DNA-conducting polymer complexes: A computational study of the hydrogen bond between building blocks

David Zanuy* **and Carlos Alemán***

Departament d'Enginyeria Química, E. T. S. d'Enginyeria Industrial de Barcelona, Universitat Politècnica de Catalunya, Diagonal 647, Barcelona E-08028, Spain

* david.zanuy@upc.edu and carlos.aleman@upc.edu

Abtract

Ab initio quantum mechanical calculations at the MP2 level have been used for an extensive study about the stability of hydrogen bonded complexes formed by pyrrole and thiophene, which are the most common building blocks of conducting polymers, and DNA bases. Results indicate that very stable complexes are formed with pyrrole, which shows a clear tendency to form specific hydrogen bonding interactions with nucleic acid bases. Furthermore, the strength of such interactions depends significantly on the base, growing in the following order: thymine < adenine ≈ cytosine < guanine. On the contrary, thiophene forms complexes stabilized by non-specific interactions between the π-cloud of the ring and N-H groups of nucleic acid bases rather than specific hydrogen bonds. The overall of these results is fully consistent with experimental observations: polypyrrole is able not only to stabilize adducts with DNA but also to interact specifically, while the interactions of the latter with polythiophene and their derivatives are weaker and non-specific.

Introduction

The interaction of conducting electroactive polymers, as polythiophene, polypyrrole and their derivatives, with selected bioentities, *e.g.* with amino acids,[1-3] proteins, [4-7] DNA and oligonucleotides,[7-15] living cells,[16-19] etc., is a subject of increasing interest.[20,21] The quest to interact more efficiently with biosystems, to obtain information related to system performance and to control that performance remains not only an exciting but also an essential area of research. The development of biotechnological applications based conducting polymers greatly depends on the control of such interactions.

We are particularly interested in the interaction of conducting polymers with DNA sequences, which may have great implications in numerous medical applications ranging from diagnosis to gene therapy.[7-15,20,21] The interaction of p-doped electroactive materials with DNA has been traditionally attributed to the tendency of the latter to interact with positively charged molecules. However, in recent studies we found that a given polythiophene derivative[10] as well as polypyrrole[22] are able to form specific interactions with well-defined nucleotide sequences of plasmid DNA. Thus, gel electrophoresis assays of a series of polymer:DNA complexes prepared considering different mass ratios were performed in presence of restriction enzymes, which cut off at specific nucleotide sequences. We observed that these polymers were able to prevent DNA digestion indicating that the restriction sites are inaccessible to the restriction enzyme within the polymer:DNA complexes. Taking into account that the plasmid DNA used in such experiments only contains a single restriction site, the selective binging mode polymer:DNA was clearly reflected. These results suggest that the formation of such polymer:DNA complexes is based not only on electrostatic interactions but also on other kind of interactions, *i.e.* hydrogen bonds, stacking, van der

Waals, charge-transfer, etc. The selective affinity between conducting polymers and DNA opens an intriguing research field based on the design of well-controlled complexes for well-defined applications. However, this requires a previous detailed analysis, which should be based on simplified models, of the different interactions that may be involved in the formation of complexes.

Nowadays, the chance to analyze interactions between the chemical repeating units of polymers and DNA bases comes from quantum chemical calculations. Thus, substantial computer advances in recent years allow apply high-level theoretical methods, which are able to describe molecular systems very accurately. For example, examination and comparison of the different interaction modes between DNA bases using such theoretical methods have provided very valuable information about the structure and dynamics of this biomacromolecule.[23-31] In this work we evaluate the ability of pyrrole (Py) and thiophene (Th), which are the most common building blocks of conducting polymers, to interact with the methylated analogues of DNA bases [9-methyladenine (mA), 9-methylguanine (mG), 1-methylcytosine (mC) and 1-methylthymine (mT)] through specific hydrogen bonding interactions. The importance of hydrogen bonds in polymer:DNA complexes with specific interactions is expected to be significantly greater than those that are of non-specific, *i.e.* stacking, van der Waals, electrostatic and charge transfer. Accordingly, we concentrate on the ability of Py and Th to form specific hydrogen bonding interactions with DNA bases rather than on the detailed description of the potential energy surfaces for the complexes under investigation.

Calculations have been performed considering both Py and Th in the neutral (reduced) state rather than in the doped (oxidized) one. In way all the structural and energetic features reported in this work must be attributed exclusively to hydrogen

bonds, no contamination due to electrostatic effects typically produced by charged species being possible. It should be emphasized that this is a right approximation since, in doped polymers, charges are not uniformly distributed along the whole molecular chains.[32-34] On the contrary, positive charges in oxidized polyheterocyclic conducting polymers, as polythiophene and polypyrrole, are localized in small segments that contain a few number of monomering units (typically a few tenths of monomeric rings with a quinoid-like electronic structure). These segments are separated among them by blocks of rings with a benzenoid-like electronic structure, which is characterisric of neutral aromatic species.[32-34] Neutral Py and Th rings belonging to non-charged blocks are those expected to participate in the formation of specific hydrogen bonding interactions between doped polymer chains and DNA bases.

Methods

Calculations were performed using the Gaussian 03[35] computer program. The structures of both complexes and isolated monomers were determined in the gas-phase by full geometry optimization at the MP2 level[36] with the 6-31G(d) basis set,[37] frequency calculations being performed to obtain the zero-point vibrational energies and both the thermal and entropic corrections. Single point energy calculations were performed on the MP2/6-31G(d) geometries at both the MP2/6-311G(d,p)[38] and MP2/6-311++G(d,p)[39] levels. In order to estimate the free energies in the gas-phase, the statistical corrections obtained at the MP2/6-31G(d) level were added to the electronic energies computed at the MP2/6-311G(d,p) and MP2/6-311++G(d,p) levels.

The counterpoise correction method was applied to correct the basis set superposition error.[40] The binding energy (ΔE_b) was calculated according to Eqn (1):

$$\Delta E_b = E_{Py/Th\cdot\cdot mNA} - E_{Py/Th,comp} - E_{mNA,comp} \qquad (1)$$

where $E_{Py/Th\cdots mNA}$ corresponds to the total energy of the optimized complex, and $E_{Py/Th,comp}$ and $E_{mNA,comp}$ are the energies of the isolated monomers with the geometries obtained from the optimization of the complex.

The distortion energy (ΔE_{dis}), which estimates the relaxation of the monomers on the dimer formation, was computed using Eqn (2):

$$\Delta E_{dis} = \left(E_{Py/Th,comp} + E_{mNA,comp}\right) - \left(E_{Py/Th,opt} + E_{mNA,opt}\right) \qquad (2)$$

where $E_{Py/Th,opt}$ and $E_{mNA,opt}$ are the energies obtained from the optimization of the isolated monomers. It should be noted that the difference between ΔE_b and ΔE_{dis} corresponds to the net binding energy.[41,42]

The effect of the solvent on the relative stability of the complexes was estimated following the polarizable continuum model (PMC) developed by Miertus, Scrocco and Tomasi.[43,44] This SCRF method involves the generation of a solvent cavity from spheres centered at each atom in the molecule and the calculation of virtual point charges on the cavity surface representing the polarization of the solvent. The magnitude of these charges is proportional to the derivative of the solute electrostatic potential at each point calculated from the molecular wavefunction. The point charges may, then, be included in the one-electron Hamiltonian, thus inducing polarization of the solute. An iterative calculation is carried out until the wavefunction and the surface charges are self-consistent.

