J|A|C|S
ARTICLES
Published on Web 11/22/2002

Binding to an RNA Aptamer Changes the Charge Distribution and Conformation of Malachite Green

Dat H. Nguyen, Steven C. DeFina,† William H. Fink, and Thorsten Dieckmann*

Contribution from the Department of Chemistry, University of California at Davis, One Shields Ave., Davis, California 95616

Received July 10, 2002

Abstract: RNA plays a central role in many biological processes and is therefore an important target for drug development. In recent years an increasing wealth of structural and functional information about RNA−ligand complexes has been obtained using in vitro selected RNAs (aptamers). However, all those studies focused on structure and changes of the nucleic acid and mostly considered the ligand as a rigid target. To develop a detailed picture of ligand structure and dynamics in RNA−small molecule complexes, the malachite green binding aptamer was studied. Isotopically labeled ligand in complex with RNA was analyzed by NMR spectroscopy in solution. The surprisingly asymmetric changes in the ^{13}C chemical shift of the ligand methyl groups indicate that the dye undergoes changes in its conformation and charge distribution upon binding. The role of the RNA electrostatic field in this interaction was explored using ab initio calculations of the ligand structure and charge distribution. The results indicate that the uneven charge distribution in the RNA binding pocket provides a major contribution to the driving force of the ligand structural changes. The observation that not only the RNA adapts to the ligand, in what is called adaptive binding, but that the ligand itself also undergoes conformational changes ("induced fit") is crucial for the rational design of RNA ligands and for understanding the properties of RNA−ligand complexes.

Introduction

Binding and recognition of small molecule ligands by nucleic acids is thought to be based mainly on stacking interactions and hydrogen bonding.[1−4] These types of interactions are crucial for RNA function in nature as well as in the context of RNA as a drug target.[5−8] RNA and DNA molecules that were selected in vitro to specifically recognize and bind a target molecule[9,10] (aptamers) are ideal systems for detailed studies of these interactions. Several structures of such systems have been determined using both X-ray crystallography[11,12] and NMR spectroscopy.[13−15] The structures of RNA complexes with ATP, biotin, or FMN, among others, demonstrated that specificity is most often achieved by specific hydrogen-bonding patterns and stacking of aromatic ring systems in the ligands with bases in the RNA.[1,14,16,17] Another characteristic feature of these complexes is that the RNA binding pocket in its ligand free form is largely unstructured and folding occurs simultaneously with ligand binding. This process has been termed ligand-dependent folding or adaptive binding.[3,15]

The malachite green binding RNA aptamer was identified by in vitro selection with the goal of providing a target for laser-mediated RNA cleavage.[18] The 38-nucleotide consensus sequence binds malachite green (MG) and several related organic dyes with K_D values in the nanomolar range. The structure of the RNA in complex with the dye tetramethylrosamine (TMR) was originally determined by X-ray crystallography.[11] The RNA binding pocket forms a complex three-dimensional structure around the ligand, including two base triples and one base quadruple. The ligand is intercalated between a Watson−Crick base pair (G8-C28) and the base quadruple formed by C7, G24, G29, and A31 (Figure 1). The MG−RNA complex did not produce useable crystals and its structure was determined by NMR spectroscopy in our laboratory.[19] The overall structures of the two complexes are very similar. However, there are some small differences in the stacking arrangement within the ligand

* To whom correspondence should be addressed. E-mail: dieckman@chem.ucdavis.edu

† Current address: Signature Biochem, 1240 S. 47th St., Richmond, CA 94806.

