Quantitative Analysis of DNA Secondary Structure from Solvent-Accessible Surfaces: The B- to Z-DNA Transition as a Model[†]

Todd F. Kagawa, Donald Stoddard, Guangwen Zhou, and Pui S. Ho*

Department of Biochemistry and Biophysics, Oregon State University, Corvallis, Oregon 97331

Received December 5, 1988; Revised Manuscript Received April 25, 1989

ABSTRACT: Solvent structure and its interactions have been suggested to play a critical role in defining the conformation of polynucleotides and other macromolecules. In this work, we attempt to quantitate solvent effects on the well-studied conformational transition between right-handed B- and left-handed Z-DNA. The solvent-accessible surfaces of the hexamer sequences $d(m^5CG)_3$, $d(CG)_3$, $d(CA)_3$, and $d(TA)_3$ were calculated in their B- and Z-DNA conformations. The difference in hydration free energies between the Z and the B conformations ($\Delta\Delta G_{H(Z-B)}$) was determined from these surfaces to be −0.494 kcal/mol for C-5 methylated $d(CG)$, 0.228 kcal/mol for unmethylated $d(CG)$, 0.756 kcal/mol for $d(CA)$-$d(TG)$, and 0.896 kcal/mol for $d(TA)$ dinucleotides. These $\Delta\Delta G_{H(Z-B)}$ values were compared to the experimental B- to Z-DNA transition energies of −0.56 kcal/mol that we measured for C-5 methylated $d(CG)$, 0.69–1.30 kcal/mol reported for unmethylated $d(CG)$, 1.32–1.48 kcal/mol reported for $d(CA)$-$d(TG)$, and 2.3–2.4 kcal/mol for $d(TA)$ dinucleotides. From this comparison, we found that the calculated $\Delta\Delta G_{H(Z-B)}$ of these dinucleotides could account for the previous observation that the dinucleotides were ordered as $d(m^5CG) > d(CG) > d(CA)$-$d(TG) > d(TA)$ in stability as Z-DNA. Furthermore, we predicted that one of the primary reasons for the inability of $d(TA)$ sequences to form Z-DNA results from a decrease in exposed hydrophilic surfaces of adjacent base pairs due to the C-5 methyl group of thymine; thus, $d(UA)$ dinucleotides should be more stable as Z-DNA than the analogous $d(TA)$ dinucleotides. This prediction was tested and confirmed by the finding that the hexamer sequence $d(m^5CGUAm^5CG)$ crystallized as Z-DNA in 2-fold lower $MgCl_2$ concentrations than the analogous $d(m^5CGTAm^5CG)$ sequence.

Since the first spectroscopic evidence for its existence (Jovin & Pohl, 1972) and the subsequent determination of its structure from single-crystal X-ray diffraction (Wang et al., 1979), left-handed Z-DNA has been an interesting and useful structure for physical studies on macromolecular stability. Solution and high-resolution single-crystal studies of Z-DNA sequences have shown that a variety of environmental conditions stabilize DNA in the Z conformation, including the presence of high salt concentrations, transition-metal complexes, base alkylation, and negative supercoiling [reviewed by Rich et al. (1984) and Jovin et al. (1987)]. Underlying these effects is an intrinsic sequence dependence for the stability of Z-DNA.

Z-DNA is characterized not only by the left-handed twist of the double helix but also by an alternating anti–syn configuration of its base pairs. Since purine bases are sterically able to adopt the syn conformation more readily than are pyrimidines, sequences that are alternating pyrimidine and purine have been reported to be more stable as Z-DNA than sequences that do not follow this alternation rule (Wang et al., 1981; Drew & Dickerson, 1981; Jovin et al., 1983). This alternation in the configuration of the base pairs also requires that the basic repeat unit of Z-DNA be the dinucleotide, or two base pairs, rather than single base pairs as is found in the known right-handed conformations (Jovin et al., 1987). Of the sequences that follow this alternation rule, there exists a hierarchy of pyrimidine–purine dinucleotides that adopt the left-handed conformation, with $d(CG) > d(CA)$-$d(TG) > d(TA)$ in stability as Z-DNA (Jovin et al., 1983). This difference in stability of dinucleotides can be quantitatively re-lated to experimental B- to Z-DNA transition energies. The transition energies for various types of dinucleotides have been calculated by determining the degree of negative superhelicity required to induce Z-DNA within well-defined sequences inserted in closed circular plasmid DNAs. The B- to Z-DNA transition energies for $d(CG)$, $d(CA)$-$d(TG)$, and $d(TA)$ dinucleotides have been measured at 0.66–1.2 kcal/mol (Peck & Wang, 1983; Nordheim et al., 1982; Frank-Kamenetskii & Vologodskii, 1984; Vologodskii & Frank-Kamenetskii, 1984), 1.32–1.48 kcal/mol (Vologodskii & Frank-Kamenetskii, 1984; Mirkin et al., 1987), and 2.3–2.4 kcal/mol (Ellison et al., 1986; Mirkin et al., 1987), respectively.

The availability of this type of experimentally derived structural and thermodynamic information on the stabilities of various sequences as B- versus Z-DNA makes this transition an excellent system for studying the contribution of various thermodynamic forces on macromolecular conformations. A previous calculation of the differences in intramolecular energy for DNA sequences in the B- and Z-conformations using the energy minimization program AMBER could account for the stabilizing effect of methylating cytosines of $d(CG)$ dinucleotides in enthalpic terms but was unable to account for the relative inability of $d(TA)$ dinucleotides to form Z-DNA (Kollman et al., 1982). Similarly, Karplus and co-workers (Tidor et al., 1983; Irikura et al., 1985), using related calculations, were unable to adequately model the salt-induced B–Z transition (Jovin et al., 1987). The shortcomings of these methods fall primarily in their inability to incorporate interactions of DNA with water networks, bulk water, and ions (Jovin et al., 1987). Thus, the stability of Z-DNA and DNA conformations in general may be better understood by considering differences in solvent interactions with their structures.

The contribution of solvent interactions to the stability of alternative DNA conformations has been previously discussed

*Author to whom correspondence should be addressed.

in qualitative terms. It was shown that the stability of right-handed A-DNA under dehydrating conditions can be explained in terms of the difference in the solvent-accessible surfaces of this conformation relative to B-DNA (Alden & Kim, 1979). Sanger et al. (1986) have attributed the propensity of a sequence to adopt the B versus alternative helical forms to the concept of "economy of hydration". They suggest that the more economical hydration of the phosphates in A- and Z-DNA is responsible for their enhanced stabilities under dehydrating conditions. The question that we address here is whether the effect of bulk water structure and interactions on DNA conformation can be described in quantitative terms that account for the actual differences in abilities of sequences to adopt non-B conformations. A useful semiempirical approach to estimating the contribution of solvent interactions on molecular structures, first developed by Langmuir (1925), has been to calculate the overall hydration energy of the molecule and of its component parts. This approach was extended by McLachlan and Eisenberg (1986) to estimate the effect of solvent interactions on "properly" versus "improperly" folded proteins. The lack of experimental data on the stabilities of the various conformational states being considered, however, limited these previous studies to only qualitative evaluations of how the calculated solvation energies contribute to the thermodynamics of macromolecular stability. Recently, we have calculated the difference in hydration energies of methylated and unmethylated d(CG) dinucleotides in the B and Z conformations to study the Z-DNA stabilizing effect due to methylation of the cytosine base (Ho et al., 1988). The stabilization due to methylation was found to result from an overall increase in the hydrophobic character of B-DNA, while the exposed hydrophobic and hydrophilic surfaces of the dinucleotide were essentially identical for both the methylated and unmethylated bases in the Z form.

In this study, we set forth to quantitate the contribution of solvent interactions on the stability of various sequences as Z-DNA by calculating the hydration energies of DNA sequences in the B and Z conformations from their solvent-accessible surfaces (SAS). To circumvent any problems that may result from steric inhibitions to Z-DNA formation, we limited our studies to alternating pyrimidine–purine sequences. The differences in the hydration energies between the Z and B conformations of d(m⁵CG), d(CG), d(CA)-d(TG), and d(TA) were compared to the experimental transition energies for the dinucleotides. From this, we examined the molecular basis for the inability of d(TA) to form Z-DNA in terms of hydrated surfaces. A comparison of d(TA) with d(UA) dinucleotides as Z-DNA showed that the C-5 methyl group of the thymine base contributes significantly to the inability of d(TA) dinucleotides to adopt the left-handed conformation. We therefore predicted that d(UA) dinucleotides should be more stable in the Z conformation than d(TA). This prediction was tested by comparing the conditions for obtaining single crystals of a hexamer sequence that contains d(UA) base pairs as Z-DNA to published conditions for analogous hexamer sequences that contain d(m⁵CG), d(TA), d(CA)-d(TG), and d(CG) dinucleotides.