PCM calculations were performed in the framework of the ab initio MP2 level with the 6-31G(d) basis set and using the standard protocol and considering the dielectric constants of chloroform (ε= 4.9) and water (ε= 78.4). Calculations were performed considering the gas-phase optimized geometries. Thus, solvent–induced changes in bond lengths and angles have been proved to have little influence on the free energy of solvation (ΔG_{sol}),[45-47] *i.e.* solute geometry relaxations in solution and single point

calculations on the gas-phase optimized geometries provide almost identical values of ΔG_{sol}. It should be noted that water was chosen for calculations because this solvent was used for the experimental assays about the interaction between conducting polymers and DNA. On the other hand, calculations on chloroform were performed to examine the influence of the polarity of the solvent on the relative stability of the computed complexes. The free energy of the complexes in solution required to examine the relative stability in solution were computed using the classical thermodynamics scheme: the ΔG_{sol} provided by the PCM model was added to the gas-phase free energy.

Results and Discussion

Initial structures of Py···mNA and Th···mNA complexes (where mNA= mA, mG, mC or mT) were constructed considering that Py is donor and acceptor of hydrogen bonds while Th acts as donor only. Accordingly, 42 and 32 starting geometries were prepared for Py···mNA and Th···mNA complexes, respectively, applying the following scheme: for each interaction site of each nucleic acid base, different orientations of the Py and Th were considered. Figure 1 shows the interaction sites considered for the methylated analogues of DNA bases as well as the number of initial structures prepared for each Py···mNA and Th···mNA complexes. Geometry optimization and frequency calculations at the MP2/6-31G(d) level[36,37] of all such initial structures led to 17 and 11 Py···mNA and Th···mNA complexes of minimum energy, respectively. Single point calculations at the MP2/6-311G(d,p)[38] and MP2/6-311++G(d,p)[39] were additionally performed to provide better estimations of both the relative stabilities and the affinities.

Pyrrole···Nucleic Acid Complexes. Figure 2 shows the Py···mNA minimum energy complexes, which are distributed as follows: 4 (Py···mA), 5 (Py···mG), 3 (Py···mC)

and 5 (Py···mT). Each minimum has been labeled using a roman number followed by the two letters associated to the corresponding methylated nucleic acid base. The relative conformational energies ($\Delta E_{r,g}$) and free energies ($\Delta G_{r,g}$) estimated in the gas-phase at different levels of theory for the characterized minima of each complex are listed in Table 1.

As can be seen, the lowest energy minimum for Py···mA complex corresponds to IImA, in which the N-H group of Py acts as hydrogen bonding donor and acceptor simultaneously, the other three minima being unfavored by less than 1.5 kcal/mol only. The stability of ImA and IVmA decreases when the size of the basis set increases, even though in these minima the Py group is also involved in two hydrogen bonds. In opposition, the stability of IIImA, which surprisingly only involves one hydrogen bond, increases with the basis set. Thus, the $\Delta E_{r,g}$ predicted for such minimum decreases from 1.5 to 0.9 kcal/mol when polarization and diffuse functions are added to the basis set. On the other hand, the lowest energy minimum of Py···mG corresponds to VmG. This structure shows a hydrogen bonding interaction (Py)N-H···O(mG) and a N-H···π interaction between the mG and the π-cloud of the Py, the latter providing a significant stabilization.[48,49] However, the behavior of the five Py···mG complexes is completely different from that of the Py···mA ones. Thus, the values of $\Delta E_{r,g}$ calculated at the best level of theory for the other four Py···mG complexes are higher than 5.0 kcal/mol indicating that, in spite of all them involves one or two hydrogen bonds, their stability with respect to VmG is very low. Furthermore, comparison between the values $\Delta E_{r,g}$ and $\Delta G_{r,g}$ reveals that the influence of the zero-point vibrational energy (ZPVE) and the thermal and entropic corrections to the energy are very significant in Py···mG complexes. Thus, the destabilization of complexes ImG, IImG, IIImG and IvmG with

respect to the global minimum decreases from ~ 5 kcal/mol to ~ 3 kcal/mol after the incorporation of such corrections.

Regarding to Py···mC, the lowest energy minimum, IIImC, which is stabilized by two hydrogen bonds (Py)NH···N(mC) and (mC)N-H···N(Py), is only slightly more stable than IImC, the latter forming a bifurcated intermolecular interaction with the N-H of Py acting hydrogen bonding donor, *i.e.* (Py)NH···O(mC) and (Py)N-H···N(mC). Complex ImC forms a single hydrogen bond (Py)N-H···O(mC) and is the least stable minimum, even though it should be considered as an accessible structure because its $\Delta G_{r,g}$ is around 1.0 kcal/mol. Inspection to the results obtained for Py···mT reveals a strong dependence on the size of the basis set. Thus, the $\Delta E_{r,g}$ values displayed in Table 1 indicate that the lowest energy complex changes from IVmT to ImT when polarization functions are added to basis set. Furthermore, addition of the ZPVE, entropic and thermal corrections produces a significant increase in the relative stability of all the local minima. Thus, the values of $\Delta G_{r,g}$ obtained by adding our best estimate of the electronic energies to the thermodynamic corrections calculated at the MP2/6-31G(d) level reveal that the separation among the five minima is very small, *i.e.* lower than 0.7 kcal/mol. It should be noted that the five minima obtained for Py···mT form a single hydrogen bond, which explains their similarity in terms of stability.

The influence of both aqueous and organic solvents on the relative stability of the different complexes has been examined using a Self-Consistent Reaction-Field (SCRF) method. Table 2 lists the free energies of solvation (ΔG_{sol}) and the relative conformational free energies in solution ($\Delta G_{r,s}$) for the 17 minimum energy complexes obtained in the gas-phase. The values of ΔG_{sol} indicate that complex-solvent interactions are stronger in water than in chloroform, the strength of such interactions increasing as follows for the two solvents: Py···mA < Py···mT < Py···mC < Py···mG.

On the other hand, inspection to the values of $\Delta G_{r,sol}$ reveals that the solvent produce significant changes in the relative stability order of the different complexes. Thus, the lowest energy Py···mA complex in solution is IIImA, solvent inducing a relative free energy variation of 1.6 (chloroform) and 2.3 kcal/mol (water) with respect to the most favored complex in the gas-phase, IImA. A similar feature is observed for Py···mC: solvent stabilizes ImC with respect IIImC, the latter being the most stable in the gas-phase. Interestingly, both ImC and IImC become unfavored by about 2 and 3 kcal/mol in chloroform and water, respectively, whereas in the gas-phase they were destabilized by less than 1 kcal/mol. However, the most drastic change occurs for Py···mG. Complex VmG, which was clearly stabilized in the gas-phase, becomes the least stable in solution with $\Delta G_{r,sol}$ values of 2.2 (chloroform) and 4.4 kcal/mol (water). In contrast, the remaining four Py···mG complexes, which were clearly unfavored in the gas-phase, become stabilized in solution. Regarding to Py···mT, solvent induces a notable destabilization of IImT and IVmT reducing the number of energetically accessible complexes from five to three.

Table 3 displays the binding energy (ΔE_b) calculated for the different Py···mNA complexes using the MP2 method combined with the 6-31G(d), 6-311G(d,p) and 6-311++G(d,p) basis sets. As expected, for each class of complexes the lowest value of ΔE_b, which reflects the strongest binding, corresponds to the lowest energy minimum: IImA, VmG, IIImC and ImT for Py···mA, Py···mG, Py···mC and Py···mT, respectively. Furthermore, the influence of the size of the basis set on ΔE_b is similar to that discussed above for $\Delta E_{r,g}$.

Comparison among the different complexes indicates that the tendency of Py to interact with DNA bases through specific hydrogen bonding interactions grows in the following order: Py···mT < Py···mA ≈ Py···mC < Py···mG. Thus, the lowest value of

ΔE_b was obtained for VmG (-12.5 kcal/mol), even though this energetic parameter is significantly higher for the other four Py···mG complexes, *i.e.* it ranges from -7.5 to -8.7 kcal/mol. This feature indicates that Py shows a very high affinity but also a significant specificity towards mG. On the other hand, the values of ΔE_b predicted for Py···mC and Py···mA, which ranges from -8.3 to -9.7 kcal, reveals that the affinity of Py towards such two nucleic acid bases is similar. Thus, the value of ΔE_b calculated for the lowest energy complex of each class (IIImC and IImA) differs by only 0.4 kcal/mol, whereas these differences are smaller for the other complexes. Finally, the ΔE_b calculated for Py···mT complexes, which range from –6.6 to –7.7 kcal/mol, clearly reflects the lowest but non-negligible affinity.

Table 3 includes the distortion energy (ΔE_d), which corresponds to the repulsive energy contribution associated to the perturbation of the equilibrium parameters (structural changes) of both the isolated Py and the isolated DNA caused by the interactions in the complex.[41,42] As can be seen, ΔE_d is relatively small for all the Py···mNA complexes, *i.e.* it ranges from 0.2 to 0.7 kcal/mol, with exception of VmG, in which ΔE_d is 1.5 kcal/mol. The large value found for the latter complex gives a measure of the remarkable strength of the interaction between Py and mG when they are appropriately arranged.

The overall of the results reported in this section are fully consistent with recent experimental evidences. Thus, it was experimentally found that polypyrrole is able to bind with both plasmid and double-helical DNA forming stable adducts.[7-15] It is worth noting that the formation of polymer:DNA complexes has been typically attributed to the positive charges of the doped conducting polymers. Thus, it has been assumed that the conducting polymer can interchange its negatively charged dopant molecules easily with other negatively charges species, including DNA. However, we recently found that

the affinity of charged polypyrrole towards DNA is higher than that of a doped copolymer formed by pyrrole and N-hydroxypropypyrrole, poly(Py-*co*-NPrOHPy), with molar ratio 25:75.[22] Thus, although both polypyrrole and poly(Py-*co*-NPrOHPy) are charged systems, which explain the affinity showed by the two systems towards DNA, the ability to bind DNA bases through hydrogen bonding interactions is significantly higher for the former. Thus, although the hydroxyl groups of poly(Py-*co*-NPrOHPy) are also able to act as hydrogen bonding donors, they are relatively far from the polymer chain perturbing the formation of interactions with DNA, as was recently evidenced.[50] Moreover, the affinity of some doped polythiophene derivatives without hydrogen bonding donor groups, *e.g.* poly(3-methyl-thiophene), towards plasmid DNA was also remarkably smaller than that of polypyrrole.[10] The behavior of polypyrrole is consistent with the strength of the interactions reported in this work for Py···mNA complexes.

On the other hand, digestion experiments with restriction enzymes indicated that polypyrrole form specific interactions with well-defined nucleotide sequences protecting DNA from enzymatic digestion, while the protection imparted by poly(Py-*co*-NPrOHPy) is significantly smaller.[22] These observations are supported by the remarkable stability and the low value of ΔE_b showed by the VmG complex. Thus, our theoretical calculations indicate that, when the relative disposition between the two interacting molecules is appropriate, Py prefers G to A, C and T.

Thiophene···Nucleic Acid Complexes. The distribution of the 11 minimum energy complexes found for Th···mNA, which are represented in Figure 3, is as follows: 3 (Th···mA), 4 (Th···mG), 3 (Th···mC) and 1 (Th···mT). The nomenclature used to label these minima is identical to that used above for Py···mNA complexes. Table 4 lists the values of $\Delta E_{r,g}$ and $\Delta G_{r,g}$ calculated using the MP2 method combined with different

basis sets. As can be seen, these energetic parameters depend strongly on the size of the basis set. This is because, in general, Th···mNA complexes are stabilized by N-H···π interactions between the N-H groups of the DNA bases and the π cloud of the Th ring rather than by intermolecular hydrogen bonds. Indeed, the latter interaction was identified in only two complexes, IIImG and IIImC, the geometric parameters associated to the (mNA)N-H···S(Th) interaction being poor in both cases (Figure 3). It is worth noting that large basis sets are required to describe N-H···π interactions satisfactorily.

The molecular geometries of the three Th···mA complexes are relatively similar: in all three cases the π cloud of the Th ring interacts with the exocyclic $–NH_2$ group of mA. As a consequence, the values of $\Delta G_{r,g}$ predicted for three complexes differ by less than 1 kcal/mol, ImA being the lowest energy minimum. On the other hand, analyses of the results obtained for Th···mG complexes indicates that IIImG, which shows a hydrogen bond, only differs from IVmG in the arrangement of the Th ring, *i.e.* in the latter the Th ring is rotated by about 70º with respect to the axis defined by the exocyclic –N-H group of mG. This difference precludes the formation of the intermolecular hydrogen bond in IVmG that becomes 0.8 kcal/mol unfavored with respect to IIImG. Structures ImG and IImG, which are isoenergetic, are the Th···mG complexes of lowest energy. Again, these structures only differ in the relative orientation of the Th ring with respect to the axis defined by the N2-H bond of mG, two N-H···π interactions being detected in each one. Examination to three complexes found for Th···mC evidences a significant difference among them: the two species interact through a N-H···π, a π-stacking and an intermolecular hydrogen bond in ImC, IImC and IIImC, respectively. Interestingly, the lowest energy minimum corresponds to IImC, the

other two complexes being unfavored by more than 2 kcal/mol. Finally, the only Th···mT complex found in this work is stabilized by a N-H···π interaction.

The values of ΔG_{sol} and $\Delta G_{r,s}$ calculated for the 11 Th···mNA complexes obtained in the gas-phase are listed in Table 5. Results state the following order of solvation: Th···mT < Th···mA < Th···mC < Th···mG, which is similar to obtained for Py···mNA complexes, *i.e.* they only differ in the relative order of complexes involving mT and mA. On the other hand, examination of the $\Delta G_{r,sol}$ values reveals that no significant change is induced by the solvent in the relative stability order of Th···mA and Th···mC complexes. Within this context, the most noticeable result corresponds to the strong destabilization of IIImC when the polarity of the environment increases, *i.e.* this complex becomes 3.3 kcal/mol less favored in aqueous solution than in the gas-phase. Regarding to Th···mG complexes, solvent effects produce drastic changes in the relative energy order of the four minima. Thus, ImG and IImG, which were the most favored in the gas phase, become the less stable in both chloroform and water solution. In contrast, IIImG and IvmG are about 2 kcal/mol more stable in solution than in the gas phase, the former being the global minimum in the two solvents.

