Structure of a Copper-Mediated Base Pair in DNA

Shane Atwell,‡ Eric Meggers,‡ Glen Spraggon,§ and Peter G. Schultz*,‡,§

Contribution from the Department of Chemistry, The Scripps Research Institute, La Jolla, California 92037, and the Genomics Institute of the Novartis Foundation, La Jolla, California 92037

Received July 27, 2001

Abstract: Stable and selective DNA base pairing by metal coordination was recently demonstrated with nucleotides containing complementary pyridine-2,6-dicarboxylate (Dipic) and pyridine (Py) bases (Meggers, E.; Holland, P. L.; Tolman; W. B.; Romesberg, F. E.; Schultz, P. G. *J. Am. Chem. Soc.* **2000**, *122*, 10714−10715). To understand the structural consequences of introducing this novel base pair into DNA we have solved the crystal structure of a duplex containing the metallo-base pair. The structure shows that the bases pair as designed, but in a Z-DNA conformation. The structure also provides a structural explanation for the B- to Z-DNA transition in this duplex. Further solution studies demonstrate that the metallo-base pair is compatible with Z- or B-DNA conformations, depending on the duplex sequence.

Introduction

Watson−Crick hydrogen bonding between natural base pairs is essential for information storage and retrieval in DNA (Figure 1a). However, is this four-letter alphabet the only structural solution to the problem of information storage? Efforts to address this question have focused largely on two different strategies: altered hydrogen bonding patterns (Figure 1b)[1−5] and the replacement of hydrogen bonding with hydrophobic interactions (Figure 1c).[6−11] Recently, we demonstrated that metal coordination can also effectively drive base pairing[12−14] (Figure 1d).

When the bases pyridine-2,6-dicarboxylate (Dipic) and pyridine (Py) are introduced in the middle of a 15 nucleotide duplex, the duplex displays comparable thermal stability in the presence of equimolar Cu^{2+} to a duplex with an A:T pair.[12] To study the structural consequences of replacing a natural pair by a metallo-pair, we have determined the crystal structure of 5′-CGCG*Dipic*AT*Py*CGCG-3′ at a resolution of 1.5 Å.

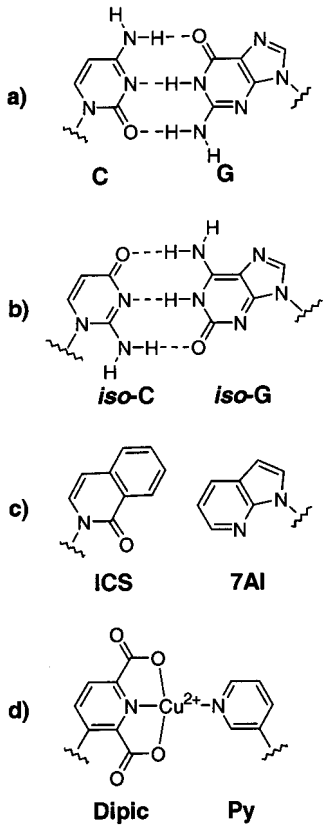

Figure 1. Examples of natural and modified DNA base pairs: (a) Watson−Crick hydrogen bonding between cytosine (C) and guanosine (G); (b) alternative hydrogen bonding between isocytosine (*iso*-C) and isoguanine (*iso*-G); (c) hydrophobic packing between isocarbostyril (ICS) and 7-azaindole (7AI); and (d) Cu^{2+} induced base pairing between 2,6-dipicolinate (Dipic) and pyridine (Py).

Experimental Section

Synthesis. Deoxyoligonucleotides were synthesized by using an Expedite nucleic acid synthesizer (PerSeptive Biosystems) on a 1.0 μm scale. A standard protocol for 2-cyanoethyl phosphoramidites was used, except that the coupling of the unnatural nucleotides was extended

* To whom correspondence should be addressed. E-mail: schultz@scripps.edu.

‡ The Scripps Research Institute.

§ The Genomics Institute of the Novartis Foundation.

(1) Switzer, S.; Moroney, S. E.; Benner, S. A. *J. Am. Chem. Soc.* **1989**, *111*, 8322−8323.

(2) Piccirilli, J. A.; Krauch, T.; Moroney, S. E.; Benner, S. A. *Nature* **1990**, *343*, 33−37.

(3) Switzer, C. Y.; Moroney, S. E.; Benner, S. A. *Biochemistry* **1993**, *32*, 10489−10496.

