

RNA Structural Forms Studied by Vibrational Circular Dichroism: Ab Initio Interpretation of the Spectra

Valery Andrushchenko,[†] Hal Wieser,[†] and Petr Bour^{*‡}

Department of Chemistry, University of Calgary, 2500 University Drive, Calgary AB, T2N 1N4, Canada, and Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nám. 2, 16610, Praha 6, Czech Republic

Received: October 14, 2003; In Final Form: January 12, 2004

Infrared (IR) and vibrational circular dichroism (VCD) spectra of single-stranded poly(rA) and poly(rU) RNA polynucleotides and double- and triple-stranded RNA helices were recorded and interpreted on the basis of density functional theory (DFT) at the BPW91/6-31G** level. Ab initio computations were performed for smaller fragments and extended to longer oligomers via transfer of atomic property tensors. The normal mode partial optimization method developed lately was found convenient for relaxation of the fragment geometries. Most of observed spectral features could be assigned to vibrations of the adenine, uracil, and sugar–phosphate chromophores; in particular, the VCD band shapes could be explained on the basis of specific spatial interactions. Thus, the VCD technique proved to be sensitive to various RNA conformational types (random coil, single strand, duplex, and triplex) in solution, spectra of which could be reliably modeled. The effect of the solvent could be only partially included, and ambiguity remains in the model structure used for the triplex RNA form.

1. Introduction

Vibrational circular dichroism (VCD) is a spectroscopic technique for monitoring the dependence of the differential absorption of left and right circularly polarized infrared light on the light frequency. The primary advantage of VCD consists of a higher sensitivity to molecular structure than is found for conventional infrared (IR) absorption spectra. During the past two decades this property has been extensively explored, and the technique has been regularly used for monitoring biopolymer conformations in solutions.^{1–4} For example, characteristic spectra were observed and assigned for standard peptide and protein secondary structures, including helical as well as extended β -sheet forms.⁵ Similarly, it was shown that the technique can discriminate between the B- and Z-helices of DNA^{3,6,7} and can also distinguish single-, double-, and triple-helical conformations of RNA.^{6,8,9} The origin of the VCD signal in the nucleic acids arises from a through-space coupling of rather localized normal mode vibrations, and the helical structures provide specific VCD patterns.^{4,10} Indeed, optimal conditions for this coupling are realized when the nucleobases as well as the sugar–phosphate backbone are stacked upon one another in a periodic structure. The high sensitivity of the VCD signal to geometry variations arises from different coupling interactions among those chromophores in specific secondary structures.

For meaningful experiments, it is desirable to establish a link between the spectral shape and the molecular structure and to understand as many features in the experimental spectra as possible. Previously, we developed a general technique to

simulate the VCD spectra of large molecules by ab initio computations on small fragments and applied it to the B- and Z-forms of DNA.¹¹ In the present work, we model the spectra of a greater variety of RNA forms, namely, of the single-, double-, and triple-stranded conformations.

Traditionally, interpretation of the VCD spectra of RNAs was based primarily on empirical observations,^{12–14} which provided a limited basis for detailed assignments and characteristics of the bands but did not allow the prediction of spectra for unknown structures. But the RNA molecule is quite flexible and the detailed geometry is often not known. Depending on experimental conditions, such as temperature, polymer concentration, and ion contents, the poly(rA) and poly(rU) polymers form a wide range of single, double, or triple helices as well as coils.^{8,9,12–14} Most of the structures have been confirmed by X-ray crystallography.¹⁵ However, many forms do not crystallize at all, or their geometrical parameters in solution and in a crystal may differ. Thus, a particular goal of the interpretation of VCD spectra is also to assess to what degree the properties of the nucleic acids determined for the solid state are transferable to solutions.

The structure of the poly(rU)*poly(rA)*poly(rU) triple helix is particularly challenging. The poly(rU)*poly(rA)*poly(rU) triplex can be formed by causing the poly(rA)*poly(rU) duplex to disproportionate into a triple helix and a single strand of poly-(rA) with increasing salt concentration^{16,17} or by the simultaneous action of salt and temperature.^{6,8,9} There are several possibilities for arranging the individual strands, and its solution geometry has not been resolved yet. It is generally accepted that the second poly(rU) strand fits into the major groove of the poly(rA)*poly(rU) double helix with a formation of the Hoogsteen-type base pairs with the poly(rA) strand.^{12,17} However, an alternative model, formation of a reverse Hoogsteen-type base pairing, better fits many IR and VCD profiles and

* Author to whom correspondence should be addressed. Fax: (420)-224-310-503. E-mail: bour@uochb.cas.cz.

[†] University of Calgary. E-mail: vandrush@ucalgary.ca (V.A.); hwieser@acs.ucalgary.ca (H.W.).

[‡] Academy of Sciences of the Czech Republic.

their interpretations by the coupled-oscillator model.^{6,9,18} So far, several attempts have been made to assign the IR and VCD observed transitions.^{6,8,9,12} Unfortunately, as shown below, this particular issue cannot be resolved in the present study either, because of the limited precision of the quantum computations for these large systems.

For rigid small molecules, ab initio simulation of the VCD spectra within the harmonic approximation is relatively straightforward. Most conveniently, the magnetic field perturbation theory (MFP) of Stephens¹⁹ combined with density functional theory and field-dependent atomic orbitals²⁰ can be applied, as implemented, for example, in the Gaussian program package.²¹ Usually, such a procedure provides reliable results in the mid-IR region; for higher-frequency vibrations anharmonic corrections must be considered. The size and conformational flexibility of the biopolymers, however, hinder such a direct modeling. Nevertheless, in these instances the locality of the VCD phenomenon can be explored;^{10,11} that is, suitably chosen fragments often provide enough information of the spectral response of the whole larger molecule. Thus, the spectra of all standard forms of peptides (α , 3_{10} , PPI, and PPII helices, parallel and antiparallel β -sheets) could be interpreted in the past.^{22,23} In peptides and proteins the interactions between neighboring amide groups defined the VCD spectral shape. Similarly for nucleic acids, the signal arises mainly from the dipolar interaction between two stacked base pairs and analogous interactions of the phosphate and sugar residues.^{4,10} Hence, the basic unit in nucleic acids determining the basic VCD pattern is composed of two stacked base pairs with two sugar-phosphate pairs. It is much larger than the basic chiral unit of peptides (two amide groups), which consequently limits the precision of the simulations. On the other hand, the vibrational coupling between two base pairs is weaker (mediated by van der Waals, dispersion, and electrostatic interactions) than between adjacent amide groups (coupled by springlike covalent bonds) and enables a perturbation treatment. Particularly in this study, we use different ab initio levels for the computation of force constants, depending on whether the associated atoms lie in the same or neighboring base pairs. The modeling becomes even more complex when hydration is taken into account. Fortunately, it has been recognized that the solvent influences mainly the band positions, whereas the spectral shape, most characteristic for the secondary structure, is quite conservative.^{22,24}

2. Methods

2.1. VCD Measurement. The methodology of preparing the sample and obtaining experimental IR and VCD spectra was described previously.⁹ The sodium salts of poly(rA)·poly(rU), poly(rA), and poly(rU) as well as NiCl₂·6H₂O were purchased from Sigma Chemicals and used without further purification. Poly(rA)·poly(rU) and metal ion solutions were prepared in sodium cacodylate buffer (Sigma Chemicals, concentration = 10⁻² M, pH \approx 6.5). All solutions were prepared in autoclaved double-distilled water. Complete deuterium exchange was achieved by lyophilizing and redissolving the H₂O solutions three times in D₂O. The concentration of the poly(rA)·poly(rU) stock solution was 74 mg/mL (0.24 M phosphate) measured by UV absorption at 260 nm ($\epsilon = 7140 \text{ M}^{-1}\text{cm}^{-1}$) after D₂O exchange. The samples were placed between two BaF₂ windows in a demountable cell (International Crystal Laboratories, Inc.) separated by a 50 μm Teflon spacer. The final poly(rA)·poly(rU) concentration in the cell was 56 mg/mL (\sim 0.18 M phosphate). Ni²⁺ solution was added dropwise to the poly(rA)·poly(rU) solution directly on the bottom window under continu-

TABLE 1: Fragments Used for the Spectral Simulations of RNA

polymer	fragments	approximation level used for the Hessian computation
(rA) ₈	two stacked adenine bases (AA) sugar-phosphate dimer (SPSP)	BPW91/6-31G** BPW91/6-31G**
(rU) ₈	two stacked uracil bases (UU) SPSP	BPW91/6-31G** BPW91/6-31G**
(rA) ₈ *(rU) ₈	A...U base pair two base pairs (A...U) ₂ SPSP	BPW91/6-31G** PM3 BPW91/6-31G**
(rU) ₈ *(rA) ₈ *(rU) ₈	single base trimer U...A...U two stacked UAU ((U...A...U) ₂) SPSP1, SPSP2, and SPSP3 (different for each strand)	BPW91/6-31G** HF/3-21G BPW91/6-31G**

ous stirring. The final metal ion concentration in the cell was 0.07 M Ni²⁺ corresponding to 0.4 molar concentration of Ni²⁺ per molar concentration of phosphate ([Ni]/[P]).

All infrared absorption and VCD spectra were measured in D₂O, in the range of 1800–800 cm⁻¹ with a homemade instrument²⁵ located at the University of Calgary. A total of 4000 ac scans (about 1 h scanning time) were accumulated for each sample and ratioed against 400 dc scans, all at 8 cm⁻¹ resolution. The VCD spectra were corrected for polarization artifacts by subtracting the spectra of the solvent obtained at the same conditions. The cell temperature was maintained to within \pm 0.1 °C with a circulating water thermostat (NESLAB Instruments Inc.) using an electronic thermometer with a copper-constantan thermocouple (OMEGA Technology Co. Inc.). A similar procedure was used for preparation and spectra acquisition of single-stranded poly(rA) and poly(rU). The double- to triple-helix transition of poly(rA)*poly(rU) was induced upon elevation of temperature in the presence of Ni²⁺ ions. The metal ions had a negligible effect on the overall spectral shape, and essentially all monitored spectral changes appeared because of structural transitions of macromolecules rather than the metal ion binding.

2.2. Calculations. Force field and the intensity tensors of larger nucleotides were constructed from smaller fragments using the Cartesian coordinate transfer technique.²⁶ A complete list of the fragments is given in Table 1. Single-stranded (rA)₈, (rU)₈, double-stranded (rA)₈*(rU)₈, and triple-stranded (rU)₈*(rA)₈*(rU)₈ octamers were used as target molecules. The geometries of (rA)₈, (rU)₈, and (rA)₈*(rU)₈ were generated using the Biosym software²⁷ with parameters based on averaged X-ray data, whereas experimental X-ray geometry¹⁵ was used for (rU)₈*(rA)₈*(rU)₈. Further optimization was not attempted because of the lack of reliable molecular dynamics (MD) models. Because experimental IR and VCD spectra of such shorter sequences are almost identical to those of longer polymers,^{7,10,11,28} the results were compared directly to the spectra of the polynucleotides. The fragments were analogous to those used in a previous modeling of DNA.¹¹ Although the geometries of the target octanucleotides were not changed during the calculations, the smaller fragments were relaxed with the aid of the normal mode optimization method.²⁹ Normal modes below 300 cm⁻¹ were frozen, while the modes in the experimentally accessible region could relax. For the optimized fragments, harmonic force fields, dipole derivatives, and the atomic axial tensors were calculated with the Gaussian 98 suite of programs²¹ and transferred²⁶ to the octamers. The DFT BPW91^{30,31}/6-31G** level was used, because of our previous good experience with this functional.^{11,24,32} Additionally, in order to account for the longer-range coupling, for larger fragments

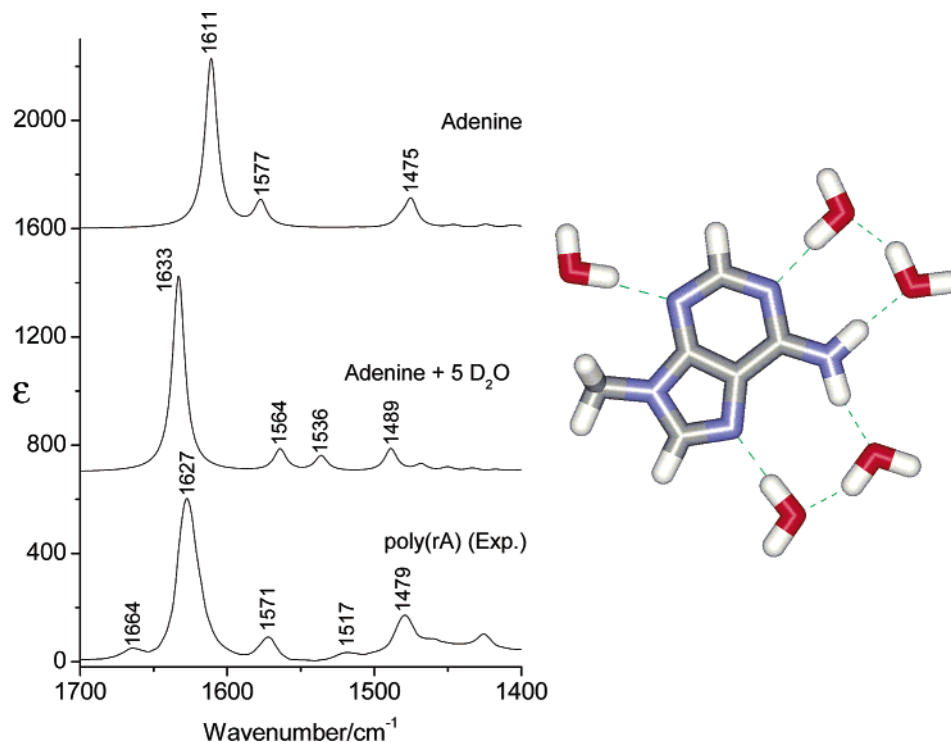


Figure 1. Effect of solvation on the absorption spectra of adenine and comparison with the experimental spectrum of poly(rA). The geometry of the solvated form is shown on the right.

comprising two chromophore units the PM3 harmonic force field was calculated. Test computations confirmed that the PM3 method provided similar longer-distance force constants and VCD patterns as those obtained from the computationally more expensive DFT methods. Homemade programs were used for the tensor transfers and computations of harmonic frequencies and intensities.²⁶ As in the experiment, acidic hydrogen atoms were always replaced by deuteria for the diagonalization of the force field matrix.

For a detailed understanding of individual chromophores and the effect of hydration (always by D₂O for all systems here), we calculated spectra of single bases, base pairs, and base triplets with and without explicit D₂O molecules. Solvent positions were chosen so that hydrogen bonds with the exposed oxygen, nitrogen, and acidic deuterium (hydrogen) atoms of RNA could be formed. The RNA units were generated as segments of the octamers. Planar reverse Hoogsteen base pairing was achieved by rotating the Hoogsteen base paired uracil by 180° and adjusting its position so that it could be stabilized by the hydrogen bonds.

3. Results and Discussion

3.1. Adenine and Uracil Bases: The Effect of the Solvent.

The calculated IR spectra of adenine in vacuum and with five explicit D₂O molecules are compared to experiment for poly(rA) in Figure 1 and Table 2. The solvent has little influence on relative peak intensities, whereas the frequency of the most intense ring stretching mode calculated in vacuum at 1611 cm⁻¹ shifts to 1633 cm⁻¹ upon solvation, much closer to the experimental band at 1627 cm⁻¹. Similarly, the vacuum band calculated at 1577 cm⁻¹ shifts to 1564 cm⁻¹ (1571 cm⁻¹ experimental). Only the shift of the 1475 cm⁻¹ band toward 1489 cm⁻¹ upon solvation moves it farther from the experimental frequency of 1479 cm⁻¹. Overall, most of these changes are rather minor because the molecule is predominantly hydrophobic. Also, water access to many binding sites may be

blocked in the poly(rA) polymer. A new signal from a ring mode appeared at 1536 cm⁻¹ as a result of the solvation, most probably corresponding to the very weak experimental absorption at 1517 cm⁻¹. This new signal is also apparent in the experimental spectra of poly(rA)*poly(rU) and the calculated spectra of the AU base pairs, both solvated and in vacuum, as shown below.

For the uracil molecule in vacuum and surrounded with four and five water molecules, the two strongest absorption bands originate from the two uracil carbonyl groups (see Figure 2 and Table 2 for comparison with the experimental spectrum of poly(rU)). Here, solvation induces more drastic changes in the spectra for uracil than for adenine. Both the C=O₂ and C=O₄ bands predicted at 1763 and 1739 cm⁻¹ in vacuum shift down significantly upon addition of four D₂O molecules to 1721 and 1700 cm⁻¹, respectively, and further to 1695 and 1671 cm⁻¹ upon addition of the fifth D₂O molecule. That means that the addition of the fifth D₂O molecule to the system still causes an additional shift by 30 cm⁻¹, for a total of 70 cm⁻¹ if related to the vacuum computation. Such sensitivity to solvation is rather exceptional and unexpected and indicates that our simplistic static cluster model of the solvation may not be adequate for uracil. Indeed, other test computations (not shown here) indicate that the spectra are quite sensitive both to the number and positions of the H₂O or D₂O molecules. Nevertheless, the band of the complex with five D₂O molecules, calculated at 1695 cm⁻¹, almost coincides with the experimental band at 1693 cm⁻¹. The other carbonyl band was calculated at a slightly higher wavenumber (1671 cm⁻¹) than was found experimentally (1656 cm⁻¹). The band at 1633 cm⁻¹ of uracil in vacuum originates from vibrations localized inside the ring and, as expected, is not affected by the solvent, shifting only to 1631 and 1626 cm⁻¹ for the complex with four and five D₂O molecules, respectively. It can be assigned convincingly to the experimental band at 1617 cm⁻¹.

3.2. The AU Base Pair. The calculated spectra of the adenine and uracil complexes and their comparison to the experimental

TABLE 2: Calculated Vibrational Frequencies (BPW91/6-31G) of the RNA Chromophores Contributing Most to the Spectra, Compared to Experimental Values (in Parentheses) for Poly(rA), Poly(rU), Poly(rA)*Poly(rU), and Poly(rU)*Poly(rA)*Poly(rU)**

