830–835 *Nucleic Acids Research, 1997, Vol. 25, No. 4*

Hydration effects on the duplex stability of phosphoramidate DNA–RNA oligomers

Daniel Barsky*, Michael E. Colvin, Gerald Zon[1] and Sergei M. Gryaznov[1]

Sandia National Labs, M.S. 9214, Livermore, CA 94551-0969, USA and [1]Lynx Therapeutics, Inc., 3832 Bay Center Place, Hayward, CA 94545, USA

Received October 14, 1996; Revised and Accepted December 16, 1996

ABSTRACT

Recent studies on uniformly modified oligonucleotides containing 3′-NHP(O)(O⁻)O-5′ internucleoside linkages (3′ amidate) and alternatively modified oligonucleotides containing 3′-O(O⁻)(O)PNH-5′ internucleoside linkages (5′ amidate) have shown that 3′ amidate duplexes, formed with DNA or RNA complementary strands, are more stable in water than those of the corresponding phosphodiesters. In contrast, 5′ amidates do not form duplexes at all. There is no steric reason that the 5′ amidate duplex should not form. We demonstrate that these differences arise from differential solvation of the sugar–phosphate backbones. By molecular dynamics calculations on models of 10mer single-stranded DNA and double-stranded DNA–RNA molecules, both with and without the phosphoramidate backbone modifications, we show that the single-stranded 3′ amidate and 5′ amidate backbones are equally well solvated, but the 5′ amidate backbone is not adequately solvated in an A-form duplex. These results are supported by quantum chemical free energy of solvation calculations which show that the 3′ amidate backbone is favored relative to the 5′ amidate backbone.

INTRODUCTION

Synthetic oligonucleotides have recently attracted considerable interest for their therapeutic and diagnostic potential. Recent studies (1,2) have focused on uniformly modified oligonucleotides containing 3′-NHP(O)(O⁻)O-5′ internucleoside linkages, where NH is substituted for O in the 3′ position along the backbone, henceforth referred to as 3′ amidate or 3′pnDNA; and modified oligonucleotides containing 3′-O(O⁻)(O)PNH-5′ internucleoside linkages, where NH is substituted for O in the 5′ position, called 5′ amidate or 5′pnDNA. These studies have shown that 3′pnDNA–RNA duplexes are more stable in water than those formed by the corresponding phosphodiester duplexes, but 5′pnDNA–RNA duplexes fail to form (2). There is no apparent steric reason that the 5′pnDNA duplex should not form. We hypothesize that these differences arise from differential solvation of the alternative backbones. More precisely, we suggest that the 3′pnDNA and 5′pnDNA backbones are equally well solvated in the single-stranded state, but the 5′pnDNA backbone is inadequately solvated in a DNA–RNA A-form duplex. Hence, there should be a solvation free energy penalty to form a 5′pnDNA–RNA duplex compared to a 3′pnDNA–RNA duplex.

Both the unmodified and the phosphoramidate-modified oligonucleotide backbones contain hydrophilic moieties which try to maintain contact with the solvent (3). For example, both the 3′ and 5′ oxygen of the unmodified backbone are expected to be hydrogen bonded to water in both the single-stranded and double-stranded forms. When an amino group (NH) is substituted for either the 3′ or 5′ oxygen, it would be energetically favorable for both the nitrogen and the hydrogen to be solvated. As will be demonstrated below, a solvent accessible surface area calculation suggests that solvation of the 5′ amide cannot occur in a 5′ pnDNA–RNA duplex. This means that formation of the 5′ amidate duplex entails the loss of favorable solvent interactions, and for this reason duplex formation should be energetically unfavorable.

By molecular dynamics calculations on models of 10mer single-stranded DNA and double-stranded DNA–RNA complexes, both with and without the backbone modifications, we will show that the backbone solvation-hypothesis is correct. These results will be further supported by quantum chemical solvation free energy calculations.