EXPERIMENTAL METHODS

The atomic coordinates for each atom, including hydrogens, of the sequences d(CGCGCG), d(CACACA)·d(TGTGTG), d(UAUAUA), and d(TATATA) were generated in their right-handed B conformation by using generalized helical parameters (Arnott, 1976) and standard distances and geometries. The coordinates of the sequence d(CGCGCG) as Z-DNA were taken from its single-crystal magnesium struc-

ture (Gessner et al., 1985). The structures of d-(m⁵CGm⁵CGm⁵CG) in the B and the Z conformations were obtained as previously described (Ho et al., 1988). The structures of d(CACACA)·d(TGTGTG) and d(TATATA) were generated by combining dinucleotides from the single-crystal structures of d(CGCGCG) (Gessner et al., 1985) and d(m⁵CGTAm⁵CG) (Wang et al., 1984) and optimizing the bond length, dihedral and torsional angles, and van der Waal's contacts. The solvent-accessible surfaces (SAS) of the DNA models were calculated by using a rolling ball method (Connally, 1983) as previously described (Ho et al., 1988). The surfaces of each atom in the molecules were assigned to different chemical groups (as either carbon, nitrogen, phosphorus, or oxygen and their associated hydrogen surfaces at the aromatic bases, the ribose sugars, or the charged phosphate group).

A set of atomic solvation parameters (ASP) (Eisenberg & MacLachlan, 1986) associated with solvating each surface type of DNA molecules was derived from the partition coefficients of a set of small organic molecules (Table I) that represent the aromatic base, ribose sugar, and phosphate components of nucleic acids. The atomic coordinates of these molecules were generated by standard distances and geometries and their solvent-exposed surfaces calculated by the rolling ball method (Connally, 1983). The calculated surfaces were related to their free energies of hydration (ΔG_H) as derived from experimentally determined partition coefficients of the molecules in aqueous solution and octanol (Hansch & Leo, 1979). From the slope of these relationships, the hydration energy per exposed area (ASP values) was calculated for each atom type for the base, sugar, and phosphate groups of nucleic acids (Table I).

The hydration free energies (ΔG_H) of the dinucleotides were calculated by applying the ASP values to the calculated areas of each surface type (SAS$_i$, where i represents a surface type listed in Table I) of the dinucleotides in the B and Z conformations and summing the resultant energies (Eisenberg & McLachlan, 1986).

$$\Delta G_H = \sum_i ASP_i SAS_i \qquad (1)$$

The surfaces of the two terminal base pairs of each hexamer sequence were excluded from the calculations to avoid errors due to end effects.
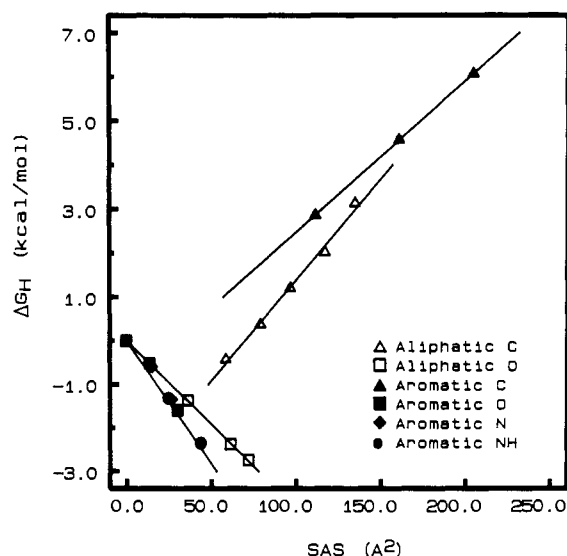
The B- to Z-DNA "transition energy" of d(m⁵CG) dinucleotides was estimated by measuring the ratio of DNA in the left- versus the right-handed conformation of poly(dG-dm⁵C). Polymeric d(m⁵CG) from Pharmacia was dialyzed against several 2-L changes of glass-distilled, deionized water. The sample was titrated with 0.1 M NaCl and then to 5.0 M NaCl in 1 M increments. The ratio of B- and Z-DNA present in solution at each titration point was calculated from the ratio of 295- versus 260-nm absorption as recorded on an HP-8451 spectrophotometer.

RESULTS

Derivation of Atomic Solvation Parameters (ASP) for DNA Surfaces. The thermodynamic weights associated with solvating the various surface types in DNA were derived by first dissecting the nucleotide components of DNA into its substituent groups. The primary components include the ribose sugar, an aromatic base, and a negatively charged phosphodiester. Each component can be further simplified to a sum of surface types, and each surface type can be mimicked by a series of small organic molecules whose energies of hydration can be calculated from their partition coefficients in an oc-

Table I: Calculated Solvent-Accessible Surfaces (SAS) and Free Energies of Transfer (ΔG_H) from Octanol to Water of Aliphatic, Aromatic, and Phosphodiester Compounds

| compound | SAS (Å^2) | | | | ΔG_H* (kcal/mol) |
	C	O	N	PO$_2$	
aliphatic surfaces					
ethanol	58.8	36.1			−0.43
1-propanol	79.3	36.1			0.37
1-butanol	96.8	36.1			1.21
1-pentanol	117.2	36.1			2.02
1-hexanol	134.9	36.1			3.14
methoxyethanol	67.9	61.6			−1.04
ethane-1,2-diol	39.3	72.1			−2.63
aromatic surfaces					
benzene	111.5				2.86
naphthalene	160.8				4.57
anthracene	204.7				6.07
pyrrole	75.19	14.97			1.02
pyridine	91.60	13.70			0.88
pyrimidine	51.53	26.15			−0.54
imidazole	64.09	26.64			−0.11
analine	98.3	24.6			1.23
o-phenylenediamine	86.6	43.8			−0.20
benzophenone	177.2	13.8			4.61
quinone	82.52	29.93			1.87
phosphodiester					
diethyl phosphate	121.2			46.9	0.477
dibutyl phosphate	195.2			52.7	3.01

Table II: Atomic Solvation Parameters (ASP) of Hydrophilic, Hydrophobic, and Charged Phosphate Surfaces in Nucleic Acids As Derived from the Partition Coefficients and Calculated Solvent-Accessible Surfaces of Small Organic Molecules

group	surface type	ASP (kcal mol^{-1} Å^{-2})
ribose	hydrophobic (C)	0.043
	hydrophilic (O)	−0.038
base	hydrophobic (C)	0.034
	methyl (C)	0.043
	hydrophilic (O/N)	−0.055
phosphate	charged (O/P)	−0.100

The ΔG_H value of each compound was corrected for the contributions from the surfaces from the carbon chains by using the ASP value for aliphatic carbons and eq 1. The linearity of this relationship with the inclusion of methoxyethanol shows that the ring oxygen (O1′) of the ribose can be treated as any other hydroxyl surface.

The aromatic bases found naturally occurring in DNA are heterocyclic ring structures composed of surfaces from four atom types: (1) unsaturated carbons, (2) ring nitrogens, (3) conjugated amides, and (4) conjugated carbonyl oxygens. The ASP values of the aromatic carbon surfaces were calculated from the SAS and the ΔG_H of benzene, naphthalene, and anthracene (Figure 1). This relationship yielded a slope of 0.034 kcal mol^{-1} Å^{-2} and a y intercept of −0.98 kcal/mol. The significance of this negative y intercept is unclear. It may be related to an intrinsic solubility of aromatic compounds in water or to a nonlinear relationship for aromatic compounds with less than six atoms. In either case, we will be comparing differences in free energies of hydrations calculated by applying the ASP values to SAS of different DNA surfaces. The y intercept, being an additive value to the surfaces of all aromatic surfaces, will be negated in these differences. The ASP values for the ring nitrogens, amides, and carbonyl oxygens were treated as a single hydrophilic surface type. The ΔG_H and SAS for the aromatic hydrophilic surfaces is linear (slope = −0.055 kcal/mol) with a y intercept of 0. The linearity of the relationship is evidence that these hydrophilic surfaces can indeed be treated as a group, and the y intercept at 0 demonstrates that the ASP value derived for the aromatic carbon surfaces is valid for aromatic molecules that are as small as five non-hydrogen atoms and consisting of only three carbon atoms.

The ASP values for the phosphodiesters were derived from diethyl and dibutyl phosphate. Because of the negative charge associated with the phosphate group, these compounds are very insoluble in octanol. The ΔG_H values were thus calculated for a partitioning of these compounds between *sec*-pentanol and 0.1 M NaClO$_4$ aqueous solution. The ΔG_H value of these two compounds was corrected for the contributions from the aliphatic carbon surfaces, averaged, and divided by the SAS of the phosphodiester group. The phosphorus and oxygen surfaces were treated as a group rather than individually. Since the aqueous solvent is ionic, some counterions are likely to be carried into the organic phase. The partition ΔG_H is likely to be underestimated in this calculation. As we will discuss later, the contribution of the phosphate surface to the differences in the ΔG_H values of the DNA conformations is not dramatic, and the overall conclusions from these studies is not strongly dependent on the precise ASP value of the phosphate groups. The ASP values for all the surface types represented in DNA are listed in Table II.

Free Energies of Hydration (ΔG_H) of Dinucleotides. The calculated solvent-accessible surfaces (SAS) of the dinucleotides d(CG), d(CA)-d(TG), and d(TA) in the B- and Z-DNA conformations show that, overall, the hydration en-

FIGURE 1: Free energies of hydration (ΔG_H) versus solvent-accessible surfaces (SAS) calculated for small organic molecules. The open symbols represent aliphatic surfaces and the solid symbols aromatic surfaces. The symbols for each surface type listed in Table I are defined in the figure.

tanol/water solvent system (Table I). The slope of the relationship between the solvent-accessible surface (SAS) of these molecules and their respective hydration energies defines the ASP value for that surface type (Figure 1).

The ribose component consists of two surface types: (1) aliphatic carbon surfaces and (2) aliphatic oxygen surfaces. These surfaces were mimicked by using a series of linear alcohols. The relationship between the SAS for a set of *n*-alcohols (for carbon lengths from 2 to 6) and their ΔG_H was essentially linear with a slope of +0.43 kcal mol^{-1} Å^{-2} and y intercept of −2.86 kcal/mol. The very negative y intercept is due to negative free energy for hydrating a hydroxyl group having no carbon surface associated with it. The actual ASP value for the aliphatic hydroxyl group was derived from the ΔG_H and SAS of ethanol, methoxyethanol, and ethane-1,2-diol.

Table III: Calculated Hydration and Experimental B-Z Transition Energies (kcal/mol) of Dinucleotides

dinucleotide sequence	hydration energy			B-Z transition energy	
	$\Delta G_{H(B)}$	$\Delta G_{H(Z)}$	$\Delta\Delta G_{H(Z-B)}$	$\Delta G_{T(B-Z)}$	ref
d(m⁵CG)[a]	−4.344	−4.838	−0.494	−0.6 ± 0.3	this work
d(CG)[a]	−5.730	−5.502	0.228	0.66	Peck and Wang (1983)
				0.90	Nordheim et al. (1982)
				1.12	Frank-Kamenetskii and Vologodskii (1984)
				1.20	Mirkin et al. (1987)
d(CG)[b]	−5.689	NA	NA	NA	NA
d(CA)-d(TG)		−5.160	0.756	1.32	Mirkin et al. (1987)
				1.38	Frank-Kamenetskii and Vologodskii (1984)
				1.40	Vologodskii and Frank-Kamenetskii (1984)
d(TA)	−5.242	−4.346	0.896	2.3	Mirkin et al. (1987)
				2.4	Ellison et al. (1986)
d(UA)	−5.032	−4.532	0.500	NA	NA

[a] The thermodynamic ASP values used to calculate the hydration energies of d(m⁵CG) and d(CG) (Ho et al., 1988) were those derived for proteins from Eisenberg and MacLachlan (1986). The values reported in this study have been updated by using ASP values that are more applicable to the methods used to calculate the solvent-accessible surfaces and the functional groups of nucleic acids. Although this affects the details of the results from the previous study, it does not alter the overall results or conclusions drawn from that study. [b] Average hydration energies calculated from the 2nd, 3rd, 10th, and 11th CG base pairs of the d(CGCGAATTCGCG) B-DNA crystal structure (Dickerson & Drew, 1981).

ergies of unmodified dinucleotides are more positive in the left-handed Z conformation than in the B conformation (Table III). This is consistent with findings that Z-DNA is stabilized in solution by dehydrating conditions such as high salt (Pohl & Jovin, 1972; Wang et al., 1981; Drew & Dickerson, 1981) or alcohol concentrations (Pohl, 1976) and is accounted for by the overall more positive energy of hydration for the phosphates in Z-DNA than in B-DNA. The hydration energies of alternating pyrimidine–purine dinucleotides increase as the ratio of TA base pairs in the dinucleotides increases, regardless of their conformation. These results indicate that alternating CG base pairs have a lower hydration energy than do TA base pairs. This base-pair dependence of the hydration energies agrees with the report that the partition coefficient (aqueous to organic solvent) is more positive for polyribo(AU) than for polyribo(CG) (Garel et al., 1972).

The hydration energies (ΔG_H) calculated for d(CG) base pairs generated in the B-DNA conformation by using standard helical parameters were compared to the hydration energies calculated for the analogous base pairs in the B-DNA crystal structure of the self-complementary sequence d-(CGCGAATTCGCG) (Dickerson & Drew, 1981). The ΔG_H calculated for individual CG base pairs within the crystal structure varied by as much as 20%. The average ΔG_H of the model CG base pairs (−5.73 kcal/mol of base pair), however, differed by less than 1% from the average ΔG_H calculated for the CG base pairs from the crystal structure (−5.69 kcal/mol of base pair). Thus, even though the hydration of individual base pairs may deviate from the average value, we believe that the structures generated for each sequence are valid representations of B-DNA in solution and that the SAS calculated provide reasonable values for determining the hydration energies of representative base pairs of each sequence in solution.

Differences in Hydration Free Energies between Z- and B-DNA ($\Delta\Delta G_{H(Z-B)}$). If we accept that the difference in stability between B- and Z-DNA is dependent on solvent interactions, then the difference in hydration energies between the conformations ($\Delta\Delta G_{H(Z-B)}$) should be more negative for sequences that favor Z-DNA formation and more positive for sequences that favor the B conformation. Thus, $\Delta\Delta G_{H(Z-B)}$ values should increase according to the hierarchy d(TA) > d(CA)-d(TG) > d(CG) established for the instability of these dinucleotides as Z-DNA. To test this, we would need to compare the calculated $\Delta\Delta G_{H(Z-B)}$ values to their stabilities as Z-DNA. The stabilities of the dinucleotides d(CG), d-(CA)-d(TG), and d(TA) in the right- versus left-handed conformations at low ionic strengths have been determined by

various groups as B–Z transition energies, $\Delta\Delta G_{T(B-Z)}$. The values of $\Delta G_{T(B-Z)}$ are a measure of the energy available from negative supercoiling required to induce Z-DNA formation from B-DNA within well-defined sequences inserted into closed circular plasmids. These values should thus be more positive for sequences that do not readily adopt the Z conformation. In all these studies, the degree of negative superhelicity that is required to induce the formation of Z-DNA was determined for well-defined inserts in closed circular plasmids. Using a two-state zipper model to analyze the transition, several laboratories derived a set of energies for initiating the B–Z transition and for propagating Z-DNA formation through the various inserts (see Table III). Since these measurements were made at low ionic strengths ([Na⁺] < 3 mM), the transitions observed are primarily supercoil induced. The propagation energies are thus a measure of the relative intrinsic stabilities of the dinucleotides as Z-DNA under low ionic strength conditions.

Stability of Poly(dG-dm⁵C) as Z-DNA. Aside from the base-sequence dependence of Z-DNA stability, methylation of cytosine bases at the C-5 position has been shown to stabilize poly(dG-dC) in the left-handed form (Behe & Felsenfeld, 1981). The B- to Z-DNA transition energy of d(m⁵CG) dinucleotides, however, has not been previously determined. To estimate the relative stability of this dinucleotide as B- and Z-DNA under conditions similar to that of the supercoil-induced transitions, the equilibrium ratio of the two conformations present in polymeric d(m⁵CG) was spectroscopically determined after extensive dialysis against deionized water. The polymer under these low ionic strength conditions was observed to exist as a 0.65:0.35 mixture of Z- and B-DNA (Figure 2), equivalent to an equilibrium constant ($K_{Z/B}$) of 1.86 or a difference in stability of −0.6 ± 0.3 kcal/mol of dinucleotide. The addition of 0.1 M sodium chloride to the solution shifted the equilibrium to an entirely B conformation. The transition to Z-DNA was not observed again until greater than 2 M NaCl had been added.