(1) Feigon, J.; Dieckmann, T.; Smith, F. W. *Chem. Biol.* **1996**, *3*, 611−617.
(2) Dieckmann, T.; Butcher, S. E.; Sassanfar, M.; Szostak, J. W.; Feigon, J. *J. Mol. Biol.* **1997**, *273*, 467−478.
(3) Patel, D. J.; Suri, A. K.; Jiang, F.; Jiang, L.; Fan, P.; Kumar, R. A.; Nonin, S. *J. Mol. Biol.* **1997**, *272*, 645−664.
(4) Heus, H. A. *Nat. Struct. Biol.* **1997**, *4*, 597−600.
(5) Ramos, A.; Gubser, C.; Varani, G. *Curr. Opin. Struct. Biol.* **1997**, *7*, 317−323.
(6) Fourmy, D.; Recht, M. I.; Puglisi, J. D. *J. Mol. Biol.* **1998**, *277*, 347−362.
(7) Yoshizawa, S.; Fourmy, D.; Eason, R. G.; Puglisi, J. D. *Biochemistry-USA* **2002**, *41*, 6263−6270.
(8) Lynch, S. R.; Puglisi, J. D. *J. Mol. Biol.* **2001**, *306*, 1037−1058.
(9) Ellington, A. D.; Szostak, J. W. *Nature* **1992**, *355*, 850−852.
(10) Schneider, D.; Gold, L.; Platt, T. *FASEB J.* **1993**, *7*, 201−207.
(11) Baugh, C.; Grate, D.; Wilson, C. *J. Mol. Biol.* **2000**, *301*, 117−128.
(12) Sussman, D.; Nix, J. C.; Wilson, C. *Nat. Struct. Biol.* **2000**, *7*, 53−57.
(13) Macaya, R. F.; Schultze, P.; Smith, F. W.; Roe, J. A.; Feigon, J. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3745−3749.
(14) Jiang, F.; Kumar, R. A.; Jones, R. A.; Patel, D. J. *Nature* **1996**, *382*, 183−186.
(15) Dieckmann, T.; Suzuki, E.; Nakamura, G. K.; Feigon, J. *RNA* **1996**, *2*, 628−640.
(16) Nix, J.; Sussman, D.; Wilson, C. *J. Mol. Biol.* **2000**, *296*, 1235−1244.
(17) Fan, P.; Suri, A. K.; Fiala, R.; Live, D.; Patel, D. J. *J. Mol. Biol.* **1996**, *258*, 480−500.
(18) Grate, D.; Wilson, C. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6131−6136.
(19) Flinders, J.; DeFina, S. C.; Baugh, C.; Wilson, C.; Dieckmann, T. Submitted for publication, 2002.

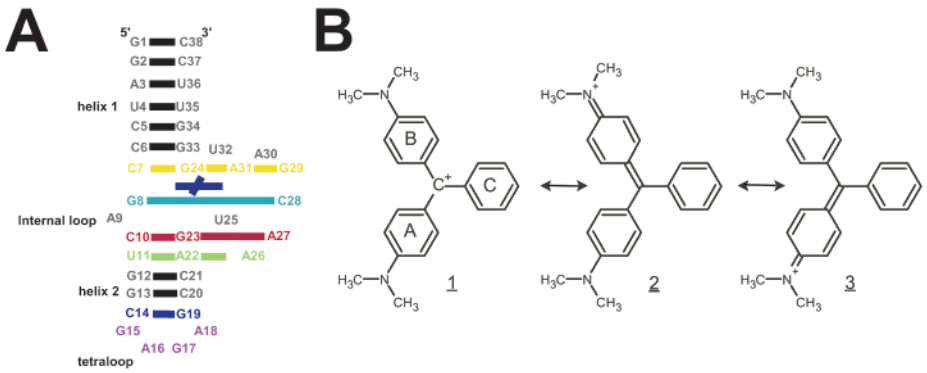

Figure 1. (A) Sequence, secondary structure, and tertiary fold of the TMR−RNA complex. The dark blue cross of bold lines in the center illustrates the position of the ligand. Bold lines connect the nucleotides of the two base triples (red and green), and yellow lines indicate the nucleotides that are part of the base quadruple. (B) Molecular structure of malachite green with resonance structures. The naming scheme for the three rings is indicated in **1**.

binding site. Two more significant differences on the periphery of the molecule are most likely due to crystal contacts in the X-ray structure. During the course of the NMR investigation, the use of specific labeling of the ligand[20] allowed a closer look at its structure and dynamics. The observed changes in ligand ^{13}C chemical shifts suggest that the ligand undergoes significant changes with respect to its charge distribution and/or conformation. Here we report an analysis of the ligand conformation and binding using NMR spectroscopy and ab initio calculations.

