5670–5678 Nucleic Acids Research, 1993, Vol. 21, No. 24 

# High-resolution NMR study of a synthetic oligoribonucleotide with a tetranucleotide GAGA loop that is a substrate for the cytotoxic protein, ricin

Masaya Orita[1], Fumiko Nishikawa, Takashi Shimayama[2], Kazunari Taira, Yaeta Endo[3] and Satoshi Nishikawa*

National Institute of Bioscience and Human Technology, Ministry of International Trade and Industry, 1-1 Higashi, Tsukuba, Ibaraki 305, [1]Yamanouchi Pharmaceutical Co. Ltd, 21 Miyukigaoka, Tsukuba, Ibaraki 305, [2]Hitachi Chemical Co. Ltd, Wadai, Tsukuba, Ibaraki 305 and [3]Department of Applied Chemistry, Ehime University, Matsuyama, Ehime 790, Japan

Received August 26, 1993; Revised and Accepted October 8, 1993

## ABSTRACT

**Ricin is a cytotoxic protein that inactivates ribosomes by hydrolyzing the N-glycosidic bond at position A4324 in eukaryotic 28S rRNA. Its substrate domain forms a double helical stem and a 17-base loop that includes the sequence GAGA, the second adenosine of which corresponds to A4324. Recently, studies of mutant RNAs have shown that the four-nucleotide loop, GAGA, can function as a substrate for ricin. To investigate the structure that is recognized by ricin, we studied the properties of a short synthetic substrate, the dodecaribonucleotide r-CUCAGAGAUGAG, which forms a RNA hairpin structure with a GAGA loop and a stem of four base pairs. The results of NMR spectroscopy allowed us to construct the solution structure of this oligonucleotide by restrained molecular-dynamic calculations. We found that the stem region exists as an A-form duplex. 5G and 8A in the loop region form an unusual G:A base pair, and the phosphodiester backbone has a turn between 5G and 6A. This turn seems to help ricin to gain access to 6A which is the only site of depurination in the entire structure. The overall structure of the GAGA loop is similar to those of the GAAA and GCAA loops that have been described but that are not recognized by ricin. Therefore, in addition to the adenosine at the depurination site, the neighboring guanosine on the 3′ side (7G) may also play a role in the recognition mechanism together with 5G and 8A.**

## INTRODUCTION

Ricin is a cytotoxic heterodimeric glycoprotein with an A-chain of 267 amino acids linked by a disulfide bond to a B-chain of 262 residues (1). This toxin is isolated from the seeds *of Ricinus communis* and inactivates mammalian ribosomes. The B-chain is a lectin that recognizes galactose-containing receptors on the surface of sensitive eukaryotic cells (2). Binding to the surface of eukaryotic cells results in the reduction of a disulfide bond, and the A-chain is translocated to the cytosol (3). The A-chain catalyzes the hydrolysis of a single N-glycosidic bond of the adenosine at position 4324 among a total of 7,000 nucleotides in 28S rRNA (4, 5). Depurination of A4324 then causes inactivation of the ribosome, apparently by disrupting the structure of the binding site for elongation factors (6, 7). To define the nature of the interaction of ricin with ribosomes, many approaches have been undertaken from two perspectives. The amino acid side-chains of the A-chain (8, 9) and the nucleotide substrate in 28S rRNA have both been examined (10, 11). However the reaction mechanisms remain obscure. The sequence of the ricin-substrate domain in 28S rRNA is highly conserved in many eukaryotes and forms a double-helical stem and 17-base loop with the sequence GAGA at its center, the second adenosine of which corresponds to A4324 (10). The sequence of the loop is UCAGUACGAGAGGAACC. The A-chain also catalyzes depurination of A1014 in naked prokaryotic 16S rRNA that contains a GAGA tetranucleotide loop at the analogous site (5). Recent studies of N-glycosidase activity with variant RNAs have clarified that the GAGA loop with a helical stem is essential for a ricin substrate (11), as follows. 1) With respect to the loop region, the tetranucleotide sequence GAGA is so critical that a GAGA-containing hexanucleotide loop is not recognized. 2) With respect to the helical stem, canonical Watson–Crick base pairs are essential and the recognition by ricin does not depend on the sequence. 3). With respect to the stability of the helical stem, more than two base pairs are needed.