These titration results show that poly(dG-m⁵dC) is stable as Z-DNA under low salt conditions but is not entirely in the left-handed conformation as observed in several previous papers (Narasimhan & Bryan, 1984; Kruger & Prairie, 1985; Latha & Brahmachari, 1985; Fuerestein et al., 1985). Devarajan and Shafer suggested that the Z-DNA observed under these conditions is due to contaminating divalent cations such as Mg²⁺ and Ca²⁺ (Devarajan & Shafer, 1986). They reported that under conditions of extremely low divalent cation concentrations poly(dG-m⁵dC) is in the B conformation. How-
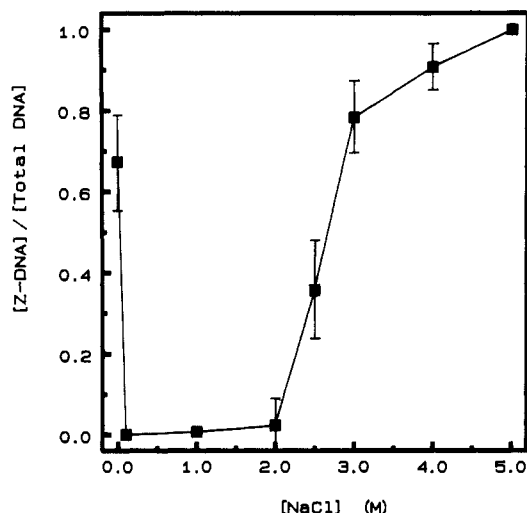
FIGURE 2: Titration of poly(dG-dm⁵C) with NaCl. The fraction of Z-DNA to the total DNA in solution was calculated from the 295- to 260-nm absorption ratios at each titration point. The data represent three independent sets of titrations.

ever, a different group (Feuerstein et al., 1985) reported that, under similar dialysis conditions, poly(dG-m⁵dC) is in the left-handed conformation. Because we attempted in our titrations to mimic the conditions of the supercoil-induced transitions, divalent cations were not rigorously depleted from the polymer. It is thus not unreasonable that this polymer is poised at an equilibrium point between the two conformations under our conditions.

Relationship of $\Delta\Delta G_{H(Z-B)}$ *to Z-DNA Stability.* Table III is a comparison of $\Delta\Delta G_{H(Z-B)}$ calculated for the dinucleotides d(m⁵CG), d(CG), d(CA)-d(TG), and d(TA) to their respective thermodynamic stabilities of Z- versus B-DNA, as defined experimentally from the B–Z transition energies ($\Delta G_{T(B-Z)}$). From the small values of $\Delta G_{T(B-Z)}$, we would expect that the differences in the hydration free energies ($\Delta\Delta G_{H(Z-B)}$) should also be small, and they were. These small values result from the difference of rather large numbers and represent at most a 16% change in the overall hydration energies calculated for each dinucleotide. One can ask then whether these small perturbations are relevant. Hydration of the phosphate groups contributes significantly to the magnitude of the negative ΔG_H values, owing to the large negative ASP for hydration of phosphate surfaces. The differences in the overall exposure of the phosphates between a dinucleotide in the B- versus the Z-conformation is 2.6 ± 2.3 Å². This translates into an overall free energy of hydration difference of 0.26 ± 0.23 kcal/mol from d(m⁵CG) to the d(TA) dinucleotides. The deviation across these differences in phosphate exposure is small. This would be expected since the overall backbones are nearly identical for alternating purine and pyrimidine sequences in each conformation. The $\Delta\Delta G_{H(Z-B)}$ ranges from –0.49 to 0.90, or a 1.39 kcal/mol of dinucleotide difference from d(m⁵CG) to d(TA). The contribution to the $\Delta\Delta G_{H(Z-B)}$ variation resulting from differences in the hydration free energies of the phosphates is thus approximately 17%. The remainder is due to differences in the exposure of the bases. A 20% error in the phosphate ASP values would increase the contribution of the phosphates from 17% to 20% toward the $\Delta\Delta G_{H(Z-B)}$ values.

In Figure 3, $\Delta G_{T(B-Z)}$ is plotted against $\Delta\Delta G_{H(Z-B)}$. Within the variation of the reported $\Delta G_{T(B-Z)}$ values, this relationship is linear. A linear least-squares fit of this plot yields a relationship having a slope of 0.491 and a *y* intercept at –0.130 kcal/mol, with a correlation coefficient (R) of 0.96. The slope < 1 of this relationship suggests that the hydration energies
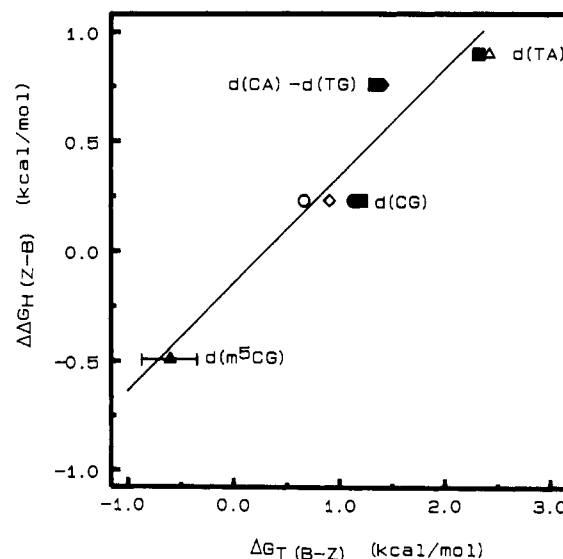

FIGURE 3: Comparison of the calculated difference in hydration energies of Z- and B-DNA with the experimentally determined B- to Z-transition energies. A linear least-squares fit of these points yielded the relationship $\Delta\Delta G_{H(Z-B)}$ = 0.491$\Delta G_{T(B-Z)}$ – 0.130, with a correlation coefficient (R) of 0.96. (▲) Data from this work; (O) data from Peck and Wang (1983); (◇) data from Nordheim et al. (1982); (●) data from Frank-Kamenetskii and Vologodskii (1984); (■) data from Mirkin et al. (1987); (◆) data from Vologodskii and Frank-Kamenetskii (1984); (△) data from Ellison et al. (1986).

Table IV: $\Delta\Delta G_{H(Z-B)}$ (kcal/mol) of Hydrophilic and Hydrophobic Surfaces of Dinucleotide Bases and Ribose Groups

dinucleotide sequence	$\Delta\Delta G_{H(Z-B)}$ hydrophilic		$\Delta\Delta G_{H(Z-B)}$ hydrophobic	
	base	ribose	base	ribose
d(m⁵CG)	0.08	–0.26	–0.06	–0.78
d(CG)	0.14	–0.00	0.12	–0.02
d(CA)-d(TG)	0.18	–0.06	0.18	–0.22
d(TA)	0.12	–0.02	0.04	0.32
d(UA)	0.04	0.02	0.10	0.12

calculated consistently underestimate the experimental transition energies. Both the calculated $\Delta\Delta G_{H(Z-B)}$ and the experimental $\Delta G_{T(B-Z)}$ values are measures of the stability of Z-DNA versus B-DNA. The observation of a one-to-one relationship between stability and hydration energies with a positive slope and a *y* intercept approximately equal to zero and the calculated and experimental energies being of the same orders of magnitude suggest that indeed the transition of a base sequence from B- to Z-DNA is at least partially accounted for by the difference in hydration of the sequences in the two conformations.

Destabilization of Z-DNA by d(TA) Dinucleotides. A better understanding of the molecular basis for the differences in the hydration of these sequences and their dependence on conformation can be attained by comparing the contributions of each component surface to the overall hydration energy calculated. The difference between solvation of the hydrophobic and of the hydrophilic surfaces of the aromatic base and ribose groups for the various dinucleotides in their B and Z conformations is summarized in Table IV. By analyzing the difference in exposed surfaces for Z- versus B-DNA of d(CG), d(CA)-d(TG), and d(TA), we can start to understand why TA base pairs destabilize Z-DNA. The average total difference in exposed hydrophobic surfaces of the sequences in Z- versus B-DNA is higher for d(TA) than for d(CG) dinucleotides, while that for d(CA)-d(TG) is intermediate between the two. This is observed as resulting from a markedly

greater exposure of the hydrophobic carbon surfaces and lower exposure of the hydrophilic oxygen surface of the ribose sugars in TA dinucleotides. The difference in exposure of both the hydrophobic and hydrophilic surfaces of the aromatic bases is lower in the d(TA) dinucleotides when compared to d(CG) base pairs.

