JCTC Journal of Chemical Theory and Computation

Intrinsic Relative Stabilities of the Neutral Tautomers of Arginine Side-Chain Models

Jan Norberg,*,† Nicolas Foloppe,‡ and Lennart Nilsson†

Center for Structural Biochemistry, Department of Biosciences at Novum, Karolinska Institutet, SE-141 57 Huddinge, Sweden

Received December 18, 2004

Abstract: The specific protonation state of amino acids is crucial for the physicochemical properties of proteins and their biological functions. These protonation states influence, for instance, properties related to hydrogen bonding, solubility, and folding. pK_a calculations for proteins are, therefore, important and require, in principle, a specification of the most stable protonated and deprotonated forms of each titratable group. This is complicated by the existence of multiple tautomers, like the five neutral tautomers of the guanidine moiety in arginine. In this study, the compounds *N*-methyl-guanidine and *N*-ethyl-guanidine were used to model the charged and all neutral protonation states of the arginine side chain. The relative stabilities of all five neutral tautomers were investigated systematically for the first time, using quantum-mechanical calculations. These relative stabilities were obtained in vacuo, water and chloroform, by combining the quantum-mechanical calculations with a continuum solvation model. The water model was used to represent arginines exposed to an aqueous solution, whereas the chloroform model has a polarity representative of a protein core or a membrane. This allowed determining the relative pK_a's associated with each neutral tautomer in these environments. A key result is that significant differences in stability are found between the neutral tautomers, in both water and chloroform. Some tautomers are consistently found to be the most stable. These findings will be helpful to refine pK_a calculations in proteins.

Introduction

The guanidine is found at the extremity of arginine side chains and is, therefore, an important biochemical building block. This guanidine functionality is usually assumed to be in the protonated guanidinium form because of its high intrinsic basicity, with an effective pK_a of about 12.0 in aqueous solution at 298 K.[1] There are several examples where arginine is found to play a key role in relation to the mechanism of action of a protein.[2−5] For instance, the amino acid compositions of the interfaces of a large number of protein−protein[6−8] and protein-nucleic acid[9] complexes have been characterized, showing that arginine occurs frequently at the interface of the complexes.[8,9] Other specific examples include the activation of G protein coupled receptors,[10] the voltage gating of ion channels,[11] a possible role for neutral arginine in porin channel selectivity,[12] and the substrate binding of the human FLAP endonuclease-1 enzyme.[13] Peptides containing arginine have an important role in proteomics projects, which use mass spectrometry for rapid and reliable identification of proteins.[14−16] The fragmentation efficiency depends strongly on the protonation state of arginine.

Despite being very basic, the guanidine functionality is included as a titratable site in methods that aim at determining the protonated state of proteins.[17−20] These methods typically rely on the solution of the Poisson−Boltzmann equation in the framework of continuum electrostatics[18,20] and calculate the pK_a shift of a titratable group when it is transferred from

* Corresponding author tel.: +46 8 608 9263, fax: +46 8 608 9290, e-mail: jan.norberg@biosci.ki.se.

† Karolinska Institutet.

‡ Present address: Vernalis (R&D) Ltd., Granta Park, Abington, Cambridge CB1 6GB, U. K.

10.1021/ct049849m CCC: $30.25
Published on Web 07/13/2005

aqueous solution to the protein. The calculations involve both the protonated and the nonprotonated states of each titratable group in the protein environment.

In these pK_a calculations, one proton has to be removed from the guanidinium functionality of each arginine side chain. Details of the positioning of the polar hydrogens can have a significant impact on the pK_a's calculated in a protein environment,[21−23] and there are, a priori, five acidic protons on the guanidinium group, corresponding to five different neutral guanidine tautomers. Because none of these five protons are rigorously chemically equivalent, it would be very useful to have a ranking of their relative intrinsic acidities. This information is, to our knowledge, not available, probably because it would be difficult to determine the corresponding microscopic pK_a values experimentally. This, indeed, would require one to distinguish experimentally between each of these five neutral tautomers in aqueous solution. Alternatively, the guanidinium and each of the five neutral tautomers can be studied independently using a computational approach.

Recent studies have shown that the relative pK_a's of a series of analogous compounds in aqueous solution can be calculated with reasonable accuracy, using high level ab initio quantum-chemistry calculations, in combination with a continuum model to represent the condensed phase dielectric properties.[24−26] One study that tested several continuum models[24] found a good agreement between experimental and calculated pK_a's when the aqueous environment was represented using the isodensity surface polarized continuum model (IPCM).[27] The present report applies a similar protocol to rank the relative intrinsic acidities associated with the arginine side chain, modeled with two compounds, *N*-methylguanidine (Me-Gua) and *N*-ethyl-guanidine (Et-Gua), to test the effect of the length of the hydrocarbon side chain on the results. The relative acidities have been investigated in water, as well as in a medium with a low dielectric constant, to represent a protein core or a membrane environment. The results do indicate that there are significant stability differences between the neutral tautomers of an arginine side chain.

Computational Methodology

The compounds used to model the arginine side chain were the protonated and neutral forms of Me-Gua (Figure 1) and Et-Gua (Figure 2). The protonated forms (guanidinium) of Me-Gua (Figure 1) and Et-Gua (Figure 2) are denoted Me-GuaP and Et-GuaP, respectively. The five neutral tautomers are referred to as Me-GuaA (Et-GuaA) to Me-GuaE (Et-GuaE) (Figure 1), with Me-GuaX/Y (Et-GuaX/Y) denoting any of the neutral forms. Although larger molecules, even arginine itself, would be tractable on modern computers, it is on purpose that the model compounds were kept as simple as possible. This is indeed needed to assess the true intrinsic properties of the guanidine group, without the complications of conformational issues that inevitably arise with larger compounds. Larger compounds also open the possibility of intramolecular interactions involving the guanidine moiety (e.g., folded structures), which would only obscure the results regarding the intrinsic preferences for the protonation states.

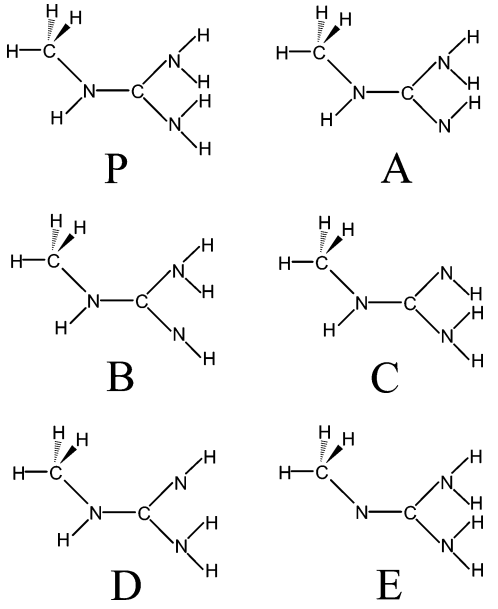

Figure 1. Structures of the protonated (Me-GuaP) and neutral (Me-GuaA to Me-GuaE) forms of Me-Gua.

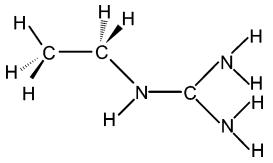

Figure 2. Structure of the protonated (Et-GuaP) form of Et-Gua. The same naming scheme was used for the Et-Gua series as for the Me-Gua series; therefore, the neutral tautomers of Et-Gua are not shown explicitly.

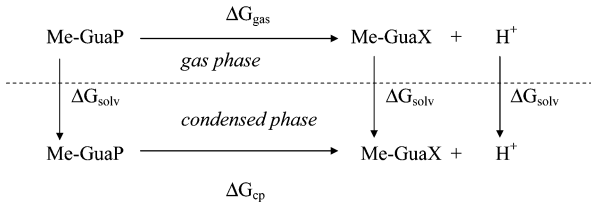

Figure 3. Thermodynamics cycle used for deriving the relative pK_a's, using Me-Gua as an example.

The dissociation constant K_a of an acid HA into its conjugate base A^- and a proton H^+ is related to the Gibbs free energy change (ΔG) of this dissociation reaction by

$$\Delta G = -RT \ln K_a = G(H^+) + G(A^-) - G(HA) \quad (1)$$

where R is the molar gas constant and T is the temperature; our calculations used $T = 298$ K. The absolute pK_a for the dissociation reaction is then

$$pK_a = \frac{1}{2.303RT} \Delta G \quad (2)$$

The absolute pK_a is difficult to calculate, in particular, because of problems associated with the determination of $G(H^+)$.[28−30] This study, therefore, focuses on the relative stabilities of the neutral tautomers of Me-Gua and Et-Gua. This can be investigated using a thermodynamic cycle (Figure 3), from which the pK_a difference, ΔpK_a, corre-

sponding to the difference in acidity associated with two different neutral tautomers can be obtained. The cycle relates the acid dissociation in the condensed phase (ΔG_{cp}) to its dissociation in the gas phase (ΔG_{gas}) and to the solvation free energies (ΔG_{solv}) of all the chemical species involved. Accordingly, if the acid Me-GuaP dissociates to produce one of its conjugate bases Me-GuaX, the corresponding free energy difference and the associated microscopic absolute pK_a in the condensed phase are given by

$$\Delta G_{cp}(\text{Me-GuaX}) = 2.303RT\, \text{p}K_a(\text{Me-GuaX}) = \Delta G_{solv}(\text{Me-GuaX}) + \Delta G_{solv}(H^+) - \Delta G_{solv}(\text{Me-GuaP}) + \Delta G_{gas}(\text{Me-GuaX}) \quad (3)$$

where ΔG_{solv}(Me-GuaX) and ΔG_{solv}(Me-GuaP) are the free energies of solvation of a neutral guanidine tautomer and the guanidinium, respectively. $\Delta G_{solv}(H^+)$ is the free energy of solvation for the proton, and ΔG_{solv}(Me-GuaX) is the deprotonation free energy of the guanidinium in the gas phase. Subtracting eq 3, obtained with a neutral tautomer, Me-GuaY, from its counterpart for another neutral tautomer, Me-GuaX, yields

$$\Delta\Delta G_{cp}(\text{Me-GuaX} - \text{Me-GuaY}) = \Delta G_{solv}(\text{Me-GuaX}) - \Delta G_{solv}(\text{Me-GuaY}) + \Delta G_{gas}(\text{Me-GuaX}) - \Delta G_{gas}(\text{Me-GuaY}) \quad (4)$$