MATERIALS AND METHODS

Several single-stranded (ss) DNA and double-stranded (ds) DNA–RNA oligomers were simulated using molecular dynamics (MD). Three 10 bp DNA–RNA duplexes—3′pnDNA–RNA, 5′pnDNA–RNA and an unmodified phosphodiester pair DNA–RNA, all with the sequence [d(TTTTTTTTTT)]—were built in the canonical A-form using the program QUANTA4.0/ CHARMm22 (4). These were phosphate terminated at the 5′ ends. Three single-stranded DNA molecules were built exactly like the duplexes, except that the complementary RNA strand was discarded. The phosphoramidate groups have not been parameterized for the CHARMm22 force field; instead we used the default choices that QUANTA4.0 provided for the bonded interactions, which assigned the nitrogen atom type NT (sp3 nitrogen) and the phosphorous type PO3 (sp3 phosphorous). The resulting P–N bond length of 1.67 Å is in close agreement with the average value of 1.662 Å for N–P bonds in the Cambridge Crystalographic database (5). Analysis of this bond length during the dynamics

*To whom correspondence should be addressed. Tel: +1 510 294 2952; Fax: +1 510 294 2234; Email: barsky@ks.uiuc.edu

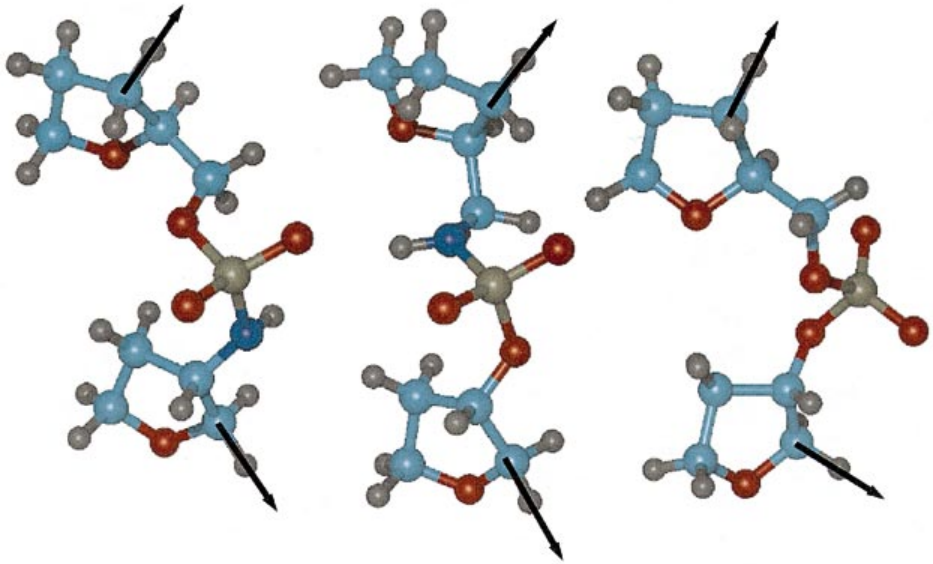

Figure 1. The sugar–phosphate 3′pn (left), 5′pn (middle) and native (right) linkages which form the backbone of the DNA molecules. These structures have been geometry-optimized in the gas phase by quantum chemistry while constraining two bond-vectors (indicated by arrows) as occurred after 40 ps of molecular dynamics simulation.

reveals a slightly shorter average length of 1.63 ± 0.05 Å. All the partial charges are the default CHARMm22 charges expect for the backbone N–H moiety where we assigned –0.56 and 0.1 charge units to the N and H respectively. The net charge on the N–H of –0.46 is thus identical to the charge on the O5′ and O3′. We confirmed heavy atom relative charges in the phosphoramide group by comparison with the Natural Atomic Orbital Population analysis of the gas phase Hartree–Fock wave function (see below). The MD partial charges are roughly half the values predicted by the *ab initio* calculation.

Using the program X-PLOR (6), a sphere of radius 48 Å of TIP3P (7) water molecules was added within 2.6 Å of the solute oligomer molecules. In all we built six molecular systems, three with DNA–RNA duplexes and three with single-stranded DNA, all at the center of a sphere of water molecules. Each total system consisted of ~6150 atoms, including ~1750 water molecules. The original water density was maintained by applying a 49.2 Å repulsive spherical shell having a harmonic force constant of 20 kcal/Å^2. (This has turned out to be stronger than necessary; the edge of the droplet remains ~0.5 Å from the shell due to surface tension.) The SETTLE algorithm (8) was used to maintain the internal equilibrium of each water molecule for improved energy conservation.

