Molecular Modeling of Acyclic Polyamide Oligonucleotide Analogues

Dwight D. Weller* and Daniel T. Daly

Department of Chemistry, Oregon State University, Corvallis, Oregon 97331-4003

Wilma K. Olson

Department of Chemistry, Rutgers—The State University of New Jersey, New Brunswick, New Jersey 08903

James E. Summerton

Antivirals, Inc., Corvallis, Oregon 97333

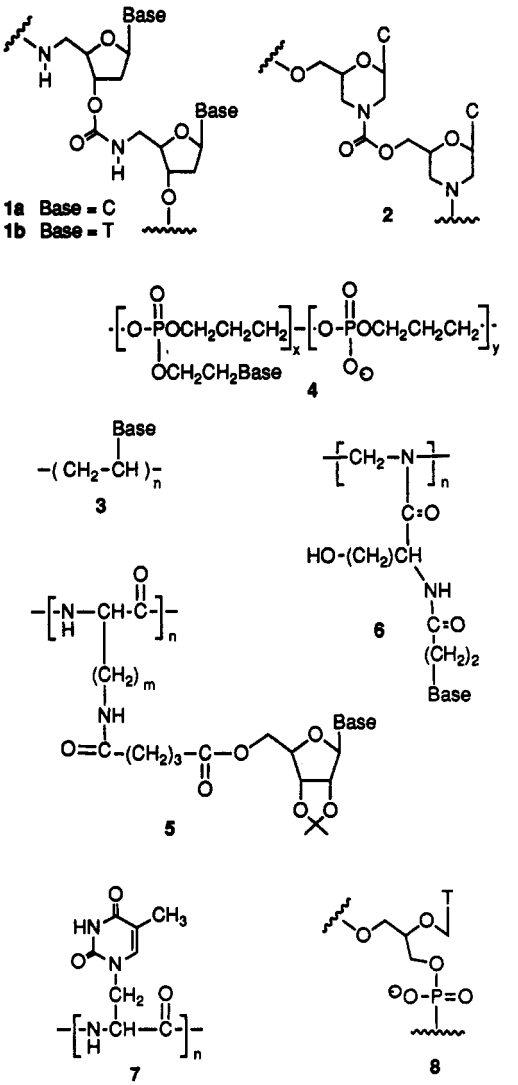
Received November 2, 1990 (Revised Manuscript Received June 24, 1991)

The feasibility of replacing the sugar–phosphate backbone of the nucleic acids with a polyamide-type backbone has been investigated by using molecular modeling techniques that examine the ability of the acyclic backbone to adopt low energy conformations that conform to the nucleic acid A- and B-form helices. Of the several backbone possibilities examined (nylons, polyurethanes, polypeptides), the most favorable appear to be those derived from a polypeptide. For most of the cases studied, the models predict a preference for binding of a given backbone type to either A- or B-form targets and, in some cases, suggest an orientational bias for direction along the helical axis, or a preferred stereochemistry at stereogenic atoms in the backbone.

Introduction

The need exists for the development of agents that can selectively attack genetic material responsible for a pathogenic state. Due to the charged phosphates, which impede travel across cell membranes, and a sensitivity to cellular nucleases, DNA and RNA themseleves are not ideal agents. Nevertheless, introduction of "antisense" DNA into cells leads to inhibition of protein synthesis from the targeted RNA, thus demonstrating the potential of the method.[1,2] We have undertaken the preparation and study of oligomeric agents that are neutral and either achiral or stereoregular. Two general approaches were considered feasible. In the first approach, the basic nucleoside unit is kept more or less intact and the phosphodiester group is replaced by a new, achiral linkage. We have previously reported on the successful binding of carbamate-linked oligonucleosides **1a** (Chart I) to complementary DNA and RNA.[3] We have also demonstrated that large structural changes in the sugar portion are tolerated as witnessed by the binding to DNA of the oligocarbamate derived from the morpholine nucleoside **2**.[4] In the second, and more radical, approach, the sugar and phosphate are replaced by a new backbone entirely. While the former strategy has the advantange of availability of the nucleoside building blocks, the latter is intriguing for the possibility of custom designing agents that might, for example, be specific for RNA vs DNA. This tack is certainly not novel, but has yet to be successfully demonstrated.[5] The early literature records the pioneering work of Pitha and Pitha on the polyvinyl compounds, e.g. **3**.[6] These species exhibit pairing with complementary homopolymers but the binding is necessarily imperfect given the unfavorable spacing of the bases and the atatic nature of the polymer.[6] More recently, binding has been found with other nonsugar backbones including the polyphosphates **4**[7] and the polylysines **5** and polyethyleneimines **6**.[8] Other studies have focused on less complex peptide backbones.[9,10] The most

Chart I



(1) (a) Stein, C. A.; Cohen, J. S. *Cancer Res.* **1988**, *48*, 2659. (b) Melton, D. A., Ed. *Current Communications in Molecular Biology: Antisense RNA and DNA*; Cold Spring Harbor Laboratory: 1988.
(2) Zon, G. *Pharm. Res.* **1988**, *5*, 539. (b) Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 544.
(3) Stirchak, E. P.; Summerton, J. E.; Weller, D. D. *J. Org. Chem.* **1987**, *52*, 4202.
(4) Stirchak, E. P.; Summerton, J. E.; Weller, D. D. *Nucl. Acids Res.* **1989**, *15*, 6129.
(5) Jones, A. S. *Int. J. Biol. Macromol.* **1979**, *1*, 194.
(6) (a) Pitha, J.; Pitha, P. M. *Biopolymers* **1970**, *9*, 965. (b) Pitha, J.; Pitha, P. M.; Ts'o, P. O. P. *Biochem. Biophys. Acta* **1970**, *204*, 39.
(7) Kropachev, V. A.; Aleksiuk, G. P.; Zaviriukha, V. L.; Kovtun, G. I. *Makromol. Chem. Suppl.* **1985**, *9*, 47–51.

0022-3263/91/1956-6000$02.50/0

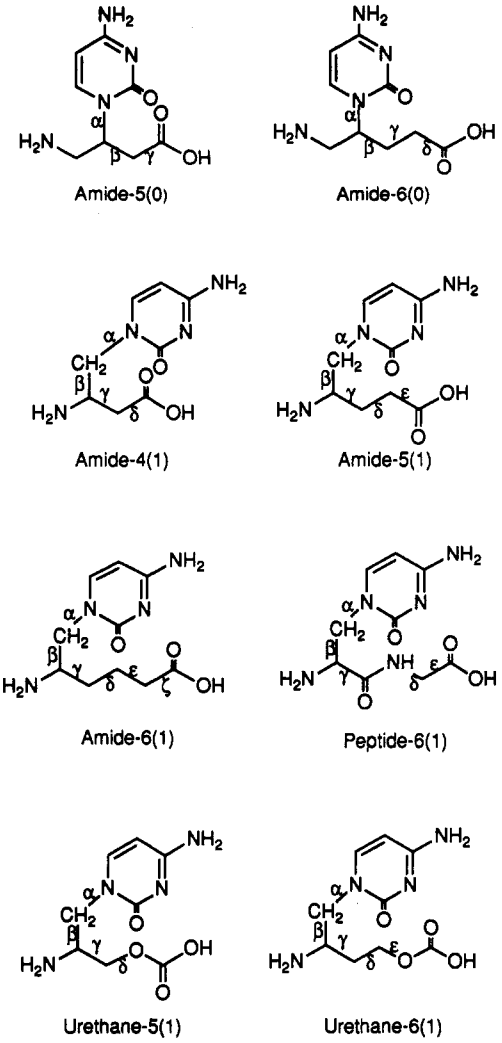


Figure 1. Backbone types. The torsions varied in this study are indicated by Greek letters.

attractive approach is that of Jones, who prepared the oligothymine derivative **7** (average $n = 9$) from optically pure amino acids.[10] With the proper set of base analogues of **7** at hand, sequence specific nucleic acid binding agents would be readily available by peptide synthesis techniques. However, **7** fails to show binding with poly(rA). It is unclear whether the length of the chain for these oligomers is sufficient, given the energetics of the A–T pairing and the fact that only every other base is roughly positioned for binding. With this background, a thorough examination of the possibilities seems in order. The present study examines the structural feasibility of using optimized nonsugar backbones of the amide type to prepare single-stranded nucleic acid binding agents.

Methods

Figure 1 illustrates the monomeric species chosen for investigation. The amide series contains an ω-amino acid backbone to which is attached a heterocyclic base. Cytosines were the chosen bases, as pyrimidines are more bulky than purines around the base nitrogen and a larger number of conformations could be excluded on steric grounds. A simple momenclature serves to distinguish the members of this series. For example, the *five*-atom long (counting from the amine to the carbonyl group) member of this group containing *one* spacing methylene group connecting

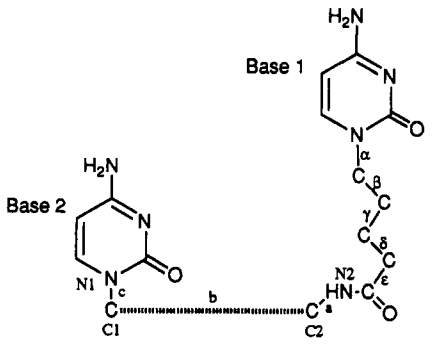


Figure 2. Construction of simplified dimeric units for the A-5(1) series. For discovery of helical dimers, torsions α–ϵ are varied according to Table II, with the bases fixed in the correct helical positions. Torsions α and c are defined to include the carbonyl carbon of the base. For discovery of the lowest energy dimer conformations, torsions α–ϵ and a–c are varied. In this case α, ϵ, a, and c are either 90° or 270° while the remainder are 60°, 180°, or 300°. This produces $2^4 \times 3^4$ (= 1296) starting conformations, each of which is minimized. The relative positions of the bases are determined solely by these torsions and the bonds and valence angles of the interconnecting chain.

the backbone and the base is the A-5(1) system. The urethane series (U-5(1), U-6(1)) represents oligomers of the aminocarbonic acids. Finally, the α-amino acid derived peptides of the type employed by Jones were modeled with the exception that a spacing glycine unit was incorporated between the base-containing monomer units (P-6(1)).

The position of attachment of the heterocyclic base to the chain was determined, in part, by practical concerns. First, it is desirable to avoid substitution of the carbon adjacent to the carboxyl group, thus disallowing racemization processes. Second, if the heterocycle is positioned on a carbon β to the carboxyl, the molecule is potentially labile with respect to elimination of the base. These guidelines are obviously difficult to obey in the Peptide series. Although there are fewer practical constraints on the position of the cytosine in the Urethane series, the position adjacent to the amino group was chosen to facilitate comparisons with the Amide series.

An additional factor in the positioning of the base along the backbone derives from the expected interaction of the base with the amide hydrogen of the backbone. CPK modeling studies suggest the potential for these series to adopt helical conformations and retain strong hydrogen bonding from the cytosine carbonyl to the amide N–H. This is an important observation as it implies that this interaction might serve as an ordering force, thus lowering the configurational entropy for the transition between unstacked and stacked (helical) forms.

In order to evaluate the capabilities for each of these series toward pairing with nucleic acids, we began by generating simplified dimers that conformed to known nucleic acid helices. The dimers contain two bases and only those parts of the backbone necessary to connect them. For their construction, two bases with their attached ribosyl C1′ carbons were positioned in either A- or B-form locations by using coordinates obtained from Arnott.[11] To one C1′ carbon was attached the component atoms of the backbone type to be examined, as shown in Figure 2 for the A-5(1) series. The torsions along the backbone were then driven over the range of available dihedral angles. Figure 1 shows the torsions varied for each backbone type. The positions of the terminus of the backbone fragment and the C1′ carbon on the adjacent base were monitored and when certain closure conditions were met, the dimeric unit was saved for further processing.[12-17] The chain

(8) Takemoto, K.; Inaki, Y. *Polym. Mat. Sci. Eng.* **1988**, *58*, 250.
(9) De Koning, H.; Pandit, U. K. *Rec. Trav. Chim.* **1971**, *91*, 1069.
(10) Buttrey, J. D.; Jones, A. S.; Walker, R. T. *Tetrahedron* **1975**, *31*, 73.

(11) Arnott, S. Personal communication.
(12) The values for all bond lengths and valence angles were taken from the parameters of the program AMBER.[13,14] Amides were evaluated only as the anti isomers. Since no carbamate structural data was at hand, the bond length, bond angle, and torsional angle values for this functional group were obtained by averaging values from related structures found in the Cambridge Crystallographic Data Bank.[15] In accord with the crystallographic data and conformational studies on esters,[16] the low energy form of the carbamate C–O torsion was assumed to be anti. The geometric values for the carbamate are given in the supplemental data.

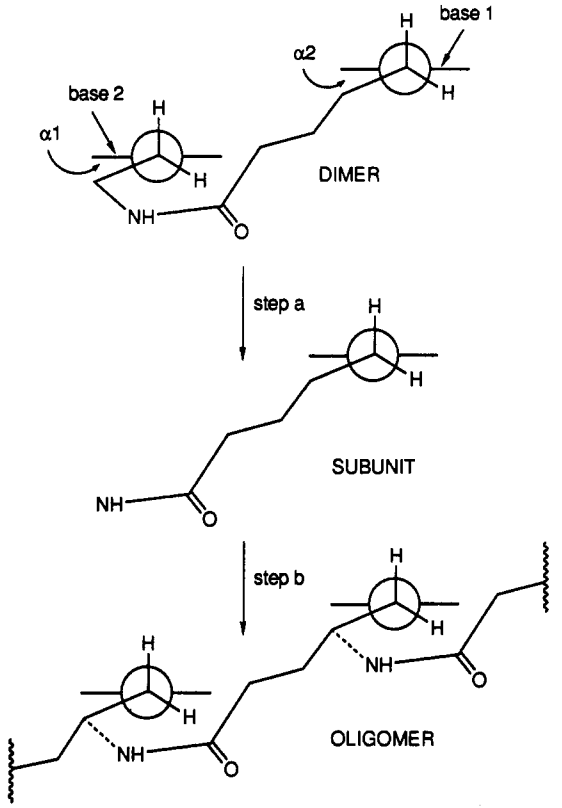


Figure 3. Construction of a helical oligomer in the A-5(1) series. Step a: A simplied dimer is converted into a subunit by removal of base 2 and two attached methylene groups. Step b: A collection of subunits is assembled with the bases in the correct helical locations. Dashed lines indicate new bonds formed by linkage of subunits. To produce an oligomer, the torsion angles $\alpha 1$ and $\alpha 2$ in the dimer normally must agree to within ±15°.

closure determinants used in generating helical dimers were (1) a distance between the fragment ends (C1---C2 in Figure 2) within 0.05 Å of the normal bond length and (2) values for the newly formed valence angles (comprising atoms N1---C1---C2 and C1---C2---N2) within 3° of the norm. This approach closely follows an earlier study of nucleic acid conformations.[17]

The proposed backbones were examined in both directions along the helical axis by reversing the position of initial attachment of the backbone fragment in Figure 2 to the lower cytosine (base 2). To distinguish the possibilities, if the amino terminus of the polymer would coincide with the 5′-terminus of the oligonucleotide it replaces in a duplex, it is labeled a type-1 backbone, and if it would coincide with the 3′-terminus, a type-2 backbone. Thus, for each backbone there are four orientations that require examination, e.g. A-5(1)A1 and A-5(1)A2 with A-form targets and A-5(1)B1 and A-5(1)B2 against B-form targets in the Amide series.

The helical dimers generated above by the proper attachment of the backbone fragment and adjacent base were further examined by application of the AMBER potential function and parameter set[18,19] to sort the structures on the basis of energy. The united atom approximation was used in the early stages and only hydrogens attached to nitrogen were explicity considered. The use of united atom conformers serves to remove structures with unacceptable steric interactions while avoiding the exaggerated effects of H---H interactions on the nonbonded energy term in these far from idealized structures.

In order to generate a helical oligomer from the helical dimers discovered above, we created hypothetical *helical subunits* from the lower energy dimer species. The helical subunit consists of a base and a portion of the backbone fragment used in the discovery of the helical dimers (see Figure 2). This is illustrated in Figure 3 for the A-5(1) system. In the case of subunits for systems that possess a spacing methylene group, the backbone fragment stops at the amide nitrogen. Helical subunits for systems that do not contain a spacing methylene group are generated in the same manner except that the C2 carbon in Figure 2 is included in the subunit. A helical oligomer is formed when the base components of several subunits are arranged in a helical array and a linkage is created from the terminus of each subunit chain to the first backbone carbon of the succeeding subunit.

An important criterion for the choice of dimers used in the construction of helices is derived by considering the torsional angle requirements at the site of attachment of the base to the backbone (referred to as terminal torsions). For example, one determinant is the matching of the two α torsional angles (Figures 1 and 3), $\alpha 1$ and $\alpha 2$, for a given dimer. To form a regular, helical chain from a single, repreating subunit conformer that is derived from this dimer would require that $\alpha 1 = \alpha 2$. A similar situation applies to the next torsion further from the base (β). If these terminal torsional requirements are not met when subunits are linked, the resulting helix will be of high energy primarily due to valence angle strain near the intersubunit linkage site.