A key point is that the free energies of solvation of the proton and the protonated guanidinium from eq 3 cancel when eq 4 is formed. Therefore, the relative pK_a difference between two distinct neutral guanidine tautomers depends only on (i) the difference in free energy of solvation between these neutral tautomers [ΔG_{solv}(Me-GuaX) − ΔG_{solv}(Me-GuaY)] and (ii) the difference in the gas-phase protonation free energies associated with these neutral tautomers [ΔG_{gas}(Me-GuaX) − ΔG_{gas}(Me-GuaY)]. The gas-phase protonation free energies were computed using ab initio quantum mechanics, and the solvation free energies have been calculated using a continuum solvation model.[27]

All calculations were performed using the Gaussian 98 suite of programs.[31] The 6-31G(d,p) basis set was used for all compounds. Further, the 6-31+G(d,p) and 6-311+G(d,p) basis sets were also applied to test the influence of the basis set on selected calculations in vacuo. First, the geometries of all the compounds were optimized by energy minimization in the gas phase to the default tolerances of the program. The gas-phase energy-minimized geometries were then used directly for all the calculations in the condensed phase. This is expected to be a reasonable approximation, given that there is no torsional freedom within the model compounds that could be drastically affected by the presence of a solvent. Also, a recent study on similar compounds tested the effect of a solvent on the internal geometry, and this influence was found to be minor, as expected.[32]

The solvation free energies were obtained using the reaction-field IPCM,[27] which previously led to satisfactory agreement with experimental results for the relative pK_a's in a series of pyridine derivatives.[24] The solvation free energies were obtained at 298 K by calculating single-point energies with the IPCM solvation model, for both water and chloroform. The dielectric constants of water and chloroform were set to 78.56 and 4.806, respectively.[33] In the IPCM model, the isodensity cutoff value was 0.001 atomic units.[24]

All calculations were performed at three levels of theory: (i) Hartree−Fock (HF), (ii) second-order Møller−Plesset (MP2),[34] and (iii) with the density functional theory using the B3LYP method.[35] Gas-phase free energy differences were calculated by adding the vibrational, rotational, and translational corrections at 298 K to the differences in heat of formation obtained directly from Gaussian.[31] This thermal free energy correction was determined from a normal-mode analysis of the previously geometry-optimized structures in vacuo at the HF and B3LYP levels of theory. The gas-phase proton affinity (PA) was computed as the change in the heat of formation and in the zero-point vibrational energy:

$$\text{PA} = H_f(\text{Me-GuaX}) - H_f(\text{Me-GuaP}) + \text{ZPVE(Me-GuaX)} - \text{ZPVE(Me-GuaP)} \quad (5)$$

where H_f(...) and ZPVE(...) are the heat of formation and the zero-point vibrational energy, respectively.

Results

The properties of interest were determined for all compounds at the HF/6-31G(d,p), B3LYP/6-31G(d,p), and MP2/6-31G(d,p) levels in the gas phase and in water and at the MP2/6-31G(d,p) level in chloroform. For the Me-GuaD, Me-GuaE, and Et-GuaE model compounds, a tendency toward a more pyramidal conformation of the amino groups was observed, as has also been reported for the amino groups of guanidine and guanidinium cations.[36] The HF and B3LYP results are presented for the sake of completeness, but our analysis focuses on the MP2 results. Overall, for a given model compound, the results for the different levels of theory were quite consistent and do reveal differences across the neutral tautomers.

Basis Set Dependence and Bond Order Analysis. Given that the amino groups were found to be planar or pyramidal in different tautomers of methyl-guanidine with the 6-31G(d,p) basis set, the influence of larger basis sets on these geometries was tested. These calculations were done with the 6-31+G(d,p) and 6-311+G(d,p) basis sets and the B3LYP and MP2 levels of theory for Me-GuaB, Me-GuaD, and Me-GuaE. These were chosen because the Me-GuaB was planar at the MP2/6−31(d,p) level, whereas Me-GuaD and Me-GuaE were slightly pyramidal. Me-GuaB remained planar for both the larger basis sets. Also, Me-GuaD and Me-GuaE showed the same pyramidal tendency for the larger basis sets as they did for the 6-31G(d,p) basis set.

Heats of formation computed for the three compounds using B3LYP and MP2 with the two larger basis sets were ranked in the same order as with the 6-31G(d,p) basis set. These differences in the heats of formation, for a given compound, computed with the three different basis sets were <0.5 kcal/mol with the MP2 method and <1.4 kcal/mol with B3LYP.

A bond order (BO) analysis was made at the MP2 level with the 6-31G(d,p) basis set for the protonation sites that showed a tendency for a pyramidal conformation. The BO

