RZ 2758 (# 89236) 11/09/95
Chemistry 4 pages

Research Report

The Full Structure of a Drug-Nucleic Acid Complex from First Principles

Paolo Carloni* and Wanda Andreoni

IBM Research Division
Zurich Research Laboratory
8803 Rüschlikon
Switzerland

* and Department of Chemistry
University of Florence
50121 Florence
Italy

IBM Research Division
Almaden · T.J. Watson · Tokyo · Zurich

The Full Structure of a Drug-Nucleic Acid Complex from First Principles

Paolo Carloni* and Wanda Andreoni

IBM Research Division, Zurich Research Laboratory, 8803 Rüschlikon, Switzerland
* *and Department of Chemistry, University of Florence, 50121 Florence, Italy*

Abstract

One of the clinically most important classes of DNA/RNA reactive agents are the intercalating drugs, such as proflavine and adryamicin. As an example, we have considered the hydrated complex of N,N–dimethylproflavine and an RNA duplex in the crystal phase. Structural information is provided exclusively by x-ray diffraction, which encounters two main difficulties: 1) the positions of the hydrogen atoms are missing; 2) the presence of water molecules is elusive, partly due to structural and dynamical disorder. We have carried out a full structural optimization from first principles, using density functional theory-based calculations. The system contains 244 atoms (134 heavy atoms and 110 hydrogens). The measured structural parameters are well reproduced. Our calculations establish the interaction between the two Watson–Crick base pairs of the RNA fragment as well as the complex H-bond network of the 16 water molecules detected in the crystal unit cell, and provide suggestions on the mode of action of the drug.

Understanding the action of nucleic-acid binding drugs at the microscopic level, while vital for progress in modern pharmacology, is far from being complete. Current knowledge relies on structural data obtained with X-ray diffraction[1] and NMR spectroscopy.[2] Theoretical modeling has so far been provided by classical approaches, which are of limited applicability[3] and unable to describe chemical bonding. Progress based on a theoretical foundation requires, however, full determination of the structural characteristics of the interacting system (e.g. the sugar-puckering mode, backbone conformation, base-pair geometry, drug conformation) as well as an accurate description of the electronic effects. This can be achieved only with *ab initio* methods, whose application has so far been hampered mainly by the large size of the systems of interest. Here we present the first such calculation of a real crystal complex,[4] namely N,N,dimethyl proflavine (DMPFL) (fig. 1) intercalated[5] into an anti-parallel Watson–Crick base-paired duplex (fig. 2). Full structural optimization provides the crucial and hitherto missing positions of the hydrogen atoms and establishes the water H-bond network as well as the base-pair interactions. Moreover, the study of the structure and of the chemical bonding should provide insight into the mode of action of this drug. As X-ray studies have indicated, its mode of action is indeed affected by methylation of the amino groups of the intercalator: passing from proflavine, where the free amino groups H-bond to phosphate oxygens ("complex intercalation"), to acridine orange (N,N-N',N' tetramethyl proflavine), where the methyl groups interact weakly with the hydrophobic pocket of DNA[4] ("simple intercalation").

The specific system we study here is the complex 2 DMPFL –2 5-iodo-cytidylyl(3'-5')guanosine (ICpG) • 16H_2O, for which an X-ray refined crystal structure is available[4] (entry PFLCPC10 of the Cambridge Structural Database[6]). One DMPFL molecule intercalates between the dinucleoside monophosphate duplex, and the other is stacked on the exterior. The positions of the hydrogen atoms are still inaccessible to experiment, so important issues concerning the actual interactions remain open. In our calculations, we take from experiment only the parameters of the crystal unit cell[7] and determine *ab initio*[8,9] the geometry of the entire complex in it. The starting configuration of the optimization procedure we assume is the X-ray structure (134 heavy atoms) complemented by an educated guess for the (110) hydrogens. The theoretical results agree to a large extent with all the structural features of the dinucleoside monophospate moiety (root mean square deviation (rmsd) = 0.03 Å). In particular: the distance between the two phosphorus atoms, which measures the width of the dinucleoside monophosphate duplex, is reproduced to within 0.1 Å; the mixed sugar-puckering mode of the X-ray structure (C3' *endo* (3',5') C2' *endo*) is retained in the optimization; the calculated backbone torsion angles in Table 1 exhibit significant deviations only in two special cases, namely the angles involving the phosphate group of one strand. The reason for this is given below.