Two approaches for generating oligomers in the all-atom form were employed. In method 1, the united-atom helical dimers within 10 kcal/mol of the lowest energy form for a given series were converted into all-atom helical dimers and minimized to a root mean square gradient of 0.1 kcal/mol·Å with constrained cytosines. This procedure afforded a much smaller selection of dimers for conversion to all-atom helical subunits. These "relaxed" subunits were linked into hexamers by following the terminal tosional criteria noted above and then minimized. In method 2, the original united-atom subunits were linked into a hexamer as before but converted into the all-atom helical subunit immediately prior to minimization. In both approaches, the bases were constrained during minimization. After energy analysis of these constrained hexamers, low energy forms were minimized without constraints on the bases, but in the presence of a complementary hexamer of oligo(deoxyguanosine).[20,21] The all-atom unconstrained helical hexamers were evaluated as tetramers by removing the terminal bases and backbones of the unconstrained hexamers and using the AMBER analysis routine.

In order to compare the relative fits of the different backbone types, it was necessary to account for the disparate electrostatic energies in each series, which results from the varied charge distributions. To this end, and to examine the energies of the helical tetramers in relation to their nonhelical, unstacked

(13) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765.

(14) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230.

(15) Allen, F. H.; Bellard, S.; Brice, M. D.; Cartwright, B. A.; Doubleday, A.; Higgs, H.; Hummelink, T.; Hummelink-Peters, B. G.; Kennard, O.; Motherwell, W. D. S.; Rodgers, J. R.; Watson, D. G. *Acta Crystallogr. B (Str. Sci)* **1979**, *35*, 2331.

(16) (a) *The Chemistry of Carboxylic Acids and Esters*; Patai, S., Ed.; Wiley; London, 1969. (b) *Tables of Interatomic Distances and Configurations in Molecules and Ions*, Special Publication No. 11 (Supplement, Special Publication No. 18); The Chemical Society: London, 1958 (1965).

(17) Srinivasan, A. R.; Olson, W. K. *J. Biomol. Struct. Dyn.* **1987**, *4*, 895.

(18) In AMBER, hydrogen bonds are represented by a 10–12 form. As suggested by Kollman,[13,14] 1,4 interactions were scaled by a factor of 2. The parameters for the carbamate function were borrowed largely from amide and ester parameters in the AMBER program. Charges for the backbones and the bases were derived independently from the programs QUEST and ESPFIT,[19] and then integrated by using the general procedure of Kollman.[13,14] The QUEST routine was performed with the STO-3G basis set. The ESPFIT routine of the QUEST module used four layers to fit the electrostatic potential. The number of points used was generally around 900 for the cytosine base and 1000–1300 for a given backbone. The dielectric constant was distance dependent and was set equal to the interatomic distance. During these energy calculations no base–base interactions were examined.

(19) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129.

(20) Parameters and charges for the nucleic acid strand were taken from the literature,[19] with the exception that the phosphate charge was reduced to –0.32e in order to mimic the effect of counterions about the phosphate.[21] For further energy analysis the terminal bases and attached backbone atoms were removed from each end of these unconstrained, minimized, hexameric all-atom chains to give helical tetramers that should be free of severe end effects.

(21) Manning, G. S. *Acc. Chem. Res.* **1979**, *12*, 443.

Table I. Discovery of Subunits

backbone[a]	torsional increment[b]	angles[c]	subunit orientation[d] B1	B2	A1	A2	total conformers[e]
A-4(0)	1.0	2	0	0	0	0	13
A-5(0)	1.0	3	2312	1235	384	780	467
A-6(0)[f]	3.0	4	282	286	160	176	2070
A-4(1)[f]	3.0	4	282	286	160	176	2070
A-5(1)	8.0	5	624	485	602	692	1850
A-6(1)	10.0	6	6419	4024	5110	4875	21800
U-5(1)	4.0	4	126	83	195	87	656
U-6(1)	10.0	5	186	134	187	127	605
P-6(1)	10.0	5	47	50	79	65	605
P-6(1)	8.0	5	158	142	235	217	1850

[a] Backbone type (see Figure 1). [b] Increment applied to the variation of torsional angles connecting the cytosine bases. [c] Number of torsional angles being varied. [d] Number of conformers found that met the chain closure conditions described in the text for the indicated helical type and orientation. See text for explanations of B1, B2, A1, A2. [e] Total number of conformations examined ($\times 10^{-4}$). [f] A-6(0) is identical with A-4(1).

counterparts, the global minimum was searched for each backbone type. The global minima were approximated by generating multiple conformations of the all-atom dimeric units in each series, varying the torsional angles of the interconnecting atoms to generate multiple initial conformations (see Figure 2), and minimizing without constraint of the cytosines. These unconstrained, generally unstacked dimers were useful in their own right in visualizing the preferred interactions of the backbone with the bases and of the bases with one another. Two units of the lowest energy dimer were then used to construct a tetramer. The torsional angles of the interconnecting atoms between the dimeric units were then varied over a range of values, and the resulting conformations were minimized. The lowest energy form from this exercise was taken to be the global minimum for the series.

Results

Table I shows the increments used to vary the backbone torsion angles and the frequency of chain closure. The torsional increment used in examining each backbone was determined by balancing the time of the calculations with the need to obtain a representive sample of the conformations available to the helical dimers in each series. In no case is the increment less than 10°. Although the number of hits for series such as A-5(0) is very large, as might be expected for the very low torsional increment, the number of fundamentally different conformers is quite small. For example, for the B1, B2, A1, and A2 forms of the A-5(0) series there were 4, 4, 5, and 6 different conformational types, respectively, while for the corresponding forms of the A-6(0) series there were 2, 3, 2, and 3 types, respectively (data not shown).