(4) Horlacher, J.; Hottiger, M.; Podust, V. N.; Hübscher, U.; Benner, S. A. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 6329−6333.

(5) Lutz, M. J.; Held, H. A.; Hottiger, M.; Hübscher, U.; Benner, S. A. *Nucleic Acids Res.* **1996**, *24*, 1308−1313.

(6) Kool, E. T. *Biopolymers* **1998**, *48*, 3−17.

(7) Guckian, K. M.; Krugh, T. R.; Kool, E. T. *Nat. Struct. Biol.* **1998**, *11*, 954−959.

(8) Kool, E. T.; Morales, J. C.; Guckian, K. M. *Angew. Chem., Int. Ed.* **2000**, *39*, 990−1009.

(9) McMinn, D. L.; Ogawa, A. K.; Wu, Y.; Liu, J.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **1999**, *121*, 11585−11586.

(10) Ogawa, A. K.; Wu, Y.; McMinn, D. L.; Liu, J.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2000**, *122*, 3274−3287.

(11) Wu, Y.; Ogawa, A. K.; Berger, M.; McMinn, D. L.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2000**, *122*, 7621−7632.

(12) Meggers, E.; Holland, P. L.; Tolman, W. B.; Romesberg, F. E.; Schultz, P. G. *J. Am. Chem. Soc.* **2000**, *122*, 10714−10715.

(13) Tanaka, K.; Shionoya, M. *J. Org. Chem.* **1999**, *64*, 5002−5003.

(14) Weizman, H.; Tor, Y. *J. Am. Chem Soc.* **2001**, *123*, 3375−3376.

10.1021/ja011822e CCC: $20.00
Published on Web 11/15/2001

Table 1. Data Collection, Phasing, and Refinement Statistics

	remote	*f″* max	*f′* min
data collection			
space group[a]	$P2_1$		
cell dimensions	$a = 25.343$, $b = 34.359$, $c = 31.093$, $\beta = 101.13$		
wavelength (Å)	1.0	1.37791	1.37991
resolution (Å)	1.5	1.8	1.8
no. of reflecns (measd/unique)	47272/8531	34393/4924	34221/4910
completeness (%) [b]	98.2 (84.1)	98.7 (87.4)	98.7 (87.2)
R_{sym} (%) [c]	9.6 (68.6)	7.8 (13.6)	7.7 (17.8)
no. of reflecns with $\langle I/\sigma(I)\rangle > 3$ (%)	66.8 (34.4)	91.6 (81.9)	86.6 (72.0)
phasing statistics (50−1.8 Å)			
R_{iso} (%)		4.6	5.4
R_{cullis} (centric/acentric)		43.4/54.1	42.2/55.9
phasing power (centric/acentric)		2.2/2.2	2.2/2.1
mean figure of merit (centric/acentric)		0.73/0.46	0.67/0.41
refinement statistics			
refinement resolution (Å)	50−1.5		
σ cutoff	0		
reflections used	8531		
R_{cryst} (%) [d]	16.9		
R_{free} (%) [e]	20.2		
rms dev			
bond lengths (Å)	0.00893		
bond angles (deg)	1.56		
B-factor bonded	2.56		
B-factor angle	3.51		
copper atoms	2		
water molecules	82		

[a] All three data sets were collected from the same crystal. [b] Values in parentheses are for the highest resolution shell. [c] $R_{sym} = \sum| I - \langle I\rangle|/\sum I$, where I is the observed intensity and $\langle I\rangle$ is the average intensity over all observations of symmetry-related reflections. [d] $R_{cryst} = ||F_{obs}| - |F_{calc}||/\sum|F_{obs}|$, where $|F_{obs}|$ and $|F_{calc}|$ are the observed and calculated structure factor amplitudes, respectively. [e] R_{free} was calculated from a randomly chosen subset of 8% of the reflections.

to 15 min. The carboxylate groups of the pyridine-2,6-dicarboxylate nucleobase were protected as methyl esters. Phosphoramidites for phenoxyacetyl protected dA, 4-isopropyl-phenoxyacetyl protected dG, and acetyl protected dC were used (ultramild reagents from Glen Research) and the oligonucleotides were deprotected for 6 h with 50 mM potassium bicarbonate in dry methanol at room temperature. The tritylated oligonucleotides were then purified by reversed phase HPLC, detritylated with 80% acetic acid, and again purified by reverse phase HPLC. Finally, the dimethyl ester of the Dipic was hydrolyzed with 20 mM sodium hydroxide for 24 h at room temperature and desalted with reverse phase HPLC. Identity was confirmed by MALDI-ToF MS.