molecule	wavenumber (cm ⁻¹)	assignment
adenine (A)	1611 (1627)	$\nu(\text{C}=\text{N})$ ring + $\nu(\text{C}-\text{ND}_2)$
	1577 (1572)	(C=N) ring
	1475 (1479)	$\nu(\text{C}=\text{N})$ ring + C-H bend
A + 5 D ₂ O	1633 (1627)	$\nu(\text{C}=\text{N})$ ring + $\nu(\text{C}-\text{ND}_2)$
	1564 (1572)	$\nu(\text{C}=\text{N})$ ring
	1536 (1518)	$\nu(\text{C}=\text{N})$ ring + C-H bend
	1489 (1479)	$\nu(\text{C}=\text{N})$ ring + C-H bend
thymin (T)	1758	$\nu(\text{C}=\text{O}_2)$
	1718	$\nu(\text{C}=\text{O}_4)$
	1649	$\nu(\text{C}=\text{N})$ ring + C-H bend
	1476	C-H ₃ (1476 and below)
uracil (U)	1763 (1693)	$\nu(\text{C}=\text{O}_2)$
	1739 (1656)	$\nu(\text{C}=\text{O}_4)$
	1633 (1618)	$\nu(\text{C}=\text{N})$ ring + C-H bend
	1488	-CH ₃ (1488 and below)
U + 4 D ₂ O	1721 (1693)	$\nu(\text{C}=\text{O}_2)$ + $\nu(\text{C}=\text{O}_4)$ in-phase
	1700 (1656)	$\nu(\text{C}=\text{O}_2)$ + $\nu(\text{C}=\text{O}_4)$ out-of-phase
	1631 (1618)	(C=N) ring + C-H bend
	1486	-CH ₃ (1486 and below)
U + 5 D ₂ O	1695 (1693)	$\nu(\text{C}=\text{O}_2)$ + N-D bend
	1671 (1656)	$\nu(\text{C}=\text{O}_4)$ + N-D bend + C-H bend
	1626 (1618)	$\nu(\text{C}=\text{N})$ ring + C-H bend
	1488	-CH ₃ (1488 and below)
AT planar	1737	$\nu(\text{C}=\text{O}_2)$ (T)
	1664	$\nu(\text{C}=\text{O}_4)$ (T)
	1643	$\nu(\text{C}=\text{N})$ (T) ring + $\nu(\text{C}=\text{O}_4)$ (T) + C-H (T) bend
	1621	$\nu(\text{C}=\text{N})$ ring (A) + $\nu(\text{C}-\text{ND}_2)$ (A)
	1569	$\nu(\text{C}=\text{N})$ ring (A)
	1518	$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A)
	1484	$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A)
AT planar + 4 D ₂ O	1722	$\nu(\text{C}=\text{O}_2)$ (T)
	1662	$\nu(\text{C}=\text{O}_4)$ (T) + $\nu(\text{C}=\text{N})$ ring (T) + C-H bend (T)
	1642	$\nu(\text{C}=\text{N})$ ring (T) + $\nu(\text{C}=\text{O}_4)$ (T) + C-H bend (T)
	1623	$\nu(\text{C}=\text{N})$ ring (A) + $\nu(\text{C}-\text{ND}_2)$ A + $\nu(\text{C}=\text{O}_4)$ (T)
	1570	$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A)
	1529	$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A) + $\nu(\text{C}-\text{ND}_2)$ (A)
	1487	$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A) + $\nu(\text{C}-\text{ND}_2)$ (A)
AU planar	1739 (1689)	$\nu(\text{C}=\text{O}_2)$ (U)
	1681 (1669)	$\nu(\text{C}=\text{O}_4)$ (U)
	1630sh (1631) ^a	$\nu(\text{C}=\text{N})$ ring (U) + C-H bend (U)
	1622 (1631)	$\nu(\text{C}=\text{N})$ ring (A) + $\nu(\text{C}-\text{ND}_2)$ (A)
	1571 (1569)	$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A)
	1517	$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A) + $\nu(\text{C}-\text{ND}_2)$ (A)
AU planar + 5 D ₂ O	1489	$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A)
	1718 (1689)	$\nu(\text{C}=\text{O}_2)$ (U)
	1671 (1669)	$\nu(\text{C}=\text{O}_4)$ (U) + $\nu(\text{C}=\text{N})$ ring (U)
	1635 (1631)	$\nu(\text{C}=\text{N})$ ring (U) + C-H bend (U) bend + $\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A) + $\nu(\text{C}-\text{ND}_2)$ (A)
	1621sh ^a	$\nu(\text{C}=\text{O}_4)$ (U) + C-H bend (U)
	1573 (1569)	$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A)
	1528	$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A) + $\nu(\text{N}-\text{D}_2)$ (A)
1489	$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A)	
UAU _h planar	1738 (1696)	$\nu(\text{C}=\text{O}_2)$ (U _{wc})
		$\nu(\text{C}=\text{O}_2)$ (U _H)
	1733 (1696)	70% $\nu(\text{C}=\text{O}_4)$ (U _H) + 30% $\nu(\text{C}=\text{O}_4)$ (U _{wc})
	1695sh (~1673) ^a	70% $\nu(\text{C}=\text{O}_4)$ (U _{wc}) + 30% $\nu(\text{C}=\text{O}_4)$ (U _H)
	1683 (~1658)	$\nu(\text{C}=\text{N})$ ring (U _H) + C-H bend (U _H)
	1632 (1627)	$\nu(\text{C}=\text{N})$ ring (U _{wc}) + C-H bend (U _{wc})
	1629 (1627)	$\nu(\text{C}=\text{N})$ ring (U _{wc}) + C-H bend (U _{wc})
	1625 (1627)	$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A)
	1567 (1573)	$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A) + $\nu(\text{C}-\text{ND}_2)$ (A)
	1524	

TABLE 2 (Continued)

molecule	wavenumber (cm ⁻¹)	assignment	
UAU _H planar + 6 D ₂ O	1722 (1696)	$\nu(\text{C}=\text{O}2)$ (U _{wc}) + N-D bend (U _{wc})	
	1715 (1696)	70% $\nu(\text{C}=\text{O}2)$ (U _H) + 30% $\nu(\text{C}=\text{O}4)$ (U _H) in-phase	
	1692 (~1673)	70% $\nu(\text{C}=\text{O}2)$ (U _H) + 30% $\nu(\text{C}=\text{O}4)$ (U _H) out-of-phase + $\nu(\text{C}=\text{N})$ ring (U _H) + C-H bend (U _H)	
	1666 (~1658)	$\nu(\text{C}=\text{O}4)$ (U _{wc}) + $\nu(\text{C}=\text{N})$ ring (U _{wc}) + C-H bend (U _{wc})	
	1639 (1627)	$\nu(\text{C}=\text{N})$ ring (U _H) + C-H bend (U _H)	
	1630 (1627)	$\nu(\text{C}=\text{N})$ ring (U _{wc}) + C-H bend (U _{wc}) + $\nu(\text{C}=\text{N})$ ring (A) + $\nu(\text{C}-\text{ND}_2)$ (A)	
	1573 (1573)	$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A)	
	1531	$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A) + $\nu(\text{N}-\text{D}_2)$ (A)	
	UAU _H planar	1738 (1696)	$\nu(\text{C}=\text{O}2)$ (U _{wc})
		1719sh (1696) ^a	50% $\nu(\text{C}=\text{O}2)$ (U _H) + 50% $\nu(\text{C}=\text{O}4)$ (U _H) in-phase
1705 (~1673)		50% $\nu(\text{C}=\text{O}2)$ (U _H) + 50% $\nu(\text{C}=\text{O}4)$ (U _H) out-of-phase	
1684 (~1658)		$\nu(\text{C}=\text{O}4)$ (U _{wc}) + C-H bend (U _{wc})	
1633 (1627)		$\nu(\text{C}=\text{N})$ ring (U _H) + C-H bend (U _H)	
1629 + 1625 (1627)		$\nu(\text{C}=\text{N})$ ring (U _{wc}) + C-H bend (U _{wc}) + $\nu(\text{C}=\text{N})$ ring (A) + $\nu(\text{C}-\text{ND}_2)$ (A)	
1568 (1573)		$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A)	
1523		$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A) + $\nu(\text{C}-\text{ND}_2)$ (A)	
UAU _H planar + 6 D ₂ O	1716 (1696)	$\nu(\text{C}=\text{O}2)$ (U _{wc})	
	1704 (1696)	$\nu(\text{C}=\text{O}2)$ (U _H)	
	1678 (~1673)	$\nu(\text{C}=\text{O}4)$ (U _H) + $\nu(\text{C}=\text{N})$ ring (U _H) + C-H bend (U _H)	
	1665 (~1658)	$\nu(\text{C}=\text{O}4)$ (U _{wc}) + $\nu(\text{C}=\text{N})$ ring (U _{wc}) + C-H bend (U _{wc})	
	1632 (1627)	$\nu(\text{C}=\text{N})$ ring + C-H bend (U _{wc} , U _H , A) + $\nu(\text{C}-\text{ND}_2)$ (A)	
	1626sh (1627) ^a	$\nu(\text{C}=\text{N})$ ring (U _H) + C-H bend (U _H) + $\nu(\text{C}=\text{O}4)$ (U _H)	
	1616 (1627)	$\nu(\text{C}=\text{N})$ ring (U _{wc}) + C-H bend (U _{wc}) + $\nu(\text{C}=\text{O}4)$ (U _{wc})	
	1567 (1573)	$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A)	
	1527	$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A) + $\nu(\text{C}-\text{ND}_2)$ (A)	

^a The "sh" in the wavenumber column indicates a band shoulder.

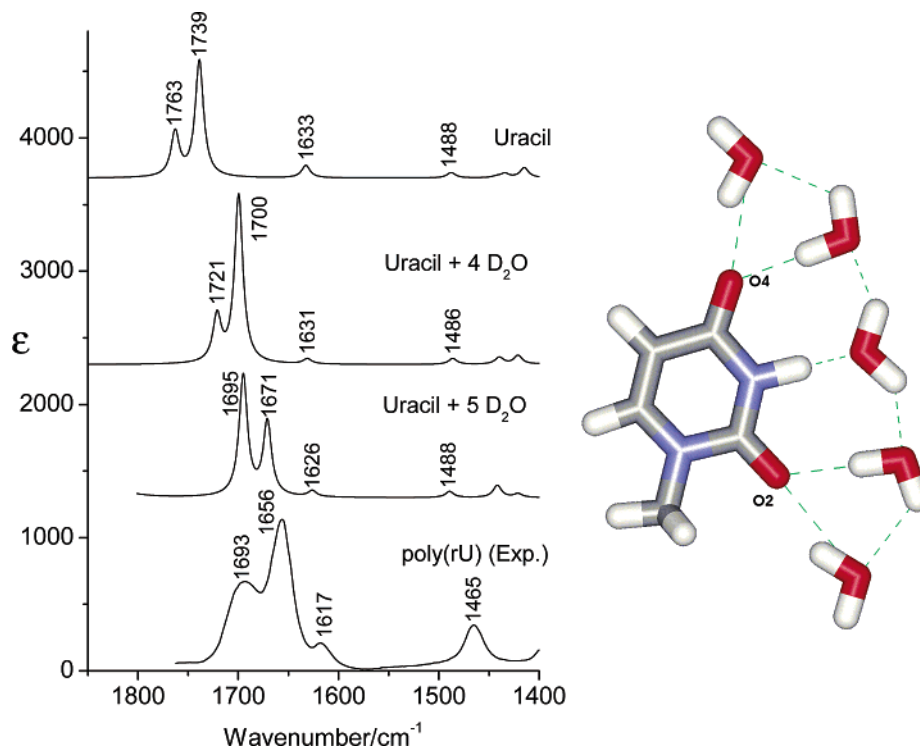


Figure 2. Effect of solvation on the absorption spectra of uracil and comparison with the experimental spectrum of poly(rU). The geometry of the second hydrated form is shown on the right.