Systems with a four-atom backbone were found not to be suitable as building blocks for helical oligomeric species. Although the A-4(1) and A-6(0) backbones are identical at this dimer stage, there is no possibility for the subunits derived from the A-4(1) backbone to form an oligomer. Backbones without spacer groups between the base and backbone are also severly restricted in their ability to form low energy conformers. For the A-5(0) and A-6(0)B1, -B2, and -A1 systems, even the lowest energy forms show high steric energies due to interaction of the base with the backbone (data not shown). Only for the A-6(0)A2 system are relatively unstrained conformations available. The frequency of total hits in the P-6(1), U-6(1), and A-6(1) series at 10° increments nicely illustrates the degree of flexibility in each of these chains. The A-6(1) series generated an enormous number of conformations by virtue of the flexible polymethylene backbone. Placing a conformationally restricted structural unit in the chain, either an amide or carbamate, greatly reduced the occurrence of successful fragment joining in the helical dimer discovery process.

The energies of the tetramers for each series are listed in Table II.[22] Column 2 of this table, which lists the method of derivation of the oligomer, reveals that the second method for all-atom oligomer generation is more robust. The first method for all-atom oligomer generation is limited, in that the large variation in structure exhibited by the united-atom conformers for a given backbone type was lost upon dimer minimization. Only a handful of relaxed all-atom subunits in a given series were found by this method and only a few of these had compatible terminal torsions as required for chain formation. Some series, in particular the urethane series, did not provide conformations suitable for oligomer formation by the first method for this reason. Additionally, not all relaxed subunits derived from dimers with compatible terminal torsions yielded low energy oligomers, as backbone–backbone interactions not possible at the dimer level often proved severe. The second all-atom approach appears to be more general, as reasonable subunits were discovered in all series, and all of the low energy forms discovered by the first method were found by the second approach.

Initially, attempts to construct the helical hexamers were performed with a distance-dependent dielectric constant ($\epsilon = 1R$). Recent studies have indicated that, for nucleic acids, a dielectric of $\epsilon = 4R$ may more closely approximate the effect of an aqueous environment when calculations are run in the gas phase.[23] We therefore repeated the second method for helix generation using a dielectric $\epsilon = 4R$.

The helical tetramers have been compared with the global minimum tetramers discovered for each series in order to interrelate the various backbone types. The energy differences between the helical forms and the unstacked tetramer were divided by 4, and the results are listed as ΔE in Table II. To put these values in perspective, the global minimum tetramer for $(dC)_4$ was determined by taking combinations of C2′-endo and C3′-endo cytidines, linking them as a phosphodiester, and driving the phosphodiester O3′–P and P–O5′ (ω', ω) torsions and C5′–C4′ (ψ) exocyclic torsion.[24] This global minimum

(22) A complete listing of energy (bonds, angles, torsions, van der Waals, hydrogen bonds, and electrostatics) is given in the supplemental data.

(23) Orzoco, M.; Laughton, C. A.; Herzyk, P.; Niedle, S. *J. Biomol. Struct. Dynamics* **1990**, *8*, 359.

(24) ω refers to the O3′–P–O5′–C5′ torsion; ω' refers to the C3′–O3′–P–O5′ torsion; ψ refers to the O5′–C5′–C4′–C3′ tosion. These were varied at 60, 180, and 300°. The values of the ϕ (P–O5′–C5′–C4′) and ϕ' (C4′–C3′–O3′–P) tosions were initially set at the values found in the A-form or B-form (from ref 15). X and X′ are the glycosyl torsions, and P is the pseudorotation parameter of Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1972**, *94*, 8205.

Table II. Energies of Tetramers[a,b]

backbone	dielectric of $\epsilon = 1R$				dielectric of $\epsilon = 4R$			
	meth[c]	con[d]	HB	ΔE	meth[c]	con[d]	HB	ΔE
A-6(0)A2	2	*S*	−0.02	9.0	2	*S*	−0.02	3.3
A-5(1)B1	1, 2	*S*	−0.42	13.1	2	*S*	−0.22	5.8
A-5(1)B2	1, 2	*R*	−1.35	8.4	2	*R*	−0.82	2.9
A-6(1)A2	1, 2	*R*	−1.51	10.9	2	*R*	−0.46	4.2
A-6(1)B1a	2	*S*	−1.08	11.7	2	*S*	−0.37	3.3
A-6(1)B1b	1, 2	*S*	−0.03	12.8	2	*S*	−0.04	3.1
A-6(1)B2	2	*S*	−0.01	14.6	2	*S*	−0.02	4.2
U-5(1)A1	2	*S*	−0.05	12.0	2	*S*	−0.05	3.9
U-5(1)B2a	2	*S*	−1.47	10.8	2	*S*	−0.29	4.9
U-5(1)B2b	2	*S*	−0.03	11.1	2	*S*	−0.03	1.4
U-6(1)A1	2	*R*	−0.02	9.6	2	*R*	−0.01	4.2
U-6(1)A2a	2	*S*	−0.01	12.2	2	*S*	−0.01	5.3
U-6(1)A2b	2	*S*	−0.07	11.2	2	*S*	−0.03	6.1
U-6(1)A2c	2	*R*	−0.03	9.8	2	*R*	−0.03	3.9
U-6(1)B2	2	*S*	−0.04	11.7	2	*S*	−0.05	4.4
P-6(1)A1	2	*R*	−1.84	11.2	2	*R*	−0.92	4.2
P-6(1)A2	2	*RS*	−1.41	11.7	2	*S*	−1.59	2.6
P-6(1)B1	2	*R*	−2.41	7.4	2	*R*	−2.11	1.4
P-6(1)B2a	2	*R*	−2.24	8.7	2	*R*	−1.85	1.8
P-6(1)B2b	2	*S*	−2.18	8.2	2	*S*	−1.96	1.3
P-6(1)B2c	1	*R*	−1.84	8.8	2	*R*	−1.55	1.6
A DNA	–	–	−0.01	4.9				
B DNA	–	–	−0.01	6.8				

[a] Helices are derived from the indicated backbone in which a hexamer has been minimized without constraint in the presence of a complementary hexamer of deoxyguanosine. The terminal subunits have been removed to produce a tetramer. [b] Energies are in kcal/mol; HB = 10–12 non-bonded energy of intramolecular hydrogen-bonded atoms; ΔE = energy per subunit. The latter is obtained from the difference in total energy for a conformer and the global minimum divided by the number of subunits in the backbone. [c] Meth = approach used to convert united atom subunits into all-atom hexamers. [d] Con = absolute configuration at stereogenic carbons.