Table 1. Relative Stabilities (kcal/mol) ($\Delta\Delta G$ from eq 4) between the Different Protonation States of Me-Gua and Et-Gua from the MP2/6-31G(d,p) Calculations

compound	water	chloroform
Me-GuaB−Me-GuaA	−1.05	−0.58
Me-GuaC−Me-GuaA	−1.92	−1.49
Me-GuaD−Me-GuaA	−4.70	−5.21
Me-GuaE−Me-GuaA	−3.35	−3.73
Me-GuaC−Me-GuaB	−0.87	−0.91
Me-GuaD−Me-GuaB	−3.65	−4.63
Me-GuaE−Me-GuaB	−2.30	−3.15
Me-GuaD−Me-GuaC	−2.78	−3.72
Me-GuaE−Me-GuaC	−1.43	−2.24
Me-GuaE−Me-GuaD	1.35	1.48
Et-GuaB−Et-GuaA	0.79	0.55
Et-GuaC−Et-GuaA	−0.69	−0.71
Et-GuaD−Et-GuaA	−1.25	−1.10
Et-GuaE−Et-GuaA	−3.21	−4.33
Et-GuaC−Et-GuaB	−1.48	−1.26
Et-GuaD−Et-GuaB	−2.04	−1.65
Et-GuaE−Et-GuaB	−4.00	−4.88
Et-GuaD−Et-GuaC	−0.56	−0.39
Et-GuaE−Et-GuaC	−2.52	−3.62
Et-GuaE−Et-GuaD	−1.96	−3.23

was about 0.99 and 1.22 for the C−N bond in the CNH_2 and CNH groups, respectively, and about 0.67−0.69 for the N−H bonds. These bond orders were observed for both Me-GuaD and Me-GuaE.

Relative Stabilities of Neutral Tautomers in Water and Chloroform. The relative stabilities of the five neutral tautomers for the two model compounds of the arginine side chain, Me-Gua and Et-Gua (Figures 1 and 2), were determined in water and in chloroform (Table 1 and Figures 4 and 5), using the thermodynamic cycle in Figure 3. The components of this cycle are presented in the sections below. For Me-Gua in water and chloroform, the A state was consistently the least stable and the D and E states were the most stable. The E state was also the most stable at the MP2 level with Et-Gua in both water and chloroform. The B rather than the A state, however, was found to be the least stable with Et-Gua. The B tautomer is also the second least stable state with Me-Gua at the MP2 level of theory. Arguably, the main differences between Me-Gua and Et-Gua are the

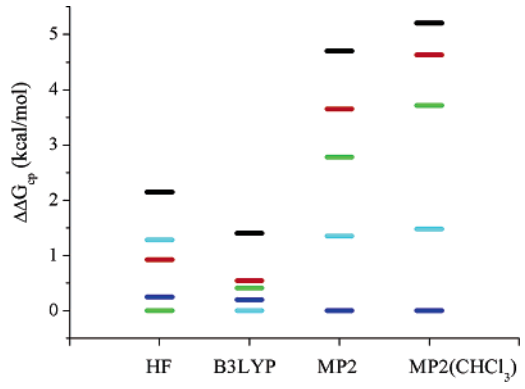

Figure 4. $\Delta\Delta G_{cp}$ in kcal/mol for the different protonation states of Me-Gua (A, black; B, red; C, green; D, blue; E, cyan).

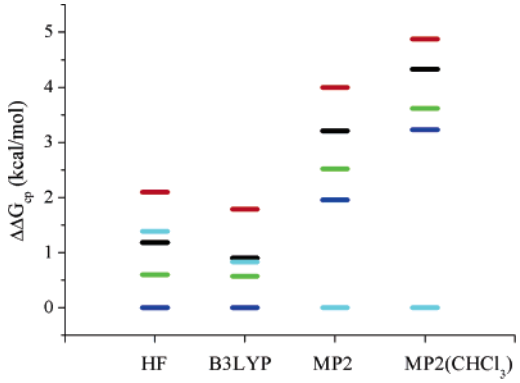

Figure 5. $\Delta\Delta G_{cp}$ in kcal/mol for the different protonation states of Et-Gua (A, black; B, red; C, green; D, blue; E, cyan).

Table 2. Gas-Phase Proton Affinity (kcal/mol) of Neutral Forms of Me-Gua and Et-Gua

compound	HF/ 6−31G(d,p)	B3LYP/ 6−31G(d,p)	MP2/ 6−31G(d,p)
Me-GuaA	252.35	249.25	247.65
Me-GuaB	252.62	249.62	247.94
Me-GuaC	251.19	248.85	246.68
Me-GuaD	251.57	249.00	241.88
Me-GuaE	250.90	246.76	243.54
Et-GuaA	253.93	251.00	249.14
Et-GuaB	254.19	251.34	249.42
Et-GuaC	252.79	250.56	248.31
Et-GuaD	253.20	250.76	248.64
Et-GuaE	253.50	250.36	242.60

relative rankings of states A and B and states D and E. The longer hydrocarbon side chain of Et-Gua, thus, has an influence on the relative stability of the neutral tautomers; however, the pattern of relative stabilities observed with the two model compounds is largely consistent.

Not only do the present results reveal differences in stability for the neutral tautomers, but these differences spread on a significant free energy range. In water, the difference between the most and least stable tautomers is at least 4.0 kcal/mol, well-above the thermal energy at 298 K. This difference is even more pronounced in chloroform. Interestingly, however, the relative ranking remains the same in chloroform as in water, for a given compound (Figures 4 and 5). Therefore, a largely consistent picture emerged regarding the relative intrinsic stabilities of the neutral tautomers of the compounds representing the arginine side chain.