absorption of poly(rA)*poly(rU) are presented in Figure 3, and the detailed assignments are given in Table 2. Contrary to our expectations, the nonplanar complex derived from the idealized crystal RNA structure (upper two traces in Figure 3) did not provide any improvement to the spectra compared to the planar arrangement. The latter is also more convenient for modeling because the C_5 symmetry speeds up the computations. Although

the spectrum of the AU base pair can be roughly considered as a sum of the spectra of adenine and uracil, a detailed comparison reveals many nonadditive effects. The experimental band at 1631 cm⁻¹ (calculated as 1635 cm⁻¹ for the AU planar + 5 D₂O complex in Figure 3) originates predominantly from the adenine vibration (calculated at 1633 cm⁻¹, cf. Figure 1). But its intensity significantly decreases upon formation of the base pair. This is

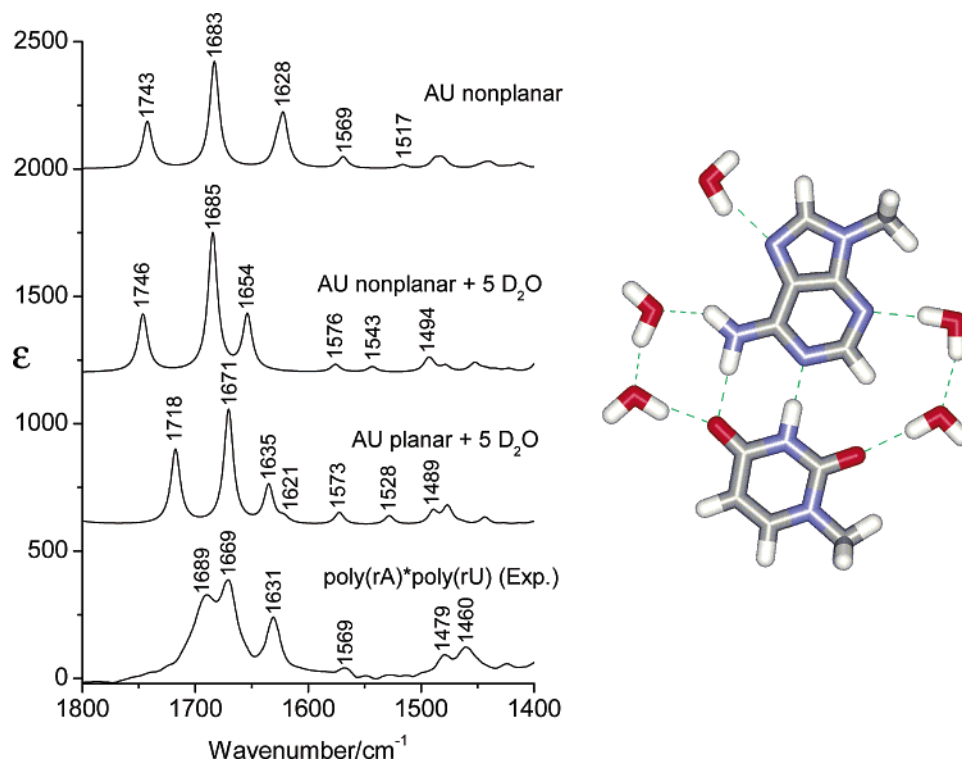


Figure 3. Effect of the formation of the AU base pair and of its nonplanarity and solvation on the absorption spectra and comparison with the experimental spectrum of poly(rA)*poly(rU). The planar hydrated complex is displayed on the right.

in agreement with experiment and was described in many previous studies.^{8,9,12} The well-established experimental upward shift of the adenine band upon forming the base pair (1627 to 1631 cm^{-1}) is also nicely reproduced theoretically (1633 to 1635 cm^{-1}) and, as follows from its higher value for the nonplanar model, it can be partially induced by the propeller twist of the base pair plane. The calculated AU bands at 1573, 1528, and 1489 cm^{-1} come from adenine ring vibrations (see the bands at 1564, 1536, and 1489 cm^{-1} in Figure 1), but only the first can be assigned with certainty to the experimental absorptions of polyA*polyU at 1569 cm^{-1} . On the other hand, the two strongest experimental bands at 1689 and 1669 cm^{-1} clearly correspond to the calculated peaks at 1718 and 1671 cm^{-1} , respectively. They originate from the carbonyl stretching of uracil. The slightly larger difference between the theoretical and experimental wavenumbers for the higher-energy band (1689/1718 cm^{-1}) can be explained by the extreme sensitivity of uracil to hydration and formation of the hydrogen bonds (cf. Figure 2).

The inclusion of the explicit D_2O molecules in the model AU base pair improves the agreement of the calculated and experimental frequencies similarly as occurred for the individual bases. The effect of hydrogen bonding combines with the effect of base pair formation. For example, the adenine ring stretching vibration at 1628 cm^{-1} in the nonsolvated pair mixes with the ring stretching of uracil at approximately the same wavenumber. As expected, the two modes involving the C=O2 and C=O4 groups of uracil are most affected shifting from 1739 and 1681 cm^{-1} , respectively, for the planar base pair in vacuum (spectrum not shown) down to 1718 and 1671 cm^{-1} with five D_2O molecules. Other modes, such as the adenine ring stretching vibration calculated around 1573 cm^{-1} , remain essentially unchanged by the solvent. The slightly better band shape for the planar base pair compared to the nonplanar model might indicate that the idealized structures do not adequately describe the base pair twist. Indeed, the model based mostly on X-ray

data may not be appropriate for solution structures, where the base pairs may be allowed to move about an equilibrium planar or semiplanar geometry.

3.3. Hoogsteen and Reversed Hoogsteen UAU Trimer. As the last building block, we analyzed the Hoogsteen (H) and reversed Hoogsteen (rH) triplets in Figure 4, because they were both proposed for triple-helix RNA structures. The uracil bases bound to adenine via Watson–Crick (WC), Hoogsteen, or reversed Hoogsteen hydrogen bonding are denoted by U_{wc} , U_{H} , and U_{rH} , respectively.

For the Hoogsteen model the highest energy absorption band at 1736 cm^{-1} consists of two C=O2 stretching modes, one of U_{wc} and the other of U_{H} . Although these two vibrations are close in wavenumber, they are not coupled, as can be seen from the dynamic displacement of the normal mode motion. Thus, in the case of the Hoogsteen base pairing the C=O2 carbonyls are not involved in hydrogen bonding. On the other hand, the C=O4 stretching vibrations of U_{wc} and U_{H} are coupled to give the bands at 1695 and 1683 cm^{-1} . Upon hydration with D_2O molecules, the carbonyl stretching vibrations are shifted to lower wavenumbers as for the previous complexes. The C=O2 signal shifts down by about 16 cm^{-1} and splits into the 1722 and 1715 cm^{-1} bands. The C=O4 vibrations exhibit a much bigger split. One of them is shifted only by 3 cm^{-1} to 1692 cm^{-1} , the other by 17 cm^{-1} to 1666 cm^{-1} . The other vibrations not containing C=O stretching (below 1630 cm^{-1}) do not change significantly upon solvation.

The spectra of the reversed Hoogsteen triplet and their response to D_2O solvation are similar. The C=O2 stretch signal is somewhat more split in the solvated spectrum (1738/1719 to 1716/1704 cm^{-1}), whereas the C=O4 bands exhibit much smaller splitting (1705/1684 to 1678/1665 cm^{-1}). Because of the limited agreement with the experimental spectra of the poly(rU)*poly(rA)*poly(rU) triple helix, the simulation does not reveal which triplex is preferred. Nevertheless, the experimental bands can still be assigned to individual chromophores. Thus,

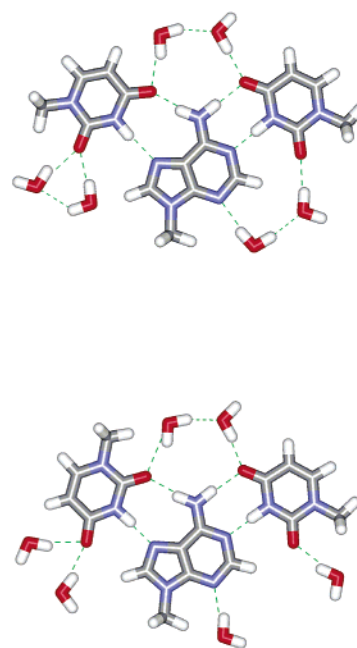
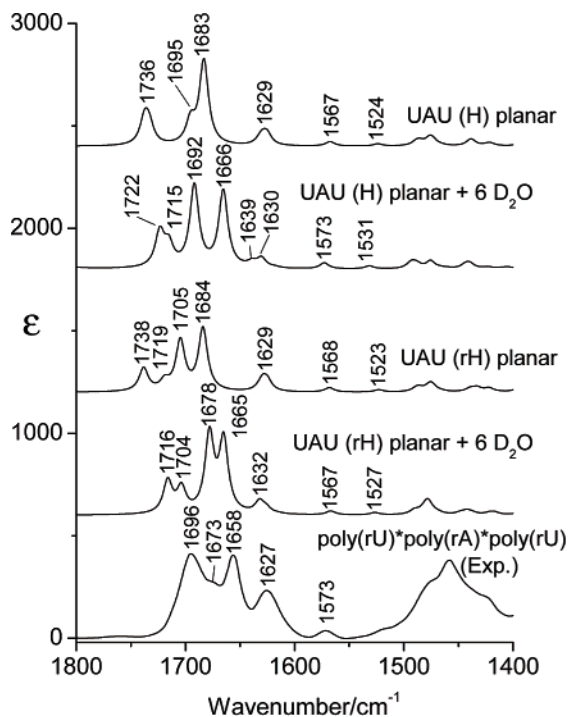