Experimental Section

Synthesis of RNA and Ligand. The RNA samples used for NMR spectroscopy were synthesized by in vitro transcription according to established protocols as previously described.[19,21,22] The specifically labeled ligand was synthesized from appropriately labeled precursors.[20]

NMR Spectroscopy. All NMR spectra were acquired on a Bruker DRX 600 MHz spectrometer at 293 or 274 K. NMR samples typically contained 0.5−1.2 mM RNA in complex with a small excess of labeled or unlabeled ligand (RNA:MG ratio ca. 1:1.1).

Calculations. All calculations were performed using the Gaussian 98 and AMBER program packages running on Compaq Alphas with EV67 processors.

Results and Discussion

NMR Spectroscopy of Bound MG. Analysis of the NOESY spectra of RNA−MG complexes clearly indicates that the two *N*-methylated rings in MG become nonequivalent upon binding to RNA, due to their drastically different environments. The two *N*-methylated rings are referred to as A and B in the following discussion (see Figure 1B). Ring A is deeply buried inside the binding pocket, stacked between G24 and C28. Its two *N*-methyl groups and the protons on both sides of the phenyl ring are nonequivalent and show distinctively different NOE patterns at short NOESY mixing time (80 ms). In contrast, ring B shows clear evidence for faster rotation of the phenyl ring and the methyl groups on the NMR time scale. The NOE patterns for both methyl groups and the protons on both sides of the ring are identical, even though their environment in the structure is quite different. Exchange peaks are observed for the methyl resonances and the aromatic protons in the 2,6- and 3,5-positions on ring B. At long NOESY mixing times (250 ms or longer), the signals of A- and B-ring protons show symmetric cross-peak patterns and exchange peaks. This is most likely due to exchange between free and bound MG. The ^{13}C *N*-methyl-labeled MG dye allowed the unambiguous identification of the ^{13}C methyl signals. For the free dye, only one signal is observed for the four methyl groups ($\delta[^{13}C] = 46.2$ ppm). Upon binding, four signals are found: two at a ^{13}C chemical shift of at 40.3 ppm (ring B) and one each at 39.7 and 36.2 ppm for ring A (Figure 2). This upfield shift of the methyl ^{13}C resonances on the A-ring relative to those on ring B is similar to that observed for the C2 in adenines upon protonation of the N1 position[23,24] and could indicate a more positive partial charge of the A-ring nitrogen compared to the B-ring. This suggests that binding to the RNA stabilizes the MG resonance structure in which the positive charge is located on the A-ring nitrogen (**3** in Figure 1B). In this structure one would expect hindered rotation of the A-phenyl ring and *N*-methyl group due to the partial double bond character of the bonds connecting the nitrogen to the ring and the ring to the central carbon. Indeed, the differences in NOE patterns (see above) and proton line widths of the methyl resonances are consistent with this model: the ^{1}H signals of the B methyl groups have much broader lines than those of the A methyls (Figure 2, ca. 43 and 32 Hz, respectively). Both dimethyl-N groups are in slow exchange with respect to the ^{1}H chemical shifts ($\tau^{-1} \ll |\delta_1 - \delta_2|$). The separation of the ^{1}H signals suggests exchange rates smaller than 45 and 222 s^{-1} for the A and B methyl groups, respectively. To further substantiate this interpretation of the NMR spectra, we undertook ab initio electronic structure calculations to gain insight into the underlying physical nature of the molecular interactions between the MG molecule and the RNA aptamer when they form the MG−RNA complex.

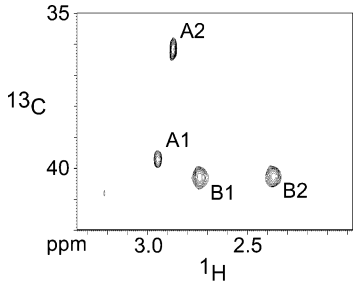

Figure 2. 2D HMQC spectrum of ^{13}C-labeled malachite green in the complex with the RNA aptamer. The spectrum was acquired at 298 K using a sample of 0.8 mM RNA aptamer in the presence of slight excess of MG (RNA:MG ratio of 1:1.1).

(20) DeFina, S. C.; Dieckmann, T. *J. Label Compd Radiopharm.* **2002**, *45*, 241−248.
(21) Nikonowicz, E. P.; Sirr, A.; Legault, P.; Jucker, F. M.; Baer, L. M.; Pardi, A. *Nucleic Acids Res.* **1992**, *20*, 4507−4513.
(22) Batey, R. T.; Inada, M.; Kujawinski, E.; Puglisi, J. D.; Williamson, J. R. *Nucleic Acids Res.* **1992**, *20*, 4515−4523.

Electronic Structure Calculations. The MG−RNA aptamer complex is too large for first-principles electronic structure calculation. Therefore, our strategy is to decompose the physical force of interactions of MG and RNA aptamer into the electrostatic and base-stacking forces, exerted by the RNA on the MG molecule. In doing so, we can account for what force would have what dominant effects on changes of physical properties of the MG molecule upon forming a complex with the RNA aptamer. Here we report the results of the calculations focusing on the effect of the electrostatic forces within the RNA.

Since the electrostatic force is a long-range force and the charge of the RNA is concentrated in the backbone and nearby counterions, we simplified the RNA molecule to a model of point charges representing the backbone RNA structure in order to examine the electrostatic effect of the RNA on the electronic structure of the MG. The backbone is modeled as a series of negative charges representing each phosphate group of the backbone and a series of positive charges representing the counterions. As mentioned before, implicit in this model is that the Coulombic interactions are the dominant influences polarizing the electron distribution of the MG and consequently its ^{13}C-chemical shift; interactions between the bases of the RNA and MG will be considered in future work. Changes between free and bound MG in our model may be attributed to responses of the electron distribution to the external field and to the resulting concomitant changes in molecular conformation.

The X-ray crystal structure of the tetramethylrosamine (TMR)−RNA aptamer complex[11] was used as our structural starting point (PDB accession no. 1F1T), because it gives more precise positions for the phosphate moieties of the RNA than the NMR structure of the MG−RNA aptamer complex (PDB accession no. 1LUY). Moreover, the NMR structure of the MG−RNA aptamer is very similar to the X-ray structure of the RNA aptamer around the TMR, especially with respect to the backbone conformation. Therefore, we feel that the X-ray structure of the RNA aptamer is more suitable for the computational studies.

The crystal structure provided the locations of the backbone phosphorus atoms, but not the positions of necessary counterions. To neutralize the phosphate groups, we first optimized the MG molecule in a vacuum quantum mechanically using the B3LYP density functional Hamiltonian with the 6-31G* basis set. Upon convergence, partial charges of each atom were computed using the Merz−Kollman charge-fitting scheme[25] in order to be consistent with the charges of the RNA bases in the AMBER force field. An AMBER model of the complex was then constructed by superimposing the optimized MG molecule onto the TMR molecule in the RNA pocket using Kabsch's algorithm[26,27] and then removing the TMR molecule from the complex. Counterions were then added to neutralize the new MG−RNA aptamer complex using the LEAP[28] program of the AMBER 5.0 software package.[29] The RNA field surrounding the MG molecule is now modeled as a collection of point charges, where a negative charge is placed on each phosphorus atom and a positive charge is placed on each counterion atom. This assembly of the MG molecule in the pocket of the model RNA point-charge field is now referred to as the MG−RNA complex system, whereas the MG molecule in a vacuum is now referred to as the MG system.

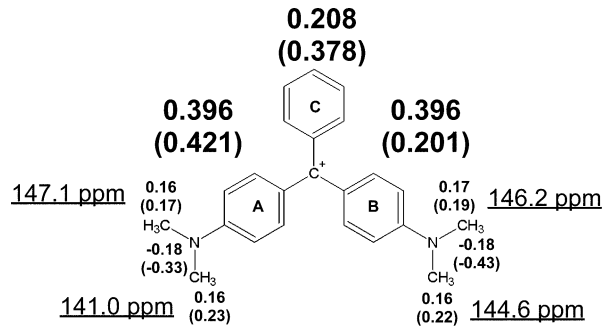

Figure 3. Charge distribution in free and bound malachite green. The numbers in brackets refer to the bound MG and those without to the free MG. The bold numbers above the rings are the charge total for the three rings. The charge of the central carbon has been summed to equal parts into the values for each ring. The underlined numbers are the calculated isotropic shielding constants for the methyl ^{13}C nuclei of the bound MG.

The molecular geometry of MG in both MG−RNA complex and the MG systems was optimized at the B3LYP theory level with the 6-31G* basis set. The MG system optimized completely to the default thresholds of Gaussian 98 (Gaussian, Inc., Pittsburgh PA). The MG−RNA optimization proved more delicate and the optimization was terminated after 100 cycles, when both RMS and Max displacement variables were converged completely to the default thresholds but forces had not dropped below the threshold. We believe the residual forces originated from a small inconsistency in the placement of the MG molecule in the model point-charge field. The resulting geometries were used to compute partial charges for each atom of the MG molecule in both systems using the CHELPG charge-fitting scheme,[30] the ^{13}C chemical shift of the methyl groups using the continuous set of gauge transformations (CSGT),[31,32] and CIS/6-31G* calculations of the first excitation energy.

Figure 3 shows the charge distribution within bound and free malachite green. For the methyl groups the hydrogen charges have been summed into the carbon atom charge. The total ring charges in the bound form are clearly different between the A and B rings. Experimentally, the carbon chemical shifts of the methyl groups on the A ring nitrogen become very different from each other upon complex formation, whereas the methyl carbon chemical shifts on the B ring remain similar.

Figure 3 also shows the correlation between charge distribution and the calculated isotropic shielding constants for each methyl carbon atom attached to the nitrogen atoms of ring A and B in the bound ligand. The isotropic shielding constants reported here are not at the saturated basis set limit, and we regard only the differences as meaningful. As shown, the chemical shift difference between two carbon atoms of *N*-methyl groups on ring A is about 6.1 ppm, whereas on ring B it is

(23) Legault, P.; Pardi, A. *J. Am. Chem. Soc.* **1997**, *119*, 6621−6628.
(24) Ravindranathan, S.; Butcher, S. E.; Feigon, J. *Biochemistry-USA* **2000**, *39*, 16026−16032.
(25) Besler, B. H.; Merz, K. M.; Kollman, P. A. *J Comput Chem* **1990**, *11*, 431−439.
(26) Kabsch, W. *Acta Crystallogr. Sect. A* **1976**, *32*.
(27) Kabsch, W. *Acta Crystallogr. Sect. A* **1978**, *32*, 827−828.
(28) Schafmeister, C.; Ross, W. S.; Romanowski, V. *In LEaP*; University of California: San Francisco, 1995.
(29) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E., III; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, P. K.; Kollman, P. A. *AMBER 5.0*; University of California: San Francisco, 1997.
(30) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361−373.
(31) Keith, T. A.; Bader, R. F. W. *Chem. Phys. Lett.* **1993**, *210*, 223−231.
(32) Cheeseman, J. R.; Trucks, G. W.; Keith, T. A.; Frisch, M. J. *J. Chem. Phys.* **1996**, *104*, 5497−5509.

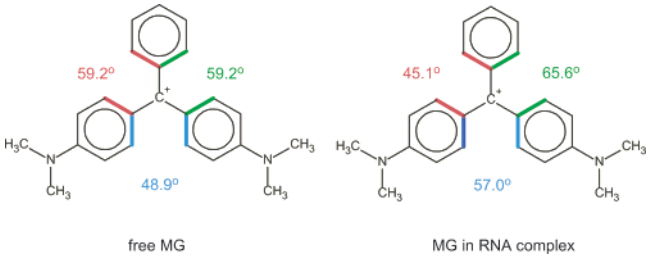

Figure 4. Conformation of free (left) and bound malachite green (right). The dihedral angles between the phenyl rings are indicated by colored numbers.

about 1.6 ppm. Experimentally, no chemical shift difference for the carbon atoms of *N*-methyl groups on ring B is observed and the corresponding chemical shift difference for the carbon atoms of ring A *N*-methyl groups is about 3.7 ppm. However, in addition to the limited basis set and limited extent of correlation in the wave function, the calculated shielding constant does not take the dynamic nature of the system into account. Given the complexity and uncertainty of the calculation, we can arguably say that the computed chemical shift distribution is in general agreement with experiment.

The optimized geometries of the MG molecule in the presence and absence of the RNA electrostatic fields are shown in Figure 4. In addition, we have included two figures (SF1 and SF2) in the Supporting Information that show the optimized valence angles and bond distances of MG in a vacuum and in the charged field representing the RNA. Comparison of the two structures shows that the only significant change between the valence angles and the bond distances is the valence angle at the ring junction carbon where the free MG has a valence angle of 121.2° and the complexed form has a valence angle of 127.6° for the angle between rings A and B. Of most interest are the dihedral angles of the three rings, these being the degrees of freedom most easily distorted by external influences. The inter-ring dihedral angles of the optimized structures are shown in Figure 4, where the effect of the RNA field upon complexation is clearly evident. As shown, ring B becomes less coplanar with respect to rings A and C in the RNA field (i.e., the dihedral angle between rings B and A increases from 48.9° in the isolated molecule to 57.0° in the complex, and the dihedral angle between rings B and C increases from 59.2° to 65.6°). The intervalence angle of rings A and B increases to help accommodate the dihedral angle changes. These dihedral angle changes provide an explanation for a structural difference found while comparing the TMR and MG complex.[19] In the MG complex the G8-C28 base pair is rotated slightly with respect to the base quadruple. This results in ring A being completely intercalated between C28, G24, and G29, whereas ring B is stacked on top of the G29-C7 base pair but has nothing above it. In contrast, TMR, in which rings A and B are coplanar, is completely intercalated in its complex with RNA. The conformational change of the MG ligand explains why the RNA aptamer adapts its structure in this way: By eliminating the stacking interactions on the top of ring B the RNA accommodates for the nonplanar nature of the ligand and the rotation of the B ring.

We undertook single excitation configuration interaction (CIS) calculations of the first excited state of MG using the optimized geometries of both the free and bound forms of MG to obtain a theoretical value for the experimentally observed 14-nm red shift of the visible absorption upon complexations. Of course the absolute values of the excitation energy with this simple treatment of the excited state are badly in error, due to neglect of both orbital relaxation effects and differential correlation; however, one can assume that the difference between the free and bound form excitation energies may give some cancellation of similar error. We obtained a value of 3 nm for the red shift (see Supporting Information). If the red shift is identified with increasing delocalization of the π-system upon complex formation, we see that the calculations suggest about 20% of that may arise from conformational changes bringing rings A and C into greater coplanarity (Figure 4) and thus can be directly attributed to the influence of the RNA electrostatic field.

Conclusions

The results discussed above demonstrate that binding to the RNA significantly affects the electronic structure of the bound dye. This is achieved without any hydrogen-bond interactions between RNA and ligand. The calculations show that the distribution of negative charges within the RNA provides a major contribution to this effect by providing a binding pocket that has greatly different electrostatic environments for the two *N*,*N*-dimethyl groups: Ring A has three phosphate groups within 6.2 Å of the nitrogen with the closest located at a distance of 4.6 Å. Ring B has only two phosphates in the same sphere with the closest at a distance of 5.6 Å. These findings could be of significance with respect to the mechanism by which the bound dye promotes specific RNA cleavage. In a more general picture, the interaction between the RNA electrostatic field and small molecule ligands could also be important for the role of RNA as a catalyst, i.e., the changes in charge distribution can promote changes in the reactivity of the involved covalent bonds. The fact that both RNA and ligand are changing during complex formation and adapt to each other in an "induced fit" manner seems to be crucial for designing drugs that target RNA and for the computational prediction of the structures of RNA−small molecule complexes.

Acknowledgment. The Bruker 600-MHz spectrometer was paid for in part by funds from NIH grant RR11973. The authors thank Dr. Dara Gilbert for helpful comments on the manuscript and Ms. Janet Trang for technical assistance. This research was supported by NSF grant MCB0110689 to T.D.

Supporting Information Available: A complete table of calculated partial charges for free and bound MG, a table of calculated excitation energies for free and bound MG, and two figures showing valence angles and bond lengths for the free and bound MG. This material is available free of charge via the Internet at http://pubs.acs.org.

JA027635D