Crystallization. An aqueous solution of the self-complementary strand 5′-CGCG*Dipic*AT*Py*CGCG-3′ (1 mM) with 1 mM $Cu(NO_3)_2$ was incubated at 55 °C for 5 min and slowly cooled to room temperature. This solution was used as the stock for setting up hanging drop crystallization trials with the Nucleic Acid Mini Screen from Hampton Research. DNA (0.75 μL) was mixed with reagent (1.5 μL) and incubated over a well solution of 35% MPD at 4 °C. The best crystals, measuring 20 μm × 20 μm × 100 μm, were obtained reproducibly within 2 weeks with 12 mM spermine tetrahydrochloride, 80 mM NaCl, 12 mM KCl, 20 mM Mg chloride, and 40 mM sodium cacodylate (pH 6.0 or 7.0).

Data Collection and Structure Determination. Crystals were mounted directly in loops and flash frozen in a liquid nitrogen stream. Data at three wavelengths were collected at the Advanced Light Source (Lawrence Berkeley Laboratories) on beamline 5.0.2 after a fluorescence scan indicated the optimal *f″* inflection and *f′* maximum. Data were reduced with Denzo and scaled with Scalepack.[15] Initial phases were calculated and density modification was performed with CNS.[16] The initial map gave clear density for most of the bases and phosphates. An initial model was constructed starting from the copper(II) coordinating bases (substituting cytosines for the metal bases) and tracing the sugars and phosphates. Refinement using CNS with data to 1.8 Å and an maximum likelihood target incorporating experimental phases and adjustment of the bases with O^{17} led to an initial model that showed good pairing for all bases and density for the missing carboxylate groups of the Dipic base. Several cycles of minimization, simulated annealing, and *B*-factor refinement with CNS improved the map and indicated several sugars had the wrong pucker in the model. A parameter file for the unnatural bases was build and the dihedral energy terms for the sugars were set to zero before several more cycles of minimization, *B*-factor refinement, and water building with data to 1.5 Å and the native data set alone using an amplitude based maximum likelihood target. The waters were double checked visually and with the water delete script of CNS.

Circular Dichroism. CD experiments were performed with an Aviv model 61 DS spectropolarimeter equipped with a Peltier thermoelectric temperature control unit. A 1 cm path length quartz cuvette was used, for scans from 200 to 320 nm, with a time constant of 3 s and a wavelength step size of 1.0 nm at 25 °C. The spectra were not baseline corrected.

Coordinates. Coordinates have been deposited in the nucleic acid database (accession ZD0006) and the protein data bank (accession 1JES).

Results and Discussion

Crystal Structure. The deoxyoligonucleotide duplex used for the structural studies is based on the well-studied palindromic dodecamer 5′-CGCGAATTCGCG-3′.[18] The duplex contains two metallo-pairs: a pyridine-2,6-dicarboxylate nucleobase replaces adenine 5 and a pyridine nucleobase replaces thymidine 8. The final refined structure contains the complete 12 nucleotide duplex in the asymmetric unit, including two copper ions and 82 water molecules (R_{fact} 17%, R_{free} 20%). The crystal structure shows that the unnatural Dipic and Py bases pair through coordination of a Cu^{2+} ion (Figure 2).

(15) Otwinowski, Z.; Minor, W. *Methods Enzymol.* **1997**, *276*, 307−326.

(16) Adams, P. E.; Pannu, N. S.; Read, R. J.; Brunger, A. T. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 5018−5023.

(17) Jones, A. T.; Zou, J. Y.; Cowan, S. W.; Kjeldgaard, M. *Acta Crystallogr. A* **1991**, *47*, 110−119.

(18) Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *Nature* **1980**, *287*, 755−758.

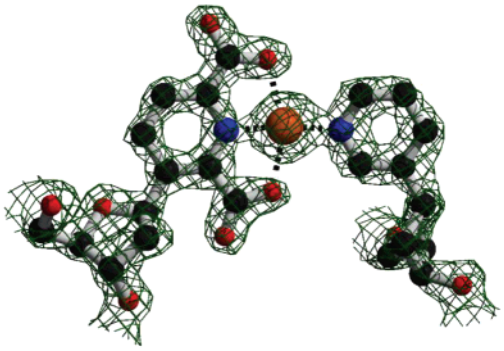

Figure 2. The Dipic:Py bases pair in the duplex structure. Electron density is from a 2fofc map using the final model, contoured at 1.3 σ. The image was prepared by bobscript.

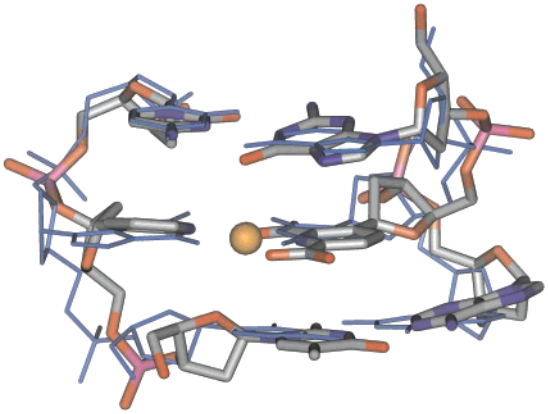

Figure 3. Local environment of the copper(II) ion. The metallo-base pair with neighboring pairs is shown with thick sticks and colored by atom type, with the same pairs of Z_I-DNA (from 2ZNA.pdb) in blue from the superposition with the whole duplex. The image was created in InsightII.

The presence of the Cu^{2+} ions is verified by their strong anomalous dispersion, yielding 20 σ peaks in the initial electron density map calculated with anomalous phases. The square-planar coordination of Cu^{2+} in the plane of the Dipic base involves two strong Cu−N coordinative bonds at a distance of 1.9 Å and two Cu−O bonds at 2.3 Å, typical distances for pyridine-2,6-dicarboxylate Cu^{2+} complexes.[19] In addition to the planar coordination of the Cu^{2+} by the metallo-pair, there appears to be additional axial coordination by neighboring nucleotides (Figure 3).

The furanose O4′ of thymidine 7 is only 3.1 Å from the copper(II) ion; the O6 of guanine 4 on the other side is also 3.1 Å from the copper(II) ion. These interactions and the planar interactions with the metallo-bases make up a Jahn−Teller distorted octahedral coordination, well documented in copper-(II) complexes.[20] For example, the structure of $CuCl_2 \cdot 2H_2O$ also has two 1.9 Å bonds (Cu−O), two 2.3 Å bonds (Cu−Cl), and two 3.0 Å bonds (Cu−Cl).[20]

The axial coordination by the furanose oxygen is only possible because the overall conformation of the duplex is that of Z-DNA. The root-mean-square difference between the sugar−phosphate backbone of the metallo-duplex and ideal Z_I-DNA[21] is 1.5 Å (1.4 Å with Z_{II}-DNA from 3ZNA.pdb) (Figure 4).

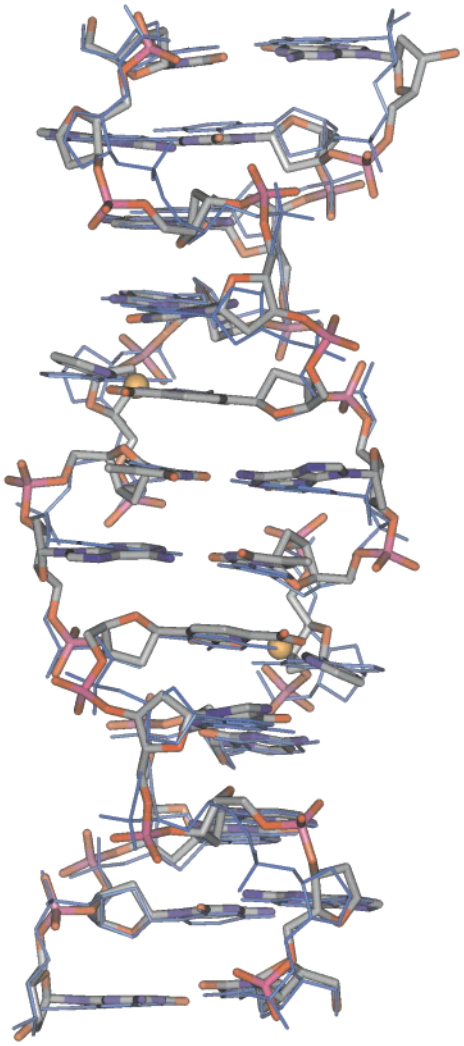

Figure 4. Superposition with ideal Z_I-DNA (from 2ZNA.pdb). Superposition was performed and root-mean-square deviation determined in InsightII using the phosphate and sugar atoms.

Z-DNA has a left-handed twist and a phosphate backbone that zigzags up the helix, one phosphate−sugar−phosphate oriented parallel with the helix axis and the other perpendicular. Bases alternate between syn and anti conformations. This usually requires an alternating purine-pyrimidine (APP) sequence to avoid steric clash of the ribose with the O2 atom of cytidine or thymidine in the *syn* conformation.[22] Fortuitously, the metallo-duplex mimics an APP sequence, since most of the duplex conforms to an APP sequence and the Py nucleotide can adopt a syn conformation without steric clash. In Z-DNA, stacking of the bases is observed in CpG steps, whereas GpC steps display considerable shearing, such that the guanine instead of packing against the cytosine base packs against the cytosine sugar (Figure 3, blue sticks). This juxtaposition of base with neighboring sugar in Z-DNA is what allows coordination of the Cu^{2+} by the furanose O4′. The only major differences between the metallo-duplex and ideal Z-DNA are the distortions around the metallo-pair. Small perturbations relative to ideal Z-DNA are attributable to the axial coordination of the Cu^{2+} ion, i.e., guanine 4 tilts, placing O6 0.7 Å closer to the Cu^{2+}

(19) Nardin, G.; Randaccio, L.; Bonomo, R. P.; Rizzarelli, E. *J. Chem. Soc. Dalton* **1980**, 369−375.

(20) Cotton, F. A.; Wilkinson, G.; Murillo, C. A.; Bochmann, M. In *Advanced Inorganic Chemistry*, 6th ed.; John Wiley & Sons: New York; 1999; pp 864−872.

(21) Wang, A. J.; Quigley, G. J.; Kolpak, F. J.; van der Marel, G.; van Boom, J. H.; Rich, A. *Science* **1981**, *211*, 171−176.

(22) Basham, B.; Eichman, B. F.; Ho, P. S. In *Oxford handbook of nucleic acid structure*; Neidle, S., Ed.; Oxford University Press: Oxford; 1999; 199−252.

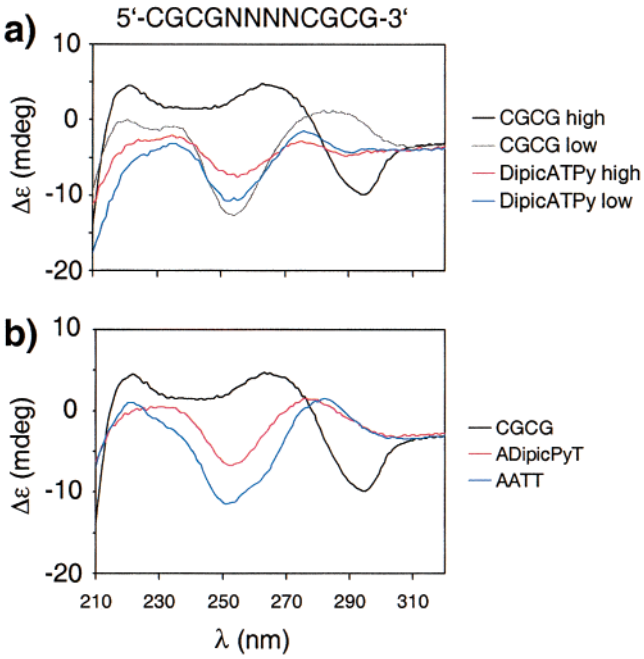

Figure 5. Circular dichroism spectra of deoxyoligonucleotide duplexes: (a) spectra of 5′-CGCGCGCGCGCG-3′ and 5′-CGCG*DipicATPy*CGCG-3′ in low (4 μM deoxyoligonucleotide, 50 mM sodium perchlorate, 5 mM sodium phosphate, pH 7.0) and high salt (4 μM deoxyoligonucleotide, 4 M sodium chloride, 20 mM sodium phosphate, pH 7.2) and (b) spectra of 5′-CGCGCGCGCGCG-3′, 5′-CGCGA*DipicPy*TCGCG-3′, and 5′-CGCGAATTCGCG-3′ in high salt.

ion. Larger perturbations are caused by the shorter C1′−C1′ distance of the designed pair. The C1′−C1′ distance in ideal Z-DNA is 10.8 Å, and is reduced by 1.5 Å in the metallo-pair to an average 9.3 Å. This compression distorts the backbone at the Py base much more than the Dipic backbone. It also contributes to placing the O4′ oxygen of thymine 7 further underneath the Cu^{2+} ion (0.7 Å closer than it would be in ideal Z-DNA).