Figure 4. Effect of solvation of the planar Hoogsteen and reversed Hoogsteen U...A...U trimers on the absorption spectra and comparison with the experimental spectrum of the poly(rU)*poly(rA)*poly(rU) triple helix. The solvated Hoogsteen (top) and reversed Hoogsteen (bottom) model complexes are shown on the right.

the experimental absorption at 1696 cm^{-1} clearly corresponds to the C=O2 carbonyl modes of the uracils, although the two components are not resolved. The experimental bands at 1673 and 1658 cm^{-1} stem from the C=O4 carbonyl stretches of U in agreement with the assignment of Ohms and Ackermann for the normal Hoogsteen pair.¹² However, their proposed order (1673 and 1658 cm^{-1} bands for the C=O4 stretching modes of U_{wc} and U_H , respectively) is opposite to our calculations. The assignments proposed for the reversed Hoogsteen pairing^{6,9,18} suggest that the 1673 cm^{-1} band corresponds to the hydrogen-bonded C=O carbonyl mode of any uracil (U_{wc} or U_{rH}), and the other band corresponds to the free (not hydrogen-bonded) C=O mode of any uracil. This was consistent with the wavenumber red-shift of the C=O4 carbonyl vibrations upon the duplex \rightarrow single strand transition. Probably the C=O4 site can be accessed by water molecules and form H-bonds upon strand breaking more easily than in a base pair. From this point of view the C=O4 (U_{rH}) is free in the reversed Hoogsteen base pair and should therefore appear at lower wavenumber than C=O4 (U_{wc}). Apparently, our computations cannot account for such small effects in wavenumber changes caused by solvation, base complex formation, triplet nonplanarity, and triple-helix dynamics, and we leave the definitive answer about the reverse and normal Hoogsteen model to future studies. Nevertheless, the band positions for the reversed Hoogsteen model appear much closer to those of the experiment than do those for the Hoogsteen model. Lower energy modes calculated around 1630 and 1570 cm^{-1} can be assigned unambiguously to ring vibrations of the bases (Table 2), because they change appreciably neither under solvation nor under the Hoogsteen to reversed Hoogsteen pairing flip. Experimentally, the corresponding peaks are at 1627 and 1573 cm^{-1} .

3.4. The Sugar–Phosphate Residue. The basic properties of the sugar–phosphate vibrations are similar to those found for DNA.¹¹ The geometry of the sugar–phosphate dimer used for RNA is displayed in Figure 5. The newly adopted optimization in the normal mode coordinates enabled a better relaxation

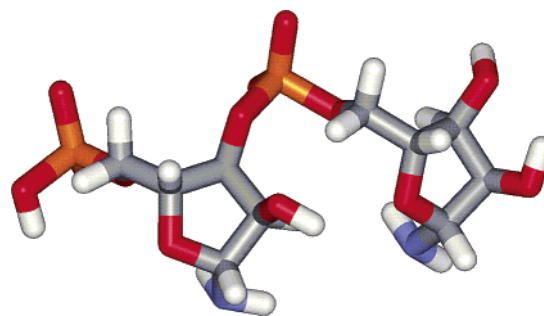


Figure 5. Geometry of the sugar–phosphate dimer fragment used in the modeling.

of the sugar–phosphate fragments than the internal coordinate method used previously for DNA, perhaps also accounting for some relaxation caused by the solvent. We therefore decided to avoid the computational problems associated with the description of the solvent and omit explicit D_2O molecules in the following models.

3.5. Single-Stranded (rA)₈. The calculated absorption and VCD spectra of the single-stranded octamer are drawn in Figure 6, together with those observed experimentally for the single-stranded poly(rA) polymer. As the VCD shape is determined by the interaction of neighboring base pairs and sugar–phosphate residues, the normal modes listed in Table 3 were classified according to the base pair vibrations. The final assignment for the polymers and the comparison to experimental numbers are given in Table 4. We observe almost perfect correspondence between the calculated and experimental spectra in the region of the base vibrations in terms of signs, relative intensities, and peak positions. The wavenumber difference between the calculated and experimental absorption of the main adenine ring mode is 10 cm^{-1} , that is, $\sim 0.6\%$. The difference for the weaker C=N ring stretching mode is even smaller (1574 vs 1572 cm^{-1}). In terms of the band shape, the biggest VCD couplet ($1633(-)/1623(+)\text{ cm}^{-1}$) as well as the negative VCD

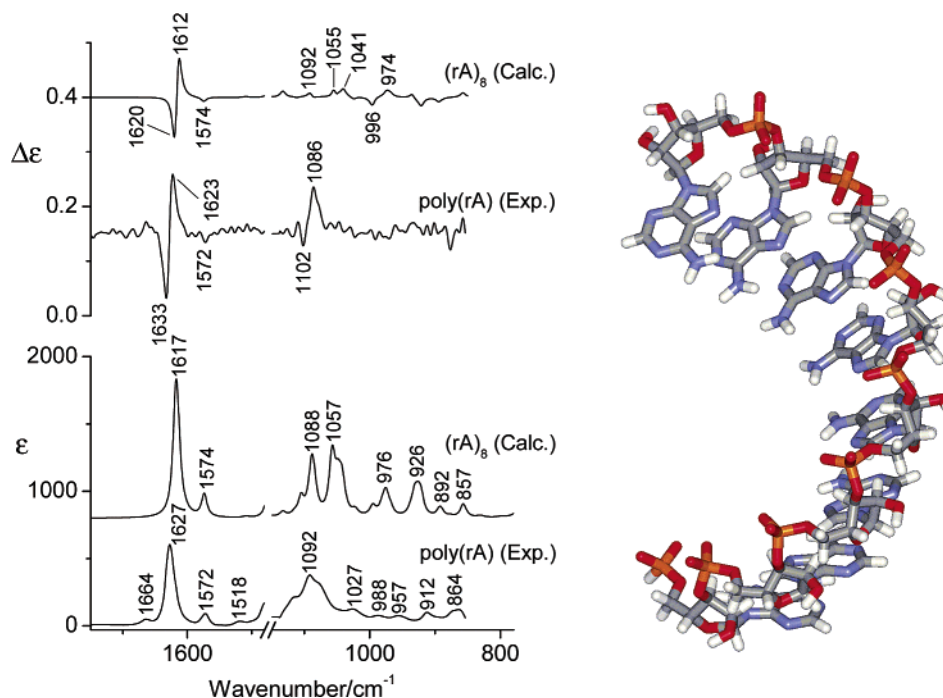


Figure 6. Calculated VCD ($\Delta\epsilon$) and absorption (ϵ) spectra of the single-stranded octamer $(rA)_8$ and the corresponding experimental VCD and absorption spectra of poly(rA). The geometry of the model octamer is displayed.

TABLE 3: Calculated Frequencies of the Main Vibrational Transitions in the RNA Chromophore Pairs

molecule	IR	VCD	assignment (BPW91/6-31G**)
2 stacked A bases	1617	1618(-)/1600(+)	$\nu(\text{C}=\text{N})$ ring + $\nu(\text{N}-\text{D}_2)$
	1601		$\nu(\text{C}=\text{N})$ ring + $\nu(\text{N}-\text{D}_2)$
	1579	1576(-)	$\nu(\text{C}=\text{N})$ ring + C-H bend (A_1)
	1578		$\nu(\text{C}=\text{N})$ ring + C-H bend (A_2)
	1515	1516(+)	$\nu(\text{C}=\text{N})$ ring + $\nu(\text{N}-\text{D}_2)$ stretch + C-H bend
	1477	1474(+)	C-H bend + C-H ₃
2 stacked U bases	1823		$\nu(\text{C}=\text{O}_2)$ (U_1)
	1820	1817(-)	$\nu(\text{C}=\text{O}_2)$ (U_2)
	1778	1778(-)/1758(+)	$\nu(\text{C}=\text{O}_4)$ (U_1+U_2)
	1759		$\nu(\text{C}=\text{O}_4)$ (U_1+U_2)
	1699		$\nu(\text{C}=\text{N})$ ring (U_1) + C-H bend (U_1) + $\nu(\text{C}=\text{O}_4)$ (U_1)
	1695		$\nu(\text{C}=\text{N})$ ring (U_2) + C-H bend (U_2) + $\nu(\text{C}=\text{O}_4)$ (U_2)
	1491		$\nu(\text{C}=\text{N})$ ring + C-H bend
	1489	1486(-)	$\nu(\text{C}=\text{N})$ ring + C-H bend
2 stacked AU base pairs	1741	1741 (+)	$\nu(\text{C}=\text{O}_2)$ (U_1)
	1733		$\nu(\text{C}=\text{O}_2)$ (U_2)
	1696	1696(-)/1672(+)	$\nu(\text{C}=\text{O}_4)$ (U_2)
	1672		$\nu(\text{C}=\text{O}_4)$ (U_1+U_2)
	1630	1625(-)/1614(+)	$\nu(\text{C}=\text{N})$ ring (all bases)
	1570		$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A)
	1522		$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A) + $\nu(\text{N}-\text{D}_2)$ stretch (A)
	1488	1483 (+)	$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A) + methyl group of A and U
S-P-S-P dimer for AU duplex	1285	1304(-)/1294(+)	$\nu(\text{PO}_2^-)$ asymmetric
	1062	1063(-)/1054(+)	(PO_2^-) symmetric
	981	981(-)	$\nu(\text{C}-\text{C})$ sugar ring
	918	918(-)	sugar ring deformations
	816	815(+)	sugar ring deformations

band at 1572 cm^{-1} and the corresponding absorption bands were clearly reproduced by the calculation. We attribute these good results also to the hydrophobicity of the adenine ring, that is, the vibrations are not influenced by the environment, as well as to the relative stability and rigidity of poly(rA) single strands. Finally, we can expect that our proposed geometry is quite close for that of the poly(rA) polymer in solution, at least as the juxtaposition of the bases is concerned.