The ΔE_b calculated at different levels of theory for the 11 Th···mNA complexes are displayed in Table 6. Again the lowest value of ΔE_b corresponds to the lowest energy complex of each family. On the other hand, it is worth noting that the magnitude of ΔE_b is drastically affected by the size of the basis set, as was also observed for $\Delta E_{r,g}$. Thus, comparison between results obtained for Py···mNA and Th···mNA (Tables 3 and 6, respectively) indicates that the influence of the basis set is significantly more important for the description of the N-H···π interaction than for the hydrogen bond. This feature is also clearly reflected in Table 6 for ImC and IIImC, which are stabilized by a N-H···π and a hydrogen bonding interaction, respectively. As can be seen, the ΔE_b of the first

complex decreases 30% (from –2.8 to –4.0 kcal/mol) when the basis set grows from the 6-31G(d) to the 6-311++G(d,p), while this reduction is only 6% (from –4.5 to –4.8 kcal/mol) for the latter complex.

The values of ΔE_b obtained for Th···mNA complexes reveals that the strength of the intermolecular interaction varies in the following order: Th···mT < Th···mA ≈ Th···mC < Th···mG, which is identical to that obtained for Py···mNA complexes. However, ΔE_b values are significantly lower for Py···mNA than for Th···mNA indicating that DNA bases prefer the interaction with Py. For example, the lowest and highest values of ΔE_b obtained for Py···mNA complexes,–12.5 (VmG) and –6.6 kcal/mol (IImT), respectively, are about two times lower than those obtained for Th···mNA, -6.8 (ImG and IImG) and –3.7 (IImA). Additionally, these values also reflect that the range of variation of ΔE_b is also significantly lower for Th···mNA complexes (3.1 kcal/mol) than for Py···mNA (5.9 kcal/mol), which is consistent with a lack of clear specificity in the former ones, *i.e.* the affinity of Th towards the four DNA bases is similar. On the other hand, the values of ΔE_d are lower for Th···mNA than for Py···mNA (Table 3), this feature being consistent with the relative strength of the intermolecular interactions that contribute to the stability of the complexes, *i.e.* N-H···π and hydrogen bond, respectively.

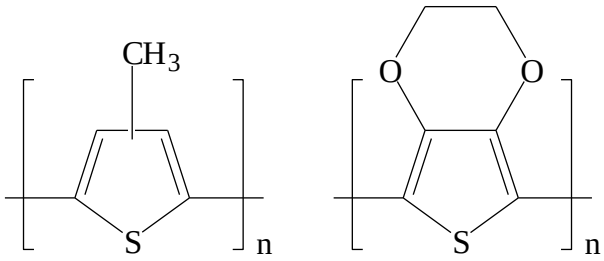
The results obtained for Th···mNA are in excellent agreement with our recently reported experimental data.[10,22] Thus, we observed that the affinity of poly(thiophene) derivatives, *e.g.* pol(3-methyltiophene), towards plasmid DNA is significantly lower than that of polypyrrole, even though the doping level of the two conducting polymers was similar. Specifically, gel electrophoresis assays were performed for a series of polypyrrole:DNA and poly(3-methylthiophene):DNA complexes considering different polymer:DNA mass ratios. For the latter complex, the bands of DNA, *i.e.* those

associated to the typical mixture of supercoiled form (I) and singly nicked form (II), were clearly identified until the concentration of polymer in the ratio increases to 100:1. In opposition, the intensity of form I and form II is weak even for 1:1 polypyrrole:DNA mass ratios, which evidenced the remarkable tendency of the homopolymer to bind DNA. The strength of electrostatic interactions between the polymer and DNA is expected to be proportional to the doping level of the former, while hydrophobic interactions are expected to be stronger for poly(3-methylthiophene) than for polypyrrole. Therefore, the different affinities observed for such two polymers are probably related with the ability of polypyrrole to form hydrogen bonds. Consistently, the ΔE_b obtained for Py···mNA is about two times more attractive than that of Th···mNA, this difference being due to the hydrogen bonding interactions that stabilize the former complexes.

On the other hand, poly(3-methylthiophene) was considerably less able to prevent plasmid DNA enzymatic digestion than polypyrrole. This observation is in agreement with the $\Delta E_{r,g}$ and ΔE_b values calculated for Py···mNA and Th···mNA complexes. Thus, the preferences of Py by each of the four DNA bases are clearly marked while those of Th are much less defined. Moreover, the preferred relative arrangement of the two interacting molecules is clearly defined for each type of Py···mNA complex, while this does not occur for the four types of Th···mNA complexes.

Finally, it should be noted that poly(3,4-ethylenedioxythiophene), with a dioxane ring fused onto each thiophene ring (Scheme 1), shows higher affinity and specificity for plasmid DNA than poly(3-methylthiophene).[10] This should be attributed to the oxygen atoms of the dioxane ring, which are more effective interaction sites than the sulfur of thiophene, *i.e.* the ability to act as hydrogen bonding acceptor is significantly higher for oxygen than for sulfur.[42] This feature again suggests that hydrogen bond is

important to explain the observed interaction patterns between conducting polymers and DNA.



Scheme 1: Poly(3-methylthiophene) (left) and poly(3,4-ethylenedioxythiophene) (right)

Conclusions

In this work we examined the ability of Py and Th, which are the monomeric units of polypyrrole and polythiophene, respectively, to interact with DNA bases through hydrogen bonding interactions. Results evidenced that Py is a strong proton donor, being able to form very stable complexes with mA, mG, mC and mT. Moreover, differences among ΔE_b values revealed that the specificity of Py to methylated nucleic acids is very remarkable. Thus, the highest and lowest affinities were for mG (VmG, ΔE_b= -12.5 kcal/mol) and mT (IImT, ΔE_b= -6.6 kcal/mol), respectively. These results are fully consistent with the high affinity of polypyrrole towards plasmid DNA as well as with the ability of this polymer to form specific interactions with well-defined nucleotide sequences protecting DNA from enzymatic digestion.

On the other hand, the sulfur of Th is a very weak proton acceptor as was revealed by the fact that no hydrogen bonded complex was formed with mA, mC and mT. Thus, Th···mNA complexes were typically stabilized by N-H···π interactions, which are about two times weaker than hydrogen bonds. Furthermore, differences between the binding energies of the different kind of complexes were significantly smaller than for Py···mNA: the highest and lowest ΔE_b values were –6.8 (ImG and IImG) and –3.8

kcal/mol (ImT), respectively. The overall of these results allow to explain the low affinity of poly(3-methylthiophene) towards DNA, as well as the lack of specificity in the interaction between the two macromolecules once the adducts have been formed.

Finally, it should be mentioned that MP2 calculations become less accourate as the basis set size is increased. Accordingly, results reported in this work at the MP2/6-311++G(d,p) level probably correspond to the upper bound to the true binding strength.

Acknowledgements

This work was supported by MEC and FEDER funds with Grant MAT2006-04029. DZ thanks financial support from the "Ramon y Cajal" program of the Spanish "*Ministerio de Educación y Ciencia*" (MEC).

Supporting Information

Pictures of all the starting geometries prepared for Py···mNA and Th···mNA complexes. This material is available free of charge via the Internet at http://pubs.acs.org.