The H-bond interaction of the two Watson–Crick C–G base-pairs (fig. 2) is characterized by H–O and H–N distances whose calculated values are in the expected range of 1.70–1.85 Å.[10] In real systems, the conformation of the base amino groups and in particular how this is affected by the presence of H-bond interactions is not known. This issue is of fundamental importance because the amino groups are involved in key interactions of information transfer processes (such as DNA duplication or RNA transcription)[10] and drug/DNA interactions.[11] Our calculations show that H-bond interactions are dominant here and render the amino group hydrogens not coplanar to the bases. Deviations from coplanarity are up to 7°, except for a guanine amino hydrogen (29°) that interacts with a water molecule (H–O distance is 1.8 Å) (see fig. 2).

The calculated structures of the intercalating drug and that located on the exterior agree fully with X-ray data (rmsd = 0.04 and 0.03 Å, respectively). For the intercalator, calculations also reflect the experimental evidence of H-bonds between the amino group and a water molecule (d(N–O) = 3.0 Å in both X-ray and theoretical structures). However, they differ in the description of the amino–phosphate interaction for which calculations would predict the presence of one hydrogen bond (POH = 122°; d(O–H)= 2.0 Å) not emerging from the refined structure. Here in fact, in the X-ray refined structure we find d(N–O) = 3.4 Å, whereas the theoretical value is 3.0 Å. This is also the origin of the discrepancy found above in the values of the phosphate backbone torsion angles. Although temperature effects are not included in our calculations, reconciling theory and experiment does not need to invoke them. In fact, we have checked that the topology of the Connolly surface[12] of the X-ray structure is consistent with the existence of hydration sites[13] located between the amino and the phosphate group. The presence of an additional water molecule interacting with the amino group may well affect the intercalator/RNA interactions. We believe that it must have eluded experimental observation owing to structural and/or dynamical disorder.[14] Defining the mode of binding from experiment is clearly not unambiguous and requires more accurate information about RNA hydration.

The nature of the drug-RNA bonding can be investigated (fig. 3) by means of the difference density $\Delta\rho(\vec{r}) = \rho_{\rm complex}(\vec{r}) - \rho_{[\rm RNA]^{2-}}(\vec{r}) - \rho_{2[\rm DMPFL]^{+}}(\vec{r})$, which describes the redistribution of the electron density $\rho(\vec{r})$ involved in the formation of the complex from the charged fragments (fig. 3). The dominant electrostatic character is apparent. The presence of H-bonds as well as delocalized charge polarization effects are also evident.

The main conclusions of this structural study are that (i) the conformation of the base amino group hydrogens, important for the biological function of DNA/RNA,[10] is affected by the environment and (ii) defining drug action from X-ray data is not trivial and may well depend on the ability to detect water molecules in the crystal. More generally, we have demonstrated the feasibility of accurate *ab initio* calculations of nucleic acids and drug complexes. This scheme is particularly desirable and overcomes some of the well-known difficulties of semiempirical force-field approaches.[3] Its potential is as a complementary tool to experiment to achieve structural refinements also for biological systems and to understand the chemical bonding.

Acknowledgments. We are grateful to M. Parrinello, J. Hutter, P. Giannozzi, and P. Guéret for useful discussions.

---

[1] Landolt–Börnstein *Numerical Data and Functional Relationships in Science and Technology*; Saenger, W., Ed., Springer: Berlin, 1989; Vol. VII/1b.

[2] Tang, P.; Juang, C.L.; Harbison, G.S. *Science* **1990**, *249*; 70 and references therein.