Gas-Phase Proton Affinities and Free Energies. The gas-phase proton affinities at the MP2 level ranged from 241.88 to 247.94 kcal/mol for Me-Gua and from 242.60 to 249.42 kcal/mol for Et-Gua (Table 2). The B3LYP method led to slightly larger gas-phase proton affinities than with MP2; however, the differences in proton affinity across the tautomers are less pronounced when obtained with B3LYP than with MP2. For all states except E, the MP2 gas-phase proton affinities were slightly larger for the compounds with the longer hydrocarbon side chain, by about 1% on average. This increase of the proton affinity with an increase in the side chain length is compatible with the slightly smaller

Table 3. Change in Heat of Formation (ΔH_f) between the Neutral and Protonated Forms and Thermal Correction (kcal/mol) of the Neutral Forms of Me-Gua and Et-Gua Used in the MP2/6-31G(d,p) Calculations

compound	ΔH_f	thermal correction[a]
Me-GuaA	258.05	−8.58
Me-GuaB	258.30	−8.74
Me-GuaC	257.09	−9.27
Me-GuaD	252.24	−9.00
Me-GuaE	252.46	−7.83
Et-GuaA	259.52	−8.52
Et-GuaB	259.75	−8.66
Et-GuaC	258.75	−9.30
Et-GuaD	258.98	−8.97
Et-GuaE	253.39	−9.41

[a] The thermal correction to get the gas-phase Gibbs free energy difference at 298 K from ΔH_f.

proton affinities (about 233−238 kcal/mol) obtained quantum-mechanically in the absence of any side chain for guanidine.[36−38] These differences with previous theoretical results may also be, in part, ascribed to the differences in the levels of theory used. These results are quite similar to the experimental value of 235.7 kcal/mol obtained for guanidine,[39] suggesting that the protonated and neutral forms of these two compounds are valid models of the protonation states of the arginine side chain.

The variation in the MP2 proton affinities in a range of about 6.0 kcal/mol clearly shows that there are intrinsic differences associated with the protonation sites and their chemistry in the guanidine moiety. For both Me-Gua and Et-Gua, the largest gas-phase proton affinities were obtained for the B state (Table 2); however, the proton affinities of states A and C were close, within 1.3 kcal/mol. The behavior of tautomer E is also consistent across the two model compounds, with its relatively low proton affinity. This peculiarity of state E is not surprising, given that it is associated with the deprotonation of the nitrogen that differs chemically most from the others. The main surprise came from state D, for which the relative proton affinity strongly depends on the model compound; that is, it is the lowest in Me-Gua but much higher in Et-Gua.

Gas-phase free energies were obtained from the heats of formation and the thermal corrections (Table 3), which, as expected,[40] were quite similar for our compounds. The correction contributed less than 0.9 kcal/mol to the $\Delta\Delta G_{cp}$ in all cases except two involving Me-GuaE, for which it was up to 1.4 kcal/mol because of the smaller correction obtained for Me-GuaE.

Solvation Free Energy in Water and Chloroform. The solvation free energy in water of the protonated species ranged from about −64 to −62 kcal/mol for Me-GuaP and from −62 to −60 kcal/mol for Et-GuaP, depending on the level of theory (Table 4). As expected, the corresponding free energies for the neutral tautomers are much less favorable, by about 50 kcal/mol. The free energies of solvation of the neutral tautomers in water vary in a range of 2.7 kcal/mol in Me-Gua and 4.1 kcal/mol in Et-Gua. In that respect, states A, B, and C are rather similar, differing by 1.1 kcal/mol at most in each model compound. Neutral tautomer E stands apart in both model compounds, with a relatively unfavorable free energy of solvation in water. Surprisingly, state D in water goes from being the least well solvated in Me-GuaD to being the best solvated in Et-Gua.

As expected, the chloroform solvation free energies were found to be significantly less favorable compared to the aqueous solvation free energies for both charged and neutral species. Et-Gua was more soluble than Me-Gua in chloroform. The ranking of the MP2 solvation free energies of the neutral states in chloroform and water was the same for Et-Gua and only slightly different for Me-Gua.

Dipole Moments. The electric dipole moments of the neutral compounds are reported in the gas phase, water, and chloroform (Table 5) to examine their relationship with solvation properties. In that context, the present work is only concerned with the magnitude of these dipoles and not their orientation. In the gas phase, the dipole varies from ∼2.5 to ∼4.5 D with both model compounds, showing that the neutral tautomers vary in overall polarity. For these dipoles, the same pattern of similarities across the tautomers is observed as those for the proton affinities and the solvation energies. Indeed, states A, B, and C of Me-Gua and states A, B, C, and D of Et-Gua have similar dipole moments. State E is an

Table 4. Solvation Free Energy in Water and Chloroform (kcal/mol) of the Protonated and Neutral Forms of Me-Gua and Et-Gua

	HF/6−31G(d,p) water	B3LYP/6−31G(d,p) water	MP2/6−31G(d,p) water	MP2/6−31G(d,p) chloroform
Me-GuaP	−64.09	−62.45	−62.92	−49.48
Et-GuaP	−62.28	−60.12	−60.52	−47.32
Me-GuaA	−12.23	−10.79	−10.43	−7.29
Me-GuaB	−13.53	−11.83	−11.58	−7.97
Me-GuaC	−12.53	−10.98	−10.71	−7.14
Me-GuaD	−12.88	−11.24	−8.90	−6.27
Me-GuaE	−10.93	−9.30	−8.95	−6.19
Et-GuaA	−15.81	−13.50	−12.78	−8.80
Et-GuaB	−14.96	−12.78	−12.08	−8.34
Et-GuaC	−14.54	−12.44	−11.92	−7.96
Et-GuaD	−15.76	−13.64	−13.04	−8.91
Et-GuaE	−14.69	−12.46	−8.97	−6.11

