345.4</td>
<td>349.2</td>
<td>325.2</td>
</tr>
<tr>
<td>pK(_a)</td>
<td>7.95</td>
<td>24.33</td>
<td>20.06</td>
<td>30.26</td>
<td>17.69</td>
<td>24.77</td>
<td>3.90</td>
</tr>
</tbody>
</table>
Figure 2. Parent and anion structures for substituted diglycines. Blue arrows indicate the site of initial protonation before optimization. Hydrogen bond distances in Å and energies in kcal/mol.

2.a Diglycine
Figure 2. Parent and anion structures for substituted diglycines. Blue arrows indicate the site of initial protonation before optimization. Hydrogen bond distances in Å and energies in kcal/mol.

2.b Diglycine (N-Me)
Figure 2. Parent and anion structures for substituted diglycines. Blue arrows indicate the site of initial protonation before optimization. Hydrogen bond distances in Å and energies in kcal/mol.

2.c Diglycine (C(O)OCH3)
**Figure 2.** Parent and anion structures for substituted diglycines. Blue arrows indicate the site of initial protonation before optimization. Hydrogen bond distances in Å and energies in kcal/mol.

**2.d** Diglycine (C(O)NH2)
Figure 3. Parent and anion structures for substituted dialanines. Blue arrows indicate the site of initial protonation before optimization. Hydrogen bond distances in Å and energies in kcal/mol.

3.a Dialanine
Figure 3. Parent and anion structures for substituted dialanines. Blue arrows indicate the site of initial protonation before optimization. Hydrogen bond distances in Å and energies in kcal/mol.

3.b Dialanine (C-Me)
**Figure 3.** Parent and anion structures for substituted dialanines. Blue arrows indicate the site of initial protonation before optimization. Hydrogen bond distances in Å and energies in kcal/mol.

**3.c** Dialanine (N-Me)
Figure 3. Parent and anion structures for substituted dialanines. Blue arrows indicate the site of initial protonation before optimization. Hydrogen bond distances in Å and energies in kcal/mol.

3.d Dialanine (C(O)OCH3)
Figure 3. Parent and anion structures for substituted dialanines. Blue arrows indicate the site of initial protonation before optimization. Hydrogen bond distances in Å and energies in kcal/mol.

3.e Dialanine (CO(NH2))
Figure 4. Parent and anion structures for trialanine. Blue arrows indicate the site of initial protonation before optimization. Energies in kcal/mol.
References


iv J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, Mass Spectrom. Rev. 9 (1990), 37-70


of Biopolymers from Venoms 47 (2006), 664-675.


xxv A. D. Becke, J. Chem. Phys. 98 (1993), 5648-5652
xxvii DZVP2

xxxiii Curtiss, L. Argonne National Lab.

http://chemistry.anl.gov/OldCHMwebsiteContent/compmat/g3theory.htm Accessed 26 Apr, 2011.

xxxiv Ochterski, J. “Thermochemistry in Gaussian.” Gaussian, Inc.