Low-Energy Tautomers and Conformers of Neutral and Protonated Arginine

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · DECEMBER 2001

Impact Factor: 12.11 · DOI: 10.1021/ja011357l · Source: PubMed

Low-Energy Tautomers and Conformers of Neutral and Protonated Arginine

Janusz Rak,[†] Piotr Skurski,[†,‡] Jack Simons,[‡] and Maciej Gutowski*[,†,§]

Contribution from the Department of Chemistry, University of Gdańsk, ul. Sobieskiego 18, 80-952 Gdańsk, Poland, Henry Eyring Center for Theoretical Chemistry, Department of Chemistry, University of Utah, Salt Lake City, Utah 84112, and Environmental Molecular Sciences Laboratory, Theory, Modeling & Simulations, Pacific Northwest National Laboratory, Richland, Washington 99352

Received June 4, 2001

Abstract: The relative stabilities of zwitterionic and canonical forms of neutral arginine and of its protonated derivative were studied by using ab initio electronic structure methods. Trial structures were first identified at the PM3 level of theory with use of a genetic algorithm to systematically vary geometrical parameters. Further geometry optimizations of these structures were performed at the MP2 and B3LYP levels of theory with basis sets of the 6-31++G** quality. The final energies were determined at the CCSD/6-31++G** level and corrected for thermal effects determined at the B3LYP level. Two new nonzwitterionic structures of the neutral were identified, and one of them is the lowest energy structure found so far. The five lowest energy structures of neutral arginine are all nonzwitterionic in nature and are clustered within a narrow energy range of 2.3 kcal/mol. The lowest energy zwitterion structure is less stable than the lowest nonzwitterion structure by 4.0 kcal/mol. For no level of theory is a zwitterion structure suggested to be the global minimum. The calculated proton affinity of 256.3 kcal/mol and gas-phase basicity of 247.8 kcal/mol of arginine are in reasonable agreement with the measured values of 251.2 and 240.6 kcal/mol, respectively. The calculated vibrational characteristics of the low-energy structures of neutral arginine provide an alternative interpretation of the IR-CRLAS spectrum (Chapo et al. *J. Am. Chem. Soc.* **1998**, *120*, 12956−12957).

I. Introduction

Zwitterions are neutral compounds that contain oppositely charged centers. Such compounds are common in solution and in solids as their dipolar character experiences strong electrostatic stabilization in such environments. Zwitterions are known to occur in amino acids, peptides, and nucleic acid bases, and they are used to construct novel materials[1−3] and as reagents in organic synthesis.[4] In water solutions, over a wide range of pH, amino acids exist primarily in their zwitterionic forms with the carboxyl group deprotonated and one of the nitrogen atoms protonated.[5] Such zwitterions are thought to play important roles in the structure and function of peptides and proteins, and are encountered in the structure of turns and bends of RNA.[6]

In the gas phase, however, the zwitterionic charge separation is not stabilized by the environment. As a result, the amino acid's canonical (**C**) tautomer H_2N-CHR-COOH is likely to be more stable than the corresponding zwitterions (**Z**). Moreover, for some amino acids (e.g., glycine), the zwitterionic structure does not even correspond to a local minimum on the gas-phase potential energy surface.[7,8]

There have been attempts to identify a zwitterionic tautomer of a neutral amino acid that is stable and the global energy minimum in the gas phase,[9] but these findings were inconclusive.[10] Other recent efforts concentrated on hydrated,[11,12] protonated,[9,13,14] and alkali cationized[13−15] amino acids. In addition, the preference of arginine to form stable positively or negatively charged aggregates that are intermolecularly bound by salt bridges has been reported.[16]

Arginine is probably the most promising common naturally occurring amino acid to form a stable zwitterion in the gas phase due to its extremely basic guanidine side chain ($-(CH_2)_3$-NH-C-(NH_2)NH). The question of which tautomeric form (**Z** or **C**) of neutral arginine is dominant in the gas phase has recently

* Address correspondence to this author: (e-mail) maciej.gutowski@pnl.gov.
† University of Gdańsk.
‡ University of Utah.
§ Pacific Northwest National Laboratory.

(1) Spencer, T. A.; Onofrey, T. J.; Cann, R. O.; Russel, J. S.; Lee, L. E.; Blanchard, D. E.; Castro, A.; Gu, P.; Jiang, G.; Shechter, I. *J. Org. Chem.* **1999**, *64*, 807−818.
(2) Biegle, A.; Mathis, A.; Galin, J.-C. *Macromol. Chem. Phys.* **1999**, *200*, 1393−1406.
(3) Hu, W.; Haddad, P. R.; Hasebe, K.; Tanaka, K. *Anal. Commun.* **1999**, *36*, 97−100.
(4) Gothelf, K. V.; Jorgensen, K. A. *Chem. Rev.* **1998**, *98*, 863−909.
(5) Cantor, C. R.; Schimmel, P. R. *Biophysical Chemistry*, Part I; W. H. Freeman and Company: New York, 1980; pp 41−53.
(6) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.
(7) Jensen, J. H.; Gordon, M. S. *J. Am. Chem. Soc.* **1991**, *113*, 7917−7924.
(8) Ding, Y.; Krogh-Jespersen, K. *Chem. Phys. Lett.* **1992**, *199*, 261−266.
(9) Price, W. D.; Jockusch, R. A.; Williams, E. R. *J. Am. Chem. Soc.* **1997**, *119*, 11988−11989.
(10) Chapo, C. J.; Paul, J. B.; Provencal, R. A.; Roth, K.; Saykally, R. J. *J. Am. Chem. Soc.* **1998**, *120*, 12956−12957.
(11) Gordon, M. S.; Jensen, J. H. *Acc. Chem. Res.* **1996**, *29*, 536−543.
(12) Nagy, P. I.; Noszal, B. *J. Phys. Chem. A* **2000**, *104*, 6834−6843.
(13) Wyttenbach, T.; Matthias, W.; Bowers, M. T. *J. Am. Chem. Soc.* **2000**, *122*, 3458−3464.
(14) Wyttenbach, T.; Bushnell, J. E.; Bowers, M. T. *J. Am. Chem. Soc.* **1998**, *120*, 5098−5103.
(15) Wyttenbach, T.; Matthias, W.; Bowers, M. T. *Int. J. Mass Spectrom.* **1999**, *182/183*, 243−252.
(16) Julian, R. R.; Hodyss, R.; Beauchamp, J. L. *J. Am. Chem. Soc.* **2001**, *123*, 3577−3583.

10.1021/ja011357l CCC: $20.00
Published on Web 11/03/2001

been addressed in experimental studies.[9,10,13,15−18] Williams and co-workers concluded, on the basis of blackbody infrared radiative dissociation plus Fourier transform-mass spectrometry measurements, that protonated dimers of arginine are bound in a salt bridge.[9] Moreover, the results of their computational study at the BLYP/6-31G* and MP2/6-31G* levels suggested that a zwitterion form of arginine is the global minimum on the potential energy surface, lower by 1 kcal/mol than the lowest canonical tautomer.

Saykally and co-workers, however, did not confirm the dominance of the zwitterion of arginine in their infrared cavity ringdown laser absorption spectroscopy (IR-CRLAS) experiments.[10] The latter workers expected transitions at ca. 1700 cm^{-1} for carbonyl stretches in the canonical arginine and transitions at ca. 1500−1600 cm^{-1} for asymmetric and symmetric stretches of the carboxylate residue in zwitterionic tautomers. Their spectrum, recorded in the 1550−1750 cm^{-1} region, revealed two peaks at 1666 and 1693 cm^{-1}, which were assigned to carbonyl stretches of carboxylic acid groups. The absence of bands in the 1500−1660 cm^{-1} region, associated with the carboxylate stretch modes, suggested a small or vanishing population of the zwitterion. It was pointed out, however, that there may be a significant barrier that separates the canonical and zwitterion forms of arginine as a result of which the thermodynamically unstable form may have a sufficiently long lifetime to be observed experimentally. Moreover, the sources of arginine used in the Williams[9] and Saykally[10] experiments are different. One employs solution electrospray at 37−149 °C; the other employs a heated pulsed beam source of pure arginine at 170 °C. In summary, the IR-CRLAS data were interpreted to suggest that the gas-phase arginine exists predominantly in its **C** tautomeric form.

The geometrical shapes of protonated, sodiated, and cesiated arginine were probed in the gas phase by using the ion mobility based ion chromatography method.[13,15] It has been suggested that the alkali-cationized arginine forms a salt bridge structure, related to the zwitterion form of this amino acid. Results from the collisionally activated dissociation experiments of Williams et al.[17] and the kinetic experiments of Cerda and Wesdemiotis[18] indicated, however, that the structure of gas-phase arginine−alkali metal cation complexes depends on the size of the alkali metal cation. For Li^+ and Na^+, a nonzwitterion arginine solvates the metal ion. For larger metal ions, a salt bridge is formed in which arginine exists as a zwitterion. The results of Beauchamp et al. confirmed that K^+ stabilizes arginine in a different way than Li^+ or Na^+.[16]

Determination of the global-minimum structure for arginine is a complex problem because it has two kinds of proton donor groups (OH and NH), six proton acceptor sites (N's and O's), and six (or seven, depending on the tautomer) bonds about which rotation may occur. Thus, its potential energy surface supports numerous tautomers and conformers stabilized by different intramolecular hydrogen bonds. The multitude of low-energy structures and the small energy differences among them complicate a definite theoretical prediction of the global minimum. Therefore, efficient minimum-energy search algorithms are required to explore the complex structural space of this molecule and large basis sets and highly correlated ab initio methods are required to determine relative energies of minimum-energy structures. Moreover, thermal energetic and entropic effects may contribute to the stability of a particular form.

Maksic and Kovacevic (MK) performed an extensive computational investigation of arginine at the second-order Møller−Plesset (MP2) and density functional theory levels.[19] They concluded that the most stable structure was canonical, but the energy difference between the lowest zwitterion and canonical structures was relatively small (within 1−3 kcal/mol depending on the theoretical model applied). Our group has recently extended their study and new zwitterion and canonical structures were identified with energies lower than those previously known.[20] However, the lowest canonical structure was still found to be lower in energy than the lowest zwitterion by 2.8 kcal/mol, although the structural space of this complex molecule was explored only in a limited range.

In the current study we perform a more extensive scan of the structural space for the neutral and protonated arginine using a genetic algorithm.[21] The low-energy structures identified in this search initially at the semiempirical PM3 level of theory are further refined at the MP2 and density functional theory (DFT) levels. Final energies are calculated at the coupled cluster level of theory with single and double excitations (CCSD) and the 6-31++G** basis set.