These results can be attributed to the difference in the stacking arrangement of TA base pairs as compared to CG base pairs. In the purine–pyrimidine stacking of base pairs in the Z conformation, the purine bases stack directly above the ribose of adjacent base pairs. A comparison of the stacking arrangements of CG:CG to TA:TA dinucleotides as Z-DNA is shown in Figure 4A,B. In Figure 4, the larger solid circles represent the contact points of solvent molecules with hydrophobic surfaces, while the smaller open circles represent the contact points with hydrophilic surfaces as calculated by the rolling ball method (Connally, 1983). The absence of the N2 amino group of adenine exposes the surface of its adjacent thymine ribose in the d(TA) dinucleotides relative to the d-(CG) dinucleotides, resulting in a greater exposure of hydrophobic surface in the minor groove surface of d(TA). The overall exposure of the carbon surfaces of the bases, particularly of adjacent base pairs, is actually lower in the TA dinucleotides than in the CG dinucleotides, even though the methyl group itself contributes to the local exposure of hydrophilic surfaces at the major groove surface of the TA base pairs. The higher overall hydrophobicity of the d(TA) dinucleotides in the Z conformation results from the greater increase in the exposed carbon surface of ribose as compared to the decrease in the exposed carbon surfaces of the bases.

The surface areas of the accessible hydrophilic groups are slightly higher for d(CG) base pairs as Z-DNA than as B-DNA, resulting in negative $\Delta\Delta G_{H(Z-B)}$ values for the oxygens of the riboses and the oxygens and nitrogens of the aromatic bases. The hydrophilic areas of the TA base pairs are less accessible in the Z versus the B conformations. $\Delta\Delta G_{H(Z-B)}$ becomes significantly more positive as the ratio of TA base pairs is increased from 0 to 1 to 2 per dinucleotide. Thus, overall, the addition of TA base pairs to a DNA sequence increases the difference in the hydration free energies of that sequence as Z- and as B-DNA.

Stability of d(UA) Dinucleotides as Z-DNA. To determine the extent to which the thymine C-5 methyl contributes to these differences in hydration free energies of d(TA) base pairs in the left-handed conformation, these same hydration energy calculations were repeated for the sequence d(UA)$_3$ as B- and Z-DNA (Table IV). This sequence is identical with d(TA)$_3$, except that the methyl groups of the pyrimidines have been removed to form the analogous deoxyuridine. The exposure of hydrophobic surfaces in the minor groove crevice are nearly identical for the d(TA) and d(UA) dinucleotides (Figure 4B,C). The accessibility of the ribose carbons is, thus, not significantly affected by the methyl group, confirming the earlier suggestion that the lack of an N2 amino group on the adenine bases is responsible for the greater exposure of the ribose to solvent in d(TA) dinucleotides. Alternatively, the major groove surface of d(UA) appears much more hydrophilic overall than hydrophobic as compared to the comparable surface of d(TA) and even the major groove surface of d(CG) (Figure 4). The $\Delta\Delta G_{H(Z-B)}$ values in Table IV show that removing this methyl group results in an increase in the accessibility of both the hydrophobic and hydrophilic groups of the aromatic bases and a decrease in the hydrophilic surface of the ribose. This is consistent with previous reports that the C-5 methyl group of pyrimidine bases sits in a pocket formed
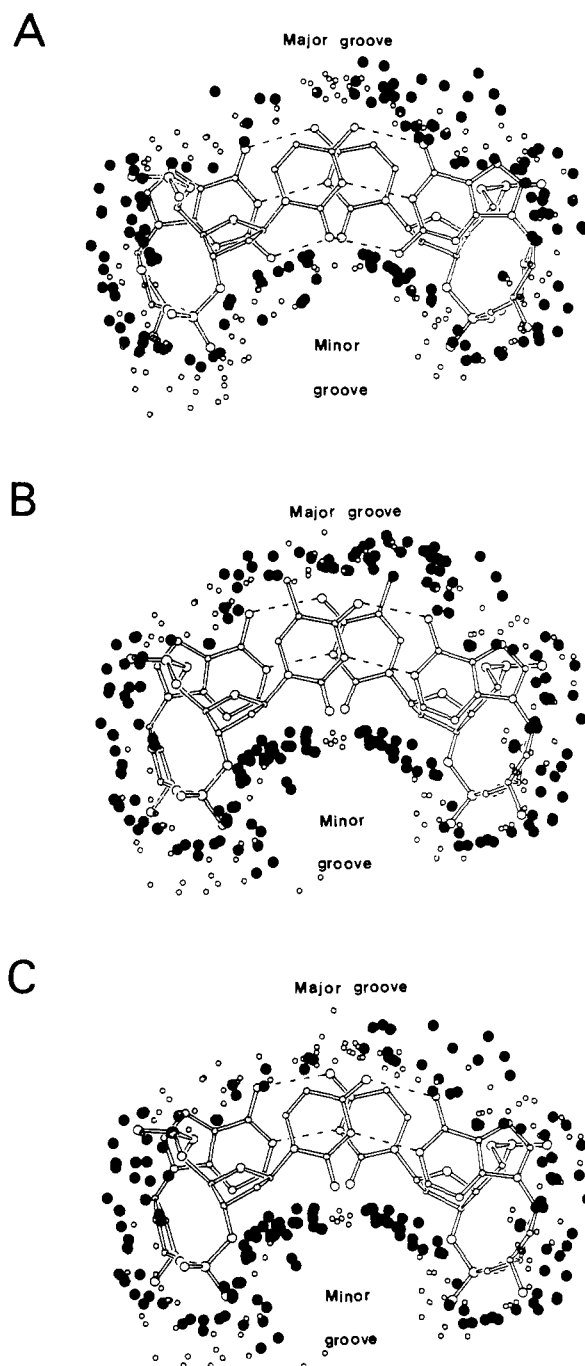

FIGURE 4: Stacking diagrams of the (A) d(CG), (B) d(TA), and (C) d(UA) dinucleotides and their solvent-accessible surfaces. Shown are the purine–pyrimidine base pairs stacked on top of a pyrimidine–purine base pair. The large solid circles represent the calculated hydrophobic surfaces of carbons, and the smaller open circles represent the calculated hydrophilic surfaces of the oxygens and nitrogens. The surfaces calculated with 1.0 dot/Å^2 are shown and the hydrogen atoms removed from the base pair models for clarity. The SAS values used to determine the ΔG_H values were calculated with a density of 5 dots/Å^2.

by the ribose and the stacked purines of adjacent base pairs (Ho et al., 1988; Fuji et al., 1982). Contrary to previous reports (Jovin et al., 1983), however, the overall effect of the C-5 methyl appears to be different for thymine as compared to cytosine.

The C-5 methyl group of the pyrimidine bases effectively increases the overall exposed surface of B-DNA. The surface in Z-DNA, however, is decreased because the methyl group sits in a pocket. The effect of the C-5 methyl on the stability of Z-DNA, therefore, is dependent on the atom types that line

the surface that makes up this pocket. For cytosine bases in alternating d(CG) dinucleotides, the C-5 pocket is lined equally by hydrophobic and hydrophilic atoms from the ribose and bases of adjacent base pairs. Thus, the decrease in exposed hydrophilic surface is accompanied by a near equivalent decrease in the hydrophobic surface, including the additional methyl surface (Ho et al., 1988). In the B conformation, this methyl group sits exposed in the minor groove. This accounts for the stabilizing effect of methylation on d(CG) dinucleotides as Z-DNA.

The C-5 pocket of d(TA) dinucleotides, as with d(CG), in the Z conformation is lined by both hydrophobic and hydrophilic surfaces. One difference in the effect of the methyl group on the stability of Z-DNA lies in how this group sits within the pocket. The C-5 methyl of d(TA) is more exposed in the major groove surface than that of d(m^{5}CG) by more than 1 Å^2 per methyl group. This is equivalent to a 0.2 kcal/mol destabilization of d(TA) dinucleotides relative to d(CG) by the methyl group. If we consider only the exposed surface of the methyl, however, we cannot account for the entire effect of the addition of one carbon. This methyl also affects the exposure of the neighboring atoms to solvent. Methylation decreases the exposure of all atoms in this pocket, but differently for the two types of dinucleotides. Methylation decreases the hydrophobicity of the d(CG) due to exposure of carbon atoms in the bases by 0.36 kcal/mol, as compared to 0.2 kcal/mol for d(TA) dinucleotides. Alternatively, methylation of d(CG) decreases the contribution of the hydrophilic oxygen surfaces to the hydration of the riboses by only 0.04 kcal/mol, while the methyl group of d(TA) buries a hydrophilic surface equivalent to 0.26 kcal/mol for each dinucleotide. These effects are thus dependent on the positioning of the methyl group relative to the adjacent base pairs and riboses resulting from the stacking arrangements in Z-DNA.

Relationship of $\Delta\Delta G_{H(Z-B)}$ to Crystallization of Hexamers as Z-DNA. From these observations, we predict that the propensity of d(UA) dinucleotides to form Z-DNA is greater than that of the analogous C-5 methylated d(TA) dinucleotides. The calculated $\Delta\Delta G_{H(Z-B)}$ value for d(UA) dinucleotides (0.500 kcal/mol) is 0.396 kcal/mol lower than that of d(TA). This translates into a 0.81 kcal/mol lower $\Delta G_{T(B-Z)}$, as calculated from the relationship from Figure 3, that would result by removing the methyl group of d(TA) dinucleotides. The d(UA) dinucleotide is still 0.55 kcal/mol less stable as Z-DNA than d(CG) and 2.02 kcal/mol less stable than the d(m^{5}CG) dinucleotides, but should have the equivalent ability to form Z-DNA as d(CA)-d(TG). To test this prediction, the hexamer sequence d(m^{5}CGUAm^{5}CG) was synthesized and crystallized. We have solved the structure of this sequence to better than 1.3-Å resolution and found it to be in the Z conformation (Zhou and Ho, unpublished results). This hexamer sequence is analogous to the sequence d-(m^{5}CGTAm^{5}CG) and d(m^{5}CGm^{5}CGm^{5}CG), both of which have been previously crystallized in the Z conformation (Wang et al., 1984; Fuji et al., 1982). A comparison of the crystallization conditions of d(m^{5}CGUAm^{5}CG) with those for d(m^{5}CGm^{5}CGm^{5}CG) and d(m^{5}CGTAm^{5}CG) would allow us to determine whether in fact d(UA) dinucleotides are more stable in the Z conformation than d(TA) and less stable than d(m^{5}CG) dinucleotides. We assume in this comparison that there is a critical equilibrium concentration of the hexamer sequences in the Z conformation that is required for initiating and propagating the crystal packing of Z-DNA in its crystal lattice and that this equilibrium concentration is strongly dependent on the salt concentration during crystallization.

Z-DNA is stabilized by both mono- and divalent cations (Pohl & Jovin, 1972; Wang et al., 1981; Drew & Dickerson, 1981). It is therefore not surprising that the conditions required to crystallize hexamer sequences as Z-DNA include high concentrations of both Na$^+$ and Mg^{2+}. The effectiveness of each type of cation at stabilizing Z-DNA in solution is dependent on the charge as well as its concentration. As expected, Mg^{2+} is much more effective at stabilizing Z-DNA than Na$^+$ (Chen et al., 1984). The ionic conditions for stabilizing Z-DNA can thus be related to the ionic strength of the solution, which is dependent on both the concentration and the square of the ionic charge (Z^2). Even though ionic strengths may not be a valid representation of specific interactions of cations with the DNA, it is a useful measure of the ionic conditions of the bulk solution. Since anions do not appear to play a significant role in Z-DNA stability, the Z-DNA crystallization conditions can be translated into cationic strengths (CS) as defined by

$$CS = Z^2_{Na^+}[Na^+] + Z^2_{Mg^{2+}}[Mg^{2+}] \qquad (2)$$

We would expect that sequences more readily stabilized as Z-DNA would require lower concentrations of Na$^+$ and Mg^{2+} to be stable in the Z conformation and, therefore, lower concentrations of these cations to form crystals as Z-DNA. Synthetic hexanucleotide sequences that crystallize as Z-DNA have been found to be isomorphous and, therefore, provide us with a consistent data set to compare the abilities of sequences to form Z-DNA. The hexamers in these crystals are packed end-to-end to form essentially continuous strands of Z-DNA aligned along the crystallographic *c* axis. The stability of the dinucleotides in the hexamers as the Z versus B conformations can thus be assessed as if the dinucleotides were all within continuous strands of Z-DNA without the need to consider end effects. The average difference in solvation of the dinucleotides in each hexamer was estimated by summing the hydration energies of each component dinucleotide and dividing this by the number of dinucleotides in the hexamer. This value can be translated into a term reflecting the average stability of the dinucleotides $\Delta G_{T(B-Z)}$ in the hexamers as Z-DNA by using the relationship derived from Figure 3.

The initial crystallization conditions and the equilibrium cationic strength calculated for crystallization of the hexamer sequences d(m^{5}CG)$_3$, d(m^{5}CGTAm^{5}CG), d(CG)$_3$, d-(CACGTG), and d(m^{5}CGUAm^{5}CG) as Z-DNA are listed in Table V, along with the $\Delta G_{T(B-Z)}$ calculated from their $\Delta\Delta G_{H(Z-B)}$ values. The hexamer sequence d(m^{5}CGUAm^{5}CG) was predicted to be more stable as Z-DNA than the analogous d(m^{5}CGTAm^{5}CG) sequence and less stable than d-(m^{5}CGm^{5}CGm^{5}CG). We found that indeed the cationic strength required to form diffraction-quality single crystals of Z-DNA of the uridine-containing hexamer was approximately 2-fold lower than that reported for the sequence d-(m^{5}CGTAm^{5}CG). When these were compared to the crystallization conditions of the sequence d(m^{5}CG)$_3$, we observed that methylated d(CG) dinucleotides require the lowest cation strengths to obtain Z-DNA crystals. Thus, the dependence of Z-DNA crystal formation on these salt conditions mirrors previous findings that methylation of cytosine at the C-5 position stabilizes the Z conformation and d(TA) dinucleotides are the least stable as Z-DNA. The d(UA) dinucleotide, as predicted, is intermediate in its ability to crystallize as Z-DNA.

An overall comparison of the log of the cationic strength for crystallization of each hexamer and their associated B-Z transition energies is shown in Figure 5. This relationship is essentially linear with a slope calculated to be 0.585 and the *y*-intercept of −0.306 M from a linear least-squares fit with

Table V: Calculated B–Z Transition Free Energies ($\Delta G_{T(B-Z)}$) and Crystallization Conditions for Synthetic Hexamers as Z-DNA

hexamer sequence	B–Z equilibrium/ $\Delta G_{T(B-Z)}$	crystallization conditions (M)				ref
		[Na$^+$]	[Mg^{2+}]	CS^a	log (CS)	
d(m^{5}CG)$_3$	−0.706	0.150	0.020	0.230	−0.638	Fuji et al. (1982)
d(m^{5}CGUAm^{5}CG)	−0.031	0.180	0.045	0.360	−0.444	Zhou and Ho (unpublished results)
d(m^{5}CGTAm^{5}CG)^b	0.238	0.188	0.094	0.563	−0.249	Wang et al. (1984)
d(CG)$_3$	0.765	0.400	0.500	2.200	0.342	Gessner et al. (1985)
d(CACGTG)	1.482	0.460	0.700	3.26	0.513	Coll et al. (1988)

^{a}CS is the cationic strength of the crystallization conditions at equilibrium as calculated by CS = $Z^2_{Na^+}[Na^+] + Z^2_{Mg^{2+}}[Mg^{2+}]$, where Z is the cation charge. ^{b}The polycation spermine was reported for the crystallization conditions of d(m^{5}CGTAm^{5}CG) but was not found in the crystal structure. For this reason, we excluded this polycation from the CS calculations. We have recently crystallized this sequence in the absence of spermine under a cationic strength of 0.62 M (Zhou and Ho, unpublished results), confirming that the presence of this polycation is not essential to obtaining Z-DNA crystals of this hexamer.
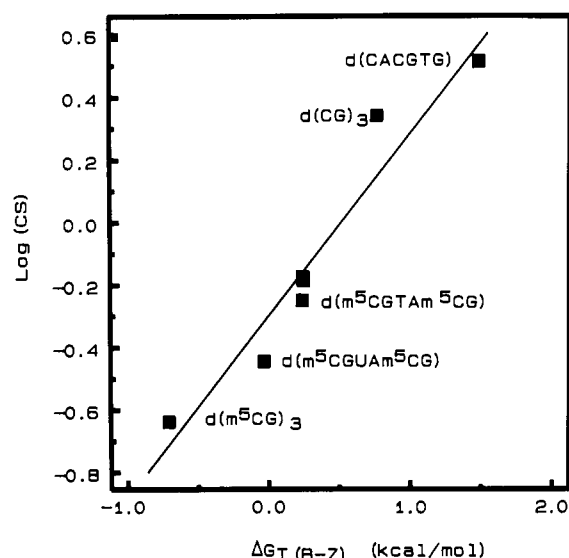

FIGURE 5: Free energies of transition calculated from the hydration energies versus logarithm of cationic strengths for crystallizing synthetic hexanucleotides as Z-DNA. Average hydration energies were calculated for the dinucleotides within each hexamer by summing the hydration energies for each type of dinucleotide present and dividing by 3. Cationic strengths were calculated from the equilibrium sodium cacodylate and MgCl$_2$ concentrations present under the conditions found to give diffraction-quality crystals of each hexamer sequence as Z-DNA (Table V). A linear least-squares fit of the data yielded a line described by log (CS) = 0.586$\Delta G_{T(B-Z)}$ − 0.306.