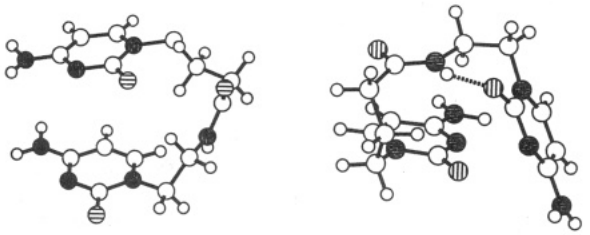


Figure 4. Lowest energy conformer for the A-6(0) series, with $\epsilon = 1R$. Empty circles are carbon and hydrogen, circles with vertical bars represent nitrogens, and circles with diagonal or horizontal bars represent oxygen. Dashed lines indicate hydrogen bonding (left). Lowest energy unconstrained dimer for the A-6(0) series, with $\epsilon = 1R$ (right).

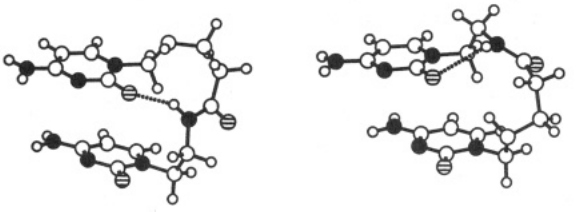


Figure 5. Lowest energy conformer for the A-5(1)B1 series, with $\epsilon = 1R$ (left). Lowest energy conformer for the A-5(1)B2 series, with $\epsilon = 1R$ (right).

species was compared to the corresponding helical tetramer obtained by following minimization of the $(dC)_6$–$(dG)_6$ systems. In the sections below, we examine the helical tetramers for each series.

Amide-6(0). The only low energy backbone in this series is the A2-form. Although the ability of this backbone to bind DNA and RNA is not yet known, the results suggest that it is unwise to assume that the B-form nucleic acid helix, which has a shorter C1′–C1′ distance (= 4.84 Å), will be easier to target with a given backbone than the A-form (C1′–C1′ = 5.50 Å). This backbone shows almost perfect staggering of torsions along the chain (Figure 4). Notably absent is hydrogen bonding from the amide to the base. Thus, contrary to our initial presumptions, it may not be possible to find helical conformations for a given backbone that are stabilized by base–backbone hydrogen bonding.

The lowest energy unconstrained dimer for this series does not show the strong base–base interactions common to the other series (vida infra). Rather, one base overlaps the nearby amide to align atoms of opposite charge favorably (Figure 4). In this form one base is held in place by the base–amide hydrogen bond. Thus, far from being a favorable orienting force for the helix, in this series the base–backbone interaction would seem to favor the unstacked form. No π-stacking of the bases was observed in the unconstrained dimers regardless of the dielectric constant.

Amide-5(1). For this species no low energy A-form subunits capable of helix formation were found and each of the B-form subunits was produced by both all-atom approaches. In these tetramers, base–amide hydrogen bonding is an important feature. For the A-5(1)B1 series the hydrogen bond is from a base to the amide N–H from the same subunit (Figure 5) while for the A-5(1)B2 species the H-bond is to the amide N–H from the adjacent subunit (Figure 5). In the former series this interaction is longer than optimal (2.72 Å vs 1.96 Å for the A-5(1)B2 form, at $\epsilon = 1R$); also the backbone adopts more torsional strain, making the B1 series of higher energy. For $\epsilon = 4R$ the degree of hydrogen bonding decreases, as does the relative stability of this backbone type. Thus, the A-5(1)B2 form has one of the lowest ΔE values for $\epsilon = 1R$ (ΔE = 8.4 kcal/mol) but not for $\epsilon = 4R$ (ΔE = 2.9 kcal/mol).

In this series, many of the low energy unconstrained dimers show the base–amide hydrogen bond found in the minimum energy conformer of the A-6(0) series. The lowest energy dimer of the A-5(1) series for $\epsilon = 1R$ does not show this behavior and is characterized instead by N4-H---N3 hydrogen bonding between the bases as well as optimized backbone torsions (Figure 6). This is the dominant pattern observed in all backbone types except A-6(0). In contrast, the lowest energy unconstrained dimer

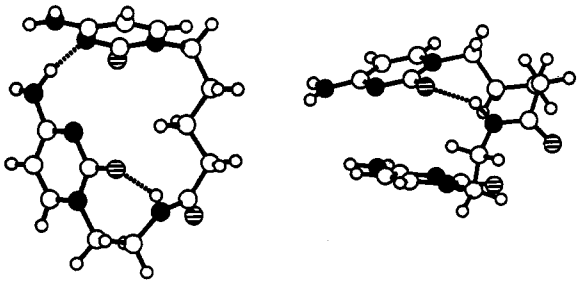


Figure 6. Lowest energy unconstrained dimer for the A-5(1) series with $\epsilon = 1R$ (left). Lowest energy unconstrained dimer for the A-5(1) series with $\epsilon = 4R$ (right).