Two new canonical structures of the neutral are identified and one of them is found to be the lowest energy structure identified so far. There are at least five canonical structures that are lower in energy than the lowest energy zwitterion, with the zwitterion being unstable with respect to the lowest energy canonical structure by 4.0 kcal/mol. The proton affinity calculated at the CCSD/6-31++G** level of 256.3 kcal/mol is in reasonable agreement with the experimental value of 251.2 kcal/mol.[22]

We also report characteristic vibrational frequencies and intensities for low-energy structures of neutral arginine to facilitate further interpretation of its IR-CRLAS spectrum.[10] Our results suggest that the two experimentally observed transitions at 1666 and 1693 cm^{-1} are related to NHX (X = H or C) bending and C═N stretching modes rather than to carbonyl group stretching modes. Moreover, we note that low-energy zwitterion structures of arginine would also have intense IR transitions in the same 1600−1700 cm^{-1} region. Therefore, we believe that the experimental results of ref 10 are not sufficient to resolve which tautomeric form of arginine dominates in the gas phase.

Our ultimate goal in this series of studies on amino acids[20,23,24] is to determine whether an excess electron can stabilize a zwitterionic molecular structure in the gas phase. So far, we have demonstrated that the instability of the zwitterion structure of glycine is reduced by 8 kcal/mol upon attachment of an excess electron and a local zwitterionic minimum develops on the anionic potential energy surface.[23] Similarly, excess electron attachment provides an extra stabilization of ca. 6 kcal/mol for betaine, which is a permanent zwitterion.[24] The current study provides us with detailed knowledge of the potential energy surface of the neutral arginine. It is very plausible that the relatively small instability of the zwitterionic structure of

(17) Jockusch, R. A.; Price, W. D.; Williams, E. R. *J. Phys. Chem. A* **1999**, *103*, 9266−9274.
(18) Cerda, B. A.; Wesdemiotis, C. *Analyst* **2000**, *125*, 657−660.
(19) Maksic, Z. B.; Kovacevic, B. *J. Chem. Soc., Perkin Trans. 2* **1999**, 2623−2629.
(20) Skurski, P.; Gutowski, M.; Barrios, R.; Simons, J. *Chem. Phys. Lett.* **2001**, *337*, 143−150.
(21) Holland, J. H. In *Adaptation in natural and artificial systems*; The University of Michigan Press: Ann Arbor, 1975.
(22) Hunter E. P. L.; Lias S. G. *J. Phys. Chem. Ref. Data* **1998**, *27*, 413−431.
(23) Gutowski, M.; Skurski, P.; Simons, J. *J. Am. Chem. Soc*. **2000**, *122*, 10159−10162.
(24) Rak, J.; Skurski, P.; Gutowski, M. *J. Chem. Phys*. **2001**, *114*, 10673−10681.

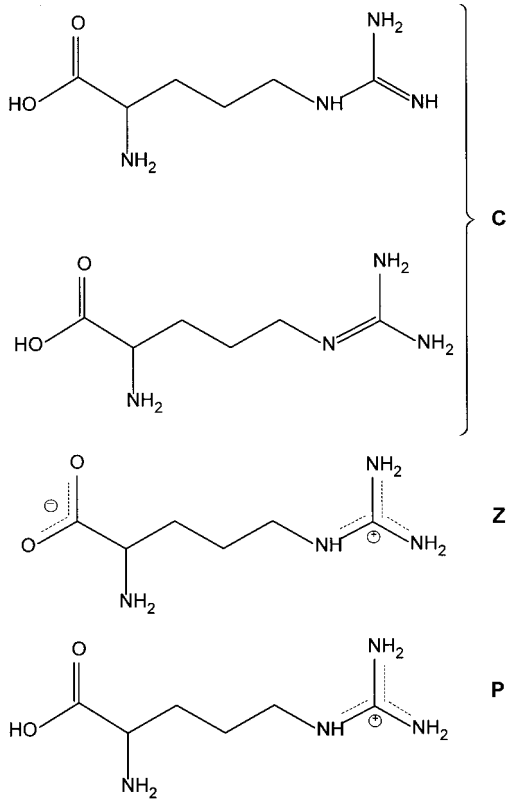

Figure 1. The structures of neutral and protonated arginine considered in this study.

arginine (4.0 kcal/mol) may indeed be suppressed by excess electron attachment. The results of our ongoing research on the anionic forms of arginine will be described in a forthcoming publication.[25]

II. Methods

A simple genetic algorithm[26] (SGA) was used to identify the lowest energy conformers for two structures of canonical arginine (**C**), the zwitterionic structure (**Z**), and the protonated structure (**P**), see Figure 1. We also considered another zwitterionic structure with the α-amino rather than the guanidine group protonated. This structure, however, proved to be geometrically unstable and collapsed in the course of geometry optimization to the canonical structure **C**. For the protonated arginine, two other structures were considered with the (i) α-amino group protonated and the COOH and guanidine groups intact and the (ii) guanidine group protonated and another proton transferred from the COOH group to the α-amino group. Both structures, however, proved to be unstable and collapsed to the protonated structure **P** depicted in Figure 1.

The number of single bonds (C−C, C−N, and C−O) considered for conformational exploration is seven for the second canonical structure and eight for other structures, see Figure 1. Consequently, we assumed a seven- or eight-element coding sequence as the natural length of a genotype that describes a conformation. Moreover, a dihedral angle can change from 0° to 360°, and a 60° increment has been assumed to ensure a complete scan of the potential energy surface. Accordingly, the coding alphabet was chosen to consist of natural numbers in the range from 1 to 6. The seven- or eight-element sequence with six different values at every position gives 6^7 or $6,^8$ i.e., 279 936 or 1 679 616, possibilities that define the size of the conformational space.

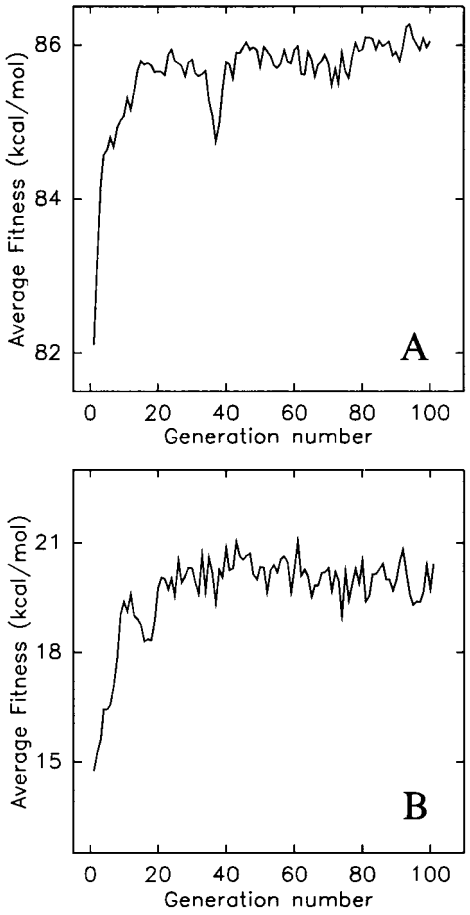

Figure 2. The dependence of average fitness on the generation number for canonical (A) and protonated (B) arginine.

For every structure (**C**, **Z**, or **P**) the population was kept constant and consisted of 30 individuals. The initial populations were selected in a random way ensuring their maximum diversity (i.e., all individuals within a given population differed one from another). Moreover, the low-energy individuals described in ref 20 and labeled here **C1**, **C2**, **C3** and **Z1**, **Z2**, **Z3** have been introduced to the initial population of the **C** and **Z** structures, respectively. The SGA program was executed several times by using various initial populations to ensure thorough exploration of the conformational space.

During execution of the SGA program, each genotype was decoded to the actual dihedral angles and the resulting geometry was optimized within the semiempirical PM3 method.[27] This method was selected because it is parametrized to reproduce formation energies of hydrogen bonds, which play an important role in the stability of different structures of arginine. The fitness of each individual was determined from its heat of formation (ΔH_f) calculated at the PM3 level. Our SGA program is formulated to maximize fitness. Thus fitness was defined as the negative of ΔH_f for the neutral arginine, for which the values of ΔH_f were always negative. For the protonated arginine, on the other hand, fitness was defined as the negative of ΔH_f shifted upward by 70 kcal/mol. The values of fitness were subsequently used to choose conformations for mating by using the roulette wheel selection scheme.[26] The crossover and mutation probabilities were assumed to be equal to 0.6 and 0.025, respectively.[28] We used the fitness linear scaling technique[26] to avoid a premature convergence and to allow the algorithm to maintain diversity and differentiate between very similar individuals. The searching process was run up to 100 generations. The dependence of the average fitness in population on the generation number is displayed in Figure 2 for the canonical and

(25) Skurski, P.; Rak, J.; Simons, J.; Gutowski, M. *J. Am. Chem. Soc*. **2001**. Published ASAP October 11, 2001.

(26) Goldberg, D. E. In *Genetic Algorithm in Search, Optimization and Machine Learning*; Addison-Wesley Publishing Company, Inc.: Reading, MA, 1989.

(27) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209−220, 221−264.

(28) De Jong, K. A. An analysis of the bahavior of a class of genetic adaptive systems; Doctoral dissertation, University of Michigan; *Dissertation Abstr. Internatl.* **1975**, *36* (10), 5140B.

protonated arginine. The average fitness clearly plateaus beyond the 50th generation.

The low-energy conformers of the **C**, **Z**, and **P** structures, selected within the SGA scheme, have been further optimized at the MP2 level as well as by applying the DFT method with a hybrid B3LYP functional.[29−31] The latter approach was also used to calculate harmonic vibrational frequencies and zero-point energies which enabled us to estimate selected thermodynamic characteristics (e.g., population of different structures in the gas-phase equilibrium, proton affinity, and gas basicity).

Moreover, single-point CCSD[32,33] calculations were performed for every such structure at its respective MP2 minimum geometry. Since the systems studied contain 12 heavy atoms and up to 15 hydrogen atoms, which leads to 320 contracted basis functions in the most elaborate basis set chosen, we had to limit the highest level of our calculations to CCSD. Although we were not able to attain a better treatment of electron correlation, recent work of Nguyen et al.[34] on glycine has shown that the energy ordering for canonical and zwitterionic forms is the same at the CCSD and CCSD(T) levels. Moreover, a new report of Fogarasi[35] on tautomers and rotamers of cytosine also shows that CCSD-level predictions are in good agreement with experimental data. For these reasons, we feel it is appropriate to use the CCSD method. Finally, the core 1s orbitals of C, N, and O were excluded from electron correlation treatments at the MP2 and CCSD levels.

The geometry optimizations and frequency calculations were performed with basis sets of 6-31++G** quality.[36,37] All protonated species were optimized at the MP2 and B3LYP levels with the 6-31++G** basis set and harmonic frequencies were determined at the B3LYP level. In calculations for the neutral molecule, the 6-31++G** sets were supplemented with a set of diffuse functions to make the predicted properties, and energies in particular, consistent with those that we are presently calculating for the corresponding anionic species. In MP2 calculations, we centered even-tempered five-term s and five-term p sets of diffuse functions on the positive pole of the molecular dipole of each species. This was either N(11) or N(12) (for atom numbering see Figure 3). The extra diffuse s and p functions shared exponent values, the geometric progression ratio was equal to 5.0, and we started to build up the exponents from the lowest sp exponent included in the 6-31++G** basis set for nitrogen. This gives lowest exponents of 2.0448×10^{-5} for both s and p symmetries. In the B3LYP calculations, the 6-31++G** basis set was supplemented with even-tempered two-term s and two-term p sets of diffuse functions centered on C(10) with the lowest exponent being 2.556×10^{-3}. We also verified that the minimum energy structures, harmonic vibrational frequencies, and relative energies of the isomers examined here are insensitive to the presence of extra diffuse sets. For instance, the relative energies of neutral structures, determined at the MP2 level, are affected by less than 0.03 kcal/mol when the extra diffuse functions are removed.

Harmonic vibrational frequencies determined at the B3LYP level are usually of high quality.[38] The errors, determined through comparison with experimental data, are frequently systematic in nature and a common procedure is to use "scaling factors" to correct for remaining discrepancies.[38] In the course of this work we recognized that the IR-CRLAS spectrum[10] may be more complicated than originally assumed and the assignment of transitions at 1666 and 1693 cm^{-1} to the carbonyl stretch of the carboxylic acid group (ultimately supporting the dominance of **C** structures) may be premature. Therefore, we calibrated our scaling factor to reproduce the measured carbonyl stretch frequency of a model compound with a carboxylic acid group; to this end, we used acetic acid.[39] This resulting scaling factor of 0.9813 was used for

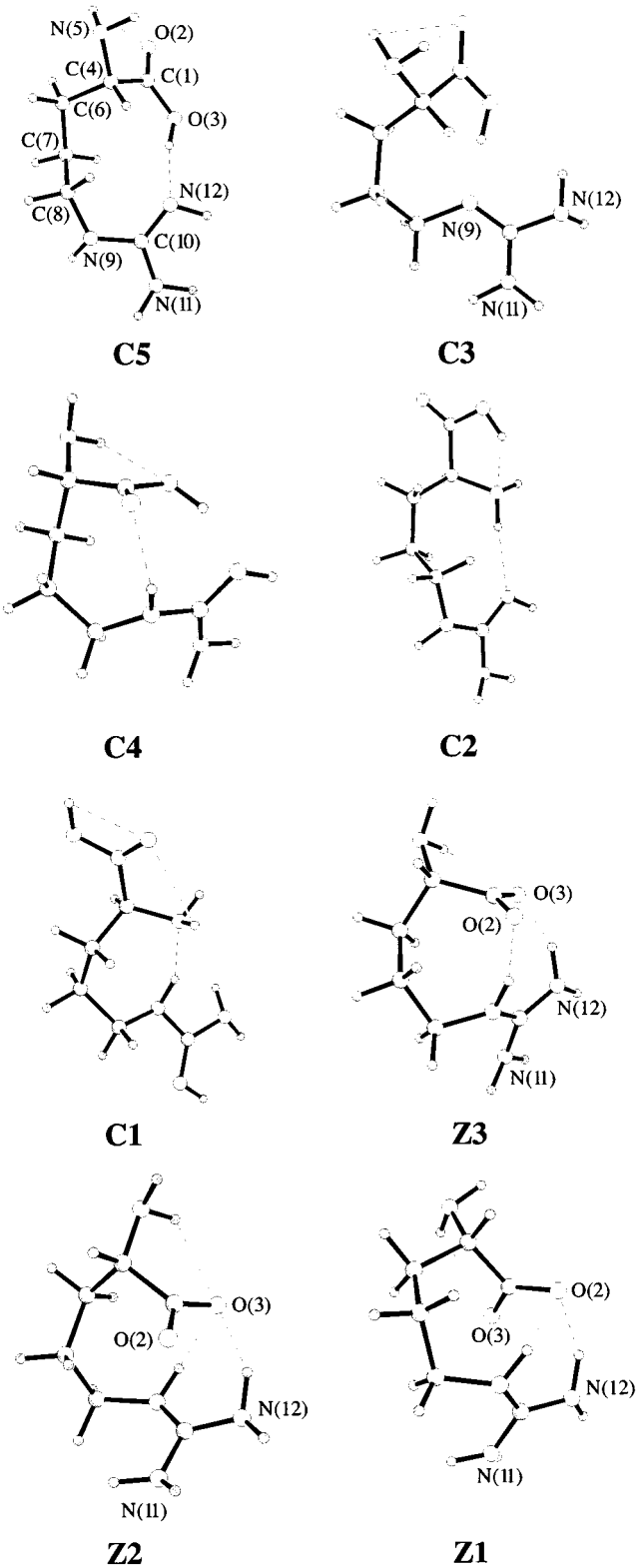

Figure 3. The equilibrium MP2 geometries for various conformers/tautomers of neutral arginine.

(29) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(30) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(31) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.
(32) Bartlett, R. J.; Stanton, J. F. In *Reviews in Computational Chemistry*; Lipkowitz, K., Boyd, D. B., Eds.; VCH: New York, 1994; Vol. V.
(33) Hampel, C.; Peterson, K.; Werner, H.-J. *Chem. Phys. Lett.* **1992**, *190*, 1−12.
(34) Nguyen, D. T.; Scheimer, A. C.; Andzelm, J. W.; Sirois, S.; Salahub, D. R.; Hagler, A. T. *J. Comput. Chem.* **1997**, *18*, 1609−1631.
(35) Fogarasi, G. *J. Mol. Struct.* **1997**, *413*−*414*, 271−278.
(36) Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724−728.
(37) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257−2261.
(38) Jensen, F. *Introduction to Computational Chemistry*; Wiley: Chichester, 1999.
(39) Shimanouchi T. *Tables of Molecular Vibrational Frequencies Consolidated;* National Bureau of Standards: Washington, DC, 1972; Vol. I, pp 1−160.

all frequencies below 2000 cm^{-1}. For higher frequency modes we used a standard scaling factor of 0.9613 recommended by Wong.[40]

The presence of many low-energy structures for neutral and protonated arginine prompted us to determine the populations of these structures in the gas-phase equilibrium. First, we selected a reference structure *R* for a given species (neutral or protonated). Next, for every structure *M* other than the reference structure *R* we determined the equilibrium constant K_M

$$K_M = [M]/[R] \tag{1}$$

from the difference in Gibbs free energies for *M* and *R*. The fraction of *M* in the equilibrated sample is given by

$$x_M = K_M/(1 + K_1 + K_2 + ...) \tag{2}$$

where the sum in the denominator goes through all structures for a given species. The fraction of *R* in the sample is

$$x_R = 1/(1 + K_1 + K_2 + ...) \tag{3}$$

The proton affinity (PA) of arginine is defined as the negative of the enthalpy change for the hypothetical gas-phase reaction

$$\mathrm{Arg(g)} + \mathrm{H^+\,(g)} \rightarrow \mathrm{ArgH^+\,(g)} \tag{4}$$

whereas gas-phase basicity (GB) is the negative of the Gibbs free energy change. Both are usually provided at $T = 298$ K. Our calculations indicate that both neutral and protonated arginine may exist in a few low-energy sructures at $T = 298$ K. Therefore, the values of the thermodynamic functions *I* (enthalpy *H* or Gibbs free energy *G*) for the species *A* (Arg or $ArgH^+$) were obtained through averaging over low-energy structures A_i:

$$I_{\mathrm{avg}}(A) = \sum_i x_i \cdot I(A_i) \tag{5}$$

where x_i is the fraction of A_i in the equilibrated sample, see eqs 2 and 3. Then,

$$\mathrm{PA} = H_{\mathrm{avg}}(\mathrm{Arg}) + H(\mathrm{H^+}) - H_{\mathrm{avg}}(\mathrm{ArgH^+}) \tag{6}$$

$$\mathrm{GB} = G_{\mathrm{avg}}(\mathrm{Arg}) + G(\mathrm{H^+}) - G_{\mathrm{avg}}(\mathrm{ArgH^+}) \tag{7}$$

To summarize, our final results are based on CCSD/6-31++G** energies determined at the optimal MP2 geometries. The zero-point vibrational corrections and thermal contributions to energy and entropy are determined from the B3LYP/6-31++G** molecular geometries and unscaled harmonic frequencies. The scaled frequencies are used only to interpret the IR-CRLAS spectrum. All calculations were performed with the GAUSSIAN 98[41] and MOLPRO[42] programs on SGI Origin2000 numerical servers, a DEC Alpha 533au two-processor workstation, and Cray SV1 computers.

III. Results

The low-energy structures of neutral and protonated arginine are displayed in Figures 3 and 4, respectively. Two new structures of canonical arginine, labeled **C4** and **C5** in Figure 3, were identified in the course of this work, and **C5** proved to be the lowest energy structure for every level of theory considered here. These structures were overlooked in our earlier studies on tautomers and conformers of gas-phase arginine.[20] Clearly, the genetic algorithm employed by us proved to be very useful as it also provided the lowest energy structure, labeled **P1** in Figure 4, for protonated arginine. This structure was *not* identified in a more conventional approach in which we protonated the most basic site for the hitherto determined variety of low-energy structures of canonical and zwitterionic arginine (**C1**−**C5** and **Z1**−**Z3**). The previously lowest energy conformer of protonated arginine identified by MK[19] corresponds to our structure **P3**, see Figure 4.