$R = 0.96$. The hexamer sequences that did not contain methylated cytosines were found to be overall less stable as Z-DNA than sequences containing methylated cytosines, regardless of the number of d(TA) base pairs in the structure, according to both the hydration energy calculations and the cationic strength values. This can be attributed to the very strong Z-DNA stabilizing ability of the flanking d(m^{5}CG) dinucleotides in the three methylated hexamers. Thus, the hydration energy calculations were found to be useful not only for predicting *whether* alternating pyrimidine and purine dinucleotides form Z-DNA but also for predicting the *extent* to which they will adopt the left-hand conformation.

DISCUSSION

In this study, we have examined the quantitative relationship between calculated solvent-accessible surfaces of dinucleotide sequences in B- and Z-DNA and the experimental transition energies for converting B- to Z-DNA for these dinucleotides. We found that a linear relationship can be drawn between the calculated energies and the experimentally determined transition energies. In addition, the magnitude of the differences in stabilities calculated from the hydration energies of the dinucleotide as Z- and B-DNA is similar to the B- to Z-DNA transition energies. This comparison showed that a simple model in which the stability of Z-DNA is attributed to dif-

ferences in hydration of the DNA dinucleotides can account for the experimentally observed hierarchy of d(m^{5}CG) > d(CG) > d(CA)-d(TG) > d(TA) dinucleotides in stability as Z-DNA.

The hydration energies calculated assume that all solvent interactions are averaged over the DNA molecule. To accommodate specific cationic interactions in these calculations, we would need to locate the positions of the cations from high-resolution crystal structures and treat these as essentially integral parts of the DNA structures. The cations in even the best-studied crystal structures of B- and Z-DNA sequences, however, have not all been located. It is also very unlikely that cations would be fixed in solution at all of these crystallographically determined positions. Soumpasis has quantitatively accounted for the high-salt transition of B- to Z-DNA in poly(dC-dG) by treating the main Coulombic interactions as an ion–ion potential in a homogeneous electrolyte (Soumpasis, 1984). Given our inability to accurately include specific counterions into the hydration energy calculations, the cation interactions were treated as one parameter in the solvent interactions that is essentially constant during the transition. This simplification appears to be valid primarily because the $\Delta\Delta G_{T(B-Z)}$ used as the measure of Z-DNA stability in these studies are restricted to low ionic strength conditions and are primarily supercoil induced.

By including the interactions of the cations into a general solvent interaction term, we have ignored the electrostatic contribution of the aquo–metal complexes to the stabilization of Z-DNA. The solvent interactions are thus underestimated by the hydration energy calculations. Since the strongest electrostatic stabilizations from the cations would be expected to be at the phosphate backbone of both the B and the Z conformations of DNA (Soumpasis, (1984), the electrostatic contribution to the hydration energies of the bases should be minimal. The electrostatic terms should therefore contribute a systematic error to the hydration energy calculations. Despite this, a relationship between the calculated differences in hydration and transition energies for B- and Z-DNA was derived that appears to be useful in predicting the relative stabilities of left- versus right-handed DNA.

An analysis of how the hydrophilic and hydrophobic groups contribute to the overall hydration energies of each dinucleotide allows us to understand the differences in stability of d(m^{5}CG), d(CG), d(CA)-d(TG), and, in particular, d(TA) dinucleotides as Z-DNA at the atomic level. We find that the decreased overlap between the adenine and adjacent ribose resulting from the stacking of TA and UA base pairs in Z-DNA is responsible for the greater exposure of the ribose carbon surfaces when compared to that of d(CG) dinucleotides. In addition, the methyl group of thymine in d(TA) base pairs destabilizes Z-DNA, contrary to the effect of C-5 methylation of cytosines by effectively decreasing the accessibility of the hydrophilic

base and ribose surfaces to solvent. Thus, removal of this methyl group should help to stabilize d(UA) dinucleotides as Z-DNA. By comparing d(TA) to the d(UA) dinucleotides in their Z- and B-DNA conformations, we can assign a 0.81 kcal/mol contribution of the C-5 methyl group to the instability of d(TA) dinucleotides as Z-DNA. The effect of the C-5 methyl group on the stability of DNA conformations is not limited to the B and Z conformations. The thymine methyl group in poly(dA-dT) has, in contrast, been reported to help stabilize an alternating B and an X form of DNA (Vorlickova & Jaroslav, 1984). Since, in these studies, we dealt only with two states for the conformation of each dinucleotide, the contribution of other forms of DNA to the overall equilibrium state of these sequences has been ignored. A significant contribution of other forms of DNA to the equilibrium, for example, the alternating B and the X forms of d(TA), would tend to effectively destabilize the Z conformation of that sequence. This two-state approximation may thus contribute to underestimation of the relative instability of Z-DNA for the dinucleotides d(CA)-d(TG) and d(TA) by the hydration energy calculations.

A comparison of the crystallization conditions required to obtain single crystals of the hexamer sequence d-(m^{5}CGUAm^{5}CG) in the Z conformation demonstrated that indeed the methyl group of thymine contributes to the inability of d(TA) dinucleotides to form Z-DNA, as predicted from the calculations. The concentration of cations present at equilibrium in the conditions required to crystallize d-(m^{5}CGUAm^{5}CG) as DNA is approximately 2-fold lower than that of the sequence d(m^{5}CGTAm^{5}CG). In fact, a linear relationship can be obtained by comparing the cationic strength required to crystallize the sequences d(m^{5}CGm^{5}CGm^{5}CG), d(m^{5}CGUAm^{5}CG), d(m^{5}CGTAm^{5}CG), and d(CGCGCG) with the Z- to B-DNA equilibria calculated from the solvent-accessible surfaces. This may ultimately be a useful relationship for helping to define the optimum conditions for crystallizing synthetic hexamers as Z-DNA.

Our prediction that d(UA) dinucleotides are as stable in the Z conformation as are d(CA)-d(TG) dinucleotides implies that the probability of forming Z-DNA is greatly increased in a genome that has uridine extensively misincorporated in place of thymines. The increased probability of forming an alternative helical conformation would be expected to adversely affect the normal pathways of genomic function and may be partially responsible for the mechanisms present in most organisms to ensure that deoxyuridine is not incorporated into the genome and to replace uridines that are misincorporated. For example, *Escherichia coli* strains that are combined *dut* (deoxyuridine triphosphatase) and *ung* (uracil-DNA *N*-glycosylase) mutants have been reported to have up to 20% of thymine residues in their DNA replaced by uridines (Warner et al., 1981). Such mutants have been further shown to be *Lac* constitutive (Duncan, 1981), a consequence of interference of the *Lac* operon at an essential thymine in its operator sequence (Fisher & Caruthers, 1979), possibly resulting from an altered helical structure in this region. Evidence that protein recognition is dependent on DNA conformation can be seen in the requirement that the sequence d-(GCGCGC) be in the B and not the Z conformation to be restricted by the enzyme *Bss*HII (Azorin et al., 1984; Ellison et al., 1987).

REFERENCES

Alden, C. J., & Kim, S. H. (1979) *J. Mol. Biol. 132*, 411–434.

Arnott, S., Smith, P. J. C., & Chandrasekaran, D. L. (1976) in *Handbook of Biochemistry and Molecular Biology* (Fasman, G. D., Ed.) pp 411–422, CRC Press, Cleveland, OH.

Azorin, F., Hahn, R., & Rich, A. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 5714–5718.

Behe, M., & Felsenfeld, G. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 1619–1623.

Chen, H. H., Behe, M. J., & Rau, D. C. (1984) *Nucleic Acids Res. 12*, 2381–2389.

Coll, M., Fita, I., Lloveras, J., Subirana, J. A., Bardella, F., Huynh-Dinh, T., & Igoleu, J. (1988) *Nucleic Acids Res. 16*, 8695–8705.

Connally, M. L. (1983) *Science 221*, 709–713.

Devarajan, S., & Shafer, R. H. (1986) *Nucleic Acids Res. 14*, 5099–5108.

Dickerson, R. E., & Drew, H. R. (1981) *J. Mol. Biol. 149*, 761–786.

Drew, H. R., & Dickerson, R. E. (1981) *J. Mol. Biol. 151*, 535–556.

Duncan, B. K. (1981) *Enzymes (3rd Ed.) 1*, 565–586.

Eisenberg, D., & McLachlan, A. D. (1986) *Nature 319*, 199–203.