[3] Beveridge, D.L.; Swaminathan, S.; Ravishanker, G.; Withka, J.M.; Srinivasan, J.; Prevost, C.; Louise-May, S.; Langley, D.R.; DiCapua, F.M.; Bolton, P.H. in: *Water and Biological Macromolecules*; Westhof, E., Ed., MacMillan: London, 1993; 165 and references therein.

[4] Bhandary, K.K.; Sakore, T.D.; Sobell, H.M.; King, D.; Gabbay, E.J. *J. Biomol. Struct. Dyn.* **1984**, *1*, 1195.

[5] Lerman, L.S. *J. Mol. Biol.* **1961**, *3*, 18.

[6] Allen, F.H.; Davies, J.E.; Galloy, J.J.; Johnson, O.; Kennard, O.; Macrae, C.F.; Mitchell, E.M.; Mitchell, G.F.; Smith, J.M.; Watson, D.G. *J. Chem. Inf. Comp. Sci.* **1991**, *31*, 187.

[7] Crystal data for the complex taken from Ref. 4: triclinic, space group P1, a = 11.78 Å, b = 14.55 Å, c = 15.50Å, α = 89.20°, β = 86.20°, γ = 96.40°. Sixteen water molecules are present in the unit cell.

[8] The electronic problem is treated with the density functional theory using the local density approximation exchange and correlation functionals, with gradient corrections necessary for the description of hydrogen-bonded systems.[a] We have performed separate calculations both with the Becke–Perdew and with the BLYP prescriptions,[b] and found that differences in the structural parameters are not significant. *Ab initio* atomic angular-momentum-dependent soft-core pseudopotentials of the Martins–Trouiller type[c] describe the core–valence interaction. The electronic wave function basis set consists of plane waves up to an energy cutoff of 70 Ry. To optimize both the electronic wave function and the geometrical parameters we use the direct inversion in iterative subspace method.[d,e] For calculations using this computational scheme, see e.g. Ref. (f).
(a) Laasonen, K.; Sprik, M.; Car, R.; Parrinello, M. *J. Chem. Phys.* **1993**, *99*, 9080.
(b) Johnson, B.G.; Gill, P.M.W.; Pople, J.A. *J. Chem. Phys.* **1993**, *98*, 5612.
(c) Troullier, N.; Martins, J.L. *Phys. Rev. B* **1991**, *43*, 1993.
(d) Pulay, P. *Chem. Phys. Lett.* **1980** *73*, 393.
(e) Hutter, J.; Lüthi, H.P.; Parrinello, M. *Comput. Mater. Sci.* **1994**, *2*, 244.
(f) Carloni P.; Andreoni, W.; Hutter, J.; Curioni, A.; Giannozzi, P.; Parrinello, M. *Chem. Phys. Lett.* **1995**, *234*, 50.
[9] For all calculations, performed on an IBM SP2 machine, we have used the parallel code developed by J. Hutter.
[10] Saenger, W. *Principles of Nucleic Acid Structure*, Springer: New York, 1984.
[11] Bailly, C.; Waring, M.J. *J. Am. Chem. Soc.* **1995**, *117*, 7311.
[12] Connolly, M.L. *Science* **1983**, *221*, 709.
[13] Clementi, E. in: *Structure and Dynamics of Nucleic Acids and Proteins*, Clementi, E.; Sarma, R.H., Eds., Adenine: New York, 1983, 321.
[14] Schneider, B.; Ginell, S.L.; Berman, H.M. *Biophys. J.* **1992**, *63*, 1572.

TABLE I. Torsion angles (in degrees) of the RNA duplex. Deviations with respect to the X-ray structure[4] are reported in brackets. Definition of the torsion angles as in Ref. 10.