However, correspondence of the model spectra to the experiment is worse in the sugar-phosphate region, where the band shapes match only approximately. This is similar to the DNA

study, where this region was left out completely. Nevertheless, most of the calculated absorptions bands could be assigned based on the absorption (Figure 6, Table 4). For example, the broad experimental band centered at 1092 cm^{-1} corresponds to two separate calculated maxima at 1088 and 1057 cm^{-1} , which are due to symmetric PO_2^- stretching and a sugar-based mode, respectively. For the VCD in this region, we attribute, as before,¹¹ the failure of the calculation to several reasons, namely, the neglect of solvation, the flexibility of the sugar-phosphate backbone, deviations from the idealized structure, and also the lower sensitivity of the spectrometer in this region suggesting

TABLE 4: Calculated and Experimental Frequencies of IR and VCD Transitions for RNA Polymers

molecule		IR		VCD		assignment	
model	exptl	calcd	exptl	calcd	exptl		
(rA) ₈	poly(rA)	1617	1627	1620(-)/1612(+)	1633(-)/1623(+)	$\nu(\text{C}=\text{N})$ ring + $\nu(\text{N}-\text{D}_2)$	
		1574	1572	1574(-)	1572(-)	$\nu(\text{C}=\text{N})$ ring	
			1518			$\nu(\text{C}=\text{N})$ ring + $\nu(\text{N}-\text{D}_2)$ + C-H bend	
			1105	1115			$\nu(\text{C}-\text{OD})$, sugar
			1088	1092	1092(+)	1102(-)/1086(+)	$\nu(\text{PO}_2^-)$ symmetric
			1057	1076	1055(+)		sugar backbone, this and below
			1049		1041(+)		
			1024	1027			
			995	988	996(-)		
			976	957	974(+)		
			926	912			
			892				
			857	864			
(rU) ₈	poly(rU)	1819	1693	1822(+)/1816(-)	1718(-)/1695(+)	$\nu(\text{C}=\text{O}_2)$ (U)	
		1784	1656	1788(-)/1778(+)	1674(-)/1649(+)	$\nu(\text{C}=\text{O}_4)$ (U)	
		1699	1618			$\nu(\text{C}=\text{N})$ ring + C-H bend + $\nu(\text{C}=\text{O}_4)$ (U)	
		1102	1116	1102(-)	1106(-)	$\nu(\text{C}-\text{OD})$, sugar	
		1073	1078	1080(-)/1070(+)	1091(-)/1072(+)	$\nu(\text{PO}_2^-)$ symmetric	
		1060					
		1047	1040	1052(-)/1043(+)	1033(-)/1023(+)	sugar backbone, this and below	
		1021					
		976	999	980(-)/973(+)			
		929	958				
		894	915				
		861	872				
		(rA) ₈ *(rU) ₈	poly(rA)* poly(rU)	1740	1689	1740(+)	1705(-)/1692(+)
1696	1669			1697(-)/1690(+)	1677(-)/1665(+)	$\nu(\text{C}=\text{O}_4)$ (U)	
1633	1631			1636(-)/1628(+)	1635(-)/1623(+)	$\nu(\text{C}=\text{N})$ ring (all bases)	
1564	1559			1564(+)		$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A)	
1523						$\nu(\text{C}=\text{N})$ ring (A) + C-H bend (A) + $\nu(\text{N}-\text{D}_2)$ (A)	
1106	1121			1107(-)/1092(+)	1124(-)/1114(+)	$\nu(\text{C}-\text{OD})$, sugar	
1084	1095			1085(-)/1080(+)		$\nu(\text{PO}_2^-)$ symmetric	
1066	1075			1073(-)/1064(+)	1100(-)/1087(+)	$\nu(\text{PO}_2^-)$ symmetric	
1053				1053(-)/1048(+)		sugar backbone, this and below	
1020	1021						
989	986				998(+)/979(-)		
977				974(+)/960(-)			
959	952						
924	913						
860	863						
(rU) ₈ *(rA) ₈ *(rU) ₈	poly(rU)* poly(rA)* poly(rU)	1750	1696	1753(-)/1748(+)	1704(-)/1689(+)	$\nu(\text{C}=\text{O}_2)$ (U _H)	
		1732	1673	1732(-)/1724(+)	1681(-)/1673(+)	$\nu(\text{C}=\text{O}_2)$ (U _{wc}) + $\nu(\text{C}=\text{O}_4)$ (U _H)	
		1702	1658	1705(-)/1695(+)	1658(-)/1650(+)	70% $\nu(\text{C}=\text{O}_4)$ (U _{wc}) + 30% $\nu(\text{C}=\text{O}_4)$ (U _H)	
		1646	1627	1649(-)/1642(+)	1631(-)/1615(+)	$\nu(\text{C}-\text{ND}_2)$ + $\nu(\text{C}=\text{N})$ (A) + $\nu(\text{C}=\text{N})$ (U)	
		1559	1573			$\nu(\text{C}=\text{N})$ + C-H bend (A)	
		1528				$\nu(\text{C}-\text{ND}_2)$ + $\nu(\text{C}=\text{N})$ + C-H bend (A)	
						$\nu(\text{C}-\text{OD})$, sugar	
		1133	1122			CH ₂ torsion, delocalized	
		1082	1095	1087(-)		$\nu(\text{PO}_2^-)$, symmetric + sugar	
		1063	1079	1061(+)	1098(-)/1083(+)	$\nu(\text{C}-\text{O})$ + OD bend, sugar	
		1052		1050(-)	1060(-)	sugar deformation	
		998	1025			O-D bend, sugar	
		944	956			$\nu(\text{C}-\text{CH}_2)$ + delocalized	
912	913						

that the experimental VCD band shape at 1102(-)/1086(+) cm⁻¹) may not be so reliable.

3.6. Single-Stranded (rU)₈. An analogous comparison as for (rA)₈ is presented in Figure 7, with the assignments in Tables 3 and 4. For this octamer, we pay a bigger price for neglecting the influence of solvent in the form of larger wavenumber differences between the calculated and experimental uracil carbonyl stretching modes. Nevertheless, all three simulated absorption bands at 1819, 1784, and 1699 cm⁻¹ can be well assigned to experimental absorptions at 1693, 1656, and 1618 cm⁻¹, respectively. More importantly, the VCD band shape characteristic for molecular conformation was also reproduced correctly. The strongest VCD couplet calculated at 1788(-)/

1778(+) cm⁻¹ arises from the C=O4 groups of uracil and corresponds to the experimental couplet at 1674(-)/1649(+) cm⁻¹. The assignment of the VCD signal from the C=O2 groups (calculated as a negatively biased couplet at 1822(+)/1816(-) cm⁻¹) is less obvious but probably corresponds to the experimental negative band at 1718 cm⁻¹.

Despite the less structured sugar-phosphate backbone of poly(rU) and, as a result, the less pronounced VCD spectrum in this region compared to poly(rA), both the band shapes and peak positions are fairly well reproduced. For poly(rU), the calculated VCD bands at 1102(-), 1070(+), and 1052(-) cm⁻¹ probably correspond to the experimental extremes at 1106, 1072, and 1033 cm⁻¹. The rest of the vibrations (with the experimental

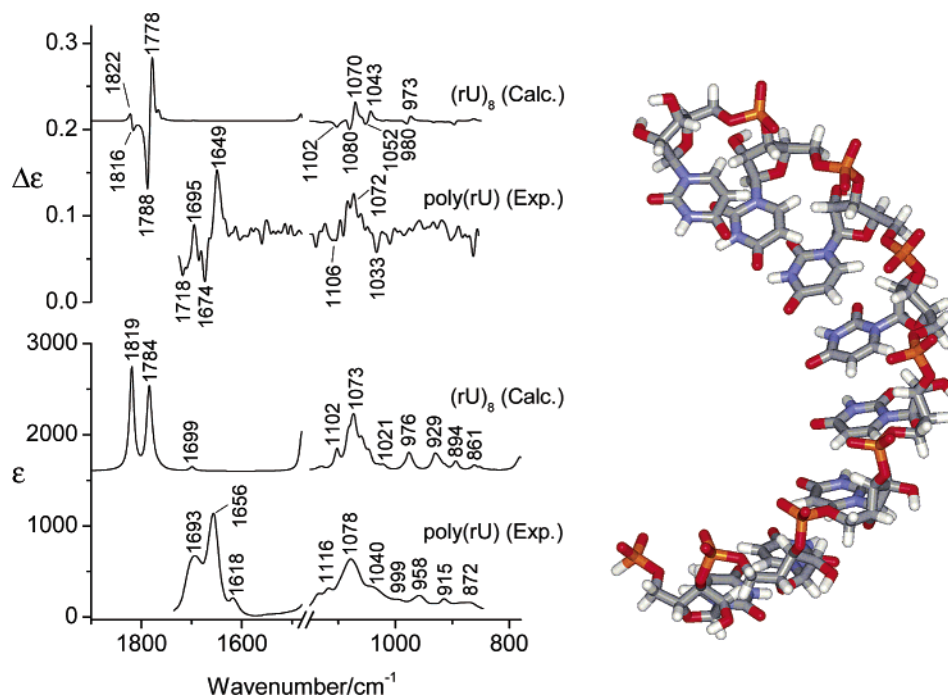


Figure 7. Calculated VCD ($\Delta\epsilon$) and absorption (ϵ) spectra of the single-stranded octamer $(rU)_8$ and the corresponding experimental VCD and absorption spectra of poly(rU). The geometry of the model octamer is displayed.