References

1. Chriswanto, H.; Wallace, G. G. *J. Liq. Chromatogr.* **1996**, *19*, 2457.

2. Miller, L. L.; Zinger, B.; Zhou, O. X. *J. Am. Chem. Soc.* **1987**, *109*, 2267.

3. Guo, H.; Knobler, C. M.; Kaner, R. B. *Synt. Met.* **1999**, *101*, 44.

4. Englebienne, P. *J. Mater. Chem.***1999**, *9*, 1043.

5. Kros, A.; van Howell, S. W. F. M.; Sommerdijk, N. A. S. M.; Nolte, R. J. M. *Adv. Mater.* **2001**, *13*, 1555.

6. Khan, G. F.; Wernet, W. *Thin Solid Films* **1997**, *300*, 265.

7. Azioune, A.; Chehimi, M. M.; Miksa, B.; Basinska, T.; Slomkowski, S. *Langmuir* **2002**, *18*, 1150.

8. Misoska, V.; Prize, W. E.; Ralph, S.; Wallace, G. G. *Synth. Met.* **2001**, *123*, 279.

9. Wang, J.; Jiang, M. *Electroanalysis* **2001**, *13*, 537.

10. Ocampo, C.; Armelin, E.; Estrany, F.; del Valle, L. J.; Oliver, R.; Sepulcre, F.; Alemán, C. *Macromol. Mat. Engin.* **2007**, *292*, 85.

11. Ho, H.-A.; Boissinot, M.; Bergeron, M. G.; Corbeil, G.; Dore, K.; Boudreau, D.; Leclerc, M. *Angew. Chem. Int. Ed.* **2002**, *41*, 1548.

12. Minehan, D. S.; Marx, K. A.; Tripathy, S. K. *Macromolecules* **1994**, *27*, 777.

13. Bae, A.-H.; Hatano, T.; Numata, M.; Takeuchi, M.; Shinkai, S. *Macromolecules* **2005**, *38*, 1609.

14. Peng, H.; Zhang, L.; Spires, J.; Soeller, C.; Travas-Sejdic, J. *Polymer* **2007**, *48*, 3413.

15. Yamamoto, T.; Shimizu, T.; Kurokawa, E. *React. Funct. Polym.* **2000**, *43*, 79.

16. Liu, B.; Bazán, G. C. *Chem. Mater.* **2004**, *16*, 4467.

17. Schmidt, C. E.; Shastri, V.; Vacanti, J. P.; Langer, R. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 8948.

18. Kotwal, A.; Schmidt, C. E. *Biomaterials* **2001**, *22*, 1055.

19. del Valle, L. J.; Aradilla, D.; Oliver, R.; Sepulcre, F.; Gamez, A.; Armelin, E.; Alemán, C.; Estrany, F. *Eur. Polym. J.* **2007**, *43*, 2342.

20. Wallace, G. G.; Kane-Maguire, L. A. P. *Adv. Mat.* **2002**, *14*, 953.

21. Adhikari, B.; Majumdar, S. *Prog. Polym. Sci.* **2004**, *29*, 699.

22. Pfeiffer, P.; Armelin, E.; Estrany, F.; del Valle, L.; Cho, L. Y.; Alemán, C. *J. Polym. Res.* **2007**, in press (DOI: 10.1007/s10965-007-9162-2).

23. Jurecka, P.; Hobza, P. *J. Am. Chem. Soc.* **2003**, *125*, 15608-15613.

24. Sponer, J.; Jurecka, P.; Hobza, P. *J. Am. Chem. Soc.* **2004**, *126*, 10142-10151.

25. Dabkowska, I.; Jurecka, P.; Hobza, P. *J. Chem. Phys.* **2005**, *122*, 204322-204331

26. Gorb, L.; Podolyan, Y.; Dziekonski, P.; Sokalski, W. A.; Leszczynski, J. *J. Am. Chem. Soc.* **2004**, *126*, 10119-10129.

27. Sponer, J. E.; Spackova, N.; Kulhanek, P.; Leszczynski, J.; Sponer, J. *J. Phys. Chem. A* **2005**, *109*, 2292-2301.

28. Orozco, M.; Hernández, B.; Luque, J. *J. Phys. Chem. B* **1998**, *102*, 5228-5233.

29. Blas, J. R.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2004**, *126*, 154-164.

30. Barsky, D.; Colvin, M. E. *J. Phys. Chem. A* **2000**, *104*, 8570-8576.

31. Sahu, P. K.; Kuo, C.-W.; Lee, S.-L. *J. Phys. Chem. B* **2007**, *111*, 2991-2998.

32. Casanovas, J.; Alemán, C. *J. Phys. Chem. C* **2007**, *111*, 4823.

33. Brédas, J. L. *J. Chem. Phys.* **1985**, *82*, 3808.

34. Hernandez, V.; Castiglioni, C.; Del Zopo, M.; Zerbi, G. *Phys. Rev. B* **1994**, *50*, 9815.

35. Gaussian 03, Revision B.02, Frisch, M. J. et al. Gaussian, Inc., Pittsburgh PA, 2003.

36. Møller, C.; Plesset, M. *Phys. Rev.* **1934**, *46*, 618.

37. Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1972**, *28*, 203.

38. McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639.

39. Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265.

40. Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.

41. Alemán, C.; Vega, M. C.; Tabernero, L.; Bella, J. *J. Phys. Chem.* **1996**, *100*, 11480.

42. Alemán, C. *J. Phys. Chem. A* **2001**, *105*, 6717.

43. Miertus, M.; Scrocco, E. ; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117.

44. Miertus, S.; Tomasi, J. *Chem. Phys.* **1982**, *65*, 239.

45. Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Phys. B* **1998,***102,* 3257.

46. Jang, Y. H.; Goddard III, W. A.; Noyes, K. T. ; Sowers, L. C. ; Hwang, S.; Chung, D. S. *J. Phys. Chem. B* **2003**, *107*, 344.

47. Iribarren, J. I.; Casanovas, J.; Zanuy, D.; Alemán, C. *Chem. Phys.* **2004**, *302*, 77.

48. Alemán, C.; Jiménez, A. I.; Cativiela, C.; Pérez, J. J.; Casanovas, J. *J. Phys. Chem. B* **2002**, *106*, 11859-11858.

49. Casanovas, J.; Jiménez, A. I.; Cativiela, C.; Pérez, J. J.; Alemán, C. *J. Phys. Chem. B* **2006**, *110*, 5762-5766.

50. Wellenzohn, B.; Loferer, M. J.; Trieb, M.; Rauch, C.; Winger, R. H.; Mayer, E.; Liedl, K. R. *J. Am. Chem. Soc.* **2003**, *125*, 1088.

Captions to Figures

Figure 1. Interaction sites of the methylated analogues of the four DNA bases. The direction of the arrows is from hydrogen bonding donor to hydrogen bonding acceptor (arrows pointing to the methylated base correspond to interaction sites with pyrrole, while arrows going out correspond to interaction sites with thiophene). The number of starting geometries considered for calculations on complexes formed between each methylated base and pyrrole (Py) or thiophene (Th) is indicated in parenthesis.

Figure 2. Geometries of 17 pyrrole···methylated nucleic acid complexes obtained from full optimisation at MP2/6-31G(d) level of theory. Pink dashed lines and green arrows indicate hydrogen bonding and N-H···π interactions, respectively. Characteristic hydrogen bonding parameters, H···acceptor distance (in Å) and N-H···acceptor angle (in degrees) are indicated.

Figure 3. Geometries of 11 thiophene···methylated nucleic acid complexes obtained from full optimisation at MP2/6-31G(d) level of theory. Green arrows and pink dashed lines indicate N-H···π and hydrogen bonding interactions, respectively. Characteristic hydrogen bonding parameters, H···acceptor distance (in Å) and N-H···acceptor angle (in degrees) are indicated.

Table 1. Relative energy ($\Delta E_{r,g}$; in kcal/mol) and free energy ($\Delta G_{r,g}$; in kcal/mol) in the gas-phase calculated for pyrrole···nucleic acid complexes at different levels of theory.[a]

	MP2/6-31G(d)		MP2/6-311G(d,p)		MP2/6-311++G(d,p)	
	$\Delta E_{r,g}$	$\Delta G_{r,g}$	$\Delta E_{r,g}$	$\Delta G_{r,g}$	$\Delta E_{r,g}$	$\Delta G_{r,g}$
			Py···mA			
ImA	1.1	1.0	1.3	1.2	1.5	1.4
IImA	0.0	0.0	0.0	0.0	0.0	0.0
IIImA	1.5	1.5	1.3	1.2	0.9	0.8
IVmA	1.1	0.5	1.3	0.7	1.5	0.9
			Py···mG			
ImG	5.5	3.6	5.2	3.3	5.3	3.4
IImG	4.5	2.5	4.0	2.0	5.0	3.0
IIImG	5.5	3.2	4.8	2.5	5.1	2.9
IVmG	4.6	2.2	4.5	2.1	5.3	2.8
VmG	0.0	0.0	0.0	0.0	0.0	0.0
			Py···mC			
ImC	2.1	1.5	1.7	1.1	1.6	0.9
IImC	0.3	0.0	0.5	0.2	0.7	0.3
IIImC	0.0	0.0	0.0	0.0	0.0	0.0
			Py···mT			
ImT	0.4	2.2	0.0	1.1	0.0	0.0
IImT	0.5	0.6	1.2	0.7	2.4	0.7
IIImT	1.2	0.5	1.4	0.0	2.5	0.0
IVmT	0.0	0.0	0.9	0.2	2.0	0.2
VmT	1.0	0.8	1.3	0.4	2.4	0.3

[a] $\Delta E_{r,g}$ and $\Delta G_{r,g}$ are relative to the most stable complex of each class. All geometry optimizations were performed at the MP2/6-31G(d) level. Values of $\Delta G_{r,g}$ at the MP2/6-311G(d,p) and MP2/6-311++G(d,p) levels were calculated by adding the ZPVE and thermodynamic corrections calculated at the MP2/6-31G(d) level to the corresponding electronic energies.

Table 2. Free energy of solvation[a] (ΔG_{sol}; in kcal/mol) and relative conformational free energy[b] ($\Delta G_{r,s}$; in kcal/mol) in chloroform and aqueous solutions for pyrrole···nucleic acid complexes. Calculated conformational free energy in the gas-phase[c] ($\Delta G_{r,g}$; in kcal/mol) are also displayed for comparison.

	ΔG_{sol}		$\Delta G_{r,g}$	$\Delta G_{r,s}$	
	Chloroform	Water	Gas-phase	Chloroform	Water
			Py···mA		
ImA	-3.1	-3.8	1.4	2.6	3.3
IImA	-3.5	-4.2	0.0	0.8	1.5
IIImA	-5.1	-6.6	0.8	0.0	0.0
IvmA	-3.1	-3.8	0.9	2.1	2.8
			Py···mG		
ImG	-12.8	-19.2	3.4	0.7	0.1
IimG	-13.1	-18.8	3.0	0.0	0.0
IIImG	-12.3	-18.5	2.9	0.7	0.2
IvmG	-12.9	-18.4	2.8	0.0	0.2
VmG	-7.9	-11.4	0.0	2.2	4.4
			Py···mC		
ImC	-10.6	-15.4	0.9	0.0	0.0
IImC	-8.0	-12.0	0.3	2.0	2.8
IIImC	-7.5	-11.4	0.0	2.2	3.1
			Py···mT		
ImT	-5.4	-7.1	0.0	0.2	0.6
IImT	-2.4	-3.8	0.7	3.9	4.6
IIImT	-5.6	-7.7	0.0	0.0	0.0
IVmT	-2.2	-3.2	0.2	3.7	4.6
VmT	-5.0	-7.9	0.3	0.0	0.1

[a] Calculations in solution were performed using the PCM model in the MP2 framework. The geometries optimised in the gas-phase at the MP2/6-31G(d) level were employed in PCM calculations. [b] Values of $\Delta G_{r,s}$ were estimated by adding the ΔG_{sol} to the $\Delta G_{r,g}$ calculated at the MP2/6-311++G(d,p) level. [c] Values calculated at the MP2/6-311++G(d,p) level (see Table 1).

Table 3. Binding energy (ΔE_b; in kcal/mol) calculated at different levels of theory and distortion energy[a] (ΔE_d; in kcal/mol) for pyrrole···nucleic acid complexes.

	ΔE_b			ΔE_d
	MP2/6-31G(d)	MP2/6-311G(d,p)	MP2/6-311++Gd,p)	
		Py···mA		
ImA	-8.1	-8.2	-8.3	0.4
IImA	-9.0	-9.2	-9.3	0.7
IIImA	-7.8	-8.4	-8.6	0.5
IVmA	-8.1	-8.1	-8.3	0.4
		Py···mG		
ImG	-7.2	-7.5	-7.5	0.4
IImG	-8.6	-8.7	-8.7	0.2
IIImG	-7.9	-8.1	-8.3	0.3
IVmG	-8.4	-8.4	-8.5	0.2
VmG	-11.9	-12.2	-12.5	1.5
		Py···mC		
ImC	-7.8	-8.1	-8.4	0.4
IImC	-9.0	-9.0	-9.2	0.5
IIImC	-9.5	-9.7	-9.7	0.6
		Py···mT		
ImT	-6.0	-7.1	-7.7	0.6
IImT	-6.5	-6.4	-6.6	0.4
IIImT	-6.2	-6.6	-6.9	0.3
IVmT	-6.8	-6.7	-6.9	0.5
VmT	-6.4	-6.7	-7.0	0.3

[a] ΔE_d was calculated at the MP2/6-31G(d) level.

Table 4. Relative energy ($\Delta E_{r,g}$; in kcal/mol) and free energy ($\Delta G_{r,g}$; in kcal/mol) in the gas-phase calculated for thiophene···nucleic acid complexes at different levels of theory.[a]

	MP2/6-31G(d)		MP2/6-311G(d,p)		MP2/6-311++G(d,p)	
	$\Delta E_{r,g}$	$\Delta G_{r,g}$	$\Delta E_{r,g}$	$\Delta G_{r,g}$	$\Delta E_{r,g}$	$\Delta G_{r,g}$
			Th···mA			
ImA	0.0	0.0	0.0	0.0	0.0	0.0
IImA	1.0	0.3	1.1	0.4	1.7	1.0
IIImA	0.9	0.2	1.1	0.4	1.6	0.9
			Th···mG			
ImG	0.0	0.0	0.0	0.0	0.0	0.0
IImG	0.0	0.0	0.0	0.0	0.0	0.0
IIImG	3.4	2.0	3.6	2.2	3.5	2.1
IVmG	3.1	2.7	3.3	2.9	3.3	2.9
			Th···mC			
ImC	2.3	1.6	1.3	1.1	2.3	2.1
IImC	1.1	0.7	0.0	0.0	0.0	0.0
IIImC	0.0	0.0	0.2	0.6	2.4	2.8
			Py···mT			
ImT	-	-	-	-	-	-

[a] $\Delta E_{r,g}$ and $\Delta G_{r,g}$ are relative to the most stable complex of each class. All geometry optimizations were performed at the MP2/6-31G(d) level. Values of $\Delta G_{r,g}$ at the MP2/6-311G(d,p) and MP2/6-311++G(d,p) levels were calculated by adding the ZPVE and thermodynamic corrections calculated at the MP2/6-31G(d) level to the corresponding electronic energies.

