CHEMISTRY
A EUROPEAN JOURNAL

DOI: 10.1002/chem.200601741

How Important Is Molecular Rigidity for the Complex Stability of Artificial Host–Guest Systems? A Theoretical Study on Self-Assembly of Gas-Phase Arginine

Sebastian Schlund, Carsten Schmuck, and Bernd Engels*[a]

Abstract: Arginine forms much less stable dimers than 2-(guanidiniocarbonyl)-1*H*-pyrrole-5-carboxylate although the principal binding interactions are very similar. The reasons for this difference are addressed in this work by state-of-the-art ab initio computations. The investigation shows that the extraordinary high stability of the 2-(guanidiniocarbonyl)-1*H*-pyrrole-5-carboxylate dimer results to about 50% from the rigidity of its monomer. Within this study monomer and dimer conformers of arginine were calculated leading to new low lying structures which have not been reported before as well as new global minima are predicted. In these structures stacking interactions with the guanidinium moiety are especially important. For the monomer we predict the energy minimum to be the canonical form with the lowest lying zwitterionic structure being only 9 kJ mol^{-1} less stable. During the course of these calculations we found that DFT did not predict the structures and their relative energy correctly in comparison to perturbation theory (MP2) and some potential reasons for the failure of DFT in these cases are discussed. Vibrational frequencies of the various structures are presented and a suitable wavenumber region for an experimental determination of the global minimum of the arginine monomer is identified. The effect of molecular rigidity on the self-assembly is probed using a local minimum of the arginine monomer which does not possess any intramolecular stabilizing effects. Our results suggest that the deliberate control of the conformational flexibility is a powerful instrument to steer the complex affinity of artificial hosts.

Keywords: ab initio calculations · artificial receptors · conformation analysis · host–guest systems · self-assembly

Introduction

Among all naturally occurring amino acids arginine is in the very focus of interest for various reasons. Due to the high basicity of the guanidine group situated in the side chain, its protonated form plays an important role in protein chemistry allowing the formation of strong salt bridge interactions with carboxylates or phosphates.[1–2] Arginine is therefore present in many reactive centres of enzymes and plays also an important role in secondary and tertiary structure formation.[3–6]

Amino acids form stable zwitterions in aqueous solution whereas their canonical tautomers are strongly favoured in the gas phase. An exception could be again arginine since the strong proton affinity of the guanidine group could outweigh the energy necessary for charge separation (see Scheme 1). The question for the tautomeric form of the global minimum of arginine in the gas phase has therefore been discussed widely by theoretical studies.[7–10] An experimental study trying to shed light on this problem was performed by Chapo et al. employing infrared cavity ringdown laser absorption spectroscopy (IR-CRLAS). They identified two peaks at $\tilde{\nu}=1666$ and 1693 cm^{-1} which were assigned to carbonyl stretches of the carboxylic acid group present in the canonical form.[11] Since symmetric and asymmetric stretches of the carboxylate group in the zwitterionic argi-

Scheme 1. Tautomeric forms of neutral arginine. Left: canonical monomer, right: zwitterionic monomer.

[a] Dipl.-Chem. S. Schlund, Prof. Dr. C. Schmuck, Prof. Dr. B. Engels
Institute for Organic Chemistry, University of Wuerzburg
Am Hubland, 97074 Wuerzburg (Germany)
Fax: (+49) 931-888-5331
E-mail: bernd@chemie.uni-wuerzburg.de

Supporting information for this article is available on the WWW under http://www.chemeurj.org/ or from the author.

nine could not be found in the expected range, they concluded that in the gas phase the canonical form of arginine is mainly populated. However, Rak et al. doubted the interpretation of the measured spectrum on the basis of new theoretical findings which predicted the carbonyl bands to occur in a region that has not be recorded.[9] Nevertheless, on the basis of the computed relative energies also their study predicts that the canonical form should dominate in the gas phase by 7 kJ mol^{-1}.

Due to its ability to form strong salt bridges, arginine is also an ideal model system to study guanidinium–carboxylate interactions.[12–16] Moreover, the strong non-covalent binding interactions of the guanidinium moiety with anionic groups is the basic concept of a research field trying to mimic biological receptor systems in order to improve ligand–receptor interactions and to understand molecular recognition processes.[17–18]

A successful rational design of new artificial receptor systems requires a profound knowledge of all inter- and intramolecular interactions as well as the ability to distinguish between molecular inherent and solvent effects. On this account theoretical studies are often used to investigate the gas-phase properties of such systems which best reflect the molecular inherent effects. Moreover, theory also allows a differentiation between various interactions. In order to study the gas phase binding properties of guanidinium-based artificial receptor systems we started to research on the dimerization of arginine monomers. A similar approach has already been performed by Goddard III and co-workers who calculated bonding energies of arginine dimers and trimers with respect to the global minimum of arginine published by Rak and co-workers.[9,15]

However, our conformational search in conjunction with accurate electron correlation computations yielded yet unknown conformers of both the arginine monomer and dimer which exhibit completely new types of geometrical arrangements not reported before. Additionally, new global minimum structures could be identified. Therefore, the first aim of the present paper is to discuss the interactions which stabilize the new conformers and to give a brief outline why these structures could not be identified in the previous works.

The second goal of this paper is to study to what extent a stiffening of a system can enhance its ability to self-aggregate, for example, to form stable dimers. The question arises from a study of the binding properties of the efficient carboxylate receptors developed by Schmuck and co-workers.[19] The 2-(guanidiniocarbonyl)-1*H*-pyrroles comprise a guanidinium group such as arginine but they are capable to complex carboxylates even in polar media.[20,21] If a terminal carboxylate group is added, self-assembly can be observed ranging from discrete dimers, for example, for the 2-(guanidiniocarbonyl)-1*H*-pyrrole-5-carboxylate (see Figure 1), to oligomers.[22–24]

Within these dimer systems the non-covalent complexation includes a variety of effects such as ionic interactions, hydrogen bonding and cooperativity which all contribute to the stabilization energy.[25] However, recent theoretical studies of the dissociation processes of the dimer indicate that not only the additional hydrogen bonds compared with the parent guanidinium cation, but also the rigidity of the monomer is of utmost decisive importance for the stability of the dimer. Due to its strong rigidity the 2-(guanidiniocarbonyl)-1*H*-pyrrole-5-carboxylate monomer cannot be stabilized by intramolecular interactions between the charged terminal groups. As a result, the formation of assemblies leads to high dimerization energies. In contrast, arginine has a significant flexibility because of the large amount of rotatable bonds in the side chain so that the oppositely charged groups strongly interact already in the monomer. Arginine is expected to show drastically reduced dimerization energies in comparison with the artificial systems introduced by Schmuck et al. This difference can not only be explained simply by the difference in the binding motifs (e.g. number of H-bonds or acidity). To estimate how an artificial stiffening of arginine would enhance its complexation ability we computed the dimerization energy of a linear conformer of arginine. These calculations indicate that the dimerization energy of an artificial rigid arginine is about twice as large as for the flexible arginine. Therefore, at least for self-complementary systems, the stiffening of a molecule seems to be a suitable instrument to steer its complexation properties. A comparison to the completely rigid 2-(guanidiniocarbonyl)-1*H*-pyrrole-5-carboxylate shows that its inherent rigidity accounts to about 50 % to the dimer stability.

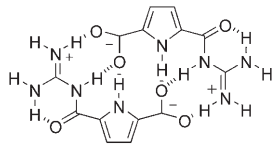

Figure 1. Zwitterionic 2-(guanidiniocarbonyl)-1*H*-pyrrole-5-carboxylate forms discrete dimers in polar media ($K = 170$ M^{-1} in water).

Computational Details

An efficient although exhaustive search of the conformational space as well as the choice of an appropriate force field is crucial for the determination of low-energy conformers of canonical and zwitterionic arginine monomers and their non-covalent assemblies. An extensive validation revealed that for the arginine monomer and dimer the Mixed Monte-Carlo Multiple Minimum/Low Mode (MCMM/LowMode) approach and the Systematic Unbound Multiple Minimum (SUMM) algorithm as implemented in the MacroModel 8.0 program package are the most effective tools for scanning the conformational space, whereas the OPLS-AA and the MMFF94 force fields gave the best structures and energy order of the conformers.[26–30] The OPLS-AA force field yielded excellent results for the zwitterionic species whereas the MMFF94 was proven to give better structures for the canonical conformers. The quality of the algorithms and of the force fields were estimated by their ability to reproduce the low lying conformers already described in the literature. All conformational searches took between 2000 to 5000 steps and were repeated from different starting structures. For the canonical monomer the number of generated conformers (MCMM/LowMode//MMFF94) within an energy range of 50 kJ mol^{-1} exceeded 400 conformers so that the dif-

ferent structures were clustered by the XCluster program based on atomic RMSD (root-mean-square distance) differences of all atoms.[31]

In the next step the most promising structures were optimised (see also Supporting Information) either on the B3LYP/3-21G or on the RI-BLYP/SV(P) level of theory employing the Jaguar 4.2 program or the TURBOMOLE program package, respectively.[28,32–33] In these calculations a great number of conformers had to be taken into account since the energy ordering resulting from the force field based conformational searches and from these DFT computations differ largely. To keep the efforts manageable, the conformers treated in this step were manually selected. From each conformational search about 30 conformers were chosen for the higher level calculations.

The lowest lying monomer structures obtained from these computations were then fully optimised on a RI-BLYP level and afterwards on a RI-MP2 level of theory. For both optimizations the TZVPP basis sets were employed. The basis set for the oxygen centres was augmented by one s and one p function with low exponents ($\zeta = 0.068$) in order to describe the diffuse shape of electrons of the carboxylate atoms in the zwitterionic conformers properly.[34] For the auxiliary basis sets the exponent was doubled ($\zeta = 0.136$). On the optimised monomer conformers CCSD(T) calculations using the MOLPRO program package were performed employing a cc-pVDZ basis and an aug-cc-pVDZ basis for the oxygen atoms.[35–36] Larger basis sets were not feasible due to hardware and software restrictions. Since MP2 calculations indicate that the double zeta basis set is insufficient, we performed an extrapolation for the CCSD(T) values based on the differences between MP2/cc-pVDZ and augmented MP2/TZVPP calculations.

To determine the lowest lying structures of the dimer system the same strategy as for the monomer was used. Due to software and hardware restriction the RI-MP2 optimizations were only feasible with a TZVP basis.[32] The final electronic energies were then calculated by single-point calculations on RI-MP2 level employing the augmented TZVPP basis mentioned above. CCSD(T) computations with reasonable basis sets were not possible. Dissociation energies of the dimer species were calculated including the counterpoise correction according to Boys and Bernardi.[37]

All optimised structures were characterized by harmonic frequency analysis and thermodynamic corrections which were obtained with TURBOMOLE on a RI-MP2/TZVP level. The free energies were calculated with a scaling factor for the wavenumbers of 0.937.[38] The IR spectra were simulated employing a Gauss fit for the line spectra. The contributions from all excitations were added according to the following equation:

$$\varepsilon(E) = \frac{\sum_i I_i \cdot \exp\left\{-\left[-\frac{(E-\Delta E_i)}{2\sigma}\right]^2\right\}}{\sqrt{2\pi\sigma}}$$

with σ as the full width of at half maximum (here $\sigma = 0.001$) and ΔE_i (in eV) and I_i as the calculated excitation energies and intensities, respectively.[39]

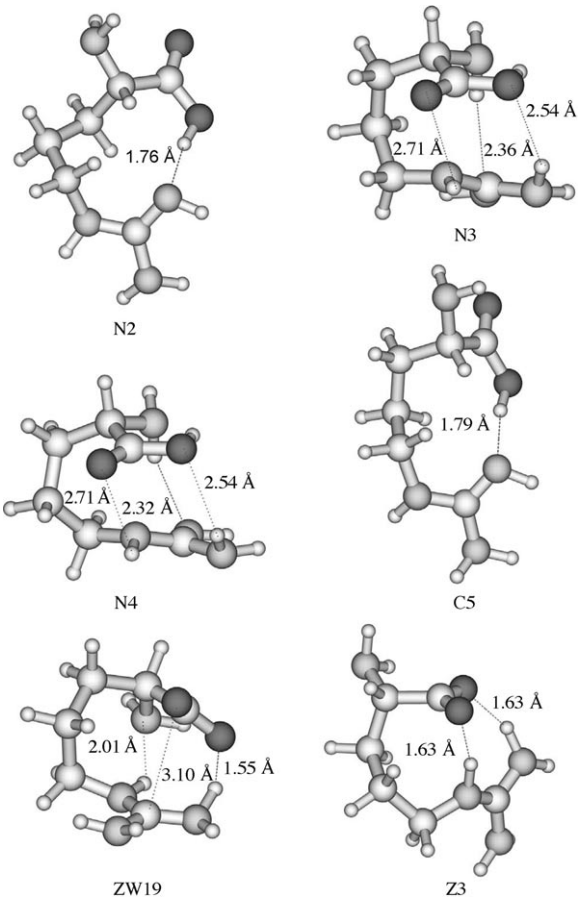

Figure 2. MP2/TZVPP optimised geometries for canonical (C5, N2-N4) and zwitterionic conformers (Z3, ZW19) of neutral arginine monomer.

Table 1. Relative electronic energies in kJ mol^{-1} of RI-MP2/TZVPP optimised geometries in dependence of basis set size and level of theory.

Method	N2	N3	N4	C5	ZW19	Z3
BLYP/TZVPP[a]	0.0	+14.8	+15.0	+1.2	+9.8	+9.8
B3LYP/TZVPP[b]	0.0	+14.1	+14.5	+2.1	+13.0	+12.3
MP2/cc-VDZ	0.0	−0.3	−0.3	+2.5	−11.2	+2.7
RI-MP2/TZVPP	0.0	+0.1	+1.6	+2.8	+1.2	+10.2
CCSD(T)/cc-VDZ	0.0	+0.8	+0.3	+1.8	−5.8	+7.6
CCSD(T)/extrapol.[c]	0.0	+1.2	+2.2	+2.1	+6.6	+15.1

[a] Relative DFT energies on BLYP/TZVPP optimised geometries. [b] Relative DFT energies on B3-LYP/TZVPP optimised geometries. [c] ΔE(CCSD(T)/extrapol.) = ΔE(CCSD(T)/cc-VDZ) + {ΔE(MP2/TZVPP) - ΔE(MP2/cc-VDZ).

Results and Discussion

Arginine monomers

Geometries and relative energies: The geometries of the lowest energy conformers of arginine are shown in Figure 2 containing new canonical and zwitterionic conformers which are lower in energy than the ones given by Rak et al. (C5, Z3). Table 1 gives the computed energies relative to conformer N2 representing the global minimum of our study. It lies about 2–3 kJ mol^{-1} below the conformer C5 representing the global minimum of the work of Rak et al. Additionally, our study reveals a new structure motif (N3, N4, ZW19) differing considerably from the already known geometrical arrangements. All structures show strong interactions between the guanidine and the carboxyl moiety. However, while the already known structure type (e.g. N2, C5, Z3) is solely stabilized by directed hydrogen bonds, the new structure motif shows strong stacking interactions between the terminal groups leading to a parallel orientation of the carboxyl and

the guanidine moiety. It is noteworthy that such stacked arrangements appear in canonical as well as zwitterionic conformers. To the best of our knowledge such stacked structures were never reported for the arginine monomer.

One reason why these structures were overlooked before is shown by Table 1. DFT strongly underestimates the stabilities of the stacked structures so that studies performing geometry optimizations only by DFT methods will neglect these structures. Using the BLYP/TZVPP approach the neutral stacked structures N3 and N4 are predicted to be about 14 kJ mol^{-1} higher in energy than the neutral structures C5 and N2, which only contain directed hydrogen bonds as stabilizing feature. In contrast RI-MP2/TZVPP predicts that all neutral structures depicted in Figure 2 possess very similar energies which is also confirmed by CCSD(T) computations. In both MP2 and CCSD(T) approaches the energy differences between the neutral structures are smaller than the expected error bars (2–4 kJ mol^{-1}). For the zwitterionic structures an analogous picture is found. Here, DFT predicts both isomers to be similar in energy whereas MP2 and CCSD(T)/extrapolated calculations compute the new stacked structure ZW19 to be lower in energy by about 9 kJ mol^{-1}.

An analysis of the situation in ZW19 is provided by Figure 3. Geometry optimization with RI-BLYP/TZVPP results in the local minimum structure given on the left-hand side of Figure 3 which lies $\approx$10 kJ mol^{-1} above the global minimum N2 on this level of theory. If this structure is used as a starting point for a MP2 geometry optimization (RI-MP2/TZVPP) the structure ZW19 shown on the right hand side of Figure 3 is obtained. RI-MP2/TZVPP predicts ZW19 to be 7 kJ mol^{-1} lower in energy than the DFT optimised structure and only about 1 kJ mol^{-1} less stable than the global minimum N2. A single point RI-BLYP/TZVPP+ calculation on the MP2 optimised structures predicts it to be 19 kJ mol^{-1} higher in energy than the BLYP optimised structure.

The reasons for the variations can be seen from an analysis of the ESP fit charges. Both approaches agree with that most of the positive charge is localized on the hydrogen atoms but they slightly differ in the charge delocalization between the heavier centres. Using MP2/TZVPP positive and negative charge is a bit more localized on the heavier centres in contrast to DFT which is in line with findings showing that DFT overestimates delocalizing effects.[40] As a consequence of the more localized charges for MP2, the Coulomb interaction between the carboxylate oxygen and the central carbon atom of the guanidinium moiety leads to a decreased oxygen-carbon distance (Figure 3). A complete parallel orientation of both groups may be impeded by repulsive interactions between the oxygen and the nitrogen centres. It is noteworthy that the strong hydrogen bond between one of the oxygen atoms and one of the NH2 groups of the guanidinium remains nearly unchanged. Comparing the two low lying zwitterionic structures Z3 and ZW19 (Figure 2) it becomes clear that the increased electrostatic attractions outweigh at least one strong hydrogen bond. Besides the differences in the localization, one can also expect that the inability of DFT to account for dispersion effects is a second reason for the discrepancies. That dispersion effects are very important for guanidinium cations was recently shown by Brady and co-workers.[41]

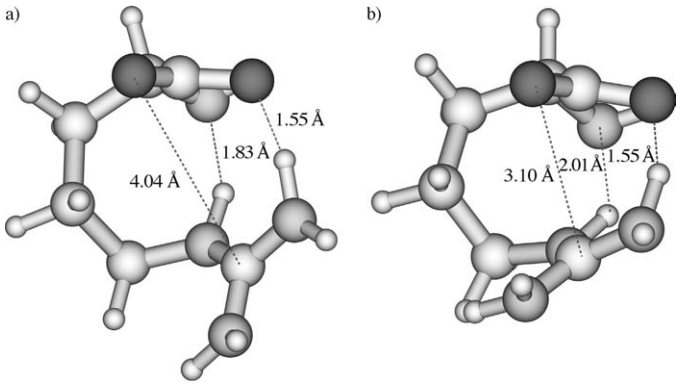

Figure 3. Geometries of the zwitterionic conformer ZW19 in dependence of the method. The a) BLYP/TZVPP optimised geometry and b) MP2/TZVPP geometry differ in energy of about 7.4 kJ mol^{-1} (MP2/TZVPP).

Despite the fact that the energies of all stacked structures are likewise underestimated by DFT the interactions stabilizing the neutral structures N3 and N4 seem to differ from those which occur in the zwitterionic conformer ZW19. This is indicated from our finding that for the canonical arginine DFT and MP2 predict quite similar structures. This points out to a smaller overall effect. Since strong electrostatic interactions are missing, we expect the dispersion effects to be the main reason for the differences.

In accordance to Saykally and co-workers, Rak et al. predicted a canonical conformer of arginine to represent the global minimum in the gas phase.[9,11] Using the MP2/6-31++G** approach the lowest lying zwitterionic structure was computed to lie only 7 kJ mol^{-1} higher but the energy difference increased to about 17 kJ mol^{-1} if CCSD was employed. Since the stacked zwitterionic structure ZW19 is found to lie considerably lower in energy than the already known zwitterionic structure Z3, the question arises if it becomes the global minimum. This is not the case as can be seen from Table 1. Using the MP2/TZVPP approach the structure ZW19 is found to lie about 1 kJ mol^{-1} above the lowest lying neutral structure N2 (Figure 2). CCSD(T) computations were performed to check the MP2 predictions. After the basis set extrapolation the zwitterionic structure ZW19 lies about 7 kJ mol^{-1} higher in energy than the global minimum N2. However, since CCSD(T) should be more influenced by basis set effects than MP2 the 7 kJ mol^{-1} represents a lower limit.[42]

Table 1 reveals that the computed energy differences between the various low lying isomers are so small that zero-point vibrational energy contributions and thermal effects are not negligible. The various contributions are summarized in Table 2 with respect to structure N2. Table 2 shows that the stacked structures are slightly destabilized by the enthalpy and entropy corrections. Due to the thermodynam-

Table 2. Thermodynamic corrections in kJ mol^{-1} for RI-MP2/TZVPP optimised geometries calculated on RI-MP2/TZVP level (T = 298.15 K).

	N2	N3	N4	C5	ZW19	Z3
ΔH_{corr}	0.0	+0.9	+0.4	−0.1	−3.0	−4.7
$-T\Delta S_{corr}$	0.0	+3.1	+2.4	+0.4	+5.1	+2.1
ΔG_{MP2}[a]	0.0	+4.1	+4.4	+3.1	+3.3	+7.6
$\Delta G_{CCSD(T)}$[b]	0.0	+5.2	+5.0	+2.4	+8.7	+12.5

[a] $\Delta G_{MP2} = \Delta E_{elec}$(RI-MP2/TZVPP) + $\Delta H^{corr} - T\Delta S^{corr}$. [b] $\Delta G_{CCSD(T)} = \Delta E_{elec}$(CCSD(T)/extrapol.) + $\Delta H^{corr} - T\Delta S^{corr}$.

ic corrections the energy difference between the global minimum (N2) and the lowest lying zwitterionic structure ZW19 increases to nearly 9 kJ mol^{-1}.

In comparison to previous works the present approach could identify new global minimum structures for both the canonical as well as the zwitterionic tautomers and finds that the expected energy gap between these structures is only about half as large as predicted by earlier calculations which employed less accurate methods.[9]

Calculated spectra: According to our study arginine possesses conformers with large structural differences already within an energy range of less than 10 kJ mol^{-1}. Taking into account remaining uncertainties resulting from the conformational search and from the notoriously difficult estimate of entropy effects only experiment can provide an unambiguous answer which structure type represents the global minimum. Due to the differences in the intramolecular interactions vibrational spectroscopy should at least be able to distinguish between the various types of conformers (stacked arrangements vs. directed hydrogen bonds) and tautomers (canonical vs zwitterionic form). This approach was utilized for the first time by Saykally and co-workers who concentrated on the region between 1500 and 1600 cm^{-1} in order to determine experimentally whether the canonical or the zwitterionic form represents the global minimum. However, as already pointed out by Rak et al. this energy range is not sufficient for a definite answer.

To estimate which interval of the spectra is best suited for an unambiguous identification the RI-MP2 method was used to compute the various IR spectra. Figure 4 shows the calculated, unscaled line spectra of the various low lying energy conformers of canonical and zwitterionic arginine with the superimpositions of Gauss fitted curves. Characteristic peaks were assigned by vibrational mode analysis. The resulting labelling is given if peaks could be related to more or less uncoupled vibrations.

All spectra can be divided into three major regions: a) the fingerprint region below 1500 cm^{-1} wavenumbers showing coupled scaffold and bending vibrations; b) the region between 1600 and 1900 cm^{-1} consisting mainly of C=O and C=N stretch vibrations; and c) the hydrogen stretch vibrations between 2600 and 4000 cm^{-1}. Chapo et al. as well as Rak et al. both concentrated on region b) in order to determine the tautomeric state of arginine. The spectra of the canonical conformers show mainly a C=O stretch band around 1860 cm^{-1} and a clear peak for one C=N stretch vibration in the guanidine part at ≈1710 cm^{-1}. Some other peaks occur representing COH bending or coupled vibrations which are quite similar for both forms of geometrical arrangements (linear H-bond vs stacking orientation). Regarding the spectra of the zwitterionic arginine the symmetric O=C=O stretch band is rather weak and coupled with other vibrations. The large C=N stretch peak is shifted to larger wavenumbers (1830–1860 cm^{-1}) in comparison to the canonical conformers and it is also coupled with NH stretch vibrations. However, although the assignments for the various conformers are different, the experimentally accessible properties such as frequencies and intensities are too similar for a definitive determination of the structure of the global minimum.

As expected region c) containing the N-H and O-H vibrations would allow an unambiguous identification which type of conformer (stacked arrangements vs directed hydrogen bonds) and tautomers (canonical vs zwitterionic form) predominates in the gas phase. The conformers C5 and N2 possess a very intense peak at about 3300 cm^{-1} assigned to the stretch vibration of the O-H group. The high intensities result from their involvement in the directed hydrogen bond to the guanidinium group. The spectra of N2 and C5 show a slightly different energy gap between the OH-stretch and the CH-stretch vibrations which may be used to differentiate between both conformers. In any case this slight difference could be used to determine if both conformers were present in gas phase.

The intensity of this vibration is drastically reduced when the arginine shows a stacked conformation (N3, N4) and also a small shift towards ≈3400 cm^{-1} is predicted. Therefore, this peak allows a differentiation between directed and stacked canonical conformers, whereas it is hardly possible to distinguish between N3 and N4.

For the zwitterionic conformers Z3 and ZW19 the N-H stretch vibrations of the zwitterionic hydrogen bond (NH···O) appear in this energy range. These vibrations all occur at wavenumbers well below 3300 cm^{-1} and should therefore be a characteristic evidence for the existence of zwitterionic conformers in the gas phase. A differentiation between stacked or directed types of conformers is also easily possible since the latter possesses two strong absorptions within a small energy range. In contrast, the stacked conformer ZW19 shows a strong peak at about 3400 cm^{-1}. It is assigned to the NH stretch vibration of the NH···N hydrogen bond between the guanidinium group and the α-amino nitrogen centre. Our calculations strongly suggest that the region around 3000 cm^{-1} can be used to experimentally differentiate between the various structures. Hence, this way allows an unambiguous determination of the global minimum of the arginine monomer.

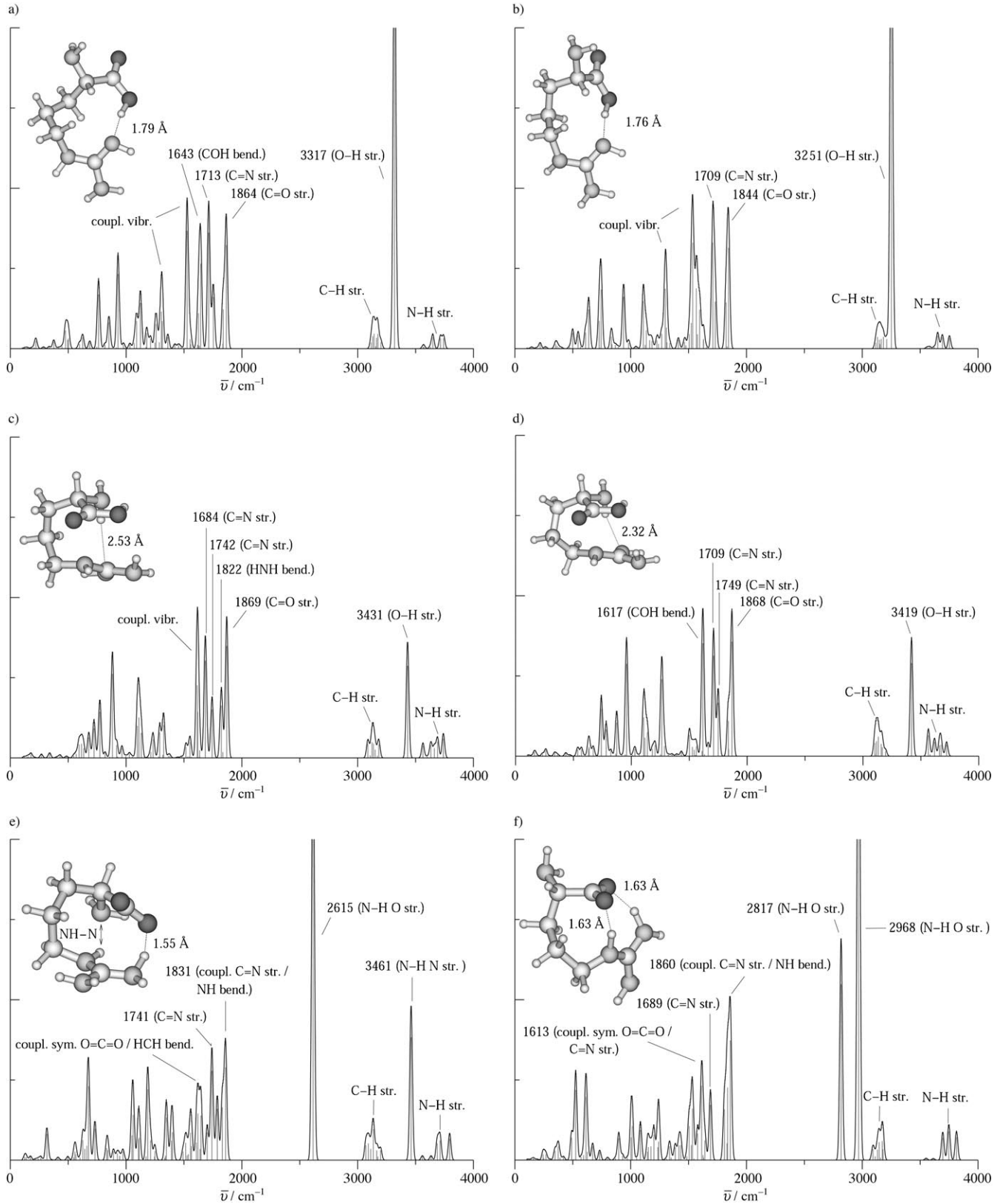

Figure 4. Gauss fitted curves of calculated vibration line spectra of neutral arginine conformers (RI-MP2/TZVP) a) N2, b) C5, c) N3, d) N4, e) ZW19, f) Z3. Some relevant unscaled vibration modes are given. The peaks of high intensity were cut for the sake of clarity (see text).

Arginine dimers

Geometries and energies: The conformational search for low lying conformers of the arginine dimer was performed with the same strategy as used for the monomer. All detected low lying structures consist of two zwitterionic species interacting through several charged H-bonds between the guanidinium and carboxylate moieties. Dimers consisting of canonical monomers are unfavourable since the zwitterionic forms are stabilized through the electric field of its counterpart. Hence, the formation of the strong bonded salt bridges outweighs the energy necessary to form the zwitterionic monomers. The deep minimum wells resulting from these strong salt bridges also explain the low number of conformers.

The three lowest lying conformers detected within our conformational search (MMFF1, MMFF2 and MMFF4), which were all predicted by the MMFF94 force field, are depicted in Figure 5. Figure 5 also contains the lowest conformer given by Goddard III and co-workers (DZ1) and the lowest structure predicted by the OPLS-AA force field (OPLS-AA1).[15] Table 3 summarizes the most important geometrical parameters. The geometrical differences between the MMFF structures and both other conformers (OPLS-AA1 and DZ1) are striking. In all structures the zwitterionic monomers form strong salt bridges between the guanidinium moiety and the carboxylate group of the counterpart through a network of directed hydrogen bonds. The striking differences result from the intramolecular interactions of the carboxylate and the guanidinium group of a given monomer. For the OPLS-AA1 and the DZ1 conformer these units also interact through one directed hydrogen bond necessitating a planar structure for the dimer. In contrast, in the MMFF structures the guanidinium and the carboxylate moiety of one monomer adopt a more parallel orientation leading to pocket-like structures.

Table 3. Selected atomic distances in zwitterionic arginine dimer conformers optimised on RI-MP2/TZVP level of theory. All values are given in Ångstrøm.

Dimer	Atoms	Distance [Å]
MMFF1	O(1)···HN(5), O(3)···HN(7)	1.79
	O(2)···HN(6), O(4)···HN(8)	1.67
	O(1)···HN(7), O(3)···HN(5)	2.12
	N(10)···HN(9), N(11)···HN(12)	1.92
	C(13)···C(14), C(15)···C(16)	3.22
MMFF2	O(1)···HN(5)	1.77
	O(2)···HN(6)	1.68
	O(1)···HN(7)	2.10
	N(10)···HN(9)	1.91
	C(13)···C(14)	3.26
	O(3)···HN(7)	1.86
	O(4)···HN(8)	1.63
	O(3)···HN(5)	1.97
	N(11)···HN(12)	1.86
	C(15)···C(16)	3.57
MMFF4	O(1)···HN(5), O(3)···HN(7)	1.82
	O(2)···HN(6), O(4)···HN(8)	1.66
	O(1)···HN(7), O(3)···HN(5)	1.98
	N(10)···HN(9), N(11)···HN(12)	1.86
	C(13)···C(14), C(15)···C(16)	3.56
OPLS-AA1	O(1)···HN(5), O(3)···HN(7)	1.71
	O(2)···HN(6), O(4)···HN(8)	1.67
	O(1)···HN(7), O(3)···HN(5)	1.80
DZ1	O(1)···HN(5), O(3)···HN(7)	1.69
	O(2)···HN(6), O(4)···HN(8)	1.71
	O(1)···HN(7), O(3)···HN(5)	1.77

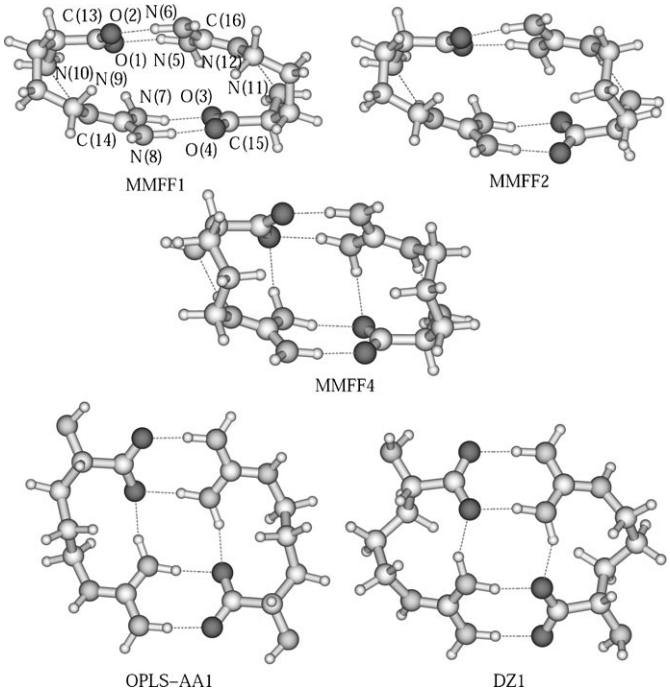

Figure 5. RI-MP2/TZVP optimised structures of zwitterionic arginine dimers.

Within Figure 5 the opening of the pocket is in the foreground (d(O(2)···N(8)) 4–5 Å). The bond lengths between O(1) and HN(7) (d = 1.98–2.12 Å) indicate moderate hydrogen-bond strengths. The structural arrangement points to interactions resembling those found in the ZW19 conformer. With dimerization energies of about 200 kJ mol^{-1} (MP2/TZVPP) these pocket-like structures are about 60 kJ mol^{-1} more stable than the planar structures DZ1 and OPLS-AA1 (see Table 4).

The stronger stabilization of the MMFF structures does not only result from the differences in the bonding network between the guanidinium and carboxylate moieties, but also

Table 4. Dimerization energies of zwitterionic arginine conformers calculated on a RI-MP2/TZVP//RI-MP2/TZVPP level of theory with the respective thermodynamic corrections (T = 298.15 K) determined on a RI-MP2/TZVP level of theory. All values are given in kJ mol^{-1}.

Dimer	MMFF1	MMFF2	MMFF4	OPLS-AA1	DZ1
ΔE	+224.7	+220.6	+218.6	+151.3	+156.5
BSSE	−16.3	−16.2	−15.8	−16.8	−16.9
$\Delta E_{(\mathrm{BSSE\ corr.})}$	+208.4	+204.4	+202.8	134.5	+139.6
ΔH_{corr}	−5.8	−5.3	−5.0	−3.8	−2.7
$-T\Delta S_{\mathrm{corr}}$	−67.2	−66.0	−65.1	−60.2	−60.1
ΔG_{corr}	+135.4	+133.1	+132.7	+70.5	+76.8

from an additional hydrogen bond between the α-amino nitrogen and the guanidinium moiety which cannot be formed in DZ1 or OPLS-AA1 due to their planarity.

The energy difference between MMFF1 and MMFF4 results from the interplay of the various bonding effects. In MMFF4 the intramolecular hydrogen bonds are shorter but the distances between the carboxylate and the guanidinium moieties (d(C(13)···C(14)) = 3.56 Å) are enlarged with respect to the MMFF1 dimer (d(C(13)···C(14)) = 3.22 Å). This indicates stronger hydrogen bonds but smaller electrostatic interactions between the negatively and positively charged terminal groups. The structural differences are caused by a flip of the alkyl backbone. MMFF2 represents a mixture between MMFF1 and MMFF4.

The interactions stabilizing the pocket structures seem to comprise electrostatic and dispersions contributions. This is indicated by the differences between the dimerization energies computed with DFT and with MP2 as shown in Table 5 and again underlines the importance of stacking effects as already discussed for the monomers. For the MMFF structures DFT always yields considerably smaller stabilization energies (16–28 kJ mol^{-1}) since it cannot account for the dispersion part. For the DZ1 and OPLS-AA1 DFT predicts slightly higher dimerization energies ($\Delta E \approx 7$ kJ mol^{-1}) since the strengths of the directed hydrogen bonds seem to be overestimated in comparison to MP2.

Table 5. Comparison between counterpoise corrected electronic dimerization energies calculated for optimised structures on DFT and MP2 level of theory (TZVPP basis). All energies are given in kJ mol^{-1}.

Dimer	MMFF1	MMFF2	MMFF4	OPLS-AA1	DZ1
ΔE_{corr}(B3LYP)	+180.7	+183.0	+187.1	+141.4	+147.2
ΔE_{corr} (MP2)	+208.4	+204.4	+202.8	+134.5	+139.6
$\Delta\Delta E$	−27.7	−21.4	−15.7	+6.9	+7.6

In conclusion, our calculations reveal new structures for the arginine dimer which are twice as stable relative to the monomer as the previously predicted structures. This shows that careful conformational searches are necessary since the lowest lying structures can be counter-intuitive even for such well-known species as arginine.

Importance of molecular rigidity for the stability of the dimer: Compared with the arginine dimer (see Figure 5) the high stability of the 2-(guanidiniocarbonyl)-1*H*-pyrrole-5-carboxylate dimer (Figure 6) can be traced back to an improved hydrogen-bonding network, the higher acidic strength of the NH atoms as well as the energy contents of the monomers. The energy contents of the arginine and the 2-(guanidiniocarbonyl)-1*H*-pyrrole-5-carboxylate monomer differ strongly since a stabilizing interaction between the oppositely charged ends can only take place in the flexible zwitterionic arginine. A comparison between arginine, an artificially stiffened arginine and the 2-(guanidiniocarbonyl)-1*H*-pyrrole-5-carboxylate allows an estimate of the importance of the various effects. As model system for the artificially stiffened arginine a conformer was chosen in which the methylene groups are arranged in an all-*trans* orientation (Figure 6). The dimerization energy of 410 kJ mol^{-1} (RI-MP2/TZVPP(aug)//B3LYP/6-311++G**) is about

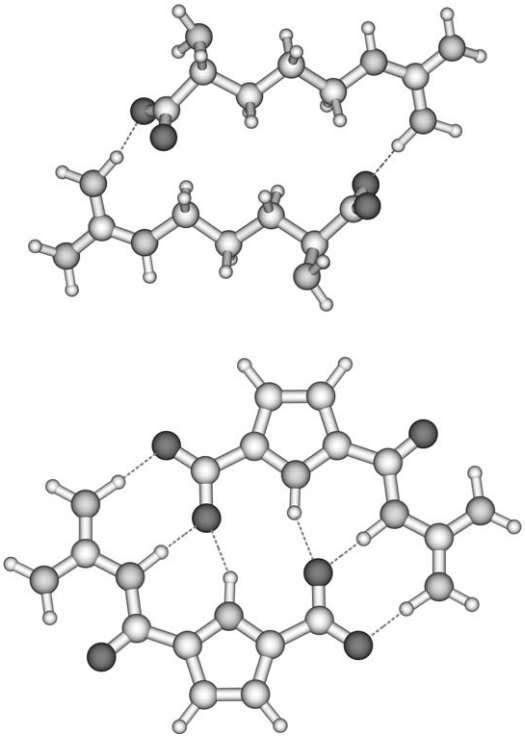

Figure 6. MP2/TZVPP calculated dimerization energies of an artificially linear arginine dimer (top) and 2-(guanidiniocarbonyl)-1*H*-pyrrole-5-carboxylate (bottom).

180 kJ mol^{-1} higher than for the regular arginine dimer (224 kJ mol^{-1}), although, due to geometrical constraints, only one hydrogen bond can be formed within one salt bridge.[43] The 2-(guanidiniocarbonyl)-1*H*-pyrrole-5-carboxylate dimer possesses a dimerization energy of 581 kJ mol^{-1} (RI-MP2/TZVPP(aug)) which is once more about 170 kJ mol^{-1} higher than for the stiffened arginine dimer system. This comparison shows that the distinct higher stability of the 2-(guanidiniocarbonyl)-1*H*-pyrrole-5-carboxylate dimer results to about 50% from the rigidity of the monomeric units. The rest is due to the improved H-bonding network and the increased acidity of the acyl guanidinium moiety.

Conclusion

The present study shows that the rigidity of a molecule significantly influences its self-assembling properties. The prevention of stabilizing intramolecular interactions within the monomers due to geometrical constraints strongly enhances the corresponding dimer stability. This was shown by comparing the self-assembly of arginine with that of 2-(guanidiniocarbonyl)-1*H*-pyrrole-5-carboxylate.

For arginine we were able to identify new global minimum structures for both the monomer and dimer in gas phase by means of a combination of force field based conformational searches and state-of-the-art electronic structure methods. We could observe new geometrical arrangements that have not been reported so far comprising a stacked orientation of the terminal groups. For the arginine monomer the explanation for this finding is twofold. In the case of the canonical structure a geometry optimization solely based on density functional theory would neglect dispersive interactions between the guanidine group and the carboxylic acid which can occur due to the high flexibility of the arginine. For the zwitterionic monomer the interaction of the charged ends can be traced back to coulomb interactions which are underestimated by DFT. It was also shown that the new zwitterionic conformer ZW19 is now energetically near the canonical global minimum implicating that a rigorous exclusion of a zwitterionic state in gas phase as it was proposed by several studies before is no longer tenable. An unambiguous proof can therefore only be given by experiment. For this purpose we computed the Gauss fitted vibrational spectra for the lowest monomer structures and we were able to show that a comparison of the hydrogen stretch vibrations between 2600 and 4000 cm^{-1} would make it possible to assign which tautomer (zwitterionic vs canonical) and which type of conformer (directed hydrogen bonds vs stacked orientation) dominates in gas phase.

The importance of rigidity was analyzed by calculating the dimerization energy of an artificially stiffened arginine dimer system and comparing it with the dimerization energy of the 2-(guanidiniocarbonyl)-1*H*-pyrrole-5-carboxylate dimer.

The analysis shows that the high binding affinity of the 2-(guanidiniocarbonyl)-1*H*-pyrrole-5-carboxylate results to about 50% from the rigidity of the monomers which cannot be stabilized by intramolecular interactions and are therefore high in energy. As a result dimerization is more favourable for rigid monomers stabilizing the terminal charges. This effect should therefore be strongly considered when optimizing the complexation ability of artificial self-complementary systems. A similar effect might be expected also for flexible non-complementary systems in which other stabilizing interactions can take place intramolecularly.

Acknowledgement

We thank C. Grassmann and R. Müller for valuable discussions. The work was financed by the DFG in the framework of the SFB 630.

[1] D. J. Barlow, J. M. Thornton, *J. Mol. Biol.* **1983**, *168*, 867–885.
[2] J. Singh, J. M. Thornton, M. Snary, S. F. Campbell, *FEBS Lett.* **1987**, *224*, 161–171.
[3] J. F. Riordan, R. D. McElvany, C. L. Borders, *Science* **1977**, *195*, 884–886.
[4] D. Gandini, L. Gogioso, M. Bolognesi, D. Bordo, *Protein J.* **1996**, *24*, 439–449.
[5] S. Sundaralingam, Y. C. Sekhadrudu, N. Yathindra, V. Ravichadran, *Protein J.* **1987**, *2*, 64–71.
[6] P. C. Lyu, P. J. Gans, N. R. Kallenbach, *J. Mol. Biol.* **1992**, *223*, 343–350.
[7] Z. B. Maksić, B. Kovačević, *J. Chem. Soc. Perkin Trans. 2* **1999**, 2623–2629.
[8] P. Skurski, M. Gutowski, R. Barrios, J. Simons, *Chem. Phys. Lett.* **2001**, *337*, 143–150.
[9] J. Rak, P. Skurski, J. Simons, M. Gutowski, *J. Am. Chem. Soc.* **2001**, *123*, 11659–11707.
[10] R. R. Julian, M. F. Jarrold, *J. Phys. Chem. A* **2004**, *108*, 10861–10864.
[11] C. J. Chapo, J. B. Paul, R. A. Provencal, K. Roth, R. J. Saykally, *J. Am. Chem. Soc.* **1998**, *120*, 12956–12957.
[12] A. Melo, M. J. Ramos, W. B. Floriano, J. A. N. F. Gomes, J. F. R. Leao, A. L. Magalhaes, B. Maigret, M. C. Nascimento, N. Reuter, *J. Mol. Struct.* (THEOCHEM) **1999**, *463*, 81–90.
[13] R. R. Julian, R. Hoyss, J. L. Beauchamp, *J. Am. Chem. Soc.* **2001**, *123*, 3577–3583.
[14] R. R. Julian, J. L. Beauchamp, *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 616–624.
[15] R. R. Julian, J. L. Beauchamp, W. A. Goddard III, *J. Phys. Chem. A* **2002**, *106*, 32–34.
[16] G. R. Kinsel, Q. Zhao, J. Narayanasamy, F. Yassin, H. V. R. Dias, B. Niesner, K. Prater, C. St. Marie, L. Ly, D. S. Marynick, *J. Phys. Chem. A* **2004**, *108*, 3153–3161.
[17] For recent reviews on anion binding by artificial guanidinium receptors see: a) K. A. Schug, W. Lindner, *Chem. Rev.* **2005**, *105*, 67–113; b) M. D. Best, S. L. Tobey, E. V. Anslyn, *Coord. Chem. Rev.* **2003**, *240*, 3–15; c) P. A. Gale, *Coord. Chem. Rev.* **2003**, *240*, 191–221; d) R. J. Fitzmaurice, G. M. Kyne, D. Douheret, J. D. Kilburn, *J. Chem. Soc. Perkin Trans. 1* **2002**, 841–864; e) T. S. Snowden, E. V. Anslyn, *Curr. Opin. Chem. Biol.* **1999**, *3*, 740–746; f) P. D. Beer, P. Schmitt, *Curr. Opin. Chem. Biol.* **1997**, *1*, 475–482; g) A. Bianchi, K. Bowman-James, E. Garcia-España, *Supramolecular Chemistry of Anions*, Wiley-VCH, New York, **1997**; h) F. P. Schmidtchen, M. Berger, *Chem. Rev.* **1997**, *97*, 1609–1646; i) C. Seel, A. Galán, J. de Mendoza, *Top. Curr. Chem.* **1995**, *175*, 101–132; j) *Functional Synthetic Receptors* (Eds.: T. Schrader, A. Hamilton), Wiley-VCH, Weinheim, **2005**.
[18] B. Kirchner, M. Reiher, *J. Am. Chem. Soc.* **2005**, *127*, 8748–8756.
[19] C. Schmuck, *Coord. Chem. Rev.* **2006**, *205*, 3053–3067.
[20] C. Schmuck, *Chem. Commun.* **1999**, 843–844.
[21] C. Schmuck, *Chem. Eur. J.* **2000**, *6*, 709–718.
[22] C. Schmuck, *Eur. J. Org. Chem.* **1999**, 2397–2403.
[23] C. Schmuck, W. Wienand, *J. Am. Chem. Soc.* **2003**, *125*, 452–459.
[24] C. Schmuck, Th. Rehm, K. Klein, F. Gröhn, *Angew. Chem.* **2007**, *119*, 1723–1727; *Angew. Chem. Int. Ed.* **2007**, *46*, 1693–1697.
[25] S. Schlund, C. Schmuck, B. Engels, *J. Am. Chem. Soc.* **2005**, *127*, 11115–11124.
[26] I. Kolossvary, W. C. Guida, *J. Comput. Chem.* **1999**, *20*, 1671–1684.
[27] J. M. Goodman, W. C. Still, *J. Comput. Chem.* **1991**, *12*, 1110–1117.
[28] Schrödinger, Inc., 101 SW Main Street, Suite 1300, Portland, OR (USA), 97204.
[29] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, *J. Am. Chem. Soc.* **1996**, *118*, 11225–11235.
[30] T. A. Halgren, *J. Comput. Chem.* **1996**, *17*, 490–512; T. A. Halgren, *J. Comput. Chem.* **1996**, *17*, 520–552; T. A. Halgren, *J. Comput. Chem.* **1996**, *17*, 553–586; T. A. Halgren, *J. Comput. Chem.* **1996**, *17*, 587–615; T. A. Halgren, *J. Comput. Chem.* **1996**, *17*, 616–641.
[31] P. S. Shenkin, D. Q. McDonald, *J. Comput. Chem.* **1994**, *15*, 899–916.
[32] a) TURBOMOLE V5.6 and V5.8, R. Ahlrichs, M. Bär, H.-P. Baron, R. Bauernschmitt, S. Böcker, M. Ehrig, K. Eichkorn, S. Elliott, F. Furche, F. Haase, M. Häser, H. Horn, C. Huber, U. Huniar, M. Kattannek, C. Kölmel, M. Kollwitz, K. May, C. Ochsenfeld, H. Whm, A. Schäfer, U. Schneider, O. Treutler, M. von Arnim, F. Weigend, P. Weis, H. Weiss, Universität Karlsruhe (Germany), **2003**; see also

http://www.turbemole.com; b) R. Ahlrichs, M. Bär, M. Häser, H. Holt, C. Kölmel, *Chem. Phys. Lett.* **1989**, *162*, 165–169.
[33] a) A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098–3100; b) C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785–789; c) A. D. Becke, *J. Chem. Phys.* **1997**, *107*, 8554–8560; d) J. S. Binkley, J. A. Pople, *J. Am. Chem. Soc.* **1980**, *102*, 939–947.
[34] R. Poirier, R. Kari, I. G. Csizmadia, *Handbook of Gaussian Basis Sets (Physical sciences data; 24)*, Elsevier, Amsterdam, **1985**.
[35] MOLPRO, version 2006.1, A package of ab initio programs, H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, M. Schütz, P. Celani, T. Korona, G. Rauhut, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, C. Hampel and G. Hetzer, A. W. Lloyd, S. J. McNicholas, W. Meyer and M. E. Mura, A. Nicklass, P. Palmieri, R. Pitzer, U. Schumann, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, **2004**.
[36] a) T. H. Dunning Jr., *J. Chem. Phys.* **1989**, *90*, 1007–1023; b) A. K. Wilson, T. van Mourik, T. H. Dunning Jr., *J. Mol. Struct.* **1996**, *388*, 339–349; c) R. A. Kendall, T. H. Dunning Jr., R. J. Harrison, *J. Chem. Phys.* **1992**, *96*, 6796–6806.
[37] S. F. Boys, F. Bernardi, *Mol. Phys.* **1970**, *19*, 553–566.
[38] A. P. Scott, L. Radom, *J. Phys. Chem.* **1996**, *100*, 16502–16513.
[39] S. Grimme, *Rev. Comput. Chem.* **2004**, *20*, 153–218.
[40] R. A. King, T. D. Crawford, J. F. Stanton, H. F. Schaefer III, *J. Am. Chem. Soc.* **1999**, *121*, 10788–10793.
[41] P. E. Mason, G. W. Neilson, J. E. Enderby, M.-L. Saboungi, C. E. Dempsey, A. D. MacKerell Jr., J. W. Brady, *J. Am. Chem. Soc.* **2004**, *126*, 11462–11470.
[42] T. Helgekar, P. Jørgensen, J. Olsen, *Molecular Electronic-Structure Theory*, Wiley, Chichester, **2000**.
[43] The linear structure of an arginine monomer is a stationary point on the potential energy surface when performing a geometry optimization on a RI-MP2/TZVPP(aug) level of theory. For the dimer system the hydrogen atoms involved in the H-bond had to be kept fix during the geometry optimization (B3LYP/6-3-11++G**) since the strong charge interaction mediated through a single hydrogen bond leads to a proton transfer from the guanidinium group towards the carboxylate oxygen. The geometry optimization for the 2-(guanidiniocarbonyl)-1*H*-pyrrole-5-carboxylate dimer was performed on a RI-MP2/TZVPP level of theory with augmented basis functions for the carboxylate oxygens. For the optimization of the monomeric 2-(guanidiniocarbonyl)-1*H*-pyrrole-5-carboxylate only a standard TZVP basis could be employed since an enlargement of the basis set causes a proton transfer from the pyrrole ring to the carboxylic group (see Figure 9) in gas phase.

Received: December 4, 2006
Revised: March 24, 2007
Published online: May 22, 2007