Since each phosphate group along the backbone carries a formal charge of –1, each oligomer had a net charge of –*n* where *n* is the number of bases in each case. For the six systems described above, counter-ions were not included because they require a long equilibration time in the presence of explicit water (more than 50 ps) (9), and a comparative study done in the absence of explicit water found that inclusion of counter-ions did not significantly alter the resulting DNA structural properties on timescales <100 ps (10). To test our assumption that on a 100 ps timescale counter-ions will not affect the essential behavior of the water vis-a-vis the DNA—but may in fact may give rise to non-equilibrium behavior of the DNA—we also simulated a 10mer duplex of 5′pnDNA–RNA, a 10mer of unmodified phosphodiester DNA–RNA, and a single-stranded 5′pnDNA 10mer. For these three simulations we employed larger spheres of 2500 water molecules with a weaker retaining force of 4 kcal/Å^2 at 55.2 Å, for a total number of ~8200 atoms. The MD simulations provide records of all the atomic coordinates over a sampling period of 120 ps. In this investigation our trajectories have been collected during dynamics where the number of atoms, the total volume and the temperature have been held constant (NVT ensemble). After the initial construction and minimization the simulations were carried out in the following steps.

1. With the oligomer (ss or ds) held fixed the water was heated and cooled twice from 20 to 300 K over equal intervals totaling 4 ps. This reduces equilibration times for the whole system (11).

2. With no atoms held fixed, 400 steps of conjugate gradient minimization were applied and heating from 0 to 300 K was accomplished by periodically rescaling the velocities over a period of 10 ps.

3a. Double-stranded oligomers: 120 ps NVT dynamics.

3b. Single-stranded oligomers: 40 ps NVT dynamics, followed by:

(i) Heating of entire system to 500 K over 20 ps period.

(ii) Cooling of entire system to 300 K over 20 ps period.

(iii) 120 ps NVT dynamics.

The simulated annealing steps (i) and (ii) move the single-stranded DNA rapidly away from the strained helical form and into an equilibrium (‘melted’) configuration.

All molecular dynamics and energy minimizations were carried out by the CCEMD program (12,13) which employs the parameters and force fields used in the CHARMm22 program and associated RTF topology files (14). A long 12 Å electrostatic cutoff for all atoms was included via a shifted-force potential as described elsewhere (15,16), according to a recent implementation (17). This avoids the artifact of local heating where the cutoff occurs. The dynamics are accomplished by Verlet integration with time steps of 1.0 fs. All analyses of the dynamics are averaged over three 40 ps consecutive trajectories, i.e., a total of 1200 structures. The simulations with counter-ions were run for an additional 40 ps.

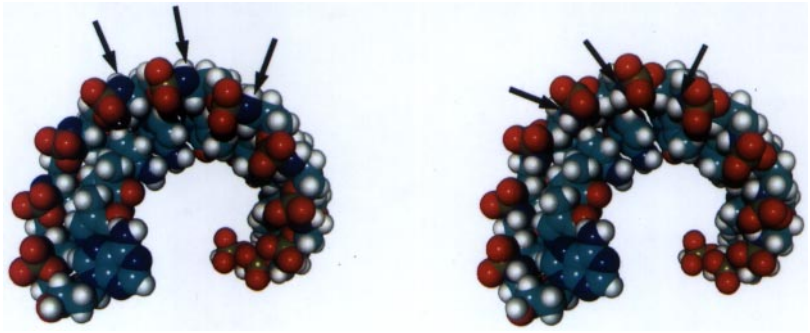

Figure 2. Canonical A-form DNA–RNA, where the complementary RNA strand is not shown. In the 3′pn DNA strand (left) the 3′ N atoms (blue) along the backbone are visibly exposed and accessible to the surrounding solvent. In contrast, the 5′ N atoms along the backbone of the 5′pn DNA strand (right) are unexposed to the solvent. The arrows indicate several backbone nitrogens.

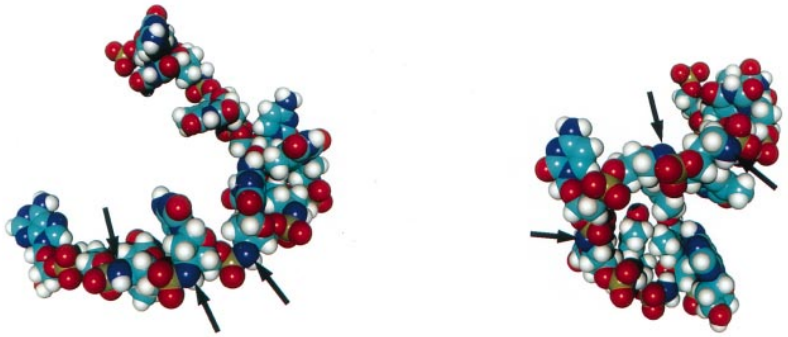

Figure 3. Melted DNA strands in the absence of a complementary strand. The 5′ N atoms (blue) along the backbone of the 5′pn DNA strand (right) are now visibly exposed to the surrounding solvent, as are the 3′ N atoms of the 3′pn DNA strand (left). The arrows indicate several backbone nitrogens.