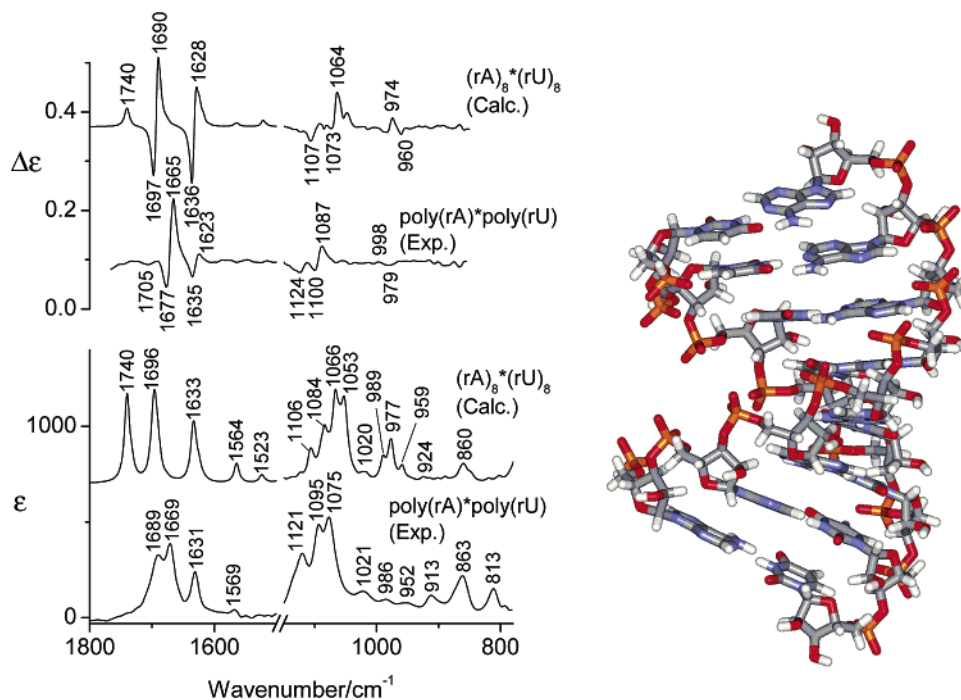


Figure 8. Calculated VCD ($\Delta\epsilon$) and absorption (ϵ) spectra of the double-stranded octamer $(rA)_8*(rU)_8$ and the corresponding experimental VCD and absorption spectra of poly(rA)*poly(rU). The geometry of the model octamer is displayed on the right.

IR bands at 999, 958, 915, and 872 cm^{-1}) can be assigned on the basis of the absorption spectra.

3.7. Double-Stranded $(rA)_8*(rU)_8$. Generally, the spectra of the helical model of double-stranded $(rA)_8*(rU)_8$ reproduce all the main features observed for poly(rA)*poly(rU) (Figure 8, Tables 3 and 4), including the sugar–phosphate signal. As before, the spectra are not single sums of the A and U components. In particular, the C=O stretching vibration is calculated much closer to experiment than for the sole poly(rU) polymer (Figure 7) because the hydrogen bonding between the bases is included in the calculation.

Because of the strong dipole–dipole interactions, conservative couplets dominate the VCD spectra as for poly(rA) and poly(rU). The sign of the highest wavenumber couplet calculated at 1740 cm^{-1} , however, is positive, contrary to the experimental couplet at $\sim 1705(-)/1692(+)$ cm^{-1} , which originates primarily from the C=O2 stretching vibration of uracil. Solvation, which influences this mode so much (Figure 2), may account partially for this difference. The absorption of the C=O4 group was calculated at ~ 1696 cm^{-1} , much closer to the observed position at ~ 1669 cm^{-1} , because H-bonding within the base pair is included in the model. The dipolar coupling of this mode gives

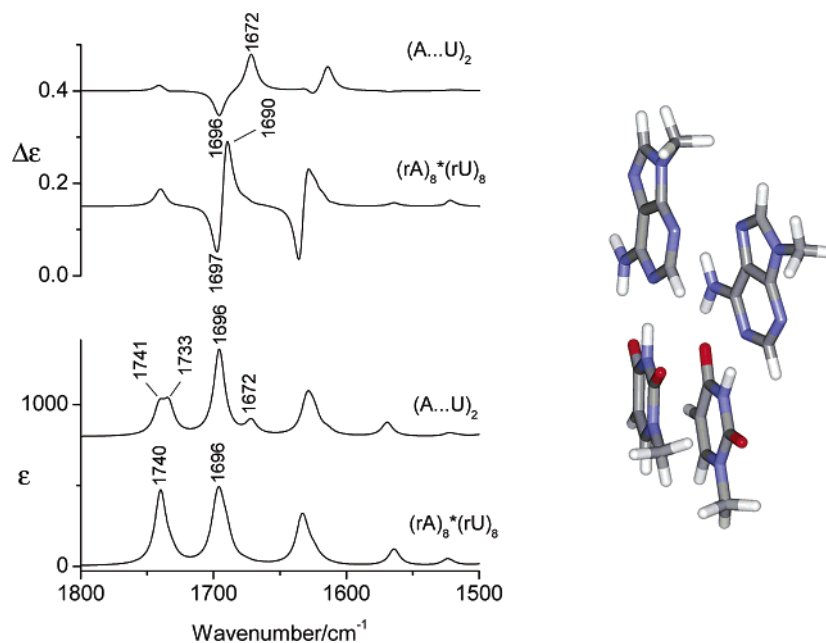


Figure 9. Calculated VCD ($\Delta\epsilon$) and absorption (ϵ) spectra of the base pair dimer $(A...U)_2$ and octamer $(rA)_8*(rU)_8$ and the geometry of the dimer.

rise to the positively biased couplet observed at $1677(-)/1665(+)$ cm^{-1} , which was perfectly reproduced by the calculation, albeit at higher wavenumber. This is true also for the second most prominent couplet observed at $1635(-)/1623(+)$ cm^{-1} and calculated at $1636(-)/1628(+)$ cm^{-1} . Here, the bias in the VCD intensities is negative and the corresponding vibrational mode is delocalized over both the adenine and uracil parts of the base pair. The calculation predicts also a weak positive VCD signal at 1564 cm^{-1} , which is not apparent in the experimental spectrum.

In the sugar–phosphate region, most of the calculated absorption bands can be assigned to the experiment (Table 4). The overall shape of the spectrum in this region is much better reproduced than for the single-stranded poly(rA) and poly(rU) or the $d(\text{GCGCGCGC})_2$ octamer.¹¹ More importantly, calculated and experimental profiles are in a reasonable agreement and consistent with the IR assignment and the structural geometry. Particularly, the calculation reproduces the weaker negative VCD signal at $1100\text{--}1124$ cm^{-1} as well as the stronger positive band at 1087 cm^{-1} ; even the weak experimental couplet at $998(+)/979(-)$ cm^{-1} is reproduced, although calculated with somewhat higher intensity at $974(+)/960(-)$ cm^{-1} .

Figure 9 illustrates the dependence of the calculated VCD and absorption spectra on the number of base pairs. Both the C=O2 and C=O4 stretching absorption bands of uracil are significantly split in the dimer ($1741\text{--}1733$ and $1696\text{--}1672$ cm^{-1} , respectively), whereas they produce apparent single bands at 1740 and 1696 cm^{-1} in the octamer. This is also reflected in the VCD spectra, where two bands with different signs at $1696(-)$ and $1672(+)$ cm^{-1} for the dimer move closer to each other and form a couplet at $1697(-)/1690(+)$ for the octamer, which is accompanied by a VCD intensity increase. The octamer's more conservative VCD signal also better corresponds to the experimental spectra of longer nucleotides. Such a decrease of the effective splitting of the absorption and VCD bands and emerging of VCD couplets for nucleic acids was first described and explained by the coupled-oscillator approach.¹⁰ Our high-level ab initio computations confirm this mechanism. In addition, they take into account the stronger-interaction correction,³³ without which the sign bias for the dimer could not be explained.

3.8. Triple-Stranded RNA. The triplex RNA is the most difficult system, because of its size and the uncertainty about its solution geometry. Still, most features in the spectra can be assigned (Figure 10, Table 4). For the geometry we chose the conventional Hoogsteen base pairing based on X-ray data (see Figure 4), since no credible model has been found for the reversed Hoogsteen pairing.

As for the double helix, the calculated absorption at 1750 cm^{-1} arises from the C=O2 vibration of the uracil bound to the adenine forming the Hoogsteen pair (U_H). The calculated wavenumber for this band is overestimated because of the lack of solvent but can be assigned clearly to the experimental absorption at 1696 cm^{-1} . Its VCD couplet $1753(-)/1748(+)$ cm^{-1} corresponds to the experimental couplet $1704(-)/1689(+)$ cm^{-1} . The stacking of the bases (similarly as the effect of solvation, Figure 4, Table 2) causes a larger normal mode coupling between the C=O2 and C=O4 vibrations, which prevents us from distinguishing reliably between the normal and reversed base pairing and makes the assignment of these modes in Table 4 rather speculative. Thus, absorptions at 1732 and 1702 cm^{-1} originate from coupled C=O2 and C=O4 vibrations of both the uracil bases bound to adenine via Hoogsteen and Watson–Crick pairing (U_H and U_{wc}). These carbonyl bands can probably be assigned to the experimental absorptions at 1673 and 1658 cm^{-1} , respectively. The calculated VCD couplets at $1732(-)/1724(+)$ cm^{-1} and $1705(-)/1695(+)$ cm^{-1} correspond to the experimental couplets at $1681(-)/1673(+)$ cm^{-1} and $1658(-)/1650(+)$ cm^{-1} , respectively. The lower wavenumber experimental absorptions at 1627 and 1573 cm^{-1} and the VCD couplet at $1631(-)/1615(+)$ cm^{-1} can be reliably assigned to ring modes of the bases, calculated at 1646 and 1559 cm^{-1} and VCD at $1649(-)/1642(+)$ cm^{-1} , respectively. As shown above, these modes do not change substantially upon solvation or switching from Hoogsteen to reversed Hoogsteen base pairing; that is, they are not influenced by hydrogen bonding. The calculated absorption spectrum in the sugar–phosphate region reproduces the experiment fairly well, whereas the correspondence of the calculated and experimental VCD spectra is not good. Nevertheless, the calculated main VCD couplet in this region at $1087(-)/1061(+)$ cm^{-1} has the same sign and a similar position as the experimental result at $1098(-)/1083(+)$ cm^{-1} ,