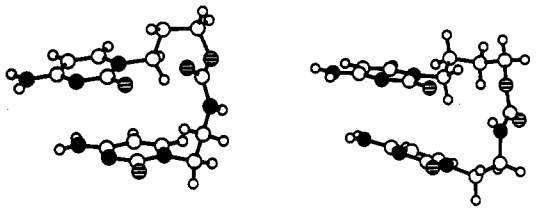


Figure 7. Lowest energy conformer for the U-5(1) series, with $\epsilon = 4R$ (left). U-6(1)A1 backbone, with $\epsilon = 1R$ (right).

at $\epsilon = 4R$ shows nearly complete overlap of the cytosine bases (Figure 6). This is the only example of π-stacking found for any backbone series. The structure resembles the A-5(1)B2 type in the nature of the base–amide hydrogen bond.

Amide-6(1). As expected, the 6-atom backbone in the Amide series provided a larger variety of reasonable helices than the previous series, lacking only lower energy members from the A1 species. Table II lists only representive examples, as multiple variations were found for most of the listed oligomers. Note that at $\epsilon = 1R$, the lower energy helical forms are those with base–amide hydrogen bonds (HB term) although these all show significantly higher strain energy (bonds, angles, torsions, van der Waals) terms than the other conformers.[22] At $\epsilon = 4R$, this is not the case. Although the results predict that binding to both A- and B-form targets is feasible, none of the conformers in this series is of very low ΔE.

The low energy unconstrained dimers in this series fail to show evidence of base–amide hydrogen bonding, and examples of species with N4–H---N3 base–base interactions are common.

Urethane-5(1). For one example of this series, the U-5(1)B2 form at $\epsilon = 4R$, one of the lowest energy backbones is observed (Figure 7). This form has no base–amide hydrogen bonding and instead shows low strain energy.[22] Base–amide hydrogen bonding is also absent in the lowest energy unconstrained dimers, which exhibit the common N4–H---N3 base–base interaction.

Urethane-6(1). The lowest energy helical forms in this series are A-form species. Although none of these shows any base–amide hydrogen bonding, the backbone shows excellent staggering of torsions, as in Figure 7 (U-6(1)A1). Despite the apparent good fit of this backbone, the ΔE values for this series at $\epsilon = 4R$ are not especially low. In fact, the van der Waals terms for this series are significantly less negative than those for most of the other series (these range from −19 to −27 for this series while most series have van der Waals terms that range from −27 to −33 kcal/mol).

In the unconstrained dimers, all of the low energy conformers show a strong base–amide hydrogen bonding interaction.

Peptide-6(1). This is the only series in which a stereochemically inhomogeneous backbone was found. The

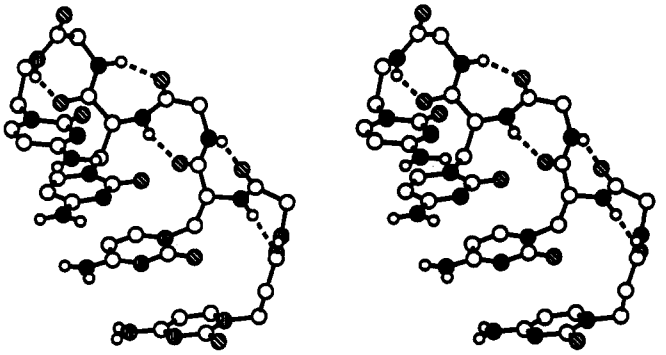


Figure 8. Stereodrawing of the P-6(1)B1 backbone with $\epsilon = 1R$.

P-6(1)A2 example has an *RRSS* configuration at the base–backbone attachment site for the four stereogenic methines in the hexamer. This was only observed for $\epsilon = 1R$ and the conformer produced at $\epsilon = 4R$ is stereochemically homogeneous.

For most of the peptide helices the backbones are strongly ordered by 1,7 hydrogen bonds (Figure 8).[25] The values for the common polypeptide backbone torsions (ϕ,ψ) for the P-6(1)B1 helix at $\epsilon = 1R$ are (77°, 298°) for the modified alanine and (86°, 305°) for the glycine moiety. For the P-6(1)B2b the corresponding values are approximately (75°, 314°) and (66°, 281°). These should be compared with the usual (ϕ,ψ) torsion angles of the C_7 conformer of (80°, 280°).[26] The B-form is predicted to be preferentially targeted with peptide backbones. The unconstrained dimers in the peptide series also reflect the intrachain amide–amide orienting forces with most forms showing 1,7 hydrogen bonding.

The B-form backbones in this series are among the lowest energy backbones for both values of the dielectric constant. The conformations found support strong hydrogen bonding without appreciable strain. When the P-6(1)B1 helical tetramer was allowed to minimize to a gradient of 0.5 kcal/Å without the complementary oligo(deoxyguanosine) strand, a new structure was produced. However, the overall conformation is only slightly changed, with a root mean square deviation of 0.5 Å relative to the original P-6(1)B1 helical form. Thus, this is a very stable backbone conformation that supports both intrastrand hydrogen bonding and base stacking.