Table 5. Dipole Moment (D) of the Neutral Forms of Me-Gua and Et-Gua in the Gas Phase, Chloroform, and Aqueous Solution

	HF/6−31G(d,p)		MP2/6−31G(d,p)			B3LYP/6−31G(d,p)	
compound	gas phase	aqueous	gas phase	chloroform	aqueous	gas phase	aqueous
Me-GuaA	3.98	5.04	4.08	4.46	4.78	3.63	4.76
Me-GuaB	4.19	5.48	4.30	4.84	5.23	3.84	5.22
Me-GuaC	3.93	5.02	4.03	4.46	4.82	3.73	4.87
Me-GuaD	4.11	5.13	3.04	3.22	3.42	3.86	4.94
Me-GuaE	2.57	3.51	2.47	2.77	3.22	2.30	3.06
Et-GuaA	4.02	5.40	4.12	4.69	5.10	3.68	5.12
Et-GuaB	4.34	5.60	4.45	4.96	5.32	4.00	5.34
Et-GuaC	3.82	5.20	3.92	4.49	4.96	3.62	5.02
Et-GuaD	4.17	5.60	4.27	4.96	5.41	3.92	5.46
Et-GuaE	3.80	4.86	2.45	2.71	3.04	3.62	4.66

outlier, with a significantly smaller dipole in both model compounds.

As expected, the dipole moments become systematically larger when going from the gas phase to a solvent, reflecting the influence of the solvent reaction field. This polarization is significantly more pronounced in water than in chloroform, as anticipated. When compared to the gas phase, the dipoles in water are on average 20.3% (Me-Gua series) and 24.1% (Et-Gua series) larger. When going from the gas phase to chloroform, the dipoles increase by 12.1% (Me-Gua) and 13.3% (Et-Gua) on average. These averages mask significant differences among the tautomers; for example, the dipole of Me-GuaD increases by only 12.5% in water, although the corresponding increase is 30.4% for Me-GuaE. It is interesting to note that Me-GuaE is both the least polar and yet the most polarizable tautomer in the Me-Gua series. Therefore, the neutral tautomers differ not only in polarity but also in polarizability.

Concluding Discussion

Although it is well-known that the guanidine group of arginine is strongly basic, the relative intrinsic stabilities of the corresponding neutral tautomers are unknown. In particular, this means that one of the neutral tautomers is chosen arbitrarily when performing pK_a calculations on proteins, using the increasingly popular continuum electrostatic techniques. This adds practical concerns to the broader and more fundamental question of the relative basicities of the guanidine tautomers. Indeed, it is now well-documented that pK_a calculations on full proteins are influenced by the detailed positioning of the polar hydrogens.[21−23] This is why the selection of a given guanidine tautomer will not only impact the calculated pK_a of this guanidine but also influence the electrostatic environment of other titratable groups in the vicinity of this guanidine.

Here, a variety of quantum-mechanical methods and model compounds were used to investigate the relative stabilities of the five neutral guanidine tautomers. This analysis was carried out in aqueous solution and in chloroform, which was used to mimic a protein core or a membrane environment. Like the presentation of the results, this discussion concentrates on the results obtained at the MP2 level. When examining each component of the free energy cycle, it is clear that the dominant contribution to $\Delta\Delta G_{cp}$ is from the heat of formation, with the thermal correction and the free energy of solvation providing a smaller modulation.

The model compounds Me-Gua and Et-Gua were selected because their size allows high level ab initio calculations to be carried out. In addition, they contain as few torsional degrees of freedom as possible, which allows derivation of properties that are truly intrinsic to the guanidinium moiety, without being confounded by conformational issues. For instance, using these compounds avoids formation of intramolecular hydrogen bonds that could be formed between the guanidine moiety and other polar groups in a full arginine,[41] which would cloud the interpretation of the results. We tested the effect of this simplification by using two model compounds, differing by the length of the aliphatic side chain. The pattern of behavior for the five tautomers is mostly similar between the Me-Gua and Et-Gua series, with the exception of state D. Me-GuaD presents surprising differences as compared to Me-GuaB, with respect to proton affinity and free energy of solvation in water and chloroform. These differences are surprising because the local topological arrangement of the protons seems quasi-equivalent in neutral states B and D and would be expected to yield similar results. This could be due to the tendency of the amino groups of the D state to form a pyramidal conformation instead of a planar conformation. States B and D, however, are much more similar in the Et-Gua series, as anticipated intuitively. This, combined with the fact that Et-Gua is, a priori, a closer analogue of the arginine side chain than Me-Gua, suggests that the interpretation of the results should put more weight on the Et-Gua series. It is, however, the similarities between both series, rather than their differences, that dominate the results.

Tautomer E is consistently an outlier, with relatively favorable proton affinities but relatively unfavorable solvation free energies in water or chloroform. In differential terms, however, the proton affinity of Et-GuaE more than compensates for its solvation free energy. That leads to tautomer E in water or chloroform being clearly the most stable in the Et-Gua series, by a distinctive margin (Figure 5). That tautomer E should differ from the other neutral tautomers is not surprising, given that it involves the deprotonation of the nitrogen that differs chemically most from the others.