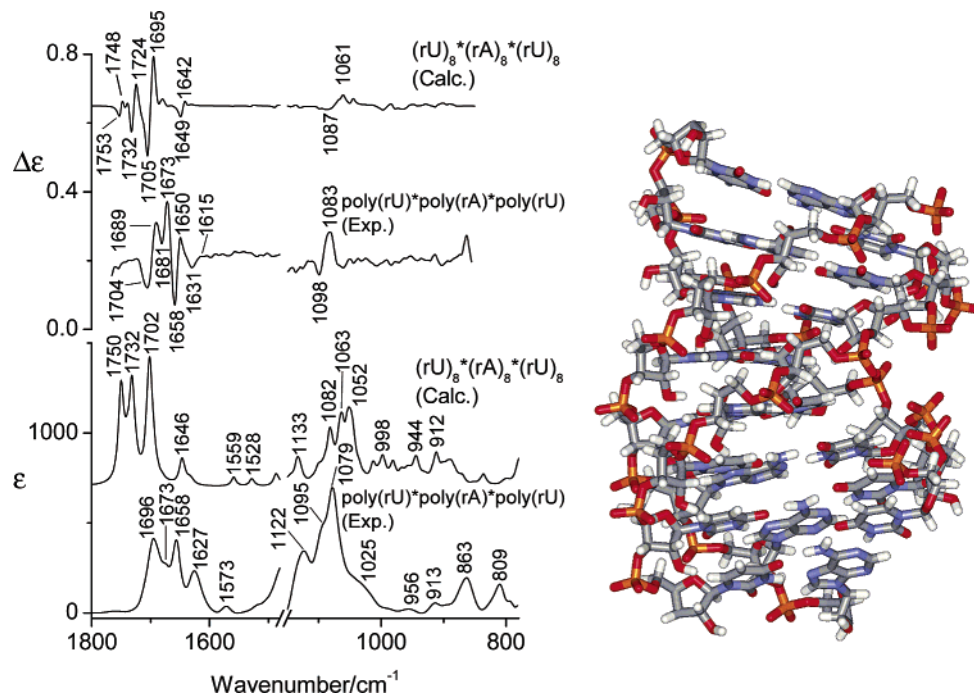


Figure 10. Calculated VCD ($\Delta\epsilon$) and absorption (ϵ) spectra of the triple-stranded octamer $(rU)_8^*(rA)_8^*(rU)_8$ and the corresponding experimental VCD and absorption spectra of $\text{poly}(rU)^*\text{poly}(rA)^*\text{poly}(rU)$. The geometry of the model octamer is displayed. Note that in the experiment an equal molar amount of single-stranded $\text{poly}(rA)$ is also present in the solution.

with rather lower intensity. Apparently, for future studies the precision of the spectral simulations will have to be further increased so that it can discriminate between finer changes of the triplex geometry.

4. Conclusions

Ab initio computations were used for the first time to interpret the absorption and circular dichroism spectra of various RNA forms. Although the computations could be done only on fragments, the overall error seems to be given by an inadequate description of the solvent and of the dynamical averaging rather than by the fragmentation. Nevertheless, most of the experimentally observed transitions could be assigned and the link between the IR and VCD spectral shapes and RNA structure, determined so far on an empirical basis, confirmed. The effects of solvation and the propeller twist on the IR spectra were investigated for the adenine and uracil bases, base pairs, and base triplets. The solvation (hydration) was found to be important mainly for the carbonyl groups not involved in the base pairing and readily accessible to H_2O (or D_2O). The ring vibrations of the bases appear to be the least affected. Surprisingly, the planar arrangement of the base pairs provided better solution spectra than the standard propeller twist derived from X-ray data. The best agreement of the calculated spectra with the experiment was obtained for the $(rA)_8$ single strand because of the weak influence of the hydration on the ring vibrations of the bases. Good agreement was found also for the $(rU)_8$ single strand and the $(rA)_8^*(rU)_8$ duplex for the base modes and the carbonyls involved in the base pairing, whereas the other $C=O_2$ carbonyl group provided $C=O$ stretching frequencies strongly dependent on the solvation model. For the triple-helical octamer $(rU)_8^*(rA)_8^*(rU)_8$ a fair agreement was obtained only for the absorption spectra because of the strong coupling of the $C=O_2$ and $C=O_4$ carbonyl vibrations of both uracil bases involved in the base triplets, as well as the lack of a reliable geometry model. Computation of the harmonic force field at the BPW91/6-31G** level was found quite efficient in terms

of accuracy and computational cost. Application of the normal mode method was superior for the partial geometry optimizations of the fragments compared to restricting the torsion angles and improved, namely, the computed spectra of the sugar-phosphate backbone. The fragment model we used is not dependent on a particular approximation and can be potentially improved in future.

Acknowledgment. This work was supported by the Grant Agency of the Academy of Sciences (A4055104) and the Grant Agency of the Czech Republic (203/01/0031) (P. Bour), a Grant from the Natural Sciences and Engineering Research Council of Canada (H. Wieser), and a Postdoctoral Fellowship from the Alberta Heritage Foundation for Medical Research (V. Andrushchenko). We thank D. Tsankov for providing absorption and VCD spectra of single-stranded $\text{poly}(rA)$ and $\text{poly}(rU)$.

References and Notes

- (1) Annamalai, A.; Keiderling, T. A. *J. Am. Chem. Soc.* **1987**, *109*, 3125.
- (2) Keiderling, T. A. In *Circular Dichroism and the Conformational Analysis of Biomolecules*; Fasman, G. D., Ed.; Plenum Press: New York, 1996; pp 555–597.
- (3) Gulotta, M.; Goss, D. J.; Diem, M. *Biopolymers* **1989**, *28*, 2047.
- (4) Zhong, W.; Gulotta, M.; Goss, D. J.; Diem, M. *Biochemistry* **1990**, *29*, 7485.
- (5) Keiderling, T. A. In *Circular Dichroism: Principles and Applications*, 2nd ed.; Berova, N.; Nakanishi, K., Woody, R. W., Eds.; John Wiley and Sons: New York, 2000; pp 621–666.
- (6) Wang, L.; Pančoška, P.; Keiderling, T. A. *Biochemistry* **1994**, *33*, 8428.
- (7) Andrushchenko, V.; van de Sande, J. H.; Wieser, H.; Kornilova, S. V.; Blagoi, Y. P. *J. Biomol. Struct. Dyn.* **1999**, *17*, 545.
- (8) Yang, L.; Keiderling, T. A. *Biopolymers* **1993**, *33*, 315.
- (9) Andrushchenko, V.; Blagoi, Y.; van de Sande, J. H.; Wieser, H. J. *Biomol. Struct. Dyn.* **2002**, *19*, 889.
- (10) Birke, S. S.; Diem, M. *Biophys. J.* **1995**, *68*, 1045.
- (11) Andrushchenko, V.; Wieser, H.; Bouř, P. *J. Phys. Chem. B* **2002**, *106*, 12623.
- (12) Ohms, J.; Ackermann, T. *Biochemistry* **1990**, *29*, 5237.
- (13) Dagneaux, C.; Liquier, J.; Taillandier, E. *Biochemistry* **1995**, *34*, 16618.

- (14) Tsuboi, M.; Kyogoku, Y.; Shimanouchi, T. *Biochim. Biophys. Acta* **1962**, *55*, 1.
- (15) Chandrasekaran, R.; Giacometti, A.; Arnott, S. *J. Biomol. Struct. Dyn.* **2000**, *17*, 1023
- (16) Felsenfeld, G.; Davies, D. R.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023.
- (17) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.
- (18) Miles, H. T. *Proc. Natl. Acad. Sci. U.S.A.* **1964**, *51*, 1104.
- (19) Stephens, P. J. *J. Phys. Chem.* **1985**, *89*, 748.
- (20) Cheeseman, J. R.; Frisch, M. J.; Devlin, F. J.; Stephens, P. J. *Chem. Phys. Lett.* **1996**, *252*, 211.
- (21) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revisions A.3 and A.7; Gaussian, Inc.: Pittsburgh, PA, 1998.
- (22) Bouř, P.; Keiderling, T. A. *J. Am. Chem. Soc.* **1993**, *115*, 9602.
- (23) Silva, R. A. G. D.; Kubelka, J.; Bouř, P.; Decatur, S. M.; Keiderling, T. A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*(15), 8318.
- (24) Bouř, P.; Kubelka, J.; Keiderling, T. A. *Biopolymers* **2000**, *53*, 380.
- (25) Tsankov, D.; Eggimann, T.; Wieser, H. *Appl. Spectrosc.* **1995**, *49*, 132.
- (26) Bouř, P.; Sopková, J.; Bednářová, L.; Maloň, P.; Keiderling, T. A. *J. Comput. Chem.* **1997**, *18*, 646.
- (27) *Insight II*; Accelrys Inc.: San Diego, CA, 1995.
- (28) Maharaj, V. Ph.D. Thesis, University of Calgary, 1996.
- (29) Bouř, P.; Keiderling, T. A. *J. Chem. Phys.* **2002**, *117*, 4126.
- (30) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
- (31) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.
- (32) Bouř, P.; McCann, J.; Wieser, H. *J. Phys. Chem. A*, **1998**, *102*, 102.
- (33) Bouř, P.; Keiderling, T. A. *J. Am. Chem. Soc.* **1992**, *114*, 9100.