Recently, phylogenetic comparisons of sequences in attempts to define the secondary structure of the large and small rRNAs have shown that tetranucleotide loops (tetraloops) account for

* To whom correspondence should be addressed

more than 55% of all loops, and nearly 70% of all tetraloops are either GNRA or UNCG (where N can be any nucleotide and R can be either G or A) (12). These tetraloop hairpins occur frequently in many other RNAs, for example, the catalytic RNA of RNase P (13) and the self-splicing RNAs (14, 15) and they display unusual thermodynamic stability (16). To identify the biological function of tetraloop RNA hairpins, several GNRA (17) and UNCG (18–20) tetraloop hairpins have been investigated by nuclear magnetic resonance (NMR) spectroscopy. Heus and Pardi reported the three-dimensional structures of the GAAA and GCAA loops that are the most frequently occurring members of the GNRA family (17). The results of NMR studies of GNRA loops indicate an unusual G:A base pair and two hydrogen bonds ($NH_2$ of G with the 5′-phosphate of A; 2′-OH of G with N7 of R). These hydrogen bonds seem to contribute to the thermostability of the tetraloop. The GAGA tetraloop, which is a member of the GNRA family, is expected to have a similar structure, but we recall here that ricin depurinates only one adenosine in a GAGA tetraloop.

To elucidate the structural basis for recognition by ricin, we synthesized a short substrate dodecaribonucleotide r-CUCAG-AGAUGAG (Figure 1), which has a potential sequence to form a hairpin structure with a stem of four base pairs and a GAGA loop, by the solid-phase phosphoramidite method, using a *tert*-butyldimethylsilyl (TBDMS) group to protect the 2′-hydroxyl group. In this report we present the results of one- and two-dimensional NMR study of this oligonucleotide, and we propose a solution structure for r-CUCAGAGAUGAG that is based on inter-proton distances and scalar coupling constants determined by NMR.

## MATERIALS AND METHODS

### Synthesis and purification of the dodecaribonucleotide

The oligonucleotide was synthesized by the phosphoramidite method on an automated DNA synthesizer (model 380B, Applied Biosystems) using a standard 1.0 μmol DNA assembly cycle with the exception that the coupling interval was increased to 10 minutes. 2′-O-TBDMS–5′-dimethoxytrityl ribonucleoside 3′-O-phosphoramidites (ABN) were dissolved at 0.1 M in anhydrous acetonitrile (Merck) and assemblies were carried out in the 'trityl off' mode.

After cleavage from the support by a mixture of concentrated ammonia and ethanol (3:1, v/v), the solution was heated at 55 °C for 8 hr to remove base-protecting groups and evaporated under reduced pressure to dryness. 2′-protecting silyl groups were removed by treatment with 0.1 M tetrabutylammonium fluoride (Aldrich) for 12 hr. The reaction mixture was quenched with 0.1 M triethylamine acetate (TEAA) buffer (pH 7.0), evaporated and desalted through an Oligo-Pak™EX (Millipore) cartridge. The crude products were purified by reversed-phase HPLC on a Pep PRC 15 μm HR10/1.0 column (Pharmacia). Elution was performed with linear gradient of $CH_3CN$ (5%–11%) in 0.1 M TEAA buffer (pH 7.0). The purified oligoribonucleotide was desalted on a column of Sephadex G-25 (Fast Desalting Column, Pharmacia). After evaporation, all counterions were replaced with sodium by successive treatments on small columns of Dowex 50W-X2 (pyridine form), Dowex 50W-X2 (sodium form), and Chelex-100 (sodium form) resins. The fractions were collected and dried by lyophilization.

### Thermodynamic analysis

UV spectra were recorded on a model UV-2100 spectrophotometer (Shimadzu). The buffer for thermodynamic studies was 0.1 M NaCl, 10 mM phosphate buffer (pH 7.0) and was the same as the buffer used for NMR. The samples were heated to above 90 °C for a few minutes and then quickly chilled on ice to begin the experiment. Curves of absorbance versus temperature were recorded at 260 nm with a heating rate of 1.0°C $min^{-1}$ and nucleotide concentrations varied from 2.54 μM to 71.5 μM.

### NMR spectroscopy

NMR spectra were recorded with an ALPHA-500 spectrometer (500 MHz for $^1H$, 202 MHz for $^{32}P$ and 125 MHz for $^{13}C$; JEOL) and an EX-400 spectrometer (400 MHz for $^1H$; JEOL). Exchangeable proton spectra and other spectra were obtained with the EX-400 spectrometer and the ALPHA-500 spectrometer, respectively. The $^1H$ chemical shifts were determined relative to internal 2-methyl-2-propanol (1.23 ppm). The $^{31}P$ chemical shifts were referenced to an external trimethyl phosphate (10% in ethanol). The $^{13}C$ chemical shifts were calibrated relative to external 3(trimethylsