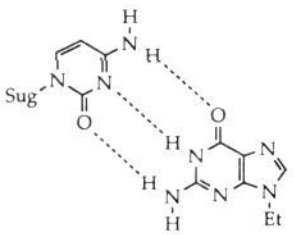
Importance of Secondary Interactions in Triply Hydrogen Bonded Complexes: Guanine–Cytosine vs Uracil–2,6-Diaminopyridine

William L. Jorgensen* and Julianto Pranata

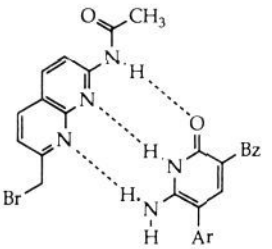
Department of Chemistry, Purdue University
West Lafayette, Indiana 47907

Received September 18, 1989

The triply hydrogen bonded complex between guanine and cytosine (**1**) is central to the structure of nucleic acids. The strength of the interaction is notable with a K_a of ca. 10^4–10^5 M^{-1} in chloroform.[1] Another complex (**2**) with a similar bonding pattern has also been found recently to have a K_a of 1.7×10^4 M^{-1} in chloroform.[2] The benefit of the third hydrogen bond is

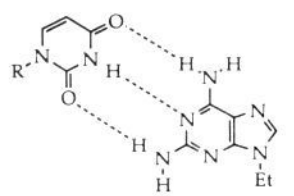


1



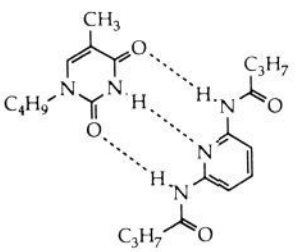
2

apparent in view of the K_a's of 40–130 M^{-1} that are observed for complexes between nucleic acid bases with two hydrogen bonds.[3] However, the situation appears not to be so simple when other triply hydrogen bonded systems are considered. Thus, the K_a's for **3** and **4** in chloroform are only 170 and 90 M^{-1}.[3-5] The origin



R = cyclohexyl

3



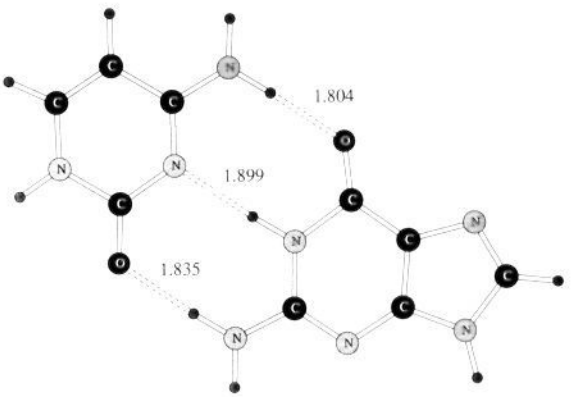
4

of the reduced binding relative to **1** and **2** is not immediately obvious; the same three basic types of hydrogen bonds, $NH_2\cdots O$, $NH\cdots N$, and $NH_2\cdots O$, are present in **1** and **3**. In view of the importance of multiple hydrogen bonding in molecular recognition,[1-6] computational studies were undertaken to investigate the discrepancies.

All-atom potential functions have been developed in the OPLS format[7] for the nucleic acid bases and 2,6-diaminopyridine (DAP) by fitting to results of ab initio 6-31G(d) calculations on the geometries and interaction energies for complexes of the bases or monocyclic fragments and a water molecule.[8] One test of the functions aimed at the present problem was to compute the relative

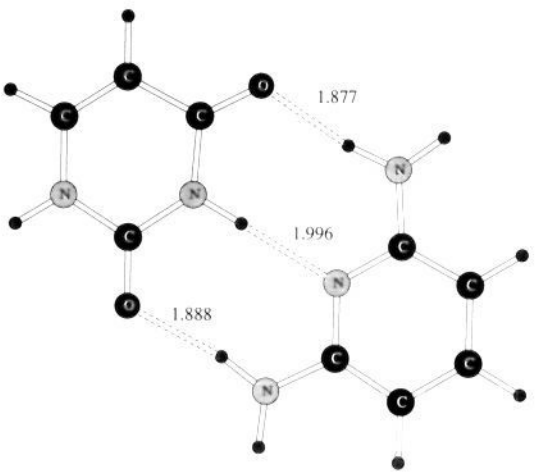
Chart I

Guanine - Cytosine



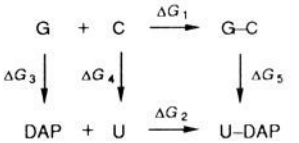
E = -22.1 kcal/mol

Uracil - 2,6-Diaminopyridine



E = -11.4 kcal/mol

free energies of binding for 9-methylguanine (G) with 1-methylcytosine (C) and 1-methyluracil (U) with DAP in chloroform. The thermodynamic cycle below was considered, and $\Delta G_1 - \Delta G_2$ was obtained by computing $\Delta G_3 + \Delta G_4 - \Delta G_5$.

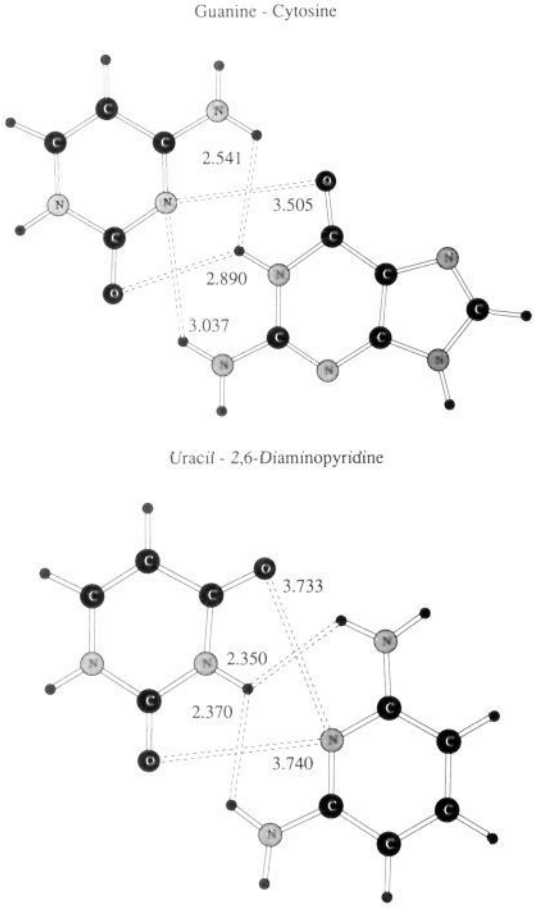


Statistical perturbation theory was used in Metropolis Monte Carlo simulations to obtain the ΔG's.[9] The simulations were all carried out on periodic systems at 25 °C and 1 atm with the BOSS program, version 2.8. If any intermolecular atom–atom distance was less than 12 Å, the entire intermolecular interaction was included. 185 chloroform molecules[10,11] were present for the G–C to U–DAP mutation in a rectangular cell ca. $26 \times 26 \times 39$ Å, while 125 solvent molecules in a cubic cell were used for G → DAP and C → U. The geometries of G, C, U, and DAP were taken from ab initio 3-21G optimizations[8,12] and were kept planar and rigid. The mutations were carried out over a series of four to seven separate

(1) Kyogoku, Y.; Lord, R. C.; Rich, A. *Biochim. Biophys. Acta* **1969**, *179*, 10.
(2) Kelly, T. R.; Zhao, C.; Bridges, G. J. *J. Am. Chem. Soc.* **1989**, *111*, 3744.
(3) Kyogoku, Y.; Lord, R. C.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *57*, 250.
(4) Hamilton, A. D.; Van Engen, D. *J. Am. Chem. Soc.* **1987**, *109*, 5035.
(5) Additional examples of imides and 1,6-diamidopyridines binding with K_a's of 125–150 M^{-1} are available: (a) Kelly, T. R.; Maguire, M. P. *J. Am. Chem. Soc.* **1987**, *109*, 6548. (b) Askew, B. C., Jr. Ph.D. Thesis, University of Pittsburgh, 1988, p 60.
(6) (a) Rebek, J., Jr. *Science* **1987**, *235*, 1478. (b) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645.
(7) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.
(8) Jorgensen, W. L.; Pranata, J.; Wierschke, S. D., to be published. See also the supplementary material.
(9) For reviews, see: (a) Beveridge, D. L.; DiCapua, F. M. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431. (b) Jorgensen, W. L. *Acc. Chem. Res.* **1989**, *22*, 184.
(10) Jorgensen, W. L.; Boudon, S.; Nguyen, T. B. *J. Am. Chem. Soc.* **1989**, *111*, 755.
(11) Jorgensen, W. L.; Briggs, J. M.; Contreras, M. L. *J. Phys. Chem.*, in press.
(12) Aida, M. *J. Comput. Chem.* **1988**, *9*, 362.

0002-7863/90/1512-2008$02.50/0

Chart II

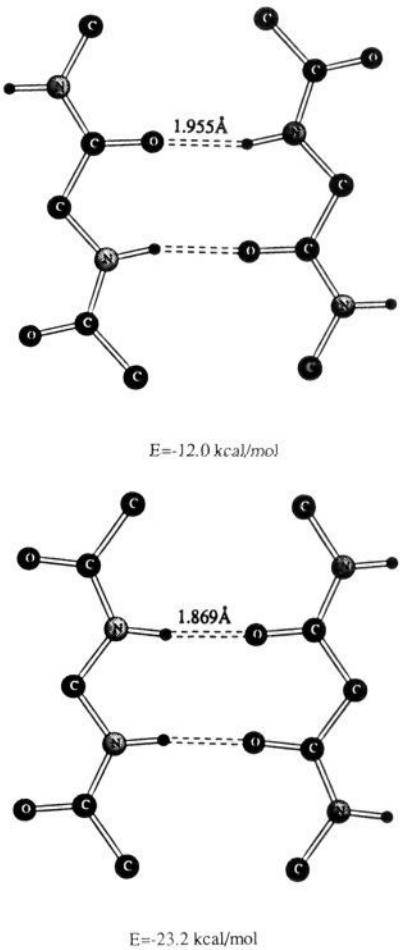


simulations yielding eight to 14 incremental ΔG's.[9] Each simulation consisted of 5×10^5 to 10^6 configurations of equilibration followed by 2×10^6 configurations of averaging. The simulations near G–C and U–DAP were initiated and remained triply hydrogen bonded, though there were no restrictions on translations or rotations. In the first perturbation step from the free G–C complex, weak harmonic constraints were introduced to limit drift at intermediate points. These constraints were then removed in the last step, yielding the free U–DAP complex.

The results for ΔG_3 (8.4 ± 0.3 kcal/mol), ΔG_4 (1.9 ± 0.2), and ΔG_5 (17.2 ± 0.3)[13] combine to yield a prediction of stronger binding for G–C than U–DAP by 7.0 ± 0.4 kcal/mol. This translates to a K_a ratio of 1.3×10^5, which shows an even stronger preference for G–C than suggested by the data for **1**–**4**. However, there are structural differences, significant experimental uncertainty for **1**,[1] and the statistical noise in the simulations. Analogous calculations were carried out for G–C versus the complex of 9-methyladenine (A) with U. Essentially the same preference (7.2 ± 0.3 kcal/mol) was found for G–C, which is consistent with the experimental observation of similar K_a's (ca. 10^2) for systems like A–U or U–DAP.[3-5]

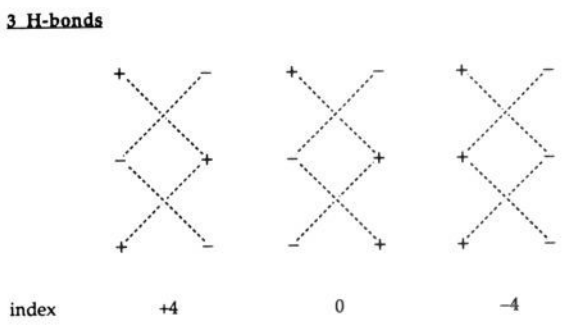
A salient component of the preference for G–C over U–DAP is the 10.5 kcal/mol lower average intersolute interaction energy computed for G–C in chloroform. In fact, full optimizations for the *isolated* complexes with the OPLS potentials yield optimal interaction energies of −22.1 kcal/mol for G–C and −11.4 kcal/mol for U–DAP (Chart I). The former value is close to experimental gas-phase results (−21.0 kcal/mol)[14] and other theoretical estimates,[15] while the computed value for U–DAP is only a little lower than for A–U (−10.5 kcal/mol).

Chart III



Since the primary hydrogen-bonding interactions arethe same, more subtle variations must be present for the 10.7 kcal/mol difference. Though a multipole analysis could be provided, a simpler approach comes from considering the secondary interactions indicated below. These separations are still short (2.3–3.7 Å), owing to the proximal hydrogen bonds, and involve substantial electrostatic interactions (Chart II). In view of the partial positive charges on H and partial negative charges on N and O, two of the secondary interactions are attractive and two are repulsive for G–C, while all four interactions are repulsive for U–DAP. The net of four destabilizing interactions at 2–3 kcal/mol each can account for the weaker binding for the latter complex.

The three possible arrangements of the partially charged sites for triply hydrogen bonded systems are shown below.



(13) This includes a necessary correction, RT ln 2, that reflects the change in symmetry numbers for the two complexations: Bailey, W. F.; Monahan, A. S. *J. Chem. Educ.* **1978**, *55*, 489.

(14) Yanson, I. K.; Teplitsky, A. B.; Sukhodub, L. F. *Biopolymers* **1979**, *18*, 1149.

(15) (a) Langlet, J.; Claverie, P.; Caron, F. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, 1981. (b) Pohorille, A.; Burt, S. K.; MacElroy, R. D. *J. Am. Chem. Soc.* **1984**, *106*, 402. (c) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230.

The worst situation is when the positive and negative sites alternate as in DAP and imides. G-C represents the intermediate case, while in the most favorable situation one molecule has all of the hydrogen-bond donor sites and the other has all of the acceptor sites. The general validity of this simple concept is unequivocally supported by the following dipeptide model (Chart III). The optimizations for the planar structures used the same charges in both forms.[7] The hydrogen bonding in the retro-inverso isomer is favored by 11 kcal/mol, owing again to the net difference of four secondary interactions. Such considerations have general applicability in understanding variations in hydrogen-bonding complexation in many contexts.[16]

Supplementary Material Available: Listing of parameters for nucleic acid bases and DAP in the all-atom OPLS force field, graphical summary of the charge distributions for the parent nucleic acid bases, and a plot showing the progress of the three mutations used to compute the relative free energy of binding for G-C vs U-DAP in chloroform (5 pages). Ordering information is given on any current masthead page.

(16) Gratitude is expressed to the National Institutes of Health (GM32136) for support of this research.

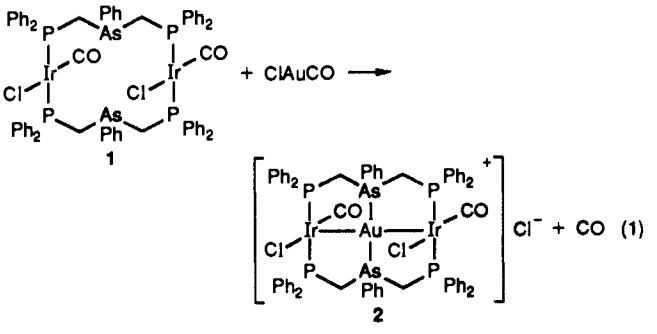
Chelate Ring Opening and Metal Ion Relocation Leading to the Formation of a Luminescent $Au^{I}Ir^{I}Au^{I}$ Chain Complex

Alan L. Balch,* Vincent J. Catalano, and Marilyn M. Olmstead

Department of Chemistry, University of California Davis, California 95616

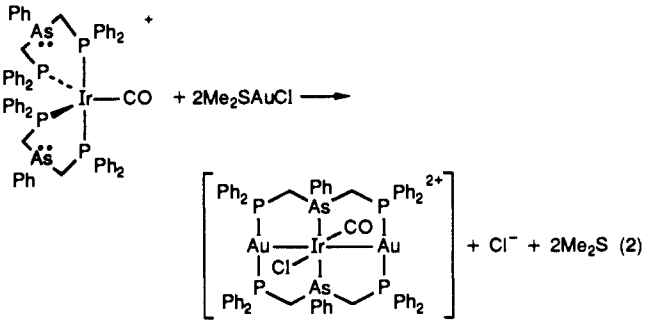
Received October 18, 1989

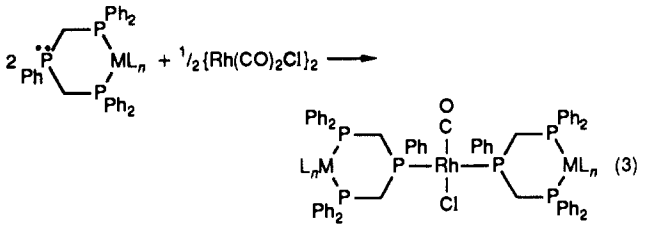
Recently we have developed rational routes to the synthesis of heterotrinuclear complexes using binuclear metallomacrocycles as precursors.[1] Equation 1 shows an example involving the incorporation of gold into a diiridium macrocycle **1**.[2] The metallomacrocycle **1** has considerable stability; for example, it is capable of extracting some metal ions from aqueous solution into dichloromethane.[3] Here we describe a new and unexpected route to the formation of a related heterotrinuclear complex. This method also uses bis[(diphenylphosphino)methyl]phenylarsine (dpma) as a bridge, but involves much more bond reorganization than encountered in eq 1.



Treatment of $Ir(CO)_2Cl$ (*p*-toluidine) in methanol with 2 equiv of dpma in dichloromethane followed by an excess of ammonium hexafluorophosphate in methanol produces, after partial evaporation of the solvent, ivory crystals of $[Ir(CO)(dpma)_2][PF_6]$ (**3**) in 76% yield. The structure of the trigonal-bipyramidal cation is shown in Figure 1. It is similar to other cations with an $Ir(CO)P_4$ core.[5]

Addition of 2 equiv of Me_2SAuCl in dichloromethane to a dichloromethane solution of **3** produces a red solution, from which red crystals of $[Au_2Ir(CO)Cl(\mu\text{-dpma})_2][PF_6]_2$ (**4**)[6] can be isolated in 84% yield by the addition of a methanol solution of ammonium hexafluorophosphate followed by partial evaporation of the solvent (eq 2). The structure of the cation of **4**, as determined by X-ray crystallography, is shown in Figure 1. Notice that the metal ion locations in **4** and **2** have an inside-out relationship. While **3** incorporates the correct Ir:dpma:CO stoichiometry for the formation of **4**, reaction 2 leaves only one metal ligand bond, the Ir–CO bond, intact.[1] All four of the Ir–P bonds in **3** and the Au–S and Au–Cl bonds are broken in this high-yield reaction. This represents quite a remarkable set of changes that contrasts with the much more conservative group of changes in eq 1. While chelate ring opening of four-membered rings, especially those involving bis(diphenylphosphino)methane, to form binuclear complexes is known,[7] opening of the less strained six-membered rings in **3** is wholly unexpected. Prior work showed that reactions of analogous triphosphine complexes resulted in formation of bi- or trinuclear complexes with widely separated metal centers by bonding to the uncoordinated donor without ring opening (eq 3).[8] We have monitored reaction 2 by ^{31}P NMR spectroscopy at −60 °C, but only the slow conversion of the starting material to the product is observed. No intermediates in this complex transformation are observable.





ML_n = $Pt(CH_3)_2$, $PtCl_2$, $PdCl_2$, $Mo(CO)_4$

(1) Balch, A. L. *Pure Appl. Chem.* **1988**, *60*, 555. Balch, A. L.; Fossett, L. A.; Olmstead, M. M.; Oram, D. E.; Reedy, P. E., Jr. *J. Am. Chem. Soc.* **1985**, *107*, 5272. Balch, A. L.; Fossett, L. A.; Olmstead, M. M.; Reedy, P. E., Jr. *Organometallics* **1986**, *5*, 1929 and references therein.

(2) Balch, A. L.; Nagle, J. K.; Oram, D. E.; Reedy, P. E., Jr. *J. Am. Chem. Soc.* **1988**, *110*, 454.

(3) Balch, A. L.; Olmstead, M. M.; Oram, D. E.; Reedy, P. E., Jr.; Reimer, S. H. *J. Am. Chem. Soc.* **1989**, *111*, 4021.

(4) IR: ν(CO), 1925 cm^{-1}. $^{31}P\{^1H\}$ NMR: δ −21.8 (t), −29.1 (t) [J(P,P) = 28.8 Hz]. Ivory crystals of $[Ir(CO)(dpma)_2]PF_6 \cdot CH_2Cl_2 \cdot (C_2H_5)_2O$ were grown by diffusion of ethyl ether into a dichloromethane solution of **3**. They belong to the orthorhombic space group *Pbnb* (No. 56) with a = 16.487 (4) Å, b = 26.297 (7) Å, and c = 30.733 (6) Å at 130 K, Z = 8, R = 0.067, and R_w = 0.065 for 5274 reflections with $I > 2\sigma(I)$ and 410 parameters.

(5) Sanger, A. R. *J. Chem. Soc., Dalton Trans.* **1977**, 1979. Miller, J. S.; Caulton, K. G. *J. Am. Chem. Soc.* **1975**, *97*, 1067.

(6) IR: ν(CO), 1971 cm^{-1}. $^{31}P\{^1H\}$ NMR: δ = 34.2 (s). Red crystals of $[Au_2Ir(CO)Cl(\mu\text{-dpma})_2][PF_6]_2 \cdot CH_2Cl_2$ were obtained with difficulty by diffusion of ethyl ether into a dichloromethane solution of the complex. They belong to the triclinic space group $P\bar{1}$ with a = 12.218 (4) Å, b = 24.768 (8) Å, c = 26.693 (9) Å, α = 112.75 (2)°, β = 91.65 (3)°, and γ = 95.50 (3)° at 130 K, Z = 4, R = 0.101, and R_w = 0.101 for 11 583 reflections with $F > 4\sigma(F)$ and 852 parameters. There are two independent cations, four independent anions, and 2.5 molecules of CH_2Cl_2 in the asymmetric unit.

(7) McEwan, D. M.; Pringle, P. G.; Shaw, B. L. *J. Chem. Soc., Chem. Commun.* **1982**, 1240. Blagg, A.; Cooper, G. R.; Pringle, P. G.; Robson, R.; Shaw, B. L. *J. Chem. Soc., Chem. Commun.* **1984**, 933. Hutton, A. T.; Pringle, P. G.; Shaw, B. L. *Organometallics* **1983**, *2*, 1889.

(8) Guimerans, R. R.; Olmstead, M. M.; Balch, A. L. *Inorg. Chem.* **1983**, *22*, 2223. Olmstead, M. M.; Guimerans, R. R.; Farr, J. P.; Balch, A. L. *Inorg. Chim. Acta* **1983**, *75*, 199. Balch, A. L.; Guimerans, R. R.; Linehan, J. *Inorg. Chem.* **1985**, *24*, 290.

0002-7863/90/1512-2010$02.50/0