CD Measurements. To determine if the Z-DNA conformation observed in the crystals is also the solution conformation of this duplex, the circular dichroism (CD) spectra of the duplex and several related sequences were measured. The metallo-duplex gives almost the same spectrum at low salt (50 mM sodium perchlorate) and high salt (4 M sodium chloride) concentrations (Figure 5a). This spectrum has a characteristic feature of Z-DNA spectra,[23] namely a minimum at 290 nm, although this feature is weaker than that measured for an all GC dodecamer in high salt (Figure 5a).

Overall, the spectrum of the metallo-duplex at low or high salt concentrations is similar to that of poly-GC DNA near the midpoint between B- and Z-DNA at medium NaCl concentrations (approximately 2 M).[24] From these spectra we conclude that the crystallized metallo-duplex has at least a significant Z-DNA content even at low salt and that it is this Z-DNA conformation that crystallizes. Furthermore, the solution conformation of the metallo-duplex is not changed upon addition of up to 12 mM spermine (not shown) or 4 M NaCl. It appears that the central *Dipic*AT*Py* sequence stabilizes a Z-DNA conformation much more than a central CGCG sequence, even though A:T pairs are known to destabilize Z-DNA.[25] For comparison, the dodecamer with a central AT surrounded by an all GC sequence, 5′-CGCGCATGCGCG-3′, remains B-DNA at high salt concentrations (not shown). CD spectra of metallo-base pair containing sequences differing from that of the crystallized oligonucleotide reveal that in non-APP sequences the overall conformation is B-form. For example, 5′-CGCGA*DypicPy*TCGCG-3′, which has the same nucleobase content as the crystallized duplex but a non-APP sequence, yields a typical B-DNA CD spectrum (Figure 5b).[26] These results demonstrate that the metallo-pair is compatible with a B-DNA conformation in a non-APP context. The structure of the metallo-pair in B-DNA most likely consists also of a planar coordination of the metal ion, stacking with the natural bases, and a C1′−C1′ distance shortened as in the Z-DNA structure.

Conclusion

The development of an extended genetic alphabet will be useful for information storage, the in vitro selection of DNA and RNA molecules with novel binding and catalytic properties, and labeling applications in biotechnology. The present study demonstrates that metal coordination can replace hydrogen bonding while preserving helical structure. The Dipic:Py pair is well-behaved in the context of duplex DNA, with base pairing parameters essentially as designed. Furthermore, the solution studies show that in non-APP sequence contexts this metallo-pair is compatible with right-handed helical conformations. The crystal structure suggests that further designs might eliminate distortions around the metallo-pair by lengthening the C1′−C1′ distance and by the use of pairs with selectivities for real square planar d^8 metal ions such as Ni^{2+}, Pd^{2+}, or Pt^{2+} instead of Cu^{2+}, which prefers tetragonal distorted coordination. Efforts are underway to design metallo-bases with different ligand−metal structures, as well as testing metal coordinating bases for enzymatic incorporation by DNA polymerases.

Acknowledgment. We thank Raymond Stevens and members of his lab for the use of his X-ray generator and help during data collection, the Peter E. Wright and Jeffery Kelly labs for the use of their spectropolarimeters, the Advanced Light Source and Thomas Earnest for synchrotron time and Gyorgy Snell for assistance with data collection, Chris Lee for help with data collection, and the Mass spectrometry facilities at TSRI. S.A. is supported by the American Cancer Society, postdoctoral fellowship; E.M. is supported by the Deutsche Forschungsgemeinschaft. P.G.S. is funded by a grant from DOE No. DE-FG03-00ER45812. This is manuscript number 14373-CH from The Scripps Research Institute.

JA011822E

(23) Rich, A.; Nordheim, A.; Wang, A. H.-J. *Ann. Rev. Biochem.* **1984**, *53*, 791−846.

(24) Pohl, F. M.; Jovin, T. M. *J. Mol. Biol.* **1972**, *67*, 375−396.

(25) Johnson, W. C. In *Circular Dichroism*; Berova, N., Nakanishi, K., Woody, R. W., Eds.; Wiley-VCH: New York, 2000; pp 703−739.

(26) This deoxyoligonucleotide and three other non-APP deoxyoligonucleotides (5′-CCDipicACGTPyGG-3′, 5′-CCADipicGCPyTGG-3′, and 5′-CGCGDipicDipicPyPyCGCG-3′) were also used in crystallization trials, but none has yielded diffracting crystals.