Ellison, M. J., Feigon, J., Kelleher, R. J., III, Wang, A. H.-J., Habener, J. F., & Rich, A. (1986) *Biochemistry 25*, 3548–3555.

Ellison, M. J., Fenton, M. J., Ho, P. S., & Rich, A. (1987) *EMBO J. 6*, 1513–1522.

Feuerstein, G., Marton, L. J., Keniry, M. A., Wade, D. L., & Shafer, R. H. (1985) *Nucleic Acids Res. 13*, 4133–4141.

Fisher, E. F., & Caruthers, M. H. (1979) *Nucleic Acids Res. 7*, 401–416.

Fuji, S., Wang, A. H.-J., van der Marel, G., van Boom, J. H., & Rich, A. (1982) *Nucleic Acids Res. 10*, 7879–7892.

Frank-Kamenetskii, M. D., & Vologodskii, A. V. (1984) *Nature 307*, 481–482.

Garel, J.-P., Jordan, J.-C., & Mandel, P. (1972) *J. Chromatogr. 67*, 277–290.

Gessner, R. V., Quigley, G. J., Wang, A. H.-J., van der Marel, G. A., van Boom, J. H., & Rich, A. (1985) *Biochemistry 24*, 237–240.

Hansch, C., & Leo, A. (1979) *Substituent Constants for Correlation Analysis in Chemistry and Biology*, Wiley, New York.

Ho, P. S., Quigley, G. J., Tilton, R. F., Jr., & Rich, A. (1988) *J. Phys. Chem. 92*, 939–945.

Irikura, K. K., Tidor, B., Brooks, B. R., & Karplus, M. (1985) *Science 229*, 571–572.

Jovin, T. M., McIntosh, L. P., Arndt-Jovin, D. J., Zarling, D. A., Robert-Nicoud, M., van de Sande, J. H., Jorgensen, K. F., & Eckstein, F. (1983) *J. Biomol. Struct. Dyn. 1*, 21–57.

Jovin, T. M., Soumpasis, D. M., & McIntosh, L. P. (1987) *Annu. Rev. Phys. Chem. 38*, 521–560.

Kollman, P., Weiner, P., Quigley, G. J., & Wang, A. (1982) *Biopolymers 21*, 1945–1969.

Kruger, W. C., & Prairie, M. D. (1985) *Biopolymers 24*, 905–910.

Langmuir, I. (1925) *Colloid Symp. Monogr. 3*, 48–75.

Latha, P. K., & Brahmachari, S. K. (1985) *FEBS Lett. 182*, 315–318.

Mirkin, S. M., Lyamichev, V. I., Kumarev, V. P., Kobzev, V. F., Nosikov, V. V., & Vologodskii, A. V. (1987) *J. Biomol. Struct. Dyn. 5*, 79–88.

Narasimhan, V., & Bryan, A. M. (1984) *Experientia 40*, 827–828.

Nordheim, A., Lafer, E. M., Peck, L. J., Wang, J. C., Stollar, B. D., & Rich, A. (1982) *Cell 31*, 309–318.

Peck, L. J., & Wang, J. C. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 6206–6210.

Pohl, F. M. (1976) *Nature 260*, 365–366.

Pohl, F. M., & Jovin, T. M. (1972) *J. Mol. Biol. 57*, 375–395.

Rich, A., Nordheim, A., & Wang, A. H.-J. (1984) *Annu. Rev. Biochem. 53*, 791–846.

Saenger, W., Hunter, W. N., & Kennard, O. (1986) *Nature 324*, 385–388.

Soumpasis, D. M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 5116–5120.

Tidor, B., Irikura, K. K., Brooks, B. R., & Karplus, M. (1983)

J. Biomol. Struct. Dyn. 1, 231–252.

Vologodskii, A. V., & Frank-Kamenetskii, M. D. (1984) *J. Biomol. Struct. Dyn. 1*, 1325–1333.

Vorlickova, M., & Jaroslav, K. (1984) *Int. J. Biol. Macromol. 6*, 77–80.

Wang, A. H.-J., Quigley, G. J., Kolpak, F. J., Crawford, J. L., van Boom, J. H., van der Marel, G. A., & Rich, A. (1979) *Nature 282*, 680–686.

Wang, A. H.-J., Quigley, G. J., Kolpak, F. J., van der Marel, G. A., van Boom, J. H., & Rich, A. (1981) *Science 211*, 171–176.

Wang, A. H.-J., Hakoshima, T., van der Marel, G., van Boom, J. H., & Rich, A. (1984) *Cell 37*, 321–331.

Warner, H. R., Duncan, B. K., Garrett, C., & Neuhard, J. (1981) *J. Bacteriol. 145*, 687–695.

Binding of Mithramycin to DNA in the Presence of Second Drugs[†]

Matilal Sarker and Fu-Ming Chen*

Department of Chemistry, Tennessee State University, Nashville, Tennessee 37209-1561

Received February 14, 1989; Revised Manuscript Received May 2, 1989

ABSTRACT: Comparative DNA equilibrium binding studies with mithramycin (MTR) and ethidium bromide in the presence and in the absence of second drugs were investigated by spectral titrations. Unusual curvatures (in contrast to those due to neighbor exclusion or anticooperativity) are found in the Scatchard plots of MTR–DNA titrations in the presence of netropsin, a minor-groove binder. Parallel studies with ethidium bromide indicate that although the presence of netropsin significantly reduces the binding ability of ethidium, no unusually curved Scatchard plots are obtained. The unusual curvature exhibited by the Scatchard plots of MTR titrations in the presence of netropsin indicates that the binding of netropsin greatly affects the MTR binding to DNA and can be simulated by an explicit incorporation of the second drug–DNA interaction in the binding formalism. Since netropsin is a minor-groove binder, its interference with the binding of MTR is in accord with the notion that MTR also binds at this groove. The observation of negligible effects on the DNA binding ability of MTR in the presence of either a major-groove or a phosphate group binder lends further support to this conclusion. Consistent with its guanine specificity, studies with synthetic polynucleotides suggest that MTR exhibits negligible affinity for poly(dA-dT)·poly(dA-dT) or poly-(dA)·poly(dT). Although the strongest bindings are demonstrated by poly(dG-m^{5}dC)·poly(dG-m^{5}dC), poly(dG-dC)·poly(dG-dC), and poly(dG)·poly(dC), the homopolymer exhibits a saturation density of approximately 1 drug molecule per every 2–3 base pairs, whereas the alternating polymers saturate at 1 drug molecule per <2 base pairs, a rather surprising finding for such an extended molecule but one consistent with the drug–dimer model on binding. The effect of sequence or conformation on the binding is further indicated by the stronger MTR affinity for poly(dA-dG)·poly(dC-dT) than for poly(dA-dC)·poly(dG-dT). These results are supported by circular dichroic measurements.

Mithramycin (MTR) is an antibiotic that contains an aureolic acid group and is closely related to chromomycin A₃ and olivomycin. Some studies on this drug, mainly clinical, have been made in recent years (Gause, 1965; Kennedy et al., 1968; Ream et al., 1968; Kennedy, 1970; Hill et al., 1972; Prasad & Nayak, 1976; Dasgupta et al., 1979; Fox & Howarth, 1985). Despite its high toxicity, this drug has proved useful for treating patients with testicular carcinomas (Ream et al., 1968; Kennedy, 1970; Hill et al., 1972) and Paget's disease (Elias et al., 1972). MTR inhibits the synthesis of both DNA and RNA in vivo, with some preference for the latter (Kersten et al., 1967; Fok & Waring, 1972). For its DNA binding, this antibiotic exhibits an absolute requirement for divalent cations such as Mg^{2+} (Behr & Hartmann, 1965;

Goldberg & Friedman, 1971). This drug is anionic above neutral pH (Illrionova et al., 1970) and exhibits guanine specificity in DNA (Ward et al., 1965; Behr et al., 1969). Its fluorescence intensity at certain wavelengths is greatly enhanced upon binding to DNA, thus enabling this compound to be used in microfluorometric analysis of cellular DNA (Coleman & Goff, 1985; Hauser-Urfer et al., 1982; Coleman et al., 1981; Hamilton et al., 1980; Groyer & Robel, 1980; Buys & Osinga, 1980). Footprinting experiments have shown that the preferred binding sites of MTR and its related antibiotics are at least 3 base pairs long (Fox & Howarth, 1985; Van Dyke & Dervan, 1983). The mode of interaction of this class of drugs, however, has not yet been unequivocally established.

It is generally believed that MTR binds nonintercalatively at the minor groove of DNA. This stems from the observations by several investigators that MTR fails to affect some prop-