Table 5. Free energy of solvation[a] (ΔG_{sol}; in kcal/mol) and relative conformational free energy[b] ($\Delta G_{r,s}$; in kcal/mol) in chloroform and aqueous solutions for thiophene···nucleic acid complexes. Calculated conformational free energy in the gas-phase[c] ($\Delta G_{r,g}$; in kcal/mol) are also displayed for comparison.

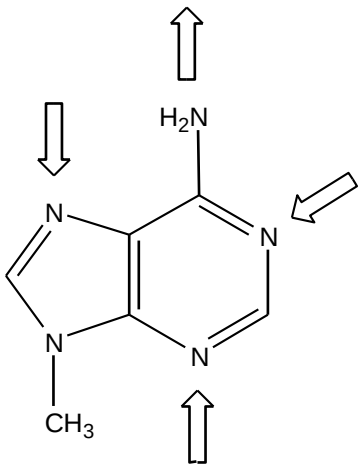
	ΔG_{sol}		$\Delta G_{r,g}$	$\Delta G_{r,s}$	
	Chloroform	Water	Gas-phase	Chloroform	Water
			Th···mA		
ImA	-3.4	-4.1	0.0	0.0	0.0
IImA	-3.4	-4.3	1.0	1.0	0.8
IIImA	-3.3	-4.2	0.9	1.0	0.8
			Th···mG		
ImG	-9.6	-14.0	0.0	1.2	3.1
IImG	-9.6	-13.9	0.0	1.3	3.2
IIImG	-13.0	-19.3	2.1	0.0	0.0
IVmG	-13.0	-19.1	2.9	0.8	0.9
			Th···mC		
ImC	-11.0	-16.0	2.1	1.3	1.4
IImC	-10.2	-15.3	0.0	0.0	0.0
IIImC	-7.7	-12.1	2.8	5.3	6.1
			Th···mT		
ImT	-2.3	-3.1	-	-	-

[a] Calculations in solution were performed using the PCM model in the MP2 framework. The geometries optimised in the gas-phase at the MP2/6-31G(d) level were employed for PCM calculations. [b] Values of $\Delta G_{r,s}$ were estimated by adding the ΔG_{sol} to the $\Delta G_{r,g}$ calculated at the MP2/6-311++G(d,p) level. [c] Values calculated at the MP2/6-311++G(d,p) level (see Table 1).

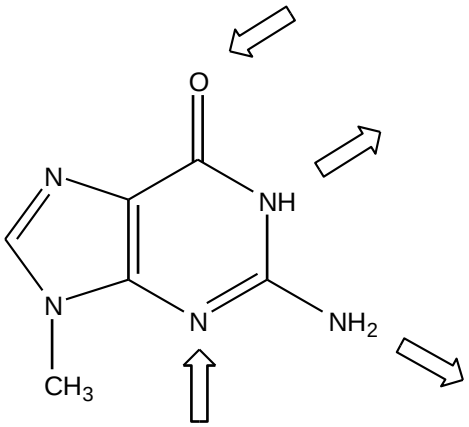
Table 6. Binding energy (ΔE_b; in kcal/mol) calculated at different levels of theory and distortion energy[a] (ΔE_d; in kcal/mol) for thiophene···nucleic acid complexes.

	ΔE_b			ΔE_d
	MP2/6-31G(d)	MP2/6-311G(d,p)	MP2/6-311++Gd,p)	
		Th···mA		
ImA	-3.5	-4.0	-4.5	0.2
IImA	-3.2	-3.5	-3.7	0.1
IIImA	-2.8	-3.2	-3.6	0.1
		Th···mG		
ImG	-5.7	-6.1	-6.8	0.4
IImG	-5.7	-6.1	-6.8	0.4
IIImG	-2.9	-3.3	-3.8	0.1
IVmG	-3.1	-3.5	-4.0	0.1
		Th···mC		
ImC	-2.8	-3.4	-4.0	0.1
IImC	-3.1	-4.0	-4.9	0.1
IIImC	-4.5	-4.5	-4.8	0.1
		Th···mT		
ImT	-2.7	-3.1	-3.8	0.1

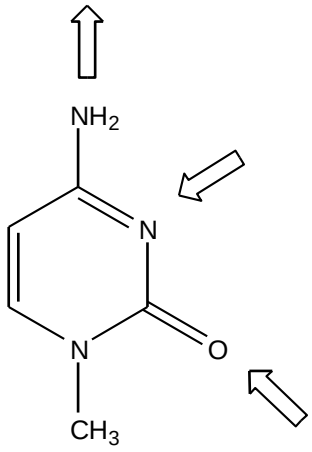
[a] ΔE_d was calculated at the MP2/6-31G(d) level.



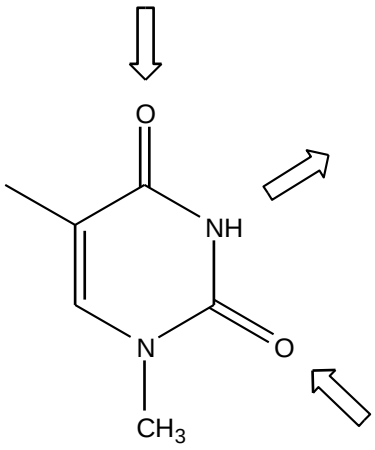
mA (Py-14/Th-8)



mG (Py-8/Th-12)



mC (Py-12/Th-8)



mT (Py-8/Th-4)