The local symmetry of the guanidinium also suggests that tautomers A and C may behave similarly, which is indeed observed in the Et-Gua series. The same reasoning would also imply a similar behavior for tautomers B and D; however, that is not observed. Instead, tautomer D is systematically more stable than B, illustrating the limits of chemical intuition and the need for detailed calculations. In sum, these calculations show that the relative intrinsic stabilities of the neutral tautomers differ, be it in a medium of high or low dielectric value.

With both media and model compounds, tautomers E and D appear to be the most stable and should, therefore, be deprotonated in priority, for example, during pK_a calculations. Relying on the Et-Gua series, tautomer E may be selected as the most likely neutral form, given that the free energy difference between Et-GuaE and its congeners in water is nearly 2.0 kcal/mol or more. The preference for Et-GuaE is even more marked in a medium of low dielectric value, which also represents the conditions in which the neutral guanidine is most likely to exist. In both media, however, the intrinsic energy differences between the neutral tautomers (Table 1) are comparable to that associated with a single hydrogen bond, and therefore, the actual tautomeric state of the guanidine will be strongly influenced by the details of its environment. The present results should provide an energetic basis for a more rational positioning of the protons around a neutral guanidine moiety, a situation encountered, at least, during most pK_a calculations with proteins and continuum electrostatic methods.

References

(1) Nozaki, Y.; Tanford, C. Examination of titration behavior. *Methods Enzymol.* **1967**, *11*, 715−734.

(2) Ballesteros, J.; Kitanovic, S.; Guarnieri, F.; Davies, P.; Fromme, B. J.; Konvicka, K.; Chi, L.; Millar, R. P.; Davidson, J. S.; Weinstein, H.; Sealfon, S. C. Functional microdomains in G-protein-coupled receptors. The conserved arginine-cage motif in the gonadotropin-releasing hormone receptor. *J. Biol. Chem.* **1998**, *273*, 10445−10453.

(3) Barth, P.; Guillouard, I.; Setif, P.; Lagoutte, B. Essential role of a single arginine of photosystem I in stabilizing the electron-transfer complex with ferredoxin. *J. Biol. Chem.* **2000**, *275*, 7030−7036.

(4) Borgstahl, G. E.; Williams, D. R.; Getzoff, E. D. 1.4 A structure of photoactive yellow protein, a cytosolic photoreceptor: unusual fold, active site, and chromophore. *Biochemistry* **1995**, *34*, 6278−6287.

(5) Demchuk, E.; Genick, U. K.; Woo, T. T.; Getzoff, E. D.; Bashford, D. Protonation states and pH titration in the photocycle of photoactive yellow protein. *Biochemistry* **2000**, *39*, 1100−1113.

(6) Jones, S.; Daley, D. T.; Luscombe, N. M.; Berman, H. M.; Thornton, J. M. Protein-RNA interactions: a structural analysis. *Nucleic Acids Res.* **2001**, *29*, 943−954.

(7) Nadassy, K.; Wodak, S. J.; Janin, J. Structural features of protein-nucleic acid recognition sites. *Biochemistry* **1999**, *38*, 1999−2017.

(8) Lo Conte, L.; Chothia, C.; Janin, J. The atomic structure of protein−protein recognition sites. *J. Mol. Biol.* **1999**, *285*, 2177−2198.

(9) Jones, S.; van Heyningen, P.; Berman, H. M.; Thornton, J. M. Protein-DNA interactions: A structural analysis. *J. Mol. Biol.* **1999**, *287*, 877−896.

(10) Oliveira, L.; Paiva, A. C. M.; Vriend, G. A low resolution model for the interaction of G proteins with G protein-coupled receptors. *Protein Eng.* **1999**, *12*, 1087−1095.

(11) Jager, H.; Grissmer, S. Regulation of a mammalian Shaker-related potassium channel, hKv1.5, by extracellular potassium and pH. *FEBS Lett.* **2001**, *488*, 45−50.

(12) Karshikoff, A.; Spassov, V.; Cowan, S. W.; Ladenstein, R.; Schirmer, T. Electrostatic Properties of 2 Porin Channels from Escherichia-Coli. *J. Mol. Biol.* **1994**, *240*, 372−384.

(13) Qiu, J.; Bimston, D. N.; Partikian, A.; Shen, B. Arginine residues 47 and 70 of human flap endonuclease-1 are involved in DNA substrate interactions and cleavage site determination. *J. Biol. Chem.* **2002**, *277*, 24659−24666.

(14) Csonka, I. P.; Paizs, B.; Suhai, S. Modeling of the gas-phase ion chemistry of protonated arginine. *J. Mass Spectrom.* **2004**, *39*, 1025−1035.

(15) Dongre, A. R.; Jones, J. L.; Somogyi, A.; Wysocki, V. H. Influence of peptide composition, gas-phase basicity, and chemical modification on fragmentation efficiency: Evidence for the mobile proton model. *J. Am. Chem. Soc.* **1996**, *118*, 8365−8374.

(16) Paizs, B.; Suhai, S.; Hargittai, B.; Hruby, V.; Somogyi, A. Ab initio and MS/MS studies on protonated peptides containing basic and acidic amino acid residues I. Solvated proton vs salt-bridged structures and the cleavage of the terminal amide bond of protonated RD-NH2. *Int. J. Mass Spectrom.* **2002**, *219*, 203−232.

(17) Antosiewicz, J.; McCammon, J. A.; Gilson, M. K. Prediction of pH-dependent properties of proteins. *J. Mol. Biol.* **1994**, *238*, 415−436.

(18) Bashford, D.; Karplus, M. pK_a's of ionizable groups in proteins: atomic detail from a continuum electrostatic model. *Biochemistry* **1990**, *29*, 10219−10225.

(19) Beroza, P.; Case, D. A. Including side chain flexibility in continuum electrostatic calculations of protein titration. *J. Phys. Chem.* **1996**, *100*, 20156−20163.

(20) Demchuk, E.; Wade, R. C. Improving the continuum dielectric approach to calculating p*K*(a)s of ionizable groups in proteins. *J. Phys. Chem.* **1996**, *100*, 17373−17387.

(21) Alexov, E. G.; Gunner, M. R. Incorporating protein conformational flexibility into the calculation of pH-dependent protein properties. *Biophys. J.* **1997**, *72*, 2075−2093.

(22) You, T. J.; Bashford, D. Conformation and hydrogen ion titration of proteins: a continuum electrostatic model with conformational flexibility. *Biophys. J.* **1995**, *69*, 1721−1733.

(23) Nielsen, J. E.; Andersen, K. V.; Honig, B.; Hooft, R. W. W.; Klebe, G.; Vriend, G.; Wade, R. C. Improving macromolecular electrostatics calculations. *Protein Eng.* **1999**, *12*, 657−662.

(24) Chen, I. J.; MacKerell, A. D. Computation of the influence of chemical substitution on the p*K*(a) of pyridine using semiempirical and ab initio methods. *Theor. Chem. Acc.* **2000**, *103*, 483−494.

(25) Shapley, W. A.; Bacskay, G. B.; Warr, G. G. Ab initio quantum chemical studies of the p*K*(a)'s of hydroxybenzoic acids in aqueous solution with special reference to the hydrophobicity of hydroxybenzoates and their binding to surfactants. *J. Phys. Chem. B.* **1998**, *102*, 1938−1944.

(26) Tran, N. L.; Colvin, M. E. The prediction of biochemical acid dissociation constants using first principles quantum chemical simulations. *THEOCHEM* **2000**, *532*, 127−137.

(27) Foresman, J. B.; Keith, T. A.; Wiberg, K. B.; Snoonian, J.; Frisch, M. J. Solvent effects. 5. Influence of cavity shape, truncation of electrostatics, and electron correlation ab initio reaction field calculations. *J. Phys. Chem.* **1996**, *100*, 16098−16104.

(28) Mejias, J. A.; Lago, S. Calculation of the absolute hydration enthalpy and free energy of H^+ and OH^-. *J. Chem. Phys.* **2000**, *113*, 7306−7316.

(29) Schaefer, M.; Karplus, M. A comprehensive analytical treatment of continuum electrostatics. *J. Phys. Chem.* **1996**, *100*, 1578−1599.

(30) Lim, C.; Bashford, D.; Karplus, M. Absolute pK_a calculations with continuum dielectric methods. *J. Phys. Chem.* **1991**, *95*, 5610−5620.

(31) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A. J.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T. A.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, rev. A.7; Gaussian Inc: Pittsburgh, PA, 1998.

(32) Alagona, G.; Ghio, C.; Nagy, P. I.; Durant, G. J. Theoretical studies on the continuum solvation of some *N*,*N*′-dimethyl- and *N*-methyl, *N*′-acetyl-guanidine and guanidinium conformers. *J. Phys. Chem. A* **1999**, *103*, 1857−1867.

(33) Lide, D. R. E. *Handbook of Chemistry and Physics*, 72nd ed.; CRC Press: Boca Raton, FL, 1991.

(34) Møller, C.; Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **1934**, *46*, 618−622.

(35) Becke, A. D. Density-Functional Thermochemistry. 3. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(36) Gobbi, A.; Frenking, G. Y-Conjugated Compounds—The Equilibrium Geometries and Electronic-Structures of Guanidine, Guanidinium Cation, Urea, and 1,1-Diaminoethylene. *J. Am. Chem. Soc.* **1993**, *115*, 2362−2372.

(37) Koskinen, J. T. Experimental and computational studies on aminoguanidine free base, monocation and dication. Part III: Proton affinities of guanidine, aminoguanidine and glyoxal bis(amidinohydrazone). *Z. Naturforsch., B: Chem. Sci.* **1998**, *53*, 386−392.

(38) Maksic, Z. B.; Kovacevic, B. Absolute proton affinity of some polyguanides. *J. Org. Chem.* **2000**, *65*, 3303−3309.

(39) Julian, R. R.; Beauchamp, J. L. Site specific sequestering and stabilization of charge in peptides by supramolecular adduct formation with 18-crown-6 ether by way of electrospray ionization. *Int. J. Mass Spectrom.* **2001**, *210*, 613−623.

(40) Silva, C. O.; da Silva, E. C.; Nascimento, M. A. C. Ab initio calculation of absolute pK_a values in aqueous solution. II. aliphatic, alcohols, thiols, and halogenated carboxylic acids. *J. Phys. Chem. A* **2000**, *104*, 2402−2409.

(41) Rak, J.; Skurski, P.; Simons, J.; Gutowski, M. Low-energy tautomers and conformers of neutral and protonated arginine. *J. Am. Chem. Soc.* **2001**, *123*, 11695−11707.

CT049849M