Published in final edited form as:

J Phys Chem B. 2008 December 25; 112(51): 16995–17002.

Influence of Sequential Guanidinium Methylation on the Energetics of the Guanidinium...Guanine Dimer and Guanidinium...Guanine...Cytosine Trimer: Implications For the Control of Protein...DNA Interactions By Arginine Methyltransferases

Jason Shearer*

Department of Chemistry/216, University of Nevada, Reno, NV 89557-0216

Abstract

Arginine methylation is a post-translational protein modification that is catalyzed by proteins known as arginine methyl transferases (RMTs). Recently, arginine methylation was postulated as an important modification in modulating biomolecular interactions. RMTs largely target nuclear proteins, so it is highly likely that they aid in modulating protein...DNA interactions. In this study we probe the influence that sequential guanidinium methylation has on the energetics of the guanidinium...guanine and guanidinium...guanine...cytosine complex using *ab initio* and doublehybrid DFT methods. Structures of guanidinium...guanine complexes derived at the MP2/6-31++G** level of theory show that mono-methylated, symmetrically dimethylated, and unsymmetrical dimethylated guanidiniums are all capable of forming guanidinium...guanine complexes. However, when cytosine is involved in a base pair to guanine only the mono-methylated and symmetrically dimethylated guanidinium groups are capable of forming hydrogen-bond complexes with guanine. At the B2-PLYP/6-311++G** level of theory we found that methylation of the guanidinium group stabilizes the formation of the guanidinium...guanine complex relative to the unmethylated guanidinium...guanine complex by ~ 2.5 kcal mol⁻¹. The biological implication of these findings are discussed.

1. Introduction

Arginine methyl transferases (RMTs) are enzymes that catalyze the methylation of the ω -nitrogens of arginine (Arg) residues.¹⁻⁴ RMTs can be broken down into two different groups, type I and type II RMTs, which are differentiated by the final methylated Arg product that the RMT affords.¹⁻⁴ Both type I and type II RMTs are capable of producing mono-methylated Arg. The difference between the two types of RMTs are observed in the dimethylated final products. Type I RMTs produce unsymmetrical dimethylated Arg, where only one of the ω -nitrogens are mono-methylated (Scheme 1). Many RMTs target proteins that are found in the nuclear regions of cells.⁴⁻¹¹ It is therefore not surprising that a large percentage of the proteins targeted by RMTs are capable of binding nucleic-acids.⁴⁻²⁰

*corresponding author shearer@unr.edu.

Supporting Information. Contains tables of Cartesian coordinates of all structures used in this study, Born-Haber cycles for the four EGs investigated in this study, and a table of dimer interaction energies for GO structures obtained at MP2/6-31++G** vs. B2-PLYP/6-311++G** levels of theory.

Arg residues are intimately involved in the interactions between many nucleic-acid binding proteins and nucleic-acid.²¹⁻²⁵ Among the different stabilizing interactions observed between the Arg residues of nucleic-acid binding proteins and nucleic-acids are the formation of hydrogen bonds between the Arg ω -NH₂⁺ protons and a guanine nucleic-acid base.^{21,24-28} In this case the Arg residue will be oriented within the major-groove of the DNA-helix, and the guanine...Arg interaction will be mostly stabilized through the formation of two key hydrogen bonds: those between the Arg ω -NH₂ groups and the guanine carbonyl oxygen and imine nitrogen groups (Figure 1). Arg methylation will undoubtedly influence the energetics of this key intermolecular interaction.

On the face of it, it seems intuitively obvious that the systematic methylation of Arg would decrease the Arg...guan dimer stabilization energy;²⁻⁴ the electron donating methyl groups would create a less acidic N-H proton, thus decrease the strength of the resulting hydrogen bonds. However, other factors could be influenced by Arg methylation. The increase in steric bulk about guanidinium moiety could significantly alter the hydrogen bonding geometry found in the dimer structure. In addition, other interactions (such as solvation) could influence the overall stability of the (methylated)Arg...guanine dimer. This study is aimed at probing these properties using *ab initio* and double-hybrid density functional theory methods. As best as we can determine, this represents the first such computational study investigating the influence of Arg methylation on the guanidinium group's ability to form a hydrogen bond dimer with guanine. It will therefore provide benchmark data for rationalizing experimental results of Arg-methylated DNA binding proteins.

2. Computational Methods, Basis Set, and Level of Theory

2.1 General Methods

To create a model small enough to be efficiently tackled by high level computational methods yet still contain the requisite functional groups to adequately simulate the Arg...guanine interaction we utilized N-methylguanine (NMG) as a model for guanine, N-methylcytosine (NMC) as a model for cytosine, and sequentially N-methylated N-ethylguanidinium (EG; Chart 1) as models for (methylated)arginine. All calculations were performed using the computational package ORCA 2.6.35.²⁹ Geometry optimized (GO) structures were obtained using the 6-31++G** basis-set with the following thresholds (in au): root-mean-square (RMS) and maximum forces of 0.0003 and 0.0001, respectively; RMS and maximum gradients of 0.002 and 0.001, respectively. Second-order Møller-Plesset theory (MP2) was used to account for electron correlation. In all cases vibrational analyses were undertaken on all GO structures to verify they correspond to a stable ground state geometry.

Interaction energies were subsequently calculated as the difference in total energies between dimers/trimers and monomers from single-point calculations. The Boys and Bernardi counterpoise procedure was employed to account for basis set superposition errors.³⁰⁻³² All reported interaction energies are counterpoise corrected. The dimer/trimer interaction energies were calculated using the B2-PLYP³³⁻³⁴/6-311++G**^{35,36} level of theory, as this gave the best trade off between accuracy and efficiency in the R_1N_g dimer test calculations (*vide infra*). With the exception of the (methylated)EGs, single point calculations did not employ a vibrational analysis. Solvation effects were accounted for using the conductor-like screening model (COSMO) with parameters recommended for water ($\omega = 79.8$ and $r = 1.33$ Å).³⁷

Relative (methylated)EG acidities were calculated using the methods of Chen and MacKerell.³⁸ Briefly, Born-Haber cycles for the deprotonation of the methylated-EGs and EG were constructed as depicted in Scheme 2. Subsequently, the ΔG_{gas} and ΔG_{sol} values were calculated at the B2-PLYP/6-311++G** level of theory. The pK_as were then be calculated from:

$$\text{pK}_a(\text{EG}) = \left[\Delta G_{\text{sol}}(\text{EG}) + \Delta G_{\text{sol}}(\text{H}^+) - \Delta G_{\text{sol}}(\text{EG}-\text{H}^+) + \Delta G_{\text{gas}}(\text{EG}) \right] / (2.3RT) \quad (1)$$

where R and T have their typical meanings. We report the ΔpK_a as differences between the methylated-EGs and EG at 300 K.

2.2 Choice of Basis Set and Level of Theory

Using the unmethylated R_1N_g dimer as a simple test system (*vide infra*, Figure 3) we sought to determine an appropriate basis set and level of theory for this study. First we surveyed a variety of Hartree Fock theories (HF), density functional theory (DFT), and hybrid-DFT levels of theory. While investigating the appropriate level of theory a moderately sized basis set (6-31++G**) ^{36,39,40} that contains both diffuse and polarized functions was employed. Standard HF theory provided for a dimer interaction energy of $-3.62 \text{ kcal mol}^{-1}$, which is the weakest interaction energy of all of the other levels of theory investigated by $\sim 4 \text{ kcal mol}^{-1}$. Inclusion of electron correlation using MP2 theory yielded a significantly stronger dimer interaction energy of $-8.04 \text{ kcal mol}^{-1}$. Use of coupled-cluster techniques (CCSD(T)) ⁴¹ yielded a similar dimer interaction energy of $-8.21 \text{ kcal mol}^{-1}$, although care must be taken with this value considering the relatively small size of the basis set. The MP2 calculations can be compared with DFT (BP86 functional) ⁴²⁻⁴⁶ and hybrid-DFT (B3LYP) ^{47,48} levels of theory, with afforded interaction energies of -8.34 and $-7.60 \text{ kcal mol}^{-1}$, respectively. Therefore, it appears that DFT and hybrid-DFT methods are capable of approaching the energies produced by MP2 theory for this system.

In addition to these DFT and hybrid DFT methods, we also examined Grimme's B2-PLYP double hybrid DFT method. ^{33,34} This functional combines the B88 exchange functional ⁴⁴ with the LYP correlation functional, ⁴⁸ and includes a semi-empirical MP2-type perturbation. ³⁴ The B2-PLYP functional has been shown to accurately reproduce the thermodynamic properties of several systems. ^{33,34,49} In the case of predicting hydrogen bond energies the B2-PLYP functional has an accuracy that approaches experimental values, and in many cases exceeds the accuracy of MP2 theory. ^{33,49} Application of the B2-PLYP/6-31++G** level to the R_1N_g dimer yielded an interaction energy of $-8.16 \text{ kcal mol}^{-1}$, which is within $0.08 \text{ kcal mol}^{-1}$ of that predicted by both MP2 and CCSD(T) theories. Considering the computational efficiency and performance of the B2-PLYP method compared to the MP2 and CCSD(T) technique, and the previously demonstrated utility in examining weak biological interactions, ^{50,51} we chose to utilize this double-hybrid functional throughout this study.

As noted by Grimme, in order to obtain an accurate description of the thermodynamic properties of a given system with the B2-PLYP functional, it is important to utilize relatively large basis set, preferably of at least triple- ζ quality. ³³ We therefore examined a number of basis sets from the groups of Pople and Dunning. The Pople basis sets examined included the 6-31G*, 6-31++G**, and the 6-311++G** basis sets. ³⁵⁻⁴⁰ The Dunning basis sets surveyed included the correlation consistent polarized triple- ζ basis set with diffuse functions (Aug-cc-pVTZ) and corresponding quadruple- ζ (Aug-cc-pVQZ) basis set. ^{52,53} Not surprisingly the 6-31G* gave the weakest interaction energy of $-3.62 \text{ kcal mol}^{-1}$. Inclusion of diffuse functions and polarization functions on the H atoms (6-31++G**) produced a dramatic increase in the stabilization energy of the dimer, increasing it to $-8.16 \text{ kcal mol}^{-1}$. A further expansion of the basis set and inclusion of an extra set of diffuse functions (6-311++G**) increased the stabilization energy to $-9.60 \text{ kcal mol}^{-1}$. Switching to the Dunning triple- ζ Aug-cc-pTQZ basis set resulted in a modest decrease in the dimer stabilization energy ($-9.11 \text{ kcal mol}^{-1}$) compared with the 6-311++G** basis set with an increase in the computational effort. Likewise, increasing the size of the Dunning basis set to the quadruple- ζ Aug-cc-pVQZ basis set resulted

in only a modest change in the dimer interaction energy ($-9.72 \text{ kcal mol}^{-1}$) over the 6-311++G** basis set at a large computational expense. Considering the extremely large increase in computational effort required for use of the Aug-cc-pVQZ for the small energy difference between it and the smaller basis sets, we chose to utilize the 6-311++G** basis set for all subsequent calculations.

3. Results and Discussion

3.1 N-ethyl guanidinium structures and relative pK_a values of the methylated and unmethylated N-ethyl guanidiniums

Geometry optimized (GO) structures for protonated N-ethylguanidinium (EG) and the N-methylated-EGs depicted in Chart 2 along with their deprotonated conjugate bases were calculated at the MP2/6-31++G** level of theory. As expected EG displays a planar guanidinium moiety (Figure 2), with no twisting in the torsion angle ϕ ($\phi = 0^\circ$; Scheme 3).⁵⁴ In contrast, the methylated-EGs all show a significant twist of the $\text{H}_2\text{N-CN}_2$ planes, with ϕ ranging between 18.2° for symmetrical dimethyl-EG to 34.1° for unsymmetrical dimethyl-EG. The reason for this increase in ϕ upon N-methylation has to do with the steric crowding of the N-methyl groups and the ethyl and proton groups coming off the “ ϵ ” nitrogen. This causes the MeNR group to rotate out of the CN_2 plane. In order for all the N- π orbitals making up the bonding e-type MOs of the guanidinium to achieve optimal orbital overlap the other NR_2 groups must also rotate out of the CN_3 plane. Thus, the steric induced twisting of the Me-N-R moiety is transferred to all of the other nitrogens of the guanidinium. As will be demonstrated, this has a significant influence on the dimer and trimer geometries presented below.

To estimate the relative pK_a s of EG and the methylated-EG groups, the corresponding ΔG_{gas} and ΔG_{sol} values outlined in Scheme 2 were calculated at the B2-PLYP/6-311++G** level of theory. From these ΔG values the aqueous pK_a s for each guanidinium cation investigated in this study were calculated according to equation 1. There is a relatively large inherent error in the calculation of a particular pK_a value using this method. However, the inherent errors should be largely canceled out if the difference in pK_a values for two similar deprotonation reactions are examined.³⁸ Therefore, we calculated the difference in the pK_a values (ΔpK_a) using the calculated pK_a of EG as a reference value. As expected we find that EG is the most acidic guanidinium cation studies. Depending on the nitrogen that is deprotonated, methyl-EG has a $\Delta \text{pK}_a = 0.1$ (if the NH_2 group is deprotonated) or $\Delta \text{pK}_a = 1.8$ (if the MeNH group is deprotonated). We find similar trends in the ΔpK_a values for symmetrical- and unsymmetrical dimethyl-EG, with $\Delta \text{pK}_a = 2.2$ and 0.3 , respectively. These are in line with what should be expected.⁵⁵ As the guanidinium groups are made more electron rich by methylation they become better at stabilizing the positive charge, and thus the pK_a s increase. Furthermore, we can conclude that all three EGs studied would be protonated in aqueous solution because Arg, which has a guanidinium side-chain $\text{pK}_a = 12.48$,⁵⁶ would be expected to have a similar pK_a as EG.

3.2 MP2/631++G** Geometry Optimized Structures of the Guanidinium...Guanine Dimers

The GO structure for the monomer of N-methyl guanine (NMG) as well as the corresponding NMG...(methylated)EG dimers were obtained at the MP2/6-31++G** level of theory (Chart 2, Figure 3, and Supporting Information). We discounted all GO structures that produced a NMG...(methylated)EG planar angle (θ ; Scheme 4) greater than 45° ,⁵⁷ as this would not allow a (methylated)Arg to hydrogen-bond within the major groove of DNA without major steric hindrance between it and the adjacent DNA base-pairs.

Depending on the starting geometry used for the GO minimization procedure we obtained two different GO structures for both the EG...NMG and methyl-EG...NMG dimers. One dimer

structure corresponds to what would be predicted; the two N-H groups of EG/methyl-EG are hydrogen bonded to the carbonyl-oxygen and imine-nitrogen on the NMG (Scheme 5). At approximately 1 kcal mol⁻¹ higher in energy at this level of theory we found a dimer that corresponds to one NH₂ group of EG/methyl-EG are hydrogen bonding to both the carbonyl-oxygen and imine-nitrogen of the NMG. Thus, these can both be considered as potentially energetically valid dimer structures. Because of the small energy difference between these two structures it was decided to investigate the interaction energies of both of these types of dimers in depth.

The higher energy R_1N_g dimer (EG...guanine dimer bound through one NH₂ moiety) displays a C=O...H-N hydrogen bond distance of 1.708 Å and a N...H-N hydrogen bond distance of 2.053 Å (Figure 3, Table 1). This can be compared to the MeR_1N_g dimer, which shows a lengthening of the C=O...H-N hydrogen bond to 1.771 Å and a contraction of the N...H-N hydrogen bond to 1.980 Å. These changes in hydrogen bond lengths are largely steric effects, and are the consequence of the non-planarity of methyl-EG compared to EG noted above. The increase in the torsion angle of methyl-EG allows the N group of the NMG monomer to achieve a closer approach to the N-H of methyl-EG than in the R_1N_g dimer due to a reduction in steric repulsion of the two monomers. It should be noted that there is also a significant change in the angle θ upon guanidinium methylation (Scheme 4), increasing it from 3.9° in the R_1N_g dimer to 29.3° in the MeR_1N_g dimer (Figure 4). The increase in θ is a result of the increase in the torsion angle noted in section 3.1; the methyl-EG portion of the dimer adopts this large degree of out-of-plane twisting to accommodate optimal overlap between LUMO on the guanidinium with the HOMO on the guanine necessitated by the change in the guanidinium torsion angle.

Rotation of the guanidinium monomers relative to the NMG groups allows for the formation of the dimers with the two guanidinium NH groups directed towards the imine nitrogen and carbonyl oxygen of the guanine (i.e. the predicted geometry). For the R_2N_g dimer the C=O...H-N hydrogen bond lengthens to 1.847 Å while the N...H-N hydrogen bond shortens to 2.028 Å. These are similar to the C=O...H-N and N...H-H hydrogen bond lengths in MeR_2N_g at 1.863 and 2.064 Å, respectively. The slight lengthening of the hydrogen bonds in MeR_2N_g relative to R_2N_g, especially the N...H-N hydrogen bond, is expected because methylation of the guanidinium nitrogen will produce a less acidic H-N proton, thus weakening the N...H interaction. Once again, an increase in the planar angle θ is obtained upon methylation, increasing it from 1.6° in R_2N_g to 34.1° in MeR_2N_g (Scheme 4, Figure 4). We note that attempts were made to locate a GO structure corresponding to a MeR_2N_g dimer with the MeN-H proton hydrogen-bonded to C=O group of NMG without success. In all cases these produced structures where the methyl-EG was positioned perpendicular to the NMG ring (θ approached 90°), and was therefore disregarded for further study.

Both the symmetrical and unsymmetrical dimethyl-EGs will only produce dimers with both guanidinium NH groups bound to guanine. In the case of symmetrical dimethyl-EG, the resulting dimer with NMG (sdMeR_2N_g) produced the predicted structure where both the MeN-H groups of the symmetric dimethyl-EG fragment are hydrogen-bonded to the NMG N and C=O groups with hydrogen bond lengths of 2.084 and 1.883 Å, respectively. This systematic increase in hydrogen bond lengths is expected because of the less acidic N-H protons created upon sequential methylation. Surprisingly, the only stable dimer structure that could be located for the unsymmetrical dimethyl-EG NMG dimer (udMeR_2N_g) corresponds to a hydrogen-bond between the NMG N and dimethyl-EG NH₂ group and the NMG C=O and dimethyl-EG "ε" N-H proton. We find that this dimer has an N...NH₂ hydrogen bond length of 2.054 Å, and a C=O...H-N^ε hydrogen bond length of 1.922 Å. This is by far the longest C=O...H-N hydrogen bond of all of the structures investigated, and is likely a combination of the reduced acidity of the ^εN-H proton coupled with steric crowding between the guanidinium

and guanine monomers. We note that there is significant twisting of the planar angle θ , with $\theta = 27.4^\circ$ in sdMeR_2N_g and 42.7° in udMeR_2N_g (Figure 3). Attempts were made to locate the corresponding udMeR_1N_g dimer without success. This is the result of unfavorable steric crowding of the Me2N group of the dimethyl-EG fragment with the NMG fragment when the udMeR_1N_g dimer is formed. These results are all summarized in Table 1.

3.3 Ethyl Guanidinium...Guanine Dimer Interaction Energies

The calculated interaction energy for the R_1N_g dimer employing the COSMO solvation model was found to be $-9.6 \text{ kcal mol}^{-1}$. Methylation of the NH_2 group not involved in the hydrogen-bond to the NMG molecule resulted in a slight reduction in the stabilization energy, reducing it to $-9.4 \text{ kcal mol}^{-1}$. This is within the expected error of the B2-PLYP level of theory at reproducing experimental hydrogen bond strengths, and is therefore not a significant difference in energy.

A predictable trend in dimer stabilization energies is noted for the *N*-methylated guanidiniums upon going from the udMeR_2N_g dimer to the sdMeR_2N_g to the MeR_2N_g dimer; as the degree of methylation is decreased we find an increase in the dimer stabilization energies. The udMeR_2N_g dimer has an interaction energy of $-12.0 \text{ kcal mol}^{-1}$, compared to $-12.5 \text{ kcal mol}^{-1}$ for the sdMeR_2N_g dimer and $-12.8 \text{ kcal mol}^{-1}$ for the MeR_2N_g dimer. We note that the differences in energy, especially between the MeR_2N_g vs. sdMeR_2N_g and the sdMeR_2N_g vs. udMeR_2N_g dimers have to be treated cautiously as they are within the $0.5 \text{ kcal mol}^{-1}$ estimated error between theory and experiment.

Surprisingly, the R_2N_g dimer actually gave the *weakest* dimer stabilization energy ($-11.3 \text{ kcal mol}^{-1}$) of the dimers examined with this coordination mode. This was unexpected as EG has the most acidic N-H protons of the series (*vide supra*). This result, however, can be readily rationalized in light of the fact that a solvation model approximating water was used for these calculations. It seems reasonable that the solvent would be better at stabilizing the more cationic EG monomer relative to the more electron-rich methylated-EG monomers. Therefore, when a dimer is formed the methylated-EG monomers will be stabilized to a greater degree than the EG monomer will be. This was confirmed by performing the corresponding gas phase calculations. We found that when the solvation model is removed the R_2N_g has the largest stabilization energy ($-32.8 \text{ kcal mol}^{-1}$) and udMeR_2N_g the smallest ($-29.9 \text{ kcal mol}^{-1}$). These stabilization energies are all summarized in Table 1.

A similar trend in stabilization energies is observed when discreet water molecules are utilized to solvate the monomer and dimer structures (Supporting Information, Figure S-6). Here, five or six water molecules were placed about the guanidinium...guanine hydrogen bonded dimers and the structures minimized at the MP2/6-31+G* level. Upon breaking the dimers the water molecules are divided between the guanine and guanidinium molecules, and an additional geometry optimization was performed at the MP2/6-31+G* level. The resulting counterpoise corrected stabilization energies (B2-PLYP/6-311++G**) are $-14.6 \text{ kcal mol}^{-1}$ for the R_2N_g dimer, $-15.6 \text{ kcal mol}^{-1}$ for the udMeR_2N_g dimer, $-15.9 \text{ kcal mol}^{-1}$ for the sdMeR_2N_g dimer, and $-16.2 \text{ kcal mol}^{-1}$ for the MeR_2N_g dimer. This demonstrates that for this system the far less expensive COSMO model gives results that approach those obtained when discreet solvent molecules are employed.

Although the single point calculations were carried out at the B2-PLYP/6-311++G** level of theory while the GO structures were obtained at the MP2/6-31+G* level, the above stabilization energies should be seen as realistic because the differences between GO structures obtained at the two level of theories should be minor.^{58,59} This point was emphasized by examining a random subset of EG...NMG and (methylated)EG...NMG dimers with GO structures obtained at the B2-PLYP/6-311++G** level of theory. The resulting B2-PLYP/6-311++G**

counterpoise corrected interaction energies between GO structures obtained at the MP2/6-31++G** vs. B2-PLYP/6-311++G** levels of theory are all within ~ 0.1 kcal mol⁻¹ of one another (Supporting Information). Thus, the minor penalty paid for obtaining GO structures at the lower level of theory is worth the increased computational efficiency.

3.4 MP2/631++G** Geometry Optimized Structures of the Guanidinium...Guanine...Cytosine Trimers

GO structures of the NMG...*N*-methylcytosine (NMC) dimer (gc) and (methylated)EG...gc trimers were obtained at the MP2/6-31++G** level of theory (Chart 1, Figure 5 and Supporting Information). The trimer structures were constructed by adding an NMC group to the six dimers described above. As above, only those structures with NMG...(methylated)EG planar angles (θ) of less than 45° were considered for further investigations (Scheme 4). Of the six possible trimer structures constructed from the (methylated)EG...NMG dimers, only five produced valid GO structures (Figure 5). Due to unfavorable steric interactions between the ethyl group of the unsymmetrical dimethyl-EG and NMC group in the gc dimer, a stable structure for the udMeR_2N_gc trimer with θ less than 45° could not be located.⁶⁰ As observed above with the R_1N_g and MeR_1N_g dimers, stable structures for the corresponding R_1N_gc and MeR_1N_gc trimers were located. However, these will not be discussed in depth as they are not energetically viable structures when compared to the other guanidinium coordination mode; the MP2/6-31++G** calculations place them ~ 4 kcal mol⁻¹ higher in energy than the corresponding trimer structures with two guanidinium NH groups hydrogen bound to guanine (i.e. the equilibrium should heavily favor the later).

As expected the gc dimer is planar with a 0° tilt between the ring systems of the two nucleotide bases. The three hydrogen bonding distances holding the dimer together compare well with what should be expected; the NMG...NMC hydrogen bond lengths are: 1.658 Å for the C=O...H-N hydrogen bond, 1.834 Å for the N-H...N hydrogen bond, and 1.828 Å for the N-H...O=C hydrogen bond (Table 2). For all five trimers, when the (methylated)EG groups hydrogen-bond to the gc dimer, a lengthening of the NMG...NMC C=O...H-N hydrogen bond (1.799 – 1.803 Å) is observed. This is completely expected because of the reduction in Lewis-basicity of the carbonyl oxygen, and is compensated for by an increase in the strength of the other two hydrogen bonding interactions in the gc dimer; the N-H...N hydrogen bonds shorten to 1.803 – 1.810 Å, and the N-H...O=C hydrogen bonds shorten to 1.721 – 1.726 Å.

Upon trimer formation, the guanidinium...NMG N-H...O=C hydrogen bond lengths are also elongated compared to the guanidinium...NMG dimer structures (1.702 – 1.802 Å). To compensate for the reduction in the hydrogen-bonding interaction between the (methylated) EG groups and the guanidinium carbonyl oxygen the guanidinium NH actually forms a tighter hydrogen bond with the guanidinium imine nitrogen. The corresponding N-H...N bond lengths range between 1.696 – 1.786 Å for the five trimers. As such, the overall interaction energies of the (methylated)EG groups with the gc dimers are similar to the interaction energies of the (methylated)EG groups with the NMG monomer (*vide infra*).

Binding of EG to the gc dimer in such a way that both NH₂ groups are hydrogen bonded to the NMG group, forming R_2N_gc, causes the two nucleotide bases to distort away from planarity by 2.6°. The EG portion of the trimer itself is substantially rotated out of planarity from the NMG ring, with $\theta = 38.1^\circ$, which is increased from $\theta = 2.5^\circ$ for the R_2N_g dimer (Figure 6). This is similar to what has been observed crystallographically in several DNA binding proteins.^{61,62} The twisting of the EG group out of planarity is due to a slight reorientation of the π -type orbitals on the guanine carbonyl oxygen and imine nitrogen induced by hydrogen-bonding to the cytosine. This causes the guanidinium group to tilt in order to achieve optimal HOMO/LUMO overlap with guanine.

This structure can be compared with the trimers MeR_2N_gc and sdMeR_2N_gc. The methyl-EG...guanine...cytosine trimer MeR_2N_gc displays an angle $\theta = 30.6^\circ$ while the symmetric dimethyl-EG...guanine...cytosine trimer sdMeR_2N_gc displays an angle $\theta = 28.0^\circ$ (Table 2). These are both similar to what was observed in the corresponding methyl-guanidinium...guanine dimers (*vide supra*, Table 1). The reason for the similarity in θ for these dimer vs. trimer structures is the intramolecular steric encumbrance on the guanidiniums (that induces the twist of the guanidiniums out of the guanine mean-plane) actually placed these monomers into an optimal position for orbital overlap between the LUMO of the guanidinium and HOMO of the guanine. As such, the three dimensional orientation of the all the guanidinium...gc trimers investigated are virtually identical, with methylation effecting little structural change in the 3D structure of the three complexes investigated.

3.5 Guanidinium...Guanine...Cytosine Trimer Stabilization Energies

As with the guanidinium...NMG dimers, the two trimers that have one NH_2 group from the (methylated)EG hydrogen bonding to both guanine carbonyl oxygen imine nitrogen are weakest in energy. In both cases, there is a 1.5 to 2.0 kcal mol^{-1} reduction in stabilization energy relative to the (Me)R_1N_g dimer structures, with R_1N_gc yielding an EG...gc stabilization energy of $-7.7 \text{ kcal mol}^{-1}$ and MeR_1N_gc yielding a methyl-EG...gc stabilization energy of $-7.4 \text{ kcal mol}^{-1}$. This reduction in stabilization energy is due to the reduction in hydrogen-bond energy between the guanidinium N-H and guanine C=O groups owing to the gc hydrogen bond network, which lowers the basicity of the guanine C=O group. We note that considering the rms error at this level of theory these two values are within error of one another, as was found with the dimer structures. Furthermore, due to the weak stabilization energy of this binding mode compared to the other, it is likely not to have any biological significance.

At $-11.2 \text{ kcal mol}^{-1}$, the EG...gc stabilization energy of the R_2N_gc trimer is almost identical to what was calculated for the R_2N_g dimer, suggesting that the reduction in the $\text{C=O}\cdots\text{HN}$ hydrogen bond strength is adequately compensated for by the increase in the $\text{N}\cdots\text{HN}$ hydrogen bond strength. We found that for the MeR_2N_gc and sdMeR_2N_gc trimers there is an increase in the stabilization energy by $\sim 1 \text{ kcal mol}^{-1}$ ($-13.8 \text{ kcal mol}^{-1}$ for MeR_2N_gc and $-13.5 \text{ kcal mol}^{-1}$ for sdMeR_2N_gc). This is due to the better optimized orbital overlap between the methylated-EG groups LUMO and the guanine HOMO in the trimers compared to the methylated-EG groups LUMO and the guanine HOMO in the dimer structures. Thus, the slight spatial reorientation of the guanine C=O and C=N π -orbitals upon gc dimer formation actually induces an energetically more favorable interaction between the methylated guanidiniums and guanine.

As with the dimer structures above, we find that the more acidic guanidinium monomer, EG, forms a weaker stabilizing interaction with the gc dimer than the less acidic methylated-EG monomers. This is also largely a solvent effect. When the COSMO solvent model is removed the expected trend in stabilization energies is realized. R_2N_gc possesses the strongest interaction energy of $-32.6 \text{ kcal mol}^{-1}$ and sdMeR_2N_gc possesses the weakest stabilization energy of $-31.0 \text{ kcal mol}^{-1}$.

4. Summary and Biological Implications

It has been speculated that Arg methylation events are important post-translational modifications for altering the nature of biomolecular interactions. Unlike similar post-translational modifications (such as phosphorylation) that act more like biological “on/off switches,” Arg methylation is predicted to influence biomolecular interactions more subtly.¹⁻⁴ The above results lend credence to this supposition; changes in guanidinium...guanine interaction energies of less than $\sim 2.5 \text{ kcal mol}^{-1}$ were obtained upon guanidinium methylation.

A number of other conclusions were drawn from this study. One involves the ability of the methylated guanidinium groups to form a hydrogen-bond complex with guanine when involved in a *gc* base-pair. With the exception of the unsymmetrical dimethyl-EG, all of the guanidinium groups are capable of forming a complex with the *gc* dimer. We found that the reason that the unsymmetrical dimethyl-EG will not coordinate to the *gc* dimer is not due to an inherently weak interaction energy with guanine; investigations into the udMeR_2N_g dimer shows it is energetically quite favorable to form a hydrogen-bonded complex with guanine. Instead, it has to do with the fact that the unsymmetrical dimethyl-EG will utilize its “ ω ” and “ ϵ ” NH groups to hydrogen bond to guanine. This in turn leads to unfavorable steric interactions between the guanidinium group and cytosine in the *gc* base-pair. This would imply that unsymmetrical dimethylated Arg would impair protein...DNA complex formation as the Arg...guanine interaction would be lost. In contrast, the other methylated Arg residues would still be capable of forming the Arg...guanine base-pair, and would therefore not impair protein...DNA complex formation when compared to the unmethylated Arg protein residue.

Another conclusion that was reached was that the guanidinium methylation will not dramatically alter the 3-D structure of the resulting guanidinium...guanine moiety in the methylated guanidinium...*gc* trimers when compared to the unmethylated guanidinium...*gc* trimer. In all cases it was found that there is a significant twist away from planarity of the guanidinium...guanine moiety. Thus, outside of the increased steric bulk of the additional methyl group, methylation should not impose any additional steric constraints on the formation of a protein...DNA complex.

In our view, the most surprising finding of this study is that the stabilization energies between the guanidinium group and guanine actually *increase* upon Arg methylation; the methylated-guanidinium...*gc* trimers are ~ 2.5 kcal mol⁻¹ more stable than the unmethylated-guanidinium...*gc* trimer. This is despite the less acidic H⁺ of the protonated NH₂⁺ groups afforded upon subsequent methylation. Gas vs. solvent phase calculation demonstrate that this is exclusively due to the rather efficient solvation of the unmethylated guanidinium monomer vs. the methylated monomers. Therefore dimer formation does not favor desolvation of guanidinium as much as it does the methylated guanidiniums. This would imply that Arg methylation would actually promote the formation of the Arg...guanine dimer, and thereby promote protein...DNA complex formation, which is counter to what has previously been predicted.¹⁻⁴ A similar finding has been reached in the formation of Arg...tryptophan cation/ π complexes using small model peptides.⁶³ We are currently in the process of experimentally verifying these theoretical results.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The NIH is acknowledged for financial support (P20 RR-016464) from the INBRE program of the National Center for Research Resources. This work was carried out, in part, through the use of the University of Nevada Research Computing Grid.

Abbreviations

NMG	<i>N</i> -methylguanine
NMC	<i>N</i> -methylcytosine
EG	<i>N</i> -ethylguanidinium

methyl-EG	<i>N</i> -ethyl- <i>N'</i> methylguanidinium
symmetrical dimethyl-EG	<i>N</i> -ethyl- <i>N''N'</i> -dimethylguanidinium
unsymmetrical dimethyl-EG	<i>N</i> -ethyl- <i>N''N''</i> -dimethylguanidinium
gc	NMGNMC dimer
(methylated)EG	the guanidiniums investigated in this study as a whole
R_1N_g	EG...NMG dimer bound through one EG NH ₂ group
R_2N_g	EG...NMG dimer bound through two EG NH ₂ groups
R_1N_gc	EG...NMG...NMC trimer EG...NMG bound through one EG NH ₂ group
R_2N_gc	EG...NMG...NMC trimer bound through two EG NH ₂ groups
MeR_1N_g	methyl-EG...NMG dimer bound through one methyl-EG NH ₂ group
MeR_2N_g	methyl-EG...NMG dimer bound through one methyl-EG NH ₂ group and one NMeH group
MeR_1N_gc	methyl-EG...NMG...NMC trimer bound through one methyl-EG NH ₂ group
MeR_2N_gc	methyl-EG...NMG...NMC trimer bound through one methyl-EG NH ₂ group and one NMeH group
sdMeR_2N_g	symmetrical dimethyl-EG...NMG dimer
udMeR_2N_g	unsymmetrical dimethyl-EG...NMG dimer
sdMeR_2N_gc	symmetrical dimethyl-EG...NMG...NMC trimer

References

1. McBride AE, Silver PA. Cell 2001;106:5–8. [PubMed: 11461695]
2. Bedford MT, Richard S. Mol. Cell 2005;18:263–272. [PubMed: 15866169]
3. Bedford MT. J. Cell Science 2007;120:4243–4246. [PubMed: 18057026]
4. Pahlich S, Zakaryan RP, Gehring H. Biochim. Biophys. Acta 2006;1764:1890–1903. [PubMed: 17010682]
5. Chen D, Ma H, Hong H, Koh SS, Huang SM, Schurter BT, Aswad DW, Stallcup MR. Science 1999;284:2174–2177. [PubMed: 10381882]
6. Stallcup MR. Oncogene 2001;20:3014–3020. [PubMed: 11420716]
7. Rice JC, Allis CD. Curr. Opin. Cell Biol 2001;13:263–273. [PubMed: 11343896]
8. Zhang Y, Reinberg D. Genes Dev 2001;15:2343–2360. [PubMed: 11562345]
9. Yu MC, Bachand F, McBride AE, Komili S, Casolari JM, Silver PA. Genes Dev 2004;18:2024–2035. [PubMed: 15314027]
10. Gary JD, Clarke S. Prog. Nucl. Acid Res. Mol. Biol 1998;61:65–131.
11. El-Andaloussi N, Valovka T, Touelle M, Steinacher R, Focke F, Gehrig P, Covic M, Hassa PO, Schar P, Hubscher U, Hottiger MO. Mol. Cell 2006;22:51–62. [PubMed: 16600869]
12. Cazanove O, Batut J, Scarlett G, Mumford K, Elgar S, Thresh S, Neant I, Moreau M, Guille M. Biochemistry 2008;47:8350–8357. [PubMed: 18636753]
13. Nichols RC, Wang XW, Tang J, Hamilton BJ, High FA, Herschman HR, Rigby WF. Exp. Cell Res 2000;256:522–532. [PubMed: 10772824]

14. Li H, Park S, Kilburn B, Jelinek MA, Henschen-Edman A, Aswad DW, Stallcup MR, Laird-Offringa IA. *J. Biol. Chem* 2002;277:44623–44630. [PubMed: 12237300]
15. Cote J, Boisvert FM, Boulanger MC, Bedford MT, Richard S. *Mol. Biol. Cell* 2003;14:274–287. [PubMed: 12529443]
16. Araya N, Hiraga H, Kako K, Arao Y, Kato S, Fukamizu A. *Biochem. Biophys. Res. Commun* 2005;329:653–660. [PubMed: 15737635]
17. Stetler A, Winograd C, Sayegh J, Cheever A, Patton E, Zhang X, Clarke S, Ceman S. *Hum. Mol. Genet* 2006;15:87–96. [PubMed: 16319129]
18. Fujiwara T, Mori Y, Chu DL, Koyama Y, Miyata S, Tanaka H, Yachi K, Kubo T, Yoshikawa H, Tohyama M. *Mol. Cell. Biol* 2006;26:2273–2285. [PubMed: 16508003]
19. Ostareck-Lederer A, Ostareck DH, Rucknagel KP, Schierhorn A, Moritz B, Huttelmaier S, Flach N, Handoko L, Wahle E. *J. Biol. Chem* 2006;281:11115–11125. [PubMed: 16492668]
20. Goulah CC, Read LK. *J. Biol. Chem* 2007;282:7181–7189. [PubMed: 17229732]
21. Jeffrey, GA.; Saenger, W. *Hydrogen Bonding in Biological Structures*. Springer; Berlin: 1991.
22. Wintjens R, Lievin J, Rooman M, Buisine E. *J. Mol. Biol* 2000;302(2):395–410. [PubMed: 10970741]
23. Warner DR, Weinstein LS. *Proceed. Natl. Acad. Sci. USA* 1999;96:4268–4272.
24. Kim CA, Berg JM. *Nat. Struct. Biol* 1996;3:940–945. [PubMed: 8901872]
25. Jantz D, Berg JM. *J. Am. Chem. Soc* 2003;125:4960–4961. [PubMed: 12708829]
26. Burke B, An S, Musier-Forsyth K. *Biochim. Biophys. Acta* 2008;1784:1222–1225. [PubMed: 18513497]
27. Rodriguez-Casado A, Molina M, Carmona P. *Appl. Spectrosc* 2007;61:1219–1224. [PubMed: 18028701]
28. Liu A, Majumdar A, Jiang F, Chernichenko N, Skripkin E, Patel DJ. *J. Am. Chem. Soc* 2000;122:11226–11227.
29. Neese, F. ORCA Version 2.6.35-2008. Universitat Bonn; Bonn, Germany:
30. Boys SF, Bernardi F. *Mol. Phys* 1970;19:553–566.
31. Latajka Z, Scheiner S. *J. Chem. Phys* 1987;87:1194–1204.
32. van Duijneveldt FB, van Duijneveldt-van de Ridt JGCM, van Lenthe JH. *Chem. Rev* 1994;94:1873–1885.
33. Schwabe T, Grimme S. *Acc. Chem. Res* 2008;41:569–579. [PubMed: 18324790]
34. Grimme S. *J. Chem. Phys* 2006;124:34108.
35. Krishnan R, Binkley JS, Seeger R, Pople JA. *J. Chem. Phys* 1980;72:650.
36. Clark T, Chandrasekhar J, Spitznagel GW, Schleyer PVR. *J. Comp. Chem* 1983;4:294.
37. Klamt A, Schüürmann G. *J. Chem. Soc. Perkin Trans. 2* 1993;220:799.
38. Chen I-J, MacKerell AD Jr. *theor. Chem. Acc* 2000;103:483.
39. Hehre WJ, Ditchfield R, Pople JA. *J. Chem. Phys* 1972;56:2257.
40. Hariharan PC, Pople JA. *Theoret. Chimica Acta* 1973;28:213.
41. Raghavachari K, Trucks GW, Pople JA, Head-Gordon M. *Chem. Phys. Lett* 1989;157:479–483.
42. Becke A. *J. Chem. Phys* 1986;84:4524–4529.
43. Becke A. *J. Chem. Phys* 1988;88:1053–1062.
44. Becke A. *Phys. Rev. A* 1988;38:3098–3100. [PubMed: 9900728]
45. Perdew JP. *Phys. Rev. B* 1986;34:7406.
46. Perdew JP. *Phys. Rev. B* 1986;33:8822–8824.
47. Becke AD. *J. Chem. Phys* 1993;98:5648–5652. 1372–1377.
48. Lee CT, Yang WT, Parr RG. *Phys. Rev. B: Condens. Matter* 1988;37:785–789. [PubMed: 9944570]
49. Schwabe T, Grimme S. *Phys. Chem. Chem. Phys* 2007;9:3397–3406. [PubMed: 17664963]
50. Häber T, Seefeld K, Engler G, Grimme S, Kleinermanns K. *Phys. Chem. Chem. Phys* 2008;10:2844. [PubMed: 18465002]
51. Antony J, Grimme S. *Phys. Chem. Chem. Phys* 2008;10:2722. [PubMed: 18464987]
52. Dunning TH Jr. *J. Chem. Phys* 1989;90:1007.

53. Kendall RA, Dunning TH Jr, Harrison RJ. *J. Chem. Phys* 1992;96:6796.
54. We defined the angle ϕ as the twist between the plane formed by an NHR group and the plane formed by the CN_2 triangle.
55. Kennedy KJ, Lundquist JT IV, Simandan TL, Kokko KP, Beeson CC, Dix TA. *J. Pept. Res* 2000;55:348–358. [PubMed: 10798380]
56. Dawson, RMC.; Elliot, DC.; Elliot, WH.; Jones, KM. *Data for Biochemical Research*. Clarendon Press; Oxford, U.K.: 1986.
57. We define θ as the angle of rotation between the two planes formed by the N_3 triangle of the guanidinium group and the planer guanine ring system.
58. Siegbahn PEM. *J. Comput. Chem* 2001;22:1634–1645.
59. Shearer J, Dehestani A, Abanda F. *Inorg. Chem* 2008;47:2649–2660. [PubMed: 18330983]
60. The udMeR_2N_gc trimer lies $\sim 5 \text{ kcal mol}^{-1}$ higher in energy than the separated gc dimer and unsymmetrical dimethyl-EG monomer.
61. Elrod-Erickson M, Benson TE, Pabo CO. *Structure* 1998;6:451–464. [PubMed: 9562555]
62. Flick KE, Jurica MS, Monnat RJ Jr, Stoddard BL. *Nature* 1998;394:96–101. [PubMed: 9665136]
63. Hughes RM, Waters ML. *J. Am. Chem. Soc* 2006;128:12735–12742. [PubMed: 17002367]

Scheme 1.

Figure 1.
Hydrogen bonding network formed between the Arg guanidinium group and guanine.

Chart 2.

Scheme 2.

Figure 2.
Geometry optimized structures of the four guanidinium monomers investigated in this study: EG (A), methyl-EG (B), symmetrical dimethyl-EG (C), and unsymmetrical dimethyl-EG (D).

Scheme 3.

Scheme 4.

Figure 3.
Geometry optimized structures for the six guanidinium...guanine dimers examined in this study.

Scheme 5.

Figure 4.

View of the guanidinium...guanine dimers looking along the xy plane of the guanine. The guanine group is depicted in blue and the guanidinium group is depicted in red.

Figure 5.
Geometry optimized structures for the five guanidinium...guanine...cytosine trimers examined in this study.

Figure 6. View of the guanidinium...guanine...cytosine trimers looking along the xy plane of the guanine. The guanine and cytosine groups are depicted in blue and the guanidinium group is depicted in red.

Table 1

Selected metric parameters and dimer stabilization energies for the guanidinium *N*-methylguanidine dimers.

Dimer	NH...N (Å) ^a	NH...O (Å) ^a	θ (°) ^a	ΔE _{water} (kcal/mol) ^b	ΔE _{gas} (kcal/mol) ^b
R_1N_g	2.052	1.708	3.9	-9.6	-31.4
R_2N_g	2.028	1.847	1.6	-11.3	-32.8
MeR_1N_g	1.980	1.771	29.3	-9.4	-28.7
MeR_2N_g	2.064	1.863	34.1	-12.8	-31.7
sdMe_2N_g	2.084	1.883	27.4	-12.5	-30.8
udMe_2N_g	2.054	1.992	42.7	-12.0	-29.9

^a MP2/6-31++G**

^b B2-PLYP/6-311++G**

Table 2

Selected metric parameters and trimer stabilization energies for the guanidinium *N*-methylguanine/*N*-methylcytosine trimers.

Trimer	N-H...O (Å) ^a _{bc}	N...H-N (Å) ^a _{ab}	O...H-N (Å) ^a _{ab}	N-H...N (Å) ^a _{ac}	N-H...O (Å) ^a _{ac}	θ (°) ^a _{ac}	ΔE _{water} (kcal/mol) ^d	ΔE _{gas} (kcal/mol) ^d
gc	1.658	1.834	1.828	---	---	---	---	---
R_1N_gc	1.808	1.810	1.722	1.773	1.796	10.4	-7.7	-28.5
R_2N_gc	1.803	1.798	1.722	1.696	1.702	38.1	-11.2	-32.6
MeR_1N_gc	1.799	1.810	1.722	1.786	1.802	30.3	-7.4	-26.9
MeR_2N_gc	1.803	1.803	1.721	1.710	1.734	30.6	-13.8	-32.0
sdMeR_2N_gc	1.803	1.804	1.726	1.754	1.731	28.0	-13.5	-31.0

^aMP2/6-31++G**

^bNMC...NMG hydrogen bond distance

^cNMG...guanidinium hydrogen bond distance

^dB2-PLYP/6-311++G**