Figure 1

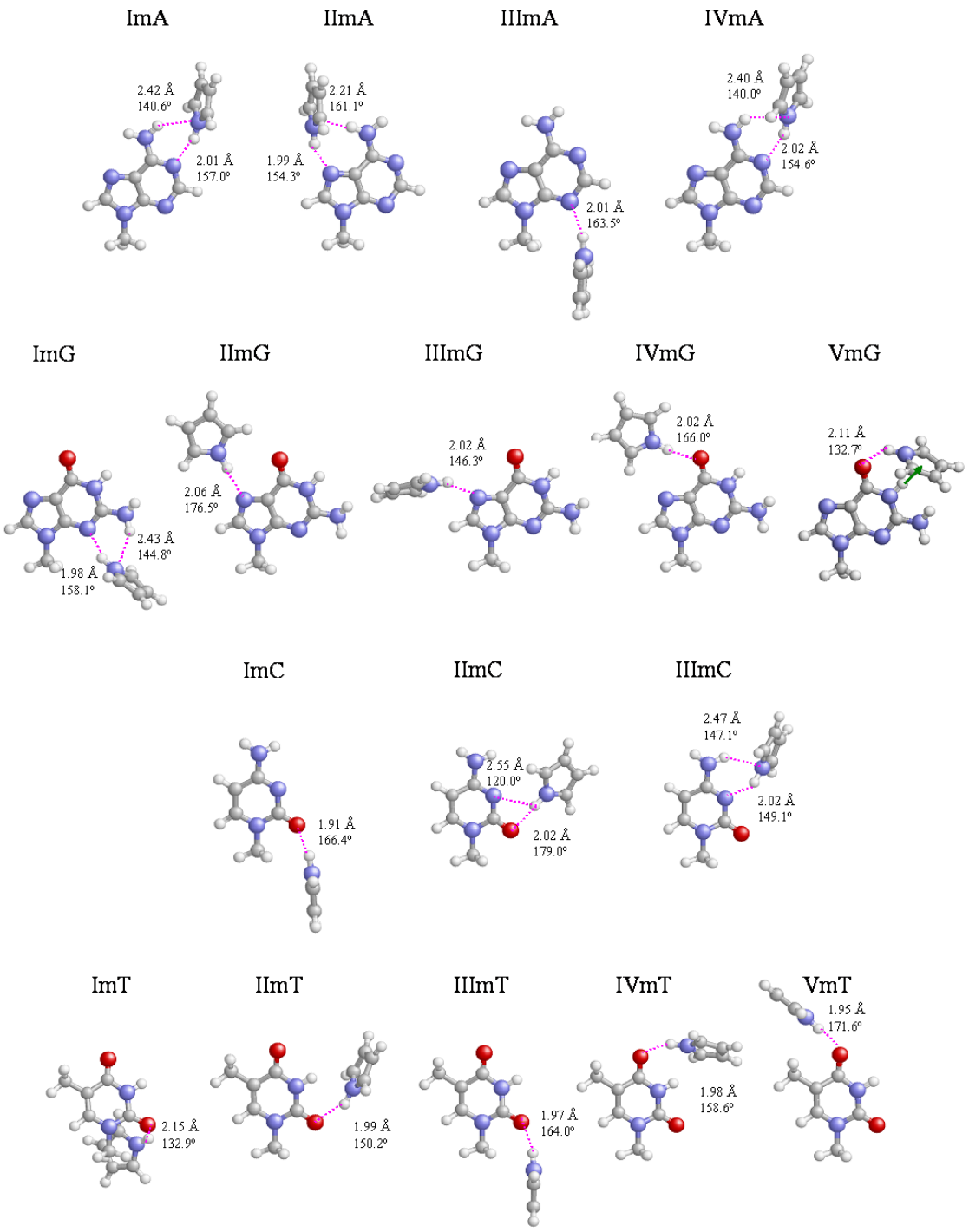

ImA
IImA
IIImA
IVmA
2.42 Å
140.6°
2.01 Å
157.0°
2.21 Å
161.1°
1.99 Å
154.3°
2.01 Å
163.5°
2.40 Å
140.0°
2.02 Å
154.6°
ImG
IImG
IIImG
IVmG
VmG
2.06 Å
176.5°
2.43 Å
144.8°
1.98 Å
158.1°
2.02 Å
146.3°
2.02 Å
166.0°
2.11 Å
132.7°
ImC
IImC
IIImC
1.91 Å
166.4°
2.55 Å
120.0°
2.02 Å
179.0°
2.47 Å
147.1°
2.02 Å
149.1°
ImT
IImT
IIImT
IVmT
VmT
2.15 Å
132.9°
1.99 Å
150.2°
1.97 Å
164.0°
1.98 Å
158.6°
1.95 Å
171.6°

Figure 2

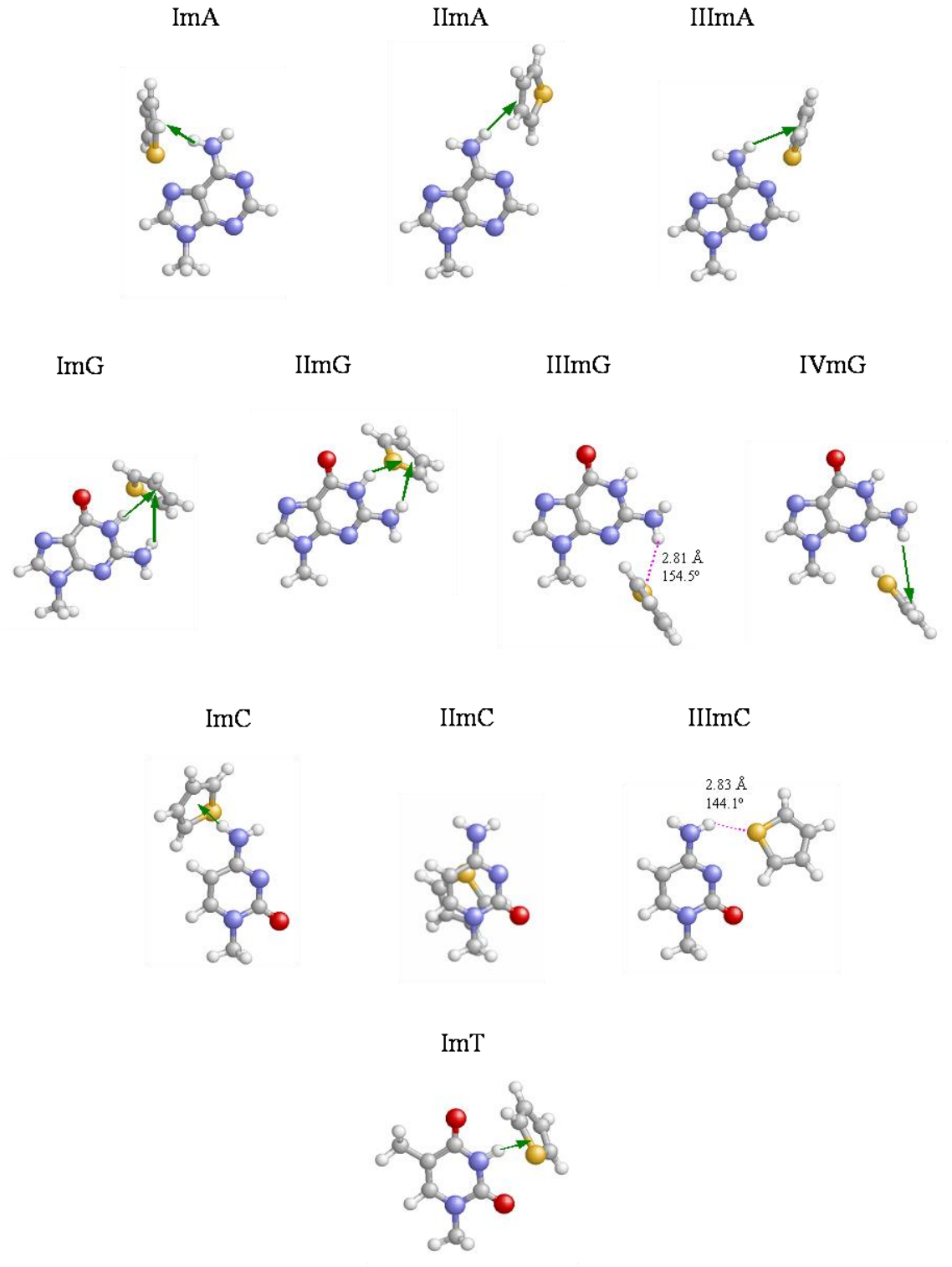

ImA
IImA
IIImA
ImG
IImG
IIImG
IVmG
2.81 Å
154.5°
ImC
IImC
IIImC
2.83 Å
144.1°
ImT

Figure 3