Solvation energy

The actual free energy of solvation cannot be reliably calculated from the MD calculations because of inaccuracies in the empirical force fields and the large fluctuations in the total energy of the system. Although *ab initio* quantum chemical calculations cannot be applied to the full DNA oligomers used in the MD simulations, representative fragments of the oligomers can be studied using *ab initio* methods. To this end, using structures from the 40 ps time step of the MD simulations for each of the three DNA–RNA oligomers (phosphodiester, 3′-amidate and 5′-amidate), we excised a small segment of the DNA strand that includes two adjacent bases and the backbone atoms linking them. We then replaced the pendant bases with a hydrogen to yield 32 atom (diester) or 33 atom (amidate) fragments (Fig. 1), which are small enough to permit the application of accurate *ab initio* quantum chemical methods. In order to provide meaningful relative energies, these structures must be minimized within the quantum chemical framework to remove any large forces in the MD structures. Since a full minimization would lose all conformational constraints imposed by the double helix, we constrained the bond vectors linking the 3′ and 4′ carbons on the deoxyribose groups to have the same relative positions and directions as in the original MD configuration. (Constrained bond vectors indicated with black arrows in Fig. 1.)

With the constraints just described, the three structures were optimized at the Hartree–Fock level of theory using a 4-31G* basis set. All calculations were performed with Gaussian 94. At these optimized structures, the ‘gas-phase’ energies were calculated using Hartree–Fock with a larger 6–31G* basis set and the ‘aqueous-phase’ energy was calculated with a polarizable continuum solvent model (PCM) coupled to the Hartree–Fock/6-31G* wave function, using the Gaussian 94 ‘scrf=SCIPCM’option. In the PCM model, the solvent is treated as an unstructured continuum outside the solvent-accessible surface of the solute and is characterized only by its dielectric constant which is 78.5 for water at 25°C. Although the PCM models are too simple to predict absolute solvations with high accuracy, relative PCM solvation energies can be accurate to within a kilocalorie per mole for similar solutes.

RESULTS

Space-filled models of the canonical forms of the 3′pnDNA–RNA and 5′pnDNA–RNA duplexes are shown in Figure 2. These structures are the A-form for DNA–RNA duplexes. The complementary phosphodiester RNA strand has been removed for better viewing of the modified DNA strand. Although the 3′-N atoms (blue) are quite exposed along the backbone, the 5′-N atoms are deeply buried in a chasm formed by the sugars, the phosphates and the bases. The 5′-N atoms become exposed to the solvent, however, when the single-stranded oligomer has been allowed to come to equilibrium with the solvent. Figure 3 shows a space-filled model of the ‘melted’ single strands after 40 ps of dynamics [during stage 3b (iii) above]. The increased flexibility of the oligomer allows more favorable electrostatic interactions between the solvent and the 5′-N atoms.

These observations concur with direct measurements of the atomic solvent accessible surface (SAS) areas. In Figure 4 we present SAS results for the first 40 structures obtained at 1 ps

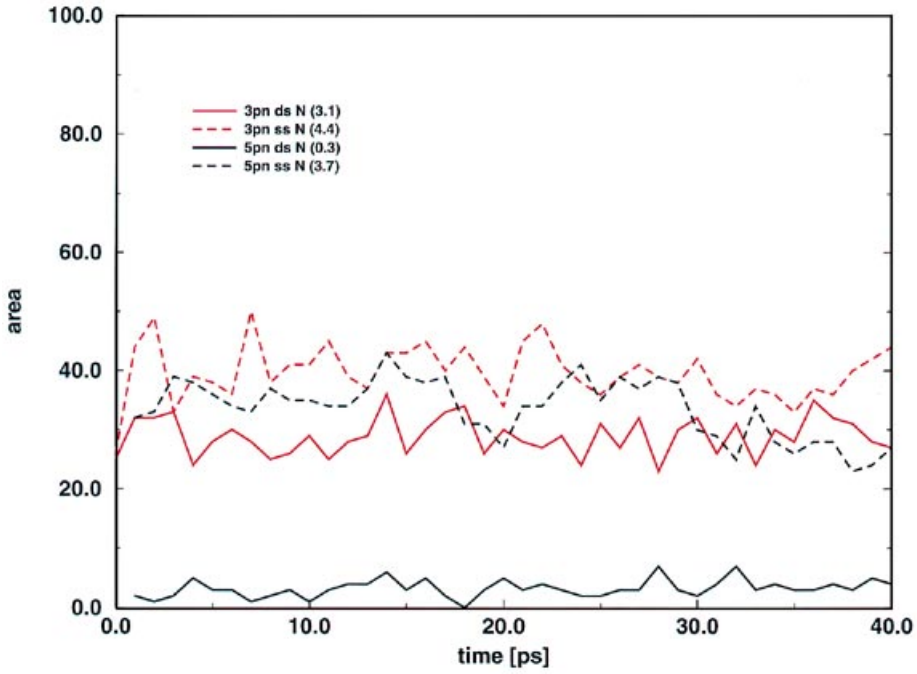

Figure 4. The total solvent accessible surface area (Å^2) of the nine nitrogens (N) for 3′ and 5′ single-stranded (ss) and double-stranded (ds) amidates over the first 40 ps of MD simulation. The average values for a single nitrogen are in parentheses (in Å^2) in the legend.

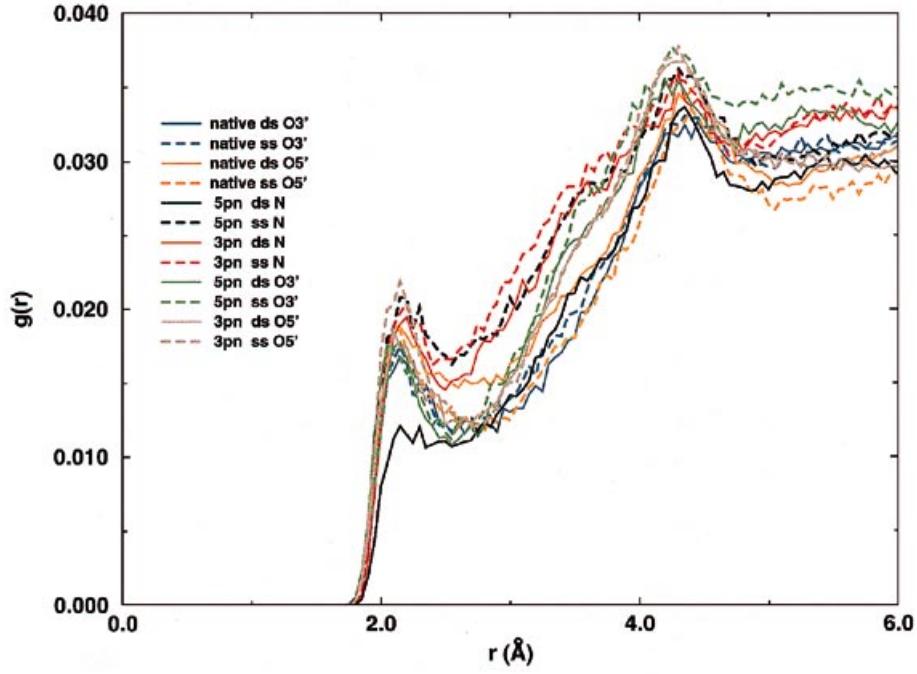

Figure 5. Unnormalized radial distributions *g(r)* of water protons around the phosphodiester oxygens and phosphoramidate nitrogens. Distributions are averaged over all nine linkages and averaged over 1200 different conformations in the course of a 120 ps free dynamics simulation.

intervals from each trajectory. The average solvent-exposed area for a single ds 5′-N is only 0.3 Å^2 while the solvent accessible surface areas for the other amino nitrogens are a factor of 10 higher.

In Table 1 we present the results from the quantum chemical calculations on the amidate and phosphodiester linkages shown in Figure 1. The 3′pnDNA fragment has a 2 kcal/mol lower (more negative) free energy of solvation than the 5′pnDNA fragment. Assuming the amidate substitution affects only the backbone–solvent interaction, the differential free energy of solvation for the decamer will be in the order of 20 kcal/mol. The phosphodiester backbone fragment has a solvation energy 0.8 kcal/mol greater than the 3′pnDNA fragment. It should be understood, however, that a 1.0 cal/mol predicted solvation energy difference is close to the expected resolution of the simple PCM solvation model.

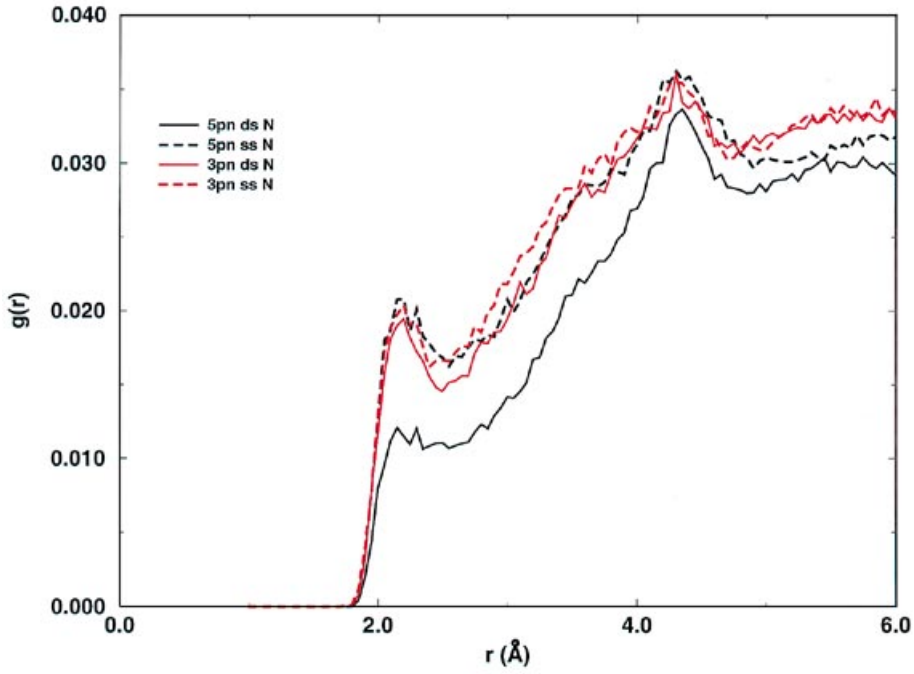

Figure 6. Unnormalized radial distributions *g(r)* of water protons around the phosphoramidate nitrogens as in Figure 5.

Table 1. Quantum chemical calculations on the oligonucleotide fragments shown in Figure 1

	Aqueous HF/6-31G* [Hartrees]	Gas phase HF/6-31G* [Hartrees]	Solv. energy (aq. – gas) [kcal/mol]	Relative solv. energy [kcal/mol]
Native DNA	–1140.222491	–1140.133878	–55.61	0.0
3′pn DNA	–1120.382294	–1120.294958	–54.80	0.81
5′pn DNA	–1120.384490	–1120.300668	–52.60	3.01

A precise description of the solvation-backbone interaction can be obtained from radial distribution functions *g(r)* of the water molecules in the vicinity of the relevant backbone atoms. In Figure 5 we present *g(r)* for water protons relative to the amidate nitrogens and the phosphodiester oxygens. These functions are computed from 1200 structures, obtained at 0.1 ps intervals in the simulations of the six DNA and DNA–RNA complexes. The corresponding coordination numbers have also been computed from the same simulations by counting the average number of water protons within a 2.5 Å radius of the chosen backbone atoms. The 3′-O, 5′-O and 3′-N atoms all have coordination numbers of 1.0 in both the ss and ds forms, while the 5′-N coordination numbers are 1.1 and 0.63 for the ss and ds forms, respectively. In Figure 7 we present *g(r)* of water protons around the 5′-N and O5′ atoms in the presence and absence of sodium counter-ions.

DISCUSSION

We have hypothesized that the observed differences in helix stabilities in the 3′ and 5′ amidates are caused by differences in solvent–backbone interactions. The hypothesis depends on two conjectures. First, that the 3′ and 5′ amidate DNA single-strand oligomers are equally well-solvated; and second, that there are differences in the solvent–backbone interactions in the 3′ and 5′DNA–RNA duplexes.

Molecular dynamics simulations of the single-stranded pnDNA oligomers confirm the first of these two conjectures. The structures in Figure 2 and the SAS results for the amidate atoms show that, in their dissociated forms, both amidate-modified backbones have comparable solvent-exposed phosphoramidate groups. Furthermore, the calculated radial distribution functions for water molecules around the amidates show similar degrees of water–amide hydrogen bonding in the dissociated forms of amidate-modified oligomers (Fig. 6).

In the duplex forms, however, this hydration symmetry disappears. The structures in Figure 2 and the SAS results for the nitrogen atoms show that, in the duplex forms, the amidate-modified backbones have very different solvent-exposed phosphoramidate groups. The SAS of the 5′ nitrogen is 10 times less than that of the 3′ counterpart (Fig. 4). The radial distribution functions presented in Figure 6 show that the density of water protons around the 3′-N looks very similar to those in the dissociated state, but there are only ~60% of the water protons around the 5′-N, indicating its lower hydration. Thus, the MD simulations of double-stranded pnDNA oligomers confirm our second conjecture.

The presence of sodium counter-ions does not change these results, as can be observed in Figure 7. The basic result, that solvation of the 5′-N increase is markedly stronger in the single-stranded form remains unaltered, even though there is somewhat less solvation in both forms in the presence of counter-ions. We emphasize that this differential solvation arises from the gross structural constraints of the A-form 3′- and 5′pnDNA, and does not depend on detailed structural or electronic differences between the two forms. Hence, this result is expected to be robust

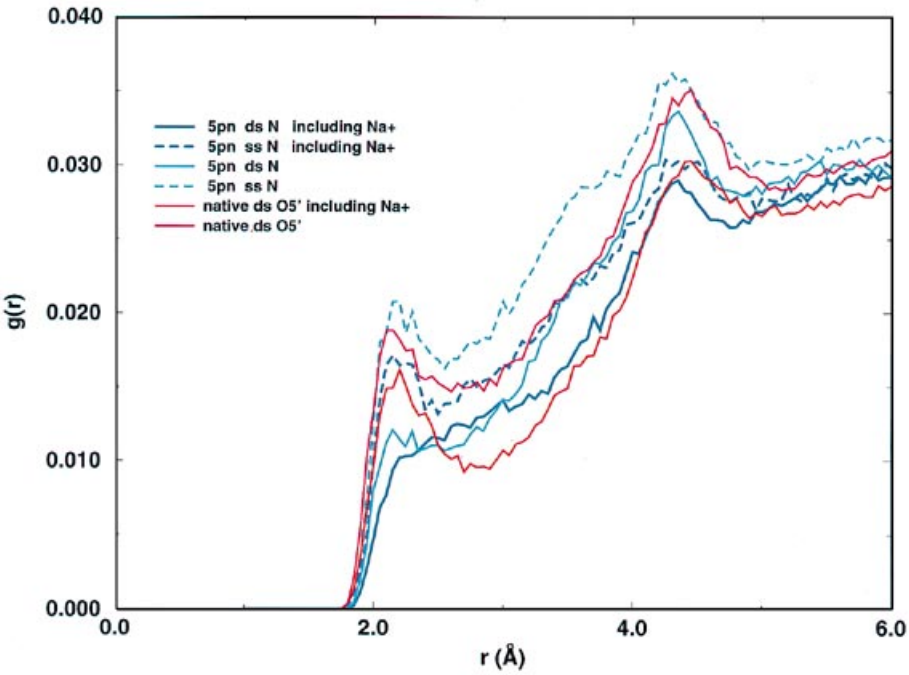

Figure 7. Unnormalized radial distributions *g(r)* of water protons around the 5′pn nitrogens and O5′ of the native DNA in the presence and absence of Na^+ ions. Distributions are averaged from 160 ps trajectories for the simulations including Na^+ ions.

to minor improvements in the parameterization of the molecular dynamics force field.

This study provides an interesting example of how a subtle intermolecular effect—the differential solvation of a relatively small moiety, the internucleoside phosphate group—can lead to a very pronounced effect on the structure and properties of large DNA molecules. This effect highlights the crucial and delicate balance of hydrophobic and hydrophilic forces in stabilizing the double helix.

It is plausible that similar backbone solvation effects mediate more subtle differences in helix stability. For example, a pronounced effect of the 3′ phosphoramidate linkages in d(CGCGAATTCGCG) is the 34°C increase in its temperature of dissociation, compared with phosphodiester DNA duplex (2). Relating this difference to the thermodynamic stability (2,21), however, reveals a stability difference of only ~0.2 kcal/mol per internucleoside linkage, which is well below the resolution of the available methods for predicting solvation free energies. Nevertheless, changes in backbone solvent accessibility should be considered in the investigation of all phenomena involving conformational or chemical modification of nucleic acid polymers and the stability of complexes they form.

ACKNOWLEDGEMENT

We wish to thank Ronald G. Schultz for helpful discussions in this work.

REFERENCES

1 Gryaznov,S. and Chen,J.–K. (1994) *J. Am. Chem. Soc.*, **116**, 3143–3144.
2 Gryaznov,S.M., Lloyd,D.H., Chen,J.-K., Schultz,R.G., DeDionisio,L.A., Ratmeyer,L., and Wilson,W.D. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 5798–5802.
3 Saenger,W. (1984) *Principles of Nucleic Acid Structure.* Springer–Verlag, New York, 368.
4 Molecular Simulations Inc., 16 New England Executive Park, Burlington, MA 01803–5297. *QUANTA version 4.0*, 1994.
5 Allen,F.H., Kennard,O., Watson,D.G., Brammer,L., Orpen,A.G. and Taylor,R. (1987) *J. Chem. Soc. Perkin Trans. II*, S1.
6 Brünger,A.T. (1992) *X–PLOR, Version S.1, A System for X–ray Crystallography and NMR*, Yale University Press.
7 Jorgensen,W., Chandrasekar,J., Madura,J., Impey,R. and Klein,M. (1983) *J. Chem. Phys.*, **79**, 925–935.
8 Miyamoto,S. and Kollman,P.A.J. (1992) *J. Comp. Chem.*, **13**, 952–962.
9 Fritsch,V., Ravishanker,G., Beveridge,D.L. and Westhof,E. (1993) *Biopolymers*, **33**, 1537–1552.
10 Singh,U.C., Weiner,S.J. and Kollman,P. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 755–759.
11 Beveridge,D.L., McConnell,K.J., Nirmala,R., Young,M.A., Vijayakumar,S. and Ravishanker,G. (1994) Molecular dynamics simulations of DNA and protein–DNA complexes including solvent. In *Structure and Reactivity in Aqueous Solutions.* American Chemical Society, pp. 381–394.
12 Judson,R.S., McGarrah,D.B., Melius,C.F., Mori,E., Barsky,D., Tan,Y.T., Windemuth,A., Tresurywala,A.M., Jaeger,E.P., Meza,J.C. and Plantenga,T. (1994) CCEMD: A Molecular Dynamics Simulation and Analysis Program. Technical Report 8258, Sandia National Labs, Livermore, CA 94551.
13 Windemuth,A. and Schulten,K. (1991) *Mol. Sim.*, **5**, 353–361.
14 Brooks,B.R., Bruccoleri,R.E., Olafson,B.D., States,D.J., Swaminathan,S. and Karplus,M.J. (1983) *J. Comp. Chem.*, **4**, 187–217.
15 Stoddard,S.D. and Ford,J. (1973) *Phys. Rer.*, A8, 1504–1512.
16 Allen,M.P. and Tildesley,D.J. (1989) *Computer Simulation of Liquids*, Oxford University Press, New York. Oxford Science Publications.
17 Levitt,M., Hirshberg,M., Sharon,R. and Daggett,V. (1995) *Comp. Phys. Comm.*, **91**, 215–231.
18 Frisch,M.J., Trucks,G.W., Schlegel,H.B., Gill,P.M.W., Johnson,B.G., Wong,M.W., Foresman,J.B., Robb,M.A., Head–Gordon,M., Replogle,E.S., *et al.* (1993) *Gaussian 92/DFT, Revision G.J.* Gaussian, Inc., Pittsburgh, PA.
19 Tomasi,J. and Persico,M. (1994) *Chem. Rer.*, **94**, 2027–2094.
20 Miertus,S., Scrocco,E., and Tomasi,J. (1981) *Chem. Phys.*, **55**, 117–129.
21 Li,Y., Zon,G. and Wilson,W.D. (1991) *Biochemistry*, **30**, 7566–7572.