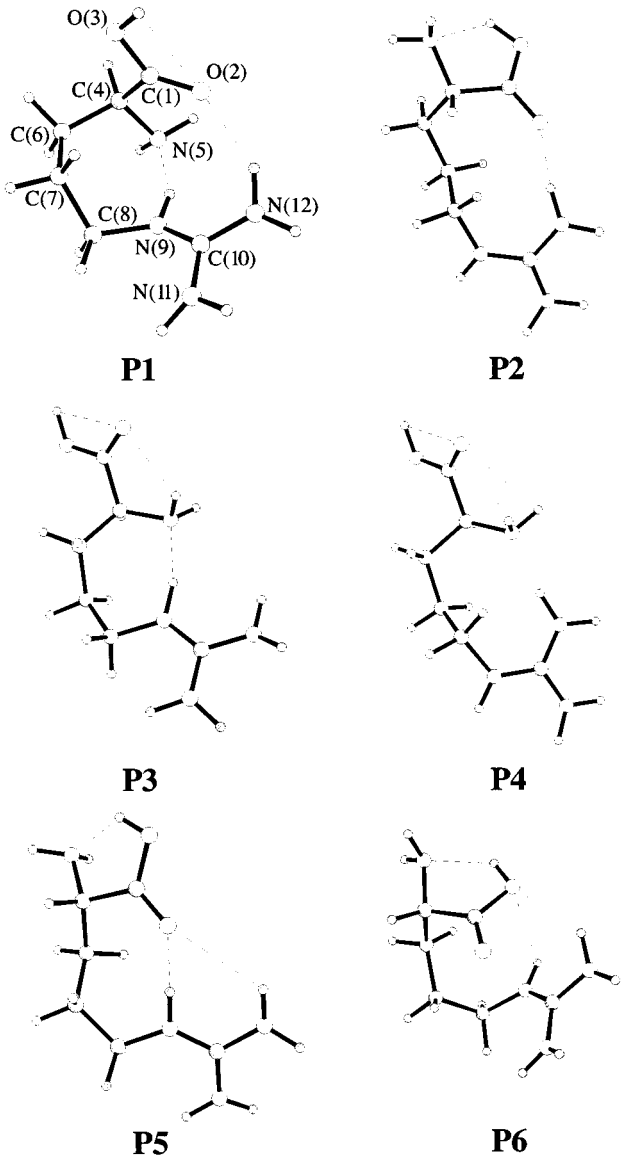

Figure 4. The equilibrium MP2 geometries for various conformers of protonated arginine.

The relative energies of different tautomers and conformers for the neutral and protonated arginine are determined by numerous factors. First, there are differences in formation energies of typical bonds that are broken and formed in the course of tautomerization. For instance, a typical OH bond is considered stronger than a typical NH bond by ca. 17 kcal/mol,[43] and this is the main reason why zwitterions of amino acids are unstable in the gas phase. Second, different sets of intramolecular hydrogen bonds develop for different tautomers and conformers of arginine. These intramolecular hydrogen bonds are formed at the expense of internal strain in the molecular framework. In addition, zwitterionic structures are stabilized by the Coulomb interaction between the negative and positive pole of the molecular dipole. The relative energies of

(40) Wong, M. W. *J. Phys. Chem.* **1996**, *100*, 391.
(41) Frisch, M. J.; et al. Gaussian, Inc.: Pittsburgh, PA, 1998.
(42) MOLPRO is a package of ab initio programs written by H.-J. Werner and P. J. Knowles.
(43) Pauling, L. *The Nature of the Chemical Bond*, 3rd ed.; Cornell University Press: Ithaca, NY, 1960; pp 85 and 453.

Table 1. The Relative Electronic Energies (E)[a] of Canonical and Zwitterionic Forms of Neutral Arginine Calculated at the B3LYP and MP2 Levels Followed by the Relative CCSD Electronic Energies Corrected for the Zero-Point Energy ($\Delta E_{0,vib}$), Enthalpy ($\Delta H_{298,corr}$), and Free Energy ($\Delta G_{298,corr}$) Contributions[b]

structure	E^{B3LYP}	E^{MP2}	E^{CCSD}	$E^{CCSD}+\Delta E_{0,vib}$	$E^{CCSD}+\Delta H_{298,corr}$	$E^{CCSD}+\Delta G_{298,corr}$
C5	0.000	0.000	0.000	0.000	0.000	0.000
C3	0.418	0.240	0.825	0.241	0.436	−0.528
C4	1.953	0.823	1.618	1.318	1.128	1.962
C2	1.830	1.820	1.695	1.999	1.984	1.480
C1	2.528	3.097	2.283	1.994	2.269	0.446
Z3	1.824	1.676	3.966	3.244	3.201	3.835
Z2	4.270	2.579	4.113	3.845	3.810	4.003
Z1	4.076	3.725	5.230	4.653	4.768	4.397

[a] The electronic energies of **C5** calculated at the B3LYP, MP2, and CCSD levels are −606.5998049, −604.8474817, and −604.9239685 au, respectively. All calculations were performed with the 6-31++G** basis set. [b] All quantities are given in kcal/mol.

Table 2. The Relative Electronic Energies (E)[a] of Protonated Forms of Arginine Calculated at the B3LYP and MP2 Levels Followed by the Relative CCSD Electronic Energies Corrected for the Zero-Point Energy ($\Delta E_{0,vib}$), Enthalpy ($\Delta H_{298,corr}$), and Free Energy ($\Delta G_{298,corr}$) Contributions[a]

structure	E^{B3LYP}	E^{MP2}	E^{CCSD}	$E^{CCSD}+\Delta E_{0,vib}$	$E^{CCSD}+\Delta H_{298,corr}$	$E^{CCSD}+\Delta G_{298,corr}$
P1	0.000	0.000	0.000	0.000	0.000	0.000
P2	2.728	5.559	4.913	4.751	4.948	4.233
P3	3.213	5.582	5.269	4.504	5.001	2.991
P4	4.735	6.408	5.779	5.484	5.790	4.525
P5	3.027	6.890	6.024	5.328	5.794	4.005
P6	9.609	8.189	7.976	7.903	8.037	7.548

[a] All quantities are given in kcal/mol. [a] The electronic energies of **P1** calculated at the B3LYP, MP2, and CCSD levels are −607.0153952, −605.2620345, and −605.3426887 au, respectively. All calculations were performed with the 6-31++G** basis set.

the canonical and zwitterionic structures of neutral arginine are collected in Table 1 and those for protonated forms in Table 2.

The ordering of neutral and protonated structures, according to their total electronic energies, varies to some extent among the DFT, MP2, and CCSD methods; see Tables 1 and 2. The two lowest structures of the canonical neutral, the lowest zwitterion, and the two lowest structures of protonated arginine are consistently **C5**, **C3**, **Z3**, **P1**, and **P2**, respectively, irrespective of the method used. Because the CCSD method is the most accurate method we employed, we will restrict further discussion to the electronic energies obtained at this level of theory.

The equilibrium geometries of different conformers, within a given tautomeric form of the neutral or protonated arginine, are determined by the interplay between intramolecular hydrogen bonds and internal strain. The hydrogen bonds develop between the proton donating OH and NH groups and the proton accepting atoms (O or N). The strength of a hydrogen bond is determined by (i) charge distribution in the proton donor (YH) and acceptor (X) fragments, (ii) the distance between H and X, and (iii) the X···HY angle. A preference to nearly linear hydrogen bonds is well established as was demonstrated for a variety of systems.[43,44] The dominance of "open" hydrogen-bonded dimers, such as $(H_2O)_2$, $(HF)_2$, or $(HCN)_2$, is a manifestation of this preference as the analogous "cyclic" structures would have an extra hydrogen bond. The hydrogen bonds in the latter structures would be, however, strongly distorted from linearity. On the other hand, intramolecular hydrogen bonds are frequently nonlinear due to steric restrictions.[45] The geometrical features

Table 3. Selected Geometrical Characteristics of Hydrogen Bonds in Neutral and Protonated Tautomers/Conformers of Arginine Optimized at the MP2 Level

structure[a]	hydrogen bond type	X···H[b] distance (Å)	X···HY[c] valence angle (deg)
C5	N(12)···HO(3)	1.775	163.2
	O(2)···HN(5)	2.509	97.9
C3	N(9)···HO(3)	1.685	177.1
	O(2)···HN(5)	2.642	90.5
	O(2)···HN(5)	2.866	78.3
C4	N(12)···HO(3)	1.731	150.1
	O(2)···HN(9)	2.077	144.1
	O(3)···HN(5)	2.451	99.3
C2	N(12)···HN(5)	2.144	172.3
	N(5)···HO(3)	1.917	124.2
C1	N(5)···HN(9)	1.966	154.0
	O(2)···HN(5)	2.602	92.5
	O(2)···HO(3)	2.312	75.5
	O(2)···HN(5)	2.913	75.5
Z3	O(3)···HN(12)	1.642	155.5
	O(2)···HN(9)	1.678	158.7
	O(3)···HN(5)	2.352	106.1
Z2	O(3)···HN(12)	1.702	149.6
	O(3)···HN(9)	2.143	123.6
	O(3)···HN(5)	2.393	105.7
Z1	O(2)···HN(12)	1.669	150.3
	O(2)···HN(9)	2.051	126.0
	O(3)···HN(5)	2.528	102.8
P1	N(5)···HN(9)	1.787	155.8
	O(2)···HN(12)	2.058	145.1
	O(2)···HO(3)	2.334	74.0
P2	O(2)···HN(12)	1.770	172.7
	N(5)···HO(3)	1.900	124.3
P3	N(5)···HN(9)	1.740	165.4
	O(2)···HN(5)	2.240	109.5
	O(2)···HO(3)	2.348	73.9
P4	N(5)···HN(12)	1.853	172.4
	O(2)···HN(5)	2.430	99.0
	O(2)···HO(3)	2.347	74.0
	O(2)···HN(5)	2.860	75.2
P5	O(2)···HN(9)	1.684	171.8
	N(5)···HO(3)	1.886	123.8
	O(2)···HN(12)	2.785	128.9
P6	N(5)···HO(3)	1.909	124.0
	O(3)···HN(12)	2.187	136.0
	O(3)···HN(9)	2.940	94.3

[a] For atomic labels see Figures 3 and 4. [b] X denotes proton acceptor (N or O). [c] X, Y denote electronegative atoms involved in the hydrogen bond.

of intramolecular hydrogen bonds that we identified for low-energy structures of neutral and protonated arginine are summarized in Table 3. The differences between the MP2 and B3LYP geometrical predictions do not exceed 0.02 Å and 1.5° for bond lengths and valence angles, respectively. It is well established that the MP2 method performs extremely well for hydrogen-bonded systems.[44] Thus, we will restrict further discussion to geometrical parameters determined at the MP2 level. The complete geometries for all structures are provided in the Supporting Information.

III.1. Structures of Canonical Arginine. The geometrical features of the five canonical structures **C1**−**C5** are displayed in Figure 3. Species **C1**, **C2**, and **C3** were characterized in our preliminary study[20] and **C1** is the same as the lowest energy canonical structure described in ref 19. Of particular interest are new species **C4** and **C5**. **C5** was identified by using the SGA algorithm whereas **C4** is closely related to the lowest energy zwitterionic structure **Z3** described below. These **C** structures will be discussed below in order of decreasing stability

(44) Scheiner, S. *Hydrogen Bonding. A Theoretical Perspective*; Oxford University Press: New York, 1997.

(45) Koll, A.; Wolschann, P. *Hydrogen Bonding Research;* Schuster, P., Mikenda, W.; Eds. Springer-Verlag: Wien, Germany, 1999.

as determined by the total CCSD/6-31++G** energies calculated at the optimal MP2 geometries.

C5 is the lowest energy structure and it is characterized by two hydrogen bonds. A stronger bond is formed between the OH group and N(12) where the H···N distance equals 1.775 Å; see Table 3. This is not surprising given that the OH hydrogen is the most acidic and N(12) the most basic site in the system. The N(12)···HO(3) bond is only slightly bent with an interatomic angle of 163°. Another hydrogen bond is formed by the carbonyl oxygen and one of the α-amino hydrogens. However, the OH distance is quite long at 2.5 Å and the O(2)···HN(5) angle is only 98°. Hence, this hydrogen bond is much weaker than the former.

C3 is the second lowest energy structure and it is less stable than **C5** by only 0.8 kcal/mol. We consider it as a canonical structure even though the guanidine group is protonated. The proton, however, is transferred not from the carboxylic acid OH group but from the N(9)H group, see Figures 1 and 3. Therefore N(9) can serve as a strong proton acceptor. The importance of such structures was also recognized in refs 9 and 19. Unfortunately, the authors of these papers focused on conformers with less favorable sets of hydrogen bonds than presented in Figure 3. In our **C3** structure there is a hydrogen bond between a strong proton acceptor N(9) and the OH group. The N(9)···H distance is only 1.685 Å and the N(9)···HO(3) angle is 177°; hence this H bond is apparently very strong. Two weaker hydrogen bonds, which also stabilize the **C3** structure, are formed by two α-amino group hydrogens and the carbonyl oxygen O(2), see Figure 3.

C4 is the third lowest energy structure and it is less stable than **C5** and **C3** by 1.6 and 0.8 kcal/mol, respectively. This structure is important because it is related to the most stable zwitterionic structure **Z3** by a single proton transfer from OH to N(12) of the guanidine group without further serious conformational adjustments. **C4** is stabilized by three hydrogen bonds. The strongest bond is between the most basic site (N(12)) and the most acidic site (OH). The N(12)···H distance is only 1.731 Å and the N(12)···HO bond angle is 150°. Hence this H-bond may be weaker than the strongest hydrogen bonds in **C5** and **C3**. In addition to the N(12)···HO bond, the carbonyl oxygen is involved in a hydrogen bond with HN(9), but neither the O(2)···H distance nor the O(2)···HN(9) angle indicates an unusual stabilization. The weakest hydrogen bond is formed between an α-amino hydrogen and the oxygen of the hydroxyl group, see Figure 3.

For both the **C2** and **C1** structures, the most acidic site (OH) and the most basic site (N(12)) are not connected by a hydrogen bond. **C2** differs from **C1** primarily in the orientation of the carboxylic acid and α-amino groups. In both structures the α-amino group is involved in a hydrogen bond with the guanidine group, but it acts as a proton acceptor in **C1** and as a proton donor in **C2**. **C2** is characterized by two hydrogen bonds: a nearly linear N(12)···HN(5) bond with an N(12)···H distance of 2.144 Å and a shorter but significantly bent bond N(5)···HO(3). **C2** is less stable than the lowest energy structure **C5** by 1.7 kcal/mol. The OH group of the **C1** structure is not involved in any hydrogen bond with nitrogens but there are five other weak hydrogen bonds that contribute to the stability of **C1**, with the strongest being the N(5)···HN(9) bond. **C1** is less stable than **C5** by 2.3 kcal/mol.

III.2. Structures of Zwitterionic Arginine. The zwitterionic structures **Z1**−**Z3** are also displayed in Figure 3. **Z1** was kindly provided to us by the Williams group.[9] The **Z** structures can be obtained from the canonical forms by a single proton transfer from the carboxyl group to the most basic nitrogen of the guanidine group. These low-energy structures are characterized by a small separation between the negative charge localized at the COO^- group and the positive charge localized on the guanidine residue. The **Z** structures will be discussed below in the order of decreasing stability as determined by the total CCSD/6-31++G** energies calculated at the optimal MP2 geometries.

The most stable structure **Z3** is characterized by two strong hydrogen bonds: O(2)···HN(9) and O(3)···HN(12). Both bonds have similar length and are extremely short, 1.678 and 1.642 Å, respectively (see Table 3). **Z3** is the lowest energy zwitterionic structure for any level of theory considered here. It is, however, less stable than the five canonical structures described above. In particular, it is less stable than **C5** by 4.0 kcal/mol.

Both **Z2** and **Z1** have two hydrogen bonds in which N(12)H and N(9)H interact with one oxygen atom (O(2) and O(3) for **Z1** and **Z2**, respectively). The shorter bond, of about 1.7 Å, involves the N(12) atom and the longer bond, of about 2.1 Å, involves the N(9) atom. In all of these zwitterionic structures one can also recognize an additional hydrogen bond between the α-amino group and the oxygen atoms. The **Z2** and **Z1** structures are less stable than the lowest energy canonical structure **C5** by 4.1 and 5.2 kcal/mol, respectively.

Any prediction of a thermodynamic property of neutral arginine can be compared to experimental data only if the thermodynamic equilibrium is attained in the time scale much shorter than that required to carry out an experiment. Indeed, the energy barriers for proton transfer between canonical and zwitterionic forms of amino acids can be very small and disappear on the free energy surface. We have calculated the activation barrier for a proton-transfer reaction **C4** ↔ **Z3**. **C4** is related to **Z3** through a single proton transfer (see Figure 3). The activation energy obtained at the B3LYP level for the **C4** → **Z3** process amounts to 1.6 kcal/mol but disappears on the free energy surface at $T = 298$ K. However, the B3LYP method is known to underestimate reaction barriers, especially for proton-transfer processes.[46] Therefore, the barrier was reevaluated at the MP2 level and it was found to be 2.6 and 0.7 kcal/mol on the energy and free energy surface, respectively. Thus, the small energy barrier ensures practically immediate attainment of the equilibrium state in gaseous arginine at room temperature.

III.3. Structures of Protonated Arginine. The low-energy conformers, labeled **P1**−**P6** in Figure 4, are ordered according to their decreasing stability as determined by the total CCSD/6-31++G** energies calculated at the optimal MP2 geometries. The lowest energy structure **P1** was identified by using the SGA algorithm whereas structures **P2**−**P6** were obtained by proton attachment to the neutral **C** and **Z** structures discussed in Sections III.1 and III.2, followed by geometry optimization.

In the latter approach, the protonated site was N(12) for **C1**, **C2**, **C4**, and **C5** and N(9) for the **C3** structure, and either O(2) or O(3) for the **Z1**−**Z3** structures. If the proton attachment did not destroy the network of preexisting hydrogen bonds in the neutral structure then the resulting structure was directly optimized at the B3LYP and MP2 levels (**C1** → **P3**, see Figures 3 and 4). Otherwise, we rotated the proton donor and/or proton acceptor groups to restore a favorable network of hydrogen bonds. These protonated and manually refined structures were further optimized at the B3LYP and MP2 levels. As an illustration, due to proton attachment to N(12) in **C2**, the preexisting hydrogen bond, N(12)···HN(5), is destroyed. Therefore, the α-amino group needs to be rotated around N(5)−C(4)

(46) Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2001**, *105*, 2936−2941.

to become a proton acceptor and to form a new hydrogen bond N(5)···HN(12). Simultaneously, the carboxyl function needs to be rotated around C(1)−C(4) to involve O(2) in hydrogen bonds with the α-amino group. Finally, the carboxylic hydrogen needs to be rotated around C(1)−O(3) to form an intracarboxyl hydrogen bond. This is how the initial guess of the **P4** conformer was created from the **C2** canonical structure. The geometry optimizations of the protonated **C3**, **C4**, **Z3**, and **Z1** structures converged to the same conformer **P6**, whereas optimization of protonated **C5**, **C2**, **C1**, and **Z2** led to **P2**, **P4**, **P3**, and **P5**, respectively.

Like neutral structures, the protonated species are stabilized by networks of hydrogen bonds. In the lowest energy structure **P1** two relatively strong H-bonds can be identified (see Figure 4): N(5)···HN(9) and O(2)···HN(12). Both hydrogen bonds are relatively short (1.787 and 2.058 Å) and the XH···Y angles are less than 35° from linearity (see Table 3). For other and less stable structures only one strong hydrogen bond can be identified.

The structures **P2** and **P5** are related in the sense that they have a very similar hydrogen bond between the carboxyl and the α-amino group, with OH acting as a proton donor. These two structures differ, however, in the second and stronger hydrogen bond. This is O(2)···HN(12) and O(2)···HN(9) for **P2** and **P5**, respectively. The **P2** and **P5** structures are less stable than the **P1** structure by 4.9 and 6.0 kcal/mol, respectively. This suggests that the O(2)···HN(12) bond is stronger than the O(2)···HN(9) bond by ca. 1 kcal/mol.

The structures **P3** and **P4** are related in the sense that they have a similar hydrogen bond between the carboxyl and the α-amino group, with NH acting as a proton donor. These two structures differ, however, in the second and stronger hydrogen bond. This is N(5)···HN(9) and N(5)···HN(12) for **P3** and **P4**, respectively. The **P3** and **P4** structures are less stable than the **P1** structure by 5.3 and 5.8 kcal/mol, respectively.

The conformer **P6** is stabilized by two hydrogen bonds: O(3)···HN(12) and N(5)···HO(3). None of them has a favorable hydrogen bond angle. Moreover, O(3) is involved in the first hydrogen bond as a proton acceptor and in the second as a proton donor. As a result both hydrogen bonds may be relatively weak, which explains why **P6** is less stable than **P1** by 8.0 kcal/mol.

The **P3** structure, identified earlier by MK,[19] has a network of intramolecular hydrogen bonds as favorable as the **P1** structure identified here; see Table 3. However, the **P1** structure is more stable than **P3** by 5.3 kcal/mol. MK concluded that the distinctively high proton affinity of arginine is related to a strong network of intramolecular hydrogen bonds. Our findings indicate that a small separation between the positively charged guanidine group and the carboxylic group may provide an additional stabilization. Indeed, this separation is much smaller for **P1** than **P3** as the O(2)−N(12) distances are 2.95 and 6.29 Å, respectively. The more compact **P1** structure is apparently not strongly destabilized by internal strain.

III.4. Dipole Moments of Neutral Tautomers and Vertical Electron Attachment Energy. The dipole moments of **C1**−**C5** and **Z1**−**Z3**, collected in Table 4, display interesting features. The dipole moment of the lowest energy zwitterion **Z3** is higher than that of the lowest energy neutral structures **C5** and **C3**. However, the dipole moment of **C2** is comparable to that of **Z3** for all computational approaches and larger than that for **Z1** and **Z2**. Thus our original assumption that the zwitterion form of the amino acid would possess a larger dipole moment than any nonzwitterion form[23] is not valid. Arginine may be unique due to the presence of a guanidine group, but other amino acids should be thoroughly examined. The fact that charge-separated structures (zwitterions **Z1**−**Z3**) have comparable or smaller dipole moments than that found for the canonical isomer **C2** is surprising but can be explained when one notices that a network of hydrogen bonds makes all zwitterionic structures very compact. In particular the COO^- group is directly involved in hydrogen bonds with the protonated guanidine group; see Figure 3. Hence, the separation between the positive and negative poles of the zwitterionic dipole becomes relatively small. For instance, the distances between the oxygen atoms and the N(11) atom are as small as 4.4−4.8 Å. In contrast, the H-bonds in structure **C2** allow the C- and N-termini to be well separated. In particular the carbonyl group of **C2** is not involved in any hydrogen bond and it is separated from the guanidine group by the α-amino group. The distance O(2)−N(11) is as large as 8.5 Å for **C2**, which maximizes the net value of the dipole moment.

Table 4. Dipole Moments (in Debye) for the Canonical and Zwitterionic Tautomers of Arginine Followed by Vertical Electron Attachment Energies (in kcal/mol) Determined at Koopmans' Theorem Level (D^{KT})

tautomer	μ^{SCF}	μ^{MP2}	μ^{B3LYP}	D^{KT}
C5	8.343	7.646	8.214	1.086
C3	8.216	7.477	7.959	1.111
C4	5.085	4.752	5.050	0.239
C2	9.678	8.854	9.203	1.255
C1	3.774	3.648	3.903	0.013
Z3	9.432	9.209	9.054	3.024
Z2	6.664	6.673	7.714	0.659
Z1	7.040	7.022	8.471	0.885

The dipole moments of canonical structures, calculated at the B3LYP level, span the range from 3.9 to 9.2 D. These differences are also related to different H-bond networks, which cause different relative orientation of the positive and negative poles of molecular dipoles. The separation between the carbonyl oxygen O(2) and the $N(11)H_2$ group roughly correlates with the magnitude of the dipole moment, but the orientation of the hydroxyl group also plays an important role. The lowest energy canonical structures **C5** and **C3** provide significant dipole moments of 8.2 and 8.0 D, respectively, which are smaller by only ∼1 D than the dipole moment of **Z3**. Hence, the stabilization of zwitterionic arginine by an excess electron cannot be taken for granted just on the basis of calculated dipole moments.

The reported values of dipole moments are larger than the critical value of ca. 2.5 D required to bind an excess electron.[47] The vertical electron attachment energies, determined at the Koopmans' theorem level and reported in Table 4, correlate only approximately with the values of the SCF dipole moments and span a range from 3.02 kcal/mol for **Z3** to 0.01 kcal/mol for the lowest dipole structure **C1**. **Z3**, with a SCF dipole moment of 9.05 D, binds the excess electron 2.4 times stronger than **C2** with a SCF dipole moment 0.15 D larger than that of **Z3**. Hence, not only the dipole moment but also other structural effects play key roles in determining the electron binding energies. The excess electron binding energy for the lowest energy canonical structures **C5** and **C3** is also significant and amounts to 1.1 kcal/mol. Determination of geometrical relaxation of the molecular frameworks upon electron attachment

(47) Gutowski, M.; Skurski, P. *Recent Res. Dev. Phys. Chem.* **1999**, *3*, 245−260.

(48) Hashiguchi, K.; Hamada, Y.; Tsuboi, M.; Koga, Y.; Kondo, S. *J. Mol. Spectrosc.* **1984**, *105*, 81−92.

and inclusion of orbital relaxation and electron correlation effects is the subject of an ongoing study.[25]

III.5. Thermodynamic Properties. The differences in the CCSD energies are very small for the different structures of neutral arginine, see Table 1. The five lowest canonical structures span an energy range of only 2.3 kcal/mol. The lowest energy zwitterion **Z3** is separated from the lowest energy canonical structure **C5** by 4.0 kcal/mol. Hence, zero-point vibrational energy contributions as well as thermal effects may contribute significantly to the population of a particular structure in the gas-phase equilibrium. The experimental values of proton affinity and gas basicity are usually provided at 298 K[22] and the IR-CRLAS spectrum of arginine was measured with the source at 443 K.[10] Therefore, we determined the composition of gas-phase arginine at 298 and 443 K and the composition of protonated arginine at 298 K.

The zero-point vibrational corrections are in general smaller for higher energy structures leading to their stabilization with respect to the lowest energy structure, as shown in Tables 1 and 2. This indicates that higher energy structures are usually less "stiff" than the lowest energy structure. The unusual stiffness of the lowest energy structures, for both protonated and neutral species, may result from the fact that they have very favorable sets of intramolecular hydrogen bonds.

The stability of the species in enthalpy scale (column $E^{CCSD} + \Delta H_{298,corr}$ in Tables 1 and 2) takes into account thermal contribution to the energy from rotations and vibrations, in addition to the electronic energy and zero-point vibrational corrections. This thermal contribution is dominated by its vibrational terms. In addition, only low-frequency modes contribute significantly to the thermal energy contribution at 298 and even 443 K. If the low-frequency modes are softer in structure *M* than in the reference structure *R*, then the thermal energy contribution destabilizes *M* relative to *R*. This is indeed the pattern that we observe in Tables 1 and 2. We note that the structures **C1**, **C3**, and **Z1** have particularly soft low-frequency modes and that structures **P2**−**P6** have softer low-frequency modes than the reference structure **P1**. This is especially evident for the forms **P3** and **P5**.

The stability of the species in Gibbs free energy scale (column $E^{CCSD} + \Delta G_{298,corr}$ in Tables 1 and 2) takes into account entropic contribution in addition to $E^{CCSD} + \Delta H_{298,corr}$. The entropic contribution is dominated by the vibrational term whereas rotational contribution is usually much smaller. Again, only low-frequency modes contribute significantly to the entropic contribution at 298 and even 443 K. Not surprisingly, one observes the largest increase of stability for those structures that were most profoundly destabilized by the vibrational thermal contribution to energy. In addition, the entropic stabilization is more important than the thermal vibrational energy destabilization.

The outcome of these small and frequently mutually competing effects is amazing. The **C3** structure is predicted to be the most stable on the Gibbs free energy scale. At 298 K, it is more stable than the closest canonical structure **C5** by 0.5 kcal/mol. The lowest zwitterion structure **Z3** is destabilized by the entropic contribution and it is 4.4 kcal/mol less stable than the **C3** structure. The **C1** structure, which was the least stable in electronic energy scale among canonical structures considered by us, becomes the third most stable structure of arginine. The three lowest canonical structures span a free energy range of 1.0 kcal/mol, which is even smaller than the range of electronic energies of 1.6 kcal/mol. The three zwitterionic structures, which span 1.26 kcal/mol in electronic energy scale, are now clustered within 0.56 kcal/mol.

Table 5. The Percent Shares of Particular Forms in Equilibrium Mixture Calculated for Neutral and Protonated Arginine, See Eqs 1−3[a]

neutral forms	percent share		protonated forms	percent share ($T = 298$)
	$T = 298$	$T = 443$		
C5	24.80	16.04	P1	99.12
C3	60.50	51.87	P2	0.08
C4	0.90	1.07	P3	0.64
C2	2.04	3.90	P4	0.05
C1	11.68	26.70	P5	0.11
Z3	0.04	0.14	P6	0.00
Z2	0.03	0.15		
Z1	0.01	0.13		

[a] The temperatures in K.

The **C5** and **P1** structures were selected as a reference structure for the neutral and protonated arginine, respectively, see eqs 1−3. The differences in Gibbs free energies, which are used to determine the values of K_M's, are provided in the last column of Tables 1 and 2 for the neutral and protonated species, respectively. The percent share of different structures in the gas-phase sample of neutral arginine at 298 and 443 K is displayed in Table 5. The most dominant at 298 K are **C3** (61%), **C5** (25%), and **C1** (12%), with **C4** and **C2** contributing less than 3% each. At 443 K, which is the temperature of the IR-CRLAS experiment,[10] these percent shares amount to 52%, 16%, and 27% for **C3**, **C5**, and **C1**, respectively, and the total contribution from **C4** and **C2** does not exceed 5%. The total percent share from zwitterionic structures does not exceed 0.45% even for the higher temperature of 443 K. The impressive stabilization of the **C1** structure is related to its open structure that leads to very soft bending and torsional modes delocalized over the whole molecular framework.

Thermal effects also seriously affect the protonated structures. **P1** remains the most favorable structure in Gibbs free energy scale but **P3** becomes the second lowest and is separated by only 3.0 kcal/mol from **P1**, see Table 2. This is still sufficient for **P1** to dominate in the gas-phase sample of protonated arginine at 298 K (>99%), see Table 5.

The values of proton affinity PA and gas-phase basicity GB, derived from eqs 4−7, are 256.3 and 247.8 kcal/mol, respectively, whereas the corresponding experimental findings are 251.2 and 240.6 kcal/mol.[22] The 5−7 kcal/mol overestimation in calculated values of PA and GB relative to the experimental data may be due to many factors. First, there are deficiencies in our model such as incomplete one-electron basis set or neglect of electron correlation energy beyond the CCSD level Also, the harmonic approximation to molecular vibrations is likely a source of error. Finally, we might still have missed a conformer of neutral arginine that is lower in energy than **C5** or **C3** by ca. 5 kcal/mol, although this possibility is quite improbable and would have to be related to a failure of the PM3 method used by us in the SGA search.

The 5−7 kcal/mol overestimation of PA and GB prompted us to benchmark our computational method against experimental data for a simpler and more rigid molecule such as glycine. The calculated values of PA and GB for glycine are 214.3 and 206.5 kcal/mol, respectively, while the corresponding experimental data are 211.9 and 203.7 kcal/mol,[22] i.e., our computational approach leads to the 2−3 kcal/mol overestimations. For arginine, the remaining overestimation by 3−4 kcal/mol may be attributed to its flexibility. On the other hand, the observed disagreement between the calculated and measured PA and GB for arginine can be also ascribed to experimental problems such as clustering reactions of protonated molecules with polar neutral

Table 6. Selected IR Frequencies[a] (in cm^{-1}) for Canonical Forms of Neutral Arginine Calculated at the B3LYP/6-31++G** Level with Scaled Values Given after Semicolon and IR Intensities in Brackets (in Km/mol) [Scaling Factor = 0.9813]

character	C5	C3	C4	C2	C1
deformation of imino amino, hydroxyl, and methylene groups, and stretching of C(10)−N bonds	1341; 1316 (23) many	1433; 1406 (180) many	1308; 1284 (87) many	1384; 1358 (96) many	1348; 1323 (46) many
	1356; 1331 (94) many	1528; 1499 (252) δO(3)H	1317; 1292 (63) many	1416; 1390 (366) many	1369; 1343 (68) many
	1385; 1359 (418) many	1644; 1613 (44) $\delta N(11)H_2$; $\delta N(12)H_2$	1355; 1330 (33) many	1559; 1530 (228) many	1394; 1368 (22) many
	1402; 1376 (57) many	1651; 1620 (172) $\delta N(11)H_2$; $\delta N(12)H_2$	1501; 1473 (142) many	1641; 1610 (96) $\delta N(11)H_2$	1424; 1397 (75) many
	1574; 1545 (218) many	1679; 1645 (25) $\delta N(5)H_2$	1510; 1482 (63) many	1695; 1663 (248) υC(10)═N(12); $\delta N(11)H_2$; $\delta N(5)H_2$	1589; 1559 (288) many
	1638; 1607 (148) $\delta N(11)H_2$; υN(12)C(10); υN(9)C(10)	1730; 1698 (750) υC(10)═N(9); $\delta N(11)H_2$ $\delta N(12)_2$	1634; 1603 (184) $\delta N(11)H_2$; νN(9)H; υN(12)C(10)	1715; 1683 (96) $\delta N(5)H_2$	1643; 1612 (104) $\delta N(11)H_2$
	1678; 1647 (25) $\delta N(5)H_2$		1671; 1640 (33) $\delta N(5)H_2$		1672; 1641 (32) $\delta N(5)H_2$
	1692; 1660 (337) υC(10)═N(12); $\delta N(11)H_2$; δN(9)H		1685; 1653 (330) υC(10)═N(12); $\delta N(11)H_2$; δN(9)H		1706; 1674 (296) υC(10)═N(12); $\delta N(11)H_2$; δN(9)H
stretching of C═O and deformation of OH in the carboxyl group	1802; 1768 (305)	1815; 1781 (345)	1749; 1716 (287)	1836; 1803 (372)	1810; 1776 (311)

[a] Localized deformation (or bending) modes are denoted with δ and stretching modes with υ.

molecules at low temperatures, pyrolysis, and isomerization reactions of molecules and ions at high temperatures, and difficulties in attaining thermodynamic equilibrium. In fact, the propensity of arginine to form unusually stable charged aggregates is well established.[16] The discrepancy between the calculated and measured proton affinity of arginine remains a challenging problem to be addressed in future studies.

III.6. Characteristic Vibrational Frequencies for Neutral and Protonated Arginine. The presence of a carboxylic acid group and a carboxylate residue in the canonical and zwitterionic arginine, respectively, was used by Saykally et al. to suggest which tautomeric form dominates in the gas phase.[10] In their spectrum, recorded in the 1550−1750 cm^{-1} region, they observed two transitions at 1666 and 1693 cm^{-1} and assigned them to carbonyl stretches of the carboxylic acid groups. They thus concluded that canonical arginine dominates in the gas phase and the upper bound to the zwitterionic population is 1%. On the other hand, the results presented below suggest that the two experimentally observed transitions correspond to NHX (X = H or C) bending modes and C═N stretching modes rather than to carbonyl group stretching modes. In addition our data suggest that zwitterionic structures, if thermodynamically populated, would provide intense IR transitions in the same 1600−1700 cm^{-1} region. Therefore, we believe that the experimental results of ref 10 alone are not sufficient to resolve which tautomeric form of arginine dominates in the gas phase, although we concur with their fundamental conclusion that the canonical structure is likely the dominant one.

The characteristic frequencies related to deformations of the imino and amino groups of canonical and zwitterionic forms of arginine occur above 1300 cm^{-1}. The selected harmonic frequencies from the 1300−1900 cm^{-1} region (unscaled and scaled), together with their intensities, are displayed in Tables 6 and 7 for the canonical and zwitterionic forms of arginine, respectively. The frequencies and intensities of selected XH (X = O, N) stretching modes, which occur above 2500 cm^{-1}, are provided in the Supporting Information. All transitions with intensities exceeding 20 Km/mol were included in Tables 6 and 7 for the lower frequency modes (bending modes and C═O and C═N stretching modes), whereas for the higher energy and usually intense XH (X = O, N) stretching modes, we selected only these transitions for which the intensities exceed 50 Km/mol.

The correctness of using scaled frequencies is clearly a crucial point in our interpretation. For this reason we provided in Table 8 the calculated and experimental[39,47] frequencies for model compounds that contain a carboxylic group (CH_3COOH), a C═N bond (CH_3CH═NH), and a carbon-bonded NH_2 group (CH_3NH_2). As discussed in Section II, the calculated transitions below 1900 cm^{-1} were scaled by a factor of 0.9813, which provides an accurate transition for the carbonyl stretch in the carboxylic group of acetic acid. The conventional scaling factor of 0.9613[40] was used for higher energy transitions. The same scaling procedure was used in Tables 6 and 7. The data presented in Table 8 indicate that the agreements between the scaled and experimental frequencies are satisfactory and calculated IR intensities are in qualitative agreement with experimental findings. The largest discrepancy in the 1300−1900 cm^{-1} region is observed for the COH bending and the C═N stretching modes (ca. 46 cm^{-1}) and the scaled frequencies usually provide overestimates. In addition, the scaling factor 0.9813 performed very well for the carbonyl stretch in formamide. The scaled frequency was found to be 1763 cm^{-1}, whereas the experimental spectrum[49] has a strong transition at ca. 1756 cm^{-1}. To the best of our knowledge, experimental data are not available for the symmetric and asymmetric stretching modes of the COO^- group in the gas phase. For zwitterionic glycine in condensed phases, the experimental frequencies of symmetric and asymmetric COO^- stretches are 1410−1417 and 1597 cm^{-1}, respectively.[50]

The original IR-CRLAS spectrum from ref 10, which was recorded in the 1550−1750 cm^{-1} region, is displayed in Figure 5A, whereas dominant transitions that result from the canonical and zwitterionic tautomers are displayed in Figure 5, parts B and C, respectively. In Figure 5B,C, the intensities of vibrational transitions that result from a given tautomer are scaled by a

(49) NIST Chemistry WebBook (http://webbook.nist.gov/chemistry).
(50) Alper, J. S.; Dothe, H.; Lowe, M. A. *Chem. Phys.* **1992**, *161*, 199−209.

Table 7. Selected IR Frequencies[a] (in cm^{-1}) for Zwitterionic Forms of Neutral Arginine Calculated at the B3LYP/6-31++G** Level with Scaled Values Given after Semicolon and IR Intensities in Brackets (in Km/mol) [Scaling Factor = 0.9813]

character	Z3	Z2	Z1
deformation of amino, and hydroxyl, and methylene groups, and stretching of C(10)−N and C(1)═O(2) bonds	1382; 1356 (26)	1321; 1296 (30)	1318; 1293 (46)
	many	many	many
	1388; 1362 (62)	1374; 1348 (160)	1379; 1353 (102)
	many	υ_sCOO + many	υ_sCOO + many
	1431; 1404 (103)	1394; 1368 (28)	1388; 1362 (31)
	υ_sCOO + many	many	many
	1474; 1446 (30)	1405; 1379 (31)	1419; 1392 (43)
	many	many	many
	1492; 1464 (32)	1459; 1432 (52)	1464; 1437 (26)
	many	many	many
	1499; 1471 (45)	1501; 1473 (28)	1491; 1463 (23)
	many	many	many
	1511; 1483 (40)	1597; 1567 (71)	1512; 1484 (30)
	many	many	δC(8)H_2
	1578; 1548 (143)	1603; 1573 (133)	1597; 1567 (57)
	υ_{as}COO + many	υ_{as}COO + many	many
	1587; 1557 (47)	1668; 1637 (54)	1613; 1583 (171)
	many	δN(5)H_2+many	υ_{as}COO + many
	1662; 1631 (274)	1672; 1641 (97)	1670; 1639 (67)
	υC(10)N(9);	υC(10)N(9);	δN(5)H_2
	υC(10)N(11) + many	υC(10)N(11) + many	
	1672; 1641 (38)	1706; 1674 (353)	1674; 1643 (152)
	δN(5)H_2	υC(10)N(11);	υC(10)N(9);
		υC(10)N(12)+many	υC(10)N(11) + many
	1689; 1657 (165)	1718; 1686 (434)	1716; 1684 (344)
	υC(10)N(11) + many	υC(10)N(9);	υC(10)N(9);
		υC(10)N(12) + many	υC(10)N(11) + many
	1726; 1694 (488)		1722; 1690 (429)
	υC(10)N(9);		υC(10)N(9);
	υC(10)N(12) + many		υC(10)N(12) + many

[a] Localized deformation (or bending) modes are denoted with δ and stretching modes with υ. The scaling factor equals 0.9813.

Table 8. The Comparison between Calculated[a] and Measured[39,48] IR Frequencies (cm^{-1}) for Selected Model Systems

system	mode[c]	calculated	scaled	measured	Δ
CH_3COOH	υC−C	862(3)	846	847(W)	−1
	υC−O	1205(210)	1182	1182(S)	0
	δOH	1335(51)	1310	1264(M)	46
	υC═O	1822(357)	1788	1788(VS)	0
	υO−H	3754(60)	3609	3583(M)	26
CH_3CH═NH	υC═N	1728(85)	1696	1651(S)	45
CH_3NH_2	$\delta_{wag}NH_2$	822(191)	807	780(VS)	27
	$\delta_{sci}NH_2$	1670(32)	1639	1623(S)	16
	υ_sNH_2	3513(0.9)	3377	3361(W)	16

[a] B3LYP/6-31++G** method. We have applied a scaling factor of 0.9813 for frequencies below 2000 cm^{-1} and 0.9613 for higher frequencies. Calculated IR intensities (in Km/mol) are given in brackets in the third column. The intensities of experimental transitions are characterized by letters.[b] Δ stands for a difference between the scaled and measured frequency. [b] W, weak; M, medium; S, strong; VS, very strong. [c] Localized deformation (or bending) modes are denoted with δ and stretching modes with υ.

factor proportional to the population of this tautomer in the gas phase at 443 K (see Table 5). Moreover, normalization constants were selected to match the intensity of the dominant experimental peak. These normalization constants are different in Figure 5B,C due to different populations of the canonical and zwitterionic structures.

The first important observation from Table 6 and Figure 5B is that the carbonyl stretch frequencies from the carboxylic group are not in the 1625−1725 cm^{-1} region, as assumed in Figure 5A from ref 10, but rather in the 1750−1800 cm^{-1} region that was not scanned experimentally. We therefore believe that the assignment of the 1666 and 1693 cm^{-1} peaks to the carbonyl stretch (see Figure 5A) is premature.

Second, both the canonical and zwitterionic tautomers provide strong transitions in the 1650−1700 cm^{-1} region, where the two experimental peaks are observed. The stronger and narrower experimental peak at 1693 cm^{-1} might be assigned either to the **C3** tautomer (the scaled frequency of 1698 cm^{-1} for the mode dominated by C(10)═N(9) stretch and N(11)H_2 and N(12)H_2 bendings) *or* to zwitterionic structures that possess C−N stretching modes in this region. Similarly, the broader peak at 1666 cm^{-1} may be due to either the **C1** and **C5** tautomers (the scaled frequency of 1674 and 1660 cm^{-1} for **C1** and **C5**, respectively, for the modes dominated by the C(10)═N(12) stretch and N(11)H_2 and N(9)H bendings) *or* zwitterionic structures that provide C−N stretching modes in this frequency region. Thus, the assignment of the 1666 and 1693 cm^{-1} transitions to the carbonyl stretch is again premature. On the other hand, our ab initio results indicate that canonical arginine dominates in the gas phase (see Table 5). Therefore, we assign the experimental peaks at 1693 and 1666 cm^{-1} to the **C3** conformer and the **C1** and **C5** conformers, respectively. We expect that relative intensities for these two peaks will be temperature dependent because the populations of these tautomers will be.

Third, the experimental peak at 1666 cm^{-1} is quite broad, asymmetric, and extends to ca. 1630 cm^{-1}. Clearly, more than one transition may be contributing to this peak. Perusal of Table 6 indicates that every important canonical tautomer has a low intensity (ca. 30 Km/mol) transition in the 1641−1647 cm^{-1} region that may contribute to the low-energy shoulder of the 1666 cm^{-1} peak. These transitions are assigned to the N(5)H_2 bending mode.

Fourth, none of the important canonical tautomers has two strong transitions in the 1650−1700 cm^{-1} region, where the two experimental peaks were observed. This suggests that the equilibrated gas-phase sample of arginine contains at least two canonical conformers, which is consistent with our prediction that **C3**, **C1**, and **C5** coexist in equilibrium conditions (see Table

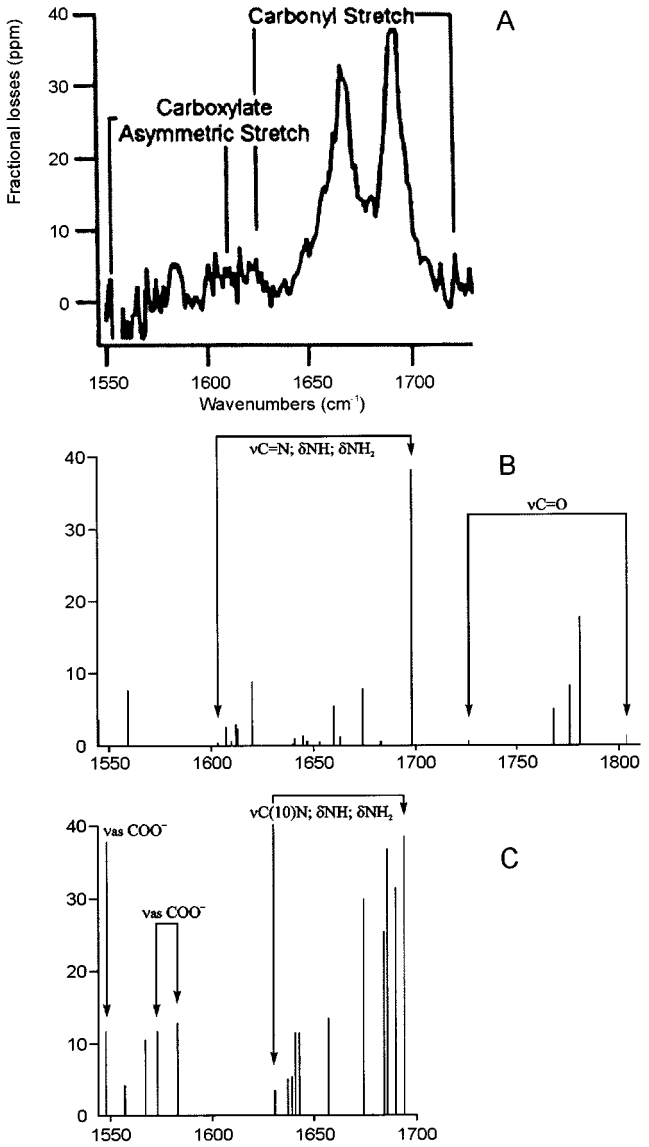

Figure 5. The comparison between the measured IR-CRLAS spectrum of arginine (from ref 10) (A) and the calculated IR spectra for its canonical (B) and zwitterionic (C) tautomers.

5). This prediction could be verified in subsequent temperature-dependent IR measurements as the relative intensities of the two peaks should depend on the temperature.

Finally, the frequencies of asymmetric stretching modes of the zwitterionic COO^- groups are in the 1548−1573 cm^{-1} region. Hence, some of them might be observed in the IR-CRLAS spectrum if the zwitterionic structures were thermodynamically populated. We should mention, however, that these transitions would have smaller intensities than the transitions in the 1650−1700 cm^{-1} region, which are related primarily to the C(10)−N stretching modes. On the other hand, the calculated frequencies of symmetric stretching modes of the zwitterionic COO^- are in the 1348−1404 cm^{-1} region. Hence, they could not be observed in the IR-CRLAS spectrum.

Future IR spectra may provide a decisive answer whether canonical or zwitterionic tautomers of arginine dominate in the gas phase. These spectra, however, should be recorded in a broader range of frequencies than has been measured up to now. The 1750−1850 cm^{-1} region is particularly interesting because we anticipate only the C═O stretches of canonical forms to be active there; see Figure 5B. Moreover, isotope labeling experiments may be required to resolve the problem. Finally, the temperature dependence of relative intensities for the recorded peaks may shed more light on the population of energetically nearly degenerate structures.

The vibrational characteristics for neutral arginine above 2000 cm^{-1} as well as the bending and stretching modes of protonated arginine are reported in the Supporting Information to facilitate future IR studies. For the neutral species, the most intense transition for the tautomers **C3** and **C5** is predicted at 2722 and 2997 cm^{-1}, respectively. As expected, these transitions are related to the O(3)H stretch. For another populous tautomer, i.e., **C1**, a stronger transition at 3275 cm^{-1} is related to the N(9)H stretch and a weaker transition at 3601 cm^{-1} to the O(3)H stretch. The zwitterions would provide strong transitions at 2587 and 2876 (**Z3**), 2565 (**Z2**), and 3012 cm^{-1} (**Z1**). However, the populations of zwitterions are probably too small (see Table 5) to make these transitions detectable.

IV. Summary

The lowest energy structures of neutral and protonated arginine were searched for using a simple genetic algorithm and the semiempirical PM3 method as well as subsequent high-level theory. This search led to the two lowest energy structures identified thus far: **C5** for the neutral and **P1** for protonated arginine. These structures were overlooked in our previous studies in which more conventional geometry optimization techniques were employed. All structures considered in this study were further refined at the B3LYP and MP2 levels by using 6-31++G** basis sets. Final electronic energies were calculated at the coupled cluster level of theory with single and double excitations.

All five of the lowest energy structures of neutral arginine are canonical and span a range of 2.3 kcal/mol of total electronic energies. The lowest zwitterionic structure is higher in electronic energy by 4.0 kcal/mol than the lowest canonical structure. Thermal effects contribute substantially to the stability of different forms of the neutral and protonated arginine due to near degeneracy of their electronic energies and softness of several vibrational modes. Three canonical structures (**C3**, **C5**, and **C1**) contribute substantially to the population of neutral arginine at $T = 298$ K. The **C1** structure, which is the least stable in its electronic energy among the canonical structures considered by us, becomes the second most stable structure at $T = 443$ K, the temperature of the IR-CRLAS experiments. The zwitterionic structures remain unstable with respect to the lowest canonical structure by more than 4.3 kcal/mol in Gibbs free energy at $T = 298$ K. The calculated proton affinity of 256.3 kcal/mol and gas basicity of 247.8 kcal/mol of arginine are in reasonable agreement with the measured values of 251.2 and 240.6 kcal/mol, respectively. We believe that the discrepancies between the calculated and experimental data are larger than the deficiencies of our theoretical model.

The calculated vibrational characteristics for different tautomers/conformers of arginine, combined with information about the relative percents of the different structures at equilibrium, provided an unexpected interpretation of the IR-CRLAS spectrum. It was recognized that the experiment is likely inconclusive about the dominance of canonical vs zwitterionic forms of arginine in the gas phase because two features assigned to carbonyl stretches from carboxylic groups (i.e., canonical structures) are most probably related to NH_2 and NHC bending modes and C═N stretching modes. In addition, it was recognized that both canonical and zwitterionic tautomers provide strong transitions in the 1650−1700 cm^{-1} region, where the two experimental peaks were observed. We recommend record-

ing the vibrational spectrum of gas-phase arginine in the 1750−1850 cm^{-1} region to resolve whether canonical arginine dominates in the gas phase. We predict that relative intensities of peaks in the 1550−1750 cm^{-1} region will be temperature dependent.

Acknowledgment. The authors thank Prof. Evan R. Williams' group for providing their minimum energy structures of canonical and zwitterion arginine and Dr. Piotr Storoniak and MSc. Monika Sobczyk for valuable technical assistance. This work was supported by the Polish State Committee for Scientific Research (KBN) Grant No. DS./8371-4-0137-1 and the NSF Grants CHE9618904 and CHE9982420 to J.S. M.G. was supported by the US DOE, Office of Basic Energy Sciences, Chemical Sciences Division. J.R is grateful to the Alexander von Humboldt Foundation for the generous gift of a DEC Alpha workstation. The computer time allocations provided by the Academic Computer Center in Gdańsk (TASK) and Center for High Performance Computing at the University of Utah are gratefully acknowledged. This research also used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098.

Supporting Information Available: The MP2 optimized geometries of canonical, zwitterionic, and protonated structures of arginine. The frequencies and intensities of selected XH (X=O, N) stretching modes, which start above 2500 cm^{-1} for canonical and zwitterionic form of the neutral arginine. The frequencies and intensities of selected transitions above 1300 cm^{-1} for the lowest energy conformer of protonated arginine. This material is available free via the Internet at *http://pubs.acs.org*. See any current masthead page for ordering information and Web access instructions.

JA011357L