Discussion

As studies of dimeric intercalating agents have shown, the nature and flexibility of the tether interconnecting the DNA binding portions of the molecule greatly influence the overall binding constant of the complex and more rigid tethers will have correspondingly higher affinities for their targets.[27] Although we have not done an exhaustive analysis of the configurational entropy for these backbones, we have two measures of the degree to which these backbone units might have ordered structures. First, we have examined the low energy conformations of the dimeric units (unconstrained dimers) in the all-atom forms that were used to generate the global minimum tetramers of the analogues. The structures reveal little tendency to form stacked structures in the manner of nucleic acid

(25) A listing of (ϕ,ψ) torsions for all the peptide systems is given in the supplemental data.
(26) Hagler, A. T. *The Peptides*, Volume 7; Academic Press: New York, 1985; pp 213–299.
(27) Jaycox, G. D.; Gribble, G. W.; Hacker, M. P. *J. Heterocycl. Chem.* 1987, *24*, 1405.

dimers.[28] In the latter case, for dCpdC, the minimum energy form is very similar to the A-form helical unit in that the bases are stacked, but twisted 60° out of phase (overwound) vs the approximately 30° twist required in the helix.[11] The bases were in close proximity in all of the low energy acyclic forms of the acyclic analogues, and, in most examples, one base is oriented so that the exocyclic amino group is directed toward the other base and minimal π-stacking interactions are present. Only for the A-5(1) series was a low energy stacked form observed.

The second ordering force observed for these backbones is the hydrogen bond between the backbone amide hydrogen and the base carbonyl group. This was expected from the study of the analogues with space-filling models, and this structural feature is present in many of the duplexes. However, in some of the lower energy helical forms, no hydrogen bond is seen, while in the global minimum dimers and tetramers, this base–amide interaction is generally present. To the extent that the modeling is valid, this ordering interaction should actually be destabilizing toward helix formation. This applies especially to the A-6(0), U-5(1), and U-6(1) series. The A-5(1) series would appear to be favorably ordered by a base–amide interaction, as the lowest energy unconstrained dimer possesses a base–amide hydrogen bond that is similar to that found in the A-5(1)B2 system.

The peptide appears to be the backbone of choice. It has strongly ordered backbones, due to the presence of the C_7 conformational type, which have low overall strain energy. The P-6(1)B1 conformation is a low energy form that accommodates base stacking interactions similar to those found in nucleic acids. The prediction that the peptide series is suitable for binding to complementary nucleic acids must be reconciled with the lack of binding observed by Jones[10] in the related oligothymine derivatives **7**. The relatively weak A–T pairing and the short length of **7** suggest this species is not the optimal test system. A recent demonstration of the effective replacemnt of T by C in the binding of oligomeric nucleoside analogues to single stranded nucleic acids is found in the carbamate-linked oligonucleoside series **1**. In one case, a carbamate-linked hexamer of 5′-amino-5′-deoxythymidine (**1b**) showed no binding to a complementary poly(A) or poly(dA),[29] while the corresponding hexamer from 5′-amino-2′,5′-dideoxycytidine (**1a**) showed strong binding to poly(rG) and $(dG)_6$.[3]

Additionally, oligomer **7** was tested for binding only against poly(rA). This RNA target will strongly prefer the A-form conformation while our modeling suggests that B-form targets will be much more effectively complexed by the peptide series. It is also possible that the presence of the additional bases in **7** might lead to strongly ordered forms that do not support helix formation, although we have not examined this. In this regard, the incorporation of glycine into the peptide series provides more than structural simplification as theoretical studies have shown glycine can stabilize C_7 forms relative to the α-helical forms,[30] a result borne out by experiment.[31]

Recently, an acylic analogue with close analogy to DNA has been studied.[32] Precursors of the glyceronucleotides **8** were prepared and inserted into DNA duplexes as a replacement for thymidylic acid. A drastic lowering of the Tm relative to the native DNA duplex was observed, leading to the conclusion that flexible nucleosides would not be suitable as nucleic acid binding agents. The glyceronucleoside system can be characterized by the long tether joining the backbone to the base and the presence of few constraints on the conformations allowable to the backbone. In the amide-based systems, the shorter tether, the presence of the rigid amide or carbamate moiety, and the possibility of ordered hydrogen bonded forms may significantly lessen the entropy cost of these acyclic analogues.

In summary, despite some literature precedent to the contrary, several of the acyclic oligonucleotide analogues examined appear to be good targets for synthesis and biophysical testing. It will be especially interesting to discover if the predicted preferences manifested in Table II will be observed. Thus, the A-5(1) and U-5(1) series should strongly prefer binding to DNA rather than RNA, as will the peptide series. The A-5(1) and U-5(1) series are also predicted to have a strong bias for the absolute stereochemistry of the stereogenic atom. Studies to test these predictions are in progress. The results will give important clues as to the suitability of this simple modeling approach for evaluating nucleic acid binding agents.

Acknowledgment. This work was supported by USPHS grants GM20861 (W.K.O.) and GM36030 (D.D.W.), by the Rutgers Center for Computational Chemistry, and by the General Research Fund of Oregon State University. We wish to thank Christine Hunter for assistance in the preparation of the drawings of molecular structure and Dr. Paul Weiner for the AMBER program.

Supplementary Material Available: Tables listing structure and parameter data for carbamates, complete energy analysis of tetramers listed in Table II and their global minima, values for the torsion angles of the tetramers listed in Table II, values of (ϕ,ψ) for the peptide backbones, minimum energy conformations of dCpdC, and partial charges for the species listed in Figure 1 (9 pages). Ordering information is given on any current masthead page.

(28) Broyde, S.; Wartell, R. M.; Stellman, S. D.; Hingerty, B. *Biopolymers* **1978**, *17*, 1485.
(29) Coull, J. M.; Carlson, D. V.; Weith, H. L. *Tetrahedron Lett.* **1987**, *28*, 745.
(30) (a) Hagler, A. T.; Stern, P. S.; Sharon, R.; Becker, J. M.; Naider, F. *J. Am. Chem. Soc.* **1979**, *101*, 6842. (b) Hagler, A. T.; Honig, B. *Proc. Nat. Acad. Sci. U.S.A.* **1978**, *75*, 554.
(31) Naider, F.; Becker, J. M.; Ribeiro, A.; Goodman, M. *Biopolymers* **1978**, *17*, 2213.
(32) Schneider, K. C.; Benner, S. A. *J. Am. Chem. Soc.* **1990**, *112*, 453.