	$\chi^{5'}$	γ	δ	ϵ	ζ	α	β	γ	δ	$\chi^{3'}$
strand 1	−159(−4)	58(+7)	78(−6)	−145(−6)	−69(+2)	−64(−8)	−133(−1)	60(+8)	144(+1)	−89(−9)
strand 2	−160(−3)	59(−7)	84(−4)	−159(−3)	−47(+2)	−92(−16)	−109(+17)	54(+5)	135(+5)	−81(+4)

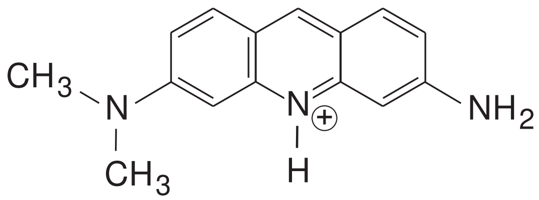

FIG. 1. Chemical structure of DMPFL.

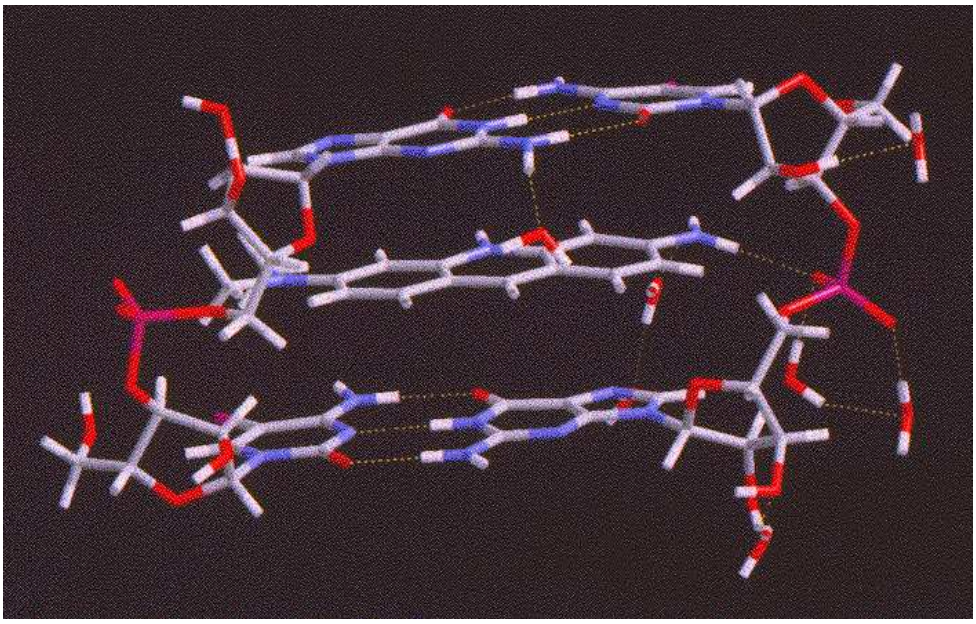

FIG. 2. Optimized structure of the DMPFL-2ICpG complex. H-bonds are indicated by dashed yellow lines. For clarity, some of the water molecules as well as DMPFL located in the exterior are not shown.

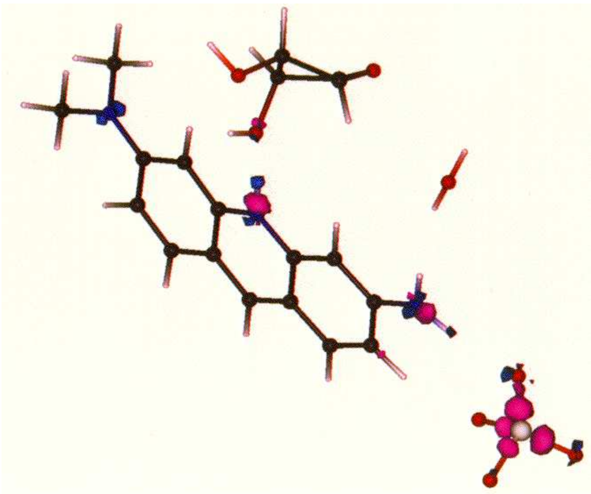

FIG. 3. Isodensity surfaces of the difference electron density $\Delta\rho(\vec{r})$ on the intercalator. Groups H-bonding to the intercalator are also shown. Red: 0.008 e/(a.u.)^3; blue: -0.008 e/(a.u.)^3.