30 March 2001

ELSEVIER

Chemical Physics Letters 337 (2001) 143–150

CHEMICAL PHYSICS LETTERS

www.elsevier.nl/locate/cplett

Non-ionic and zwitterionic forms of neutral arginine – an ab initio study

Piotr Skurski [a,b], Maciej Gutowski [b,c], Robyn Barrios [a], Jack Simons [a,*]

[a] *Henry Eyring Center for Theoretical Chemistry, Department of Chemistry, University of Utah, Salt Lake City, UT 84112, USA*
[b] *Department of Chemistry, University of Gdansk, ul. Sobieskiego 18, 80-952 Gdansk, Poland*
[c] *Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory Richland, WA 99352, USA*

Received 14 November 2000; in final form 23 January 2001

Abstract

Six low-energy structures of arginine were studied at the zero-point corrected CCSD/6-31++G(d,p)+5(sp)//MP2/6-31++G(d,p)+5(sp) level. Two new non-ionic structures were identified, one of which is 1.75 kcal/mol lower than any previously reported structure. Two new zwitterion conformers are lower in energy than any previously reported zwitterion. The lowest non-ionic structure is lower in energy than the lowest zwitterion by 2.8 kcal/mol at our highest level of theory, and for no basis or theory level is a zwitterion structure suggested to be the global minimum. Finally, we also examined, at Koopmans' theorem level, the electron binding energies of the six structures.

1. Introduction

It is well known that, in aqueous solutions at pH = 7, aminoacids exist primarily in their zwitterionic forms with the carboxyl group deprotonated and one of the nitrogen atoms protonated. In contrast, the zwitterionic forms are usually higher in energy in the gas phase than the corresponding non-zwitterion H_2N–CHR–COOH tautomers. Moreover, for some aminoacids (e.g., glycine), the zwitterionic structure does not even correspond to a local minimum on the gas-phase potential energy surface [1,2]. There have been attempts to identify a zwitterionic tautomer of a neutral aminoacid which is globally stable in the gas phase [3] but these findings were inconclusive [4]. Other recent efforts concentrated on hydrated [5,6], protonated [3,7,8], and alkali cationized [7–9] aminoacids.

We have recently suggested stabilizing the zwitterion form of an aminoacid in the gas phase with an excess electron [10] and it was via this route that we began our exploration of various arginine isomer energetics. Our reasoning is that the zwitterion form of an aminoacid would possess a larger dipole moment than the non-zwitterion form. It is well established that a molecule with a dipole moment larger than ca. 2.5 D binds an excess electron [11] with an electron binding energy roughly correlated with the magnitude of the dipole moment [12]. Therefore, we anticipated the instability of the zwitterion relative to the non-zwitterion structure might be reversed by the excess electron binding energy. It is our plan to

* Corresponding author. Fax: +1-801-581-8433.
E-mail address: simons@chemistry.utah.edu (J. Simons).

0009-2614/01/$ - see front matter
PII: S0009-2614(01)00166-X

report on our arginine *anion* findings in a future publication. However, because there is strong current interest in arginine itself, we decided to put forth our findings on the neutral species at this time.

In a recent study, we demonstrated that the instability of the zwitterion structure of glycine is significantly reduced by the attachment of an excess electron as a result of which a local minimum develops on the anionic potential energy surface [10]. However, its energy is still higher than that of the anion based on the non-zwitterion isomer of glycine. This outcome may be related to the fact that the proton affinity of glycine of 211.9 kcal/mol is the smallest among all 20 common naturally occurring amino acids [7]. The largest proton affinity of 251.2 kcal/mol is displayed by arginine [7], which possesses an extremely basic guanidine group. Therefore, we selected arginine for our study on stabilization through electron binding which leads to the results presented here.

The question of which tautomeric form of neutral arginine is dominant in the gas phase has recently been addressed in experimental studies [3,4,7,9,13,14]. Williams and co-workers [3] concluded, on the basis of black body infrared radiative dissociation plus Fourier transform-mass spectrometry measurements, that protonated dimers of arginine are bound in a salt-bridge. Moreover, the results of their extensive 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 non-ionic tautomer. Saykally and co-workers [4], however, did not confirm the dominance of the zwitterion of arginine in their infrared cavity ringdown laser absorption spectroscopy experiments. The observed band at ca. 1700 cm^{-1}, which is associated with the carbonyl stretch mode of a carboxylic acid, implied the presence of a non-ionic structure in their gas-phase sample. The absence of bands in the 1500–1660 cm^{-1} region, which are 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 neutral and zwitterion forms of arginine and the thermodynamically unstable form may have a sufficiently long lifetime to be observed experimentally. Moreover, the sources of arginine used in the Williams and Saykally experiments are different. One employes solution electro spray at 37–149°C; the other employes a heated pulsed beam source of pure arginine at 170°C.

The geometrical shapes of the protonated, sodiated, and cesiated arginine were probed in the gas phase by using the ion mobility based ion chromatography method [7,9]. Unfortunately, the qualitative structure of the protonated arginine could not be unambiguously determined from these experiments. It has been suggested that the alkali cationized arginine forms a salt bridge structure, related to the zwitterion form of this aminoacid. Results from the collisionally activated dissociation experiments of Williams et al. [13] and the kinetic experiments of Cerda and Wesdemiotis [14] 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^+, the non-zwitterion arginine solvates the metal ion. For the larger metal ions, a salt bridge is formed in which the arginine exists as a zwitterion.

Maksic and Kovacevic (MK), inspired by the variety of intriguing experimental findings, performed an extensive computational investigation in order to find the global minimum for neutral arginine [15]. They concluded from their thorough MP2 and B3LYP calculations using several basis sets, that the most stable structure is a non-ionic. These authors admitted, however, that the energy difference between the lowest non-ionic and a pair of low lying zwitterion structures is relatively small (within 1–3 kcal/mol depending on the theoretical model applied) [15].

As detailed above, our long-term goal is to determine whether an excess electron can stabilize a zwitterion structure of arginine. This goal requires extensive knowledge about the potential energy surface for the neutral species, and in this contribution we characterize what we have found to be the most promising neutral non-ionic and zwitterionic structures. In addition to the structures characterized so far [3,15], we identified two non-zwitterion and two zwitterion structures that are

promising candidates. Their relative energies have been determined at the MP2, B3LYP, and CCSD levels and their dipole moments have been calculated at the MP2 and B3LYP levels. These dipole moments will guide us as to the ability to bind an excess electron. With this goal in mind, the vertical electron attachment energies are determined at Koopmans' theorem (KT) level. An unexpected outcome of the current study on the six low energy tautomers of arginine is that the largest dipole moment of ca. 9 D is displayed by the lowest energy zwitterion's structure and by one of the non-ionic structures. However, their vertical electron detachment energies, determined at the KT level, differ considerably.

2. Methods

The equilibrium geometries of the neutral species have been optimized at the second-order Møller-Plesset (MP2) perturbation theory level as well as by applying the DFT method with a hybrid B3LYP functional. The latter method was also used to calculate harmonic vibrational frequencies that were used in zero-point vibrational energy corrections for all the structures. In addition, single point coupled-cluster calculations were performed with single and double excitations (CCSD) for every structure at its corresponding MP2 minimum geometry. Since the system studied contains 12 heavy atoms and 14 hydrogen atoms (which led to 320 contracted basis functions in the chosen basis set), we had to limit the level of our calculations to the CCSD level. Although we are not able to attain a better (e.g., CCSD(T) level) treatment of electron correlation, recent work by Nguyen et al. [16] on glycine has shown that the energy ordering for normal and zwitterion forms is the same at the CCSD and CCSD(T) levels. Moreover, recent findings of Fogarasi [17] on tautomers and rotomers of cytosine show that CCSD-level predictions are in good agreement with available experimental data. Finally, the general performances of CCSD vs CCSD(T) theories are discussed in [18–20]. For all of these reasons, we feel comfortable in reporting the CCSD data that we have been able to achieve.

In terms of atomic orbitals, we primarily used 6-31++G(d,p) [21,22] basis sets to describe the neutral molecule. We supplemented this basis with diffuse functions having lower exponents to: (i) make our electronic energies consistent with those that we are presently calculating for the corresponding anionic species, and (ii) estimate properly the vertical electron attachment energies at Koopmans' theorem level [23]. Thus, we centered even-tempered [24] five-term s and five-term p sets of diffuse functions on the positive end of the molecular dipole of each species. This was either the C atom surrounded by three nitrogen atoms or one of the N atoms belonging to the $-C(NH_2)_2$ (or $C(NH)NH_2$) functional group. The extra diffuse s and p functions share 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(d,p) basis set designed for nitrogen. As a consequence, we achieved lowest exponents of 2.0448×10^{-5} a.u. for both s and p symmetries. In the B3LYP calculations, the 6-31++G(d,p) basis set was supplemented with even-tempered two-term s and two-term p sets of diffuse functions centered on the C atom surrounded by three nitrogen atoms. We verified, at the MP2 level, that the relative energies of the isomers examined here are insensitive on the centering of the additional diffuse 5(sp) set. In particular, if the additional 5(sp) diffuse set is removed, the relative energies are affected by less than 0.03 kcal/mol.

To probe the effect of extending our valence basis, we also computed the energies by performing single point MP2 calculations with a 6-311++G(d,p) [21,22] basis set at the MP2/6-31++G(d,p) geometries.

All calculations were performed with GAUSSIAN 98 [25] and the coupled-cluster calculations were performed with MOLPRO [1] on 500 MHz dual processor Intel Pentium III computers and on SGI Origin2000 numerical servers.

[1] It is a package of ab initio programs written by H.-J. Werner, P.J. Knowles, CCSD [26].

3. Results

3.1. Identification of tautomers

The first zwitterion (Z1) and non-zwitterion structures (N1) were kindly provided to us by the Williams group (see Fig. 1) [27]. Species Z1 was identified in [3] where it was characterized as the lowest energy zwitterion structure consistent with the findings of MK who labeled this structure **3** [3,15]. Species (N1) is the same as the lowest energy non-ionic structure **4** characterized by MK [15]. The extensive search performed by the Williams group involved checking 240 arginine conformers. Molecular dynamics as well as the semiempirical AM1 and PM3 methods and the RHF/6-31G* method were used by them to select a few low lying conformers. Then MP2 and BLYP methods, with 6-31G* basis sets, were applied to determine final energies of these candidate structures [3].

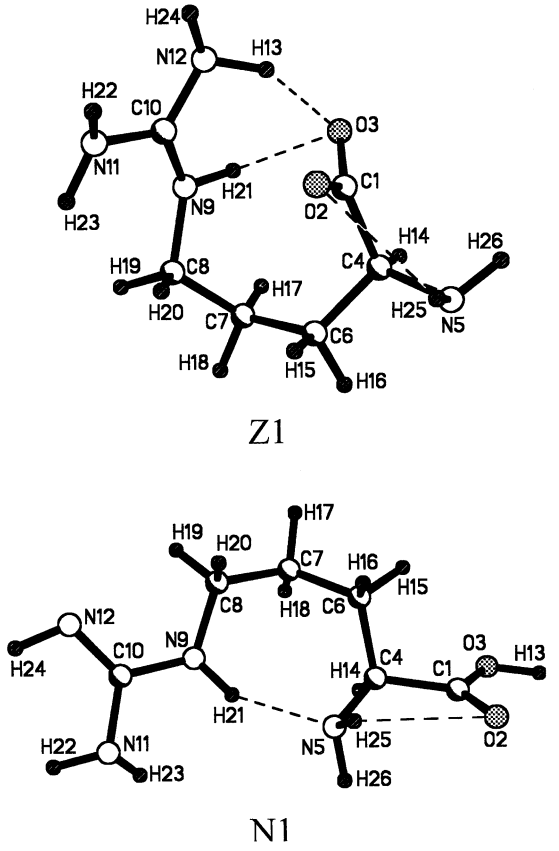

Fig. 1. The equilibrium MP2/6-31++G(d,p)+5(sp) geometries of neutral arginine corresponding to the lowest energy structures Z1 and N1.

The structures Z1 and N1 can be characterized as possessing three and two hydrogen bonds, respectively, as indicated in Fig. 1, and they differ primarily by movement of a proton from the carboxylic acid group to the terminal guanidine group.

We also examined a number of other structures suggested by our own chemical intuition. In particular, we designed a number of 'test' conformers by attempting to generate as many hydrogen bond contacts as possible while not causing too much internal strain. We also began with an open chain-like structure and allowed it to rearrange to minimize the energy. Frankly, it was through a rather exhaustive albeit non-systematic process that we were able to generate the candidate structures that we subsequently examined in greater detail. At present, we are extending this study, using a genetic algorithm (within a molecular mechanics force field) to attempt to find even more potential low-energy structures.

Nevertheless, using the intuitive procedure outlined, we succeeded in identifying four other low-lying structures labeled Z2 and N2 (see Fig. 2) and Z3 and N3 (see Fig. 3). Z2 is a zwitterion that differs from Z1 by having a different network of hydrogen bonds (i.e., only one oxygen atom is involved in all three hydrogen bonds which is a consequence of a different orientation of guanidine group with respect to Z1). Species N2 is a non-ionic structure which differs from N1 primarily in the orientations of the carboxylic acid and alpha-amino groups. The hydrogen atom (H13) belonging to the –COOH group forms a H-bond with a nitrogen (N5) lone pair in structure N2 but not in N1. In both non-ionic structures, the alpha-amino group is also involved in a hydrogen bond with the guanidine group, but it acts as a proton acceptor in N1 and as a proton donor in N2 (see Figs. 1 and 2).

For the Z3 structure, both oxygen atoms are involved in hydrogen bonds with the –NH–C–$(NH_2)_2$ terminus (see Fig. 3). The O2 atom is

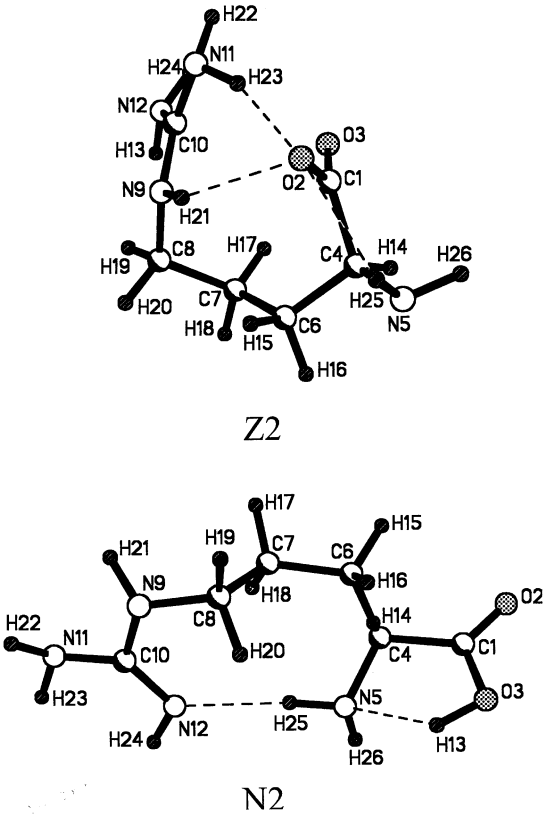

Fig. 2. The equilibrium MP2/6-31++G(d,p)+5(sp) geometries of neutral arginine corresponding to the lowest energy structures Z2 and N2.

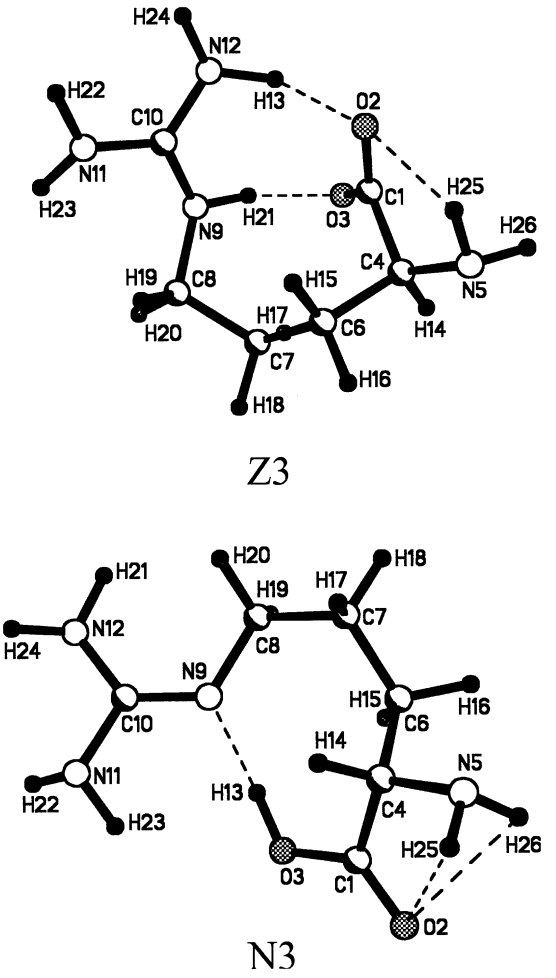

Fig. 3. The equilibrium MP2/6-31++G(d,p)+5(sp) geometries of neutral arginine corresponding to the lowest energy structures Z3 and N3.

bound to H13 and O3 to H21. There is also a hydrogen bond between O2 and H25, as it was for Z1 and Z2. In the neutral structure N3, a H atom transfers from nitrogen N9 to N12 (see Fig. 3). A short hydrogen bond is formed between N9 and H13 of the –COOH group and standard hydrogen bonds develop between O2 and H25(H26). The importance of such structures was recognized by Price et al. [3] and by MK [15] who label such a species **4**. Unfortunately, these authors focused on a higher energy structure with a different set of hydrogen bonds.

3.2. Relative energies and dipole moments

The energies of all the structures were determined relative to the energy of N3, which appears to be the lowest energy structure at every level of theory and for all basis sets (see Table 1). The MP2 results obtained with the 6-31++G(d,p)+5(sp) basis set (first two rows in Table 1) suggest that the neutral structure N3 is the most stable, and the zwitterion Z3 the second most stable. The relative energies of all six structures were also computed within a larger valence basis at the MP2/6-311++G(d,p)// MP2/6-31++G(d,p)+5(sp) level (Table 1, third row). Even though the separation between N3 and two higher in energy neutral structures N1 and N2 remains relatively stable

Table 1
Relative[a] electronic and zero-point corrected energies of the non-ionic and zwitterionic structures of the neutral arginine studied in this work[b]

Method	Z1	Z2	Z3	N1	N2	N3
MP2	3.388	2.269	1.278	2.861	1.612	0.000
MP2 + $\Delta E_{0,vib}$	3.374	2.580	1.135	3.155	2.500	0.000
MP2[c]	5.729	4.533	3.509	2.893	1.379	0.000
B3LYP[d]	3.681	3.855	1.408	2.221	1.506	0.000
B3LYP[d] + $\Delta E_{0,vib}$	3.674	4.163	1.262	2.512	2.394	0.000
CCSD	4.309	3.213	2.988	1.022	0.904	0.000
CCSD + $\Delta E_{0,vib}$	4.305	3.524	2.845	1.751	1.792	0.000

[a] The electronic energies of isomer N3 calculated at the MP2, B3LYP, and CCSD levels are −604.8242180, −606.5885150, and −604.903364 au, respectively. MP2 and CCSD calculations were performed with 6-31++G(d,p)+5(sp) basis set.
[b] Energies are given in kcal/mol.
[c] MP2 calculations with 6-311++G(d,p) basis set.
[d] B3LYP results were obtained in the 6-31++G(d,p) basis set supplemented with the 2(sp) diffuse set.

with respect to the MP2/6-31++G(d,p)+5(sp) results, the zwitterionic Z1, Z2, and Z3 structures are significantly destabilized when the large 6-311++G(d,p) valence basis is used.

At the DFT B3LYP level, the non-ionic N3 structure is still favored but Z3 is lower in energy than N2 by 0.1 kcal/mol. This difference increases to 1.13 kcal/mol after including the zero-point vibrational correction ($E_{0,vib}$).

N3 remains the lowest energy structure at the CCSD levels. At this level, N1 and N2 are very close in energy but they are less stable than N3 by 1.8 kcal/mol. Z3 proves again to be the lowest energy zwitterion and is separated from N3 by 2.8 kcal/mol.

The stability of the non-ionic structures N1, N2, and N3 relative to the zwitterion structures Z1, Z2, and Z3 supports the experimental findings of Saykally et al. that non-ionic arginine dominates in the gas phase [4,13,14]. Due to the small energy differences between N1, N2, and N3, determination of the global minimum of arginine will require computational approaches more accurate than those applied by us in the current study.

The dipole moments of Z1–Z3 and N1–N3, collected in Table 2 display interesting features. The dipole moment of the lowest energy zwitterion Z3 is higher than that of the lowest energy neutral structure N3. However, the dipole moment of N2 is comparable to that of Z3 for all computational approaches. Thus our original assumption that the zwitterion form of the aminoacid would possess a larger dipole moment than any non-zwitterion form [12] is not valid. Arginine may be unique due to the presence of a guanidine group, but other aminoacids should be thoroughly examined. There is also a significant difference of 5 D between the dipole moment of N2 and that of N1. The difference is apparently related to the different H-bond network which causes different relative orientation of the –COOH and –NH–C(NH_2)NH groups in N1 with respect to N2. The dipole moments of two zwitterion structures, Z1 and Z2, are similar (7–8 D) and slightly smaller than that of Z3 (9 D). The

Table 2
Dipole moments (in Debye) for the zwitterions and neutral tautomers of arginine followed by vertical electron attachment energies (in kcal/mol) determined at Koopmans' theorem level (D^{KT})

	Z1	Z2	Z3	N1	N2	N3
μ^{SCF}	7.040	6.664	9.432	3.774	9.678	8.216
μ^{MP2}	7.022	6.673	9.209	3.648	8.854	7.477
μ^{B3LYP}	8.471	7.714	9.054	3.903	9.203	7.959
D^{KT}	0.885	0.659	3.024	0.013	1.255	1.111

fact that charge-separated structures (zwitterions Z1 and Z2) have smaller dipole moments than that found for non-ionic isomer N2 is surprising but can be explained when one notices that a network of hydrogen bonds makes both zwitterionic structures very compact. In contrast, the H-bonds in structure N2 allow the C- and N-termini to be well separated, which maximizes the net value of the dipole moment (see Figs. 1 and 2). In zwitterions (Z1 and Z2) the partial charges (resulting from a proton transfer) are localized relatively closely to each other (due to the H-bond network) which decreases the value of total dipole moment for those structures.

The reported values of dipole moments are larger than the critical value of ca. 2.5 D required to bind an excess electron [11]. The vertical electron attachment energies, determined at Koopmans' theorem level and reported in Table 2, correlate with the values of SCF dipole moments and span a range from 3.02 kcal/mol for Z3 to 0.01 kcal/mol for the lowest-dipole structure N1. Interestingly, the Z3, with a SCF dipole moment of 9.05 D, binds the excess electron 2.4 times stronger than N2 with a SCF dipole moment 0.15 D larger than that of Z3. Determination of geometrical relaxation of the molecular frameworks upon electron attachment and inclusion of orbital relaxation and electron correlation effects is, as noted earlier, the subject of our ongoing study.

4. Conclusions

Six carefully selected structures of neutral arginine have been studied using electronic structure methods (CCSD/6-31++G(d,p)+5(sp)//MP2/6-31++G(d,p)+5(sp) results corrected for zero-point vibrational contribution determined at the B3LYP/6-31++G(d,p)+2(sp) level). Our conclusions are that:

(1) The non-ionic structure N3 is lower in energy by 1.8 kcal/mol than the lowest previously known structure N1.

(2) A new zwitterion structure Z3 was found to be lower in energy by 1.5 kcal/mol than the lowest previously known zwitterion Z1.

(3) The non-ionic structure N3 is lower in energy than the lowest zwitterion Z3 by 2.8 kcal/mol. Hence, neutral arginine should exist primarily as a non-ionic tautomer in the gas phase at room temperature.

(4) The largest dipole moment of ca. 9 D is displayed by the lowest zwitterion structure Z3 and the non-ionic structure N2. The vertical electron attachment energies of N2 and Z3 amounts to 1.3 and 3.0 kcal/mol, respectively, at Koopmans' theorem level.

(5) Further refinement in the global minimum of arginine will offer a serious challenge to computational chemists due to the small energy differences among the different tautomers and conformers.

Acknowledgements

The authors thank Prof. Evan R. Williams' group for providing their minimum energy structures of non-ionic and zwitterion arginine and Dr. Janusz Rak for his valuable comments. This work was supported by the NSF Grants CHE9618904 and CHE9982420 to J.S. and the Polish State Committee for Scientific Research (KBN) Grant No. DS/8371-4-0137-0. M.G. was supported by the US DOE, Office of Basic Energy Sciences, Chemical Sciences Division. The computer time provided by the Center for High Performance Computing at the University of Utah is also gratefully acknowledged.

References

[1] J.H. Jensen, M.S. Gordon, J. Am. Chem. Soc. 113 (1991) 7917.
[2] Y. Ding, K. Krogh-Jespersen, Chem. Phys. Lett. 199 (1992) 261.
[3] W.D. Price, R.A. Jockusch, E.R. Williams, J. Am. Chem. Soc. 119 (1997) 11988.
[4] C.J. Chapo, J.B. Paul, R.A. Provencal, K. Roth, R.J. Saykally, J. Am. Chem. Soc. 120 (1998) 12956.
[5] M.S. Gordon, J.H. Jensen, Acc. Chem. Res. 29 (1996) 536.
[6] P.I. Nagy, B. Noszal, J. Phys. Chem. A 104 (2000) 6834.
[7] T. Wyttenbach, W. Matthias, M.T. Bowers, J. Am. Chem. Soc. 122 (2000) 3458.

[8] T. Wyttenbach, J.E. Bushnell, M.T. Bowers, J. Am. Chem. Soc. 120 (1998) 5098.
[9] T. Wyttenbach, W. Matthias, M.T. Bowers, Int. J. Mass Spectrom. 183 (1999) 243.
[10] M. Gutowski, P. Skurski, J. Simons, J. Am. Chem. Soc. 122 (2000) 10159.
[11] M. Gutowski, P. Skurski, Recent Res. Devel. Phys. Chem. 3 (1999) 245, and references therein.
[12] M. Gutowski, K.D. Jordan, P. Skurski, J. Phys. Chem. A 102 (1998) 2624.
[13] R.A. Jockusch, W.D. Price, E.R. Williams, J. Phys. Chem. A 103 (1999) 9266.
[14] B.A. Cerda, C. Wesdemiotis, Analyst 125 (2000) 657.
[15] Z.B. Maksic, B. Kovacevic, J. Chem. Soc. Perkin Trans. 2 (1999) 2623.
[16] D.T. Nguyen, A.C. Scheimer, J.W. Andzelm, S. Sirois, D.R. Salahub, A.T. Hagler, J. Comput. Chem. 18 (1997) 1609–1631.
[17] G. Fogarasi, J. Mol. Struct. 413–414 (1997) 271.
[18] R.J. Bartlett, J.F. Stanton, in: K.B. Lipkowitz, D.B. Boyd (Eds.), Reviews in Computational Chemistry, vol. V, VCH Publishers Inc, New York, 1994.
[19] R.J. Bartlett, in: D.R. Yarkony (Ed.), Modern Electronic Structure Theory, Part II, World Scientific, Singapore, 1995, p. 1047.
[20] T.H. Dunning Jr., J. Phys. Chem. A 104 (2000) 9062.
[21] R. Ditchfield, W.J. Hehre, J.A. Pople, J. Chem. Phys. 54 (1971) 724.
[22] W.J. Hehre, R. Ditchfield, J.A. Pople, J. Chem. Phys. 56 (1972) 2257.
[23] T. Koopmans, Physica (Amsterdam) 1 (1934) 104.
[24] M.W. Schmidt, K.J. Ruedenberg, Chem. Phys. 71 (1979) 3961.
[25] M.J. Frisch et al., GAUSSIAN, Pittsburgh, PA, 1998.
[26] C. Hampel, K. Peterson, H.-J. Werner, Chem. Phys. Lett. 190 (1992) 1.
[27] E.R. Williams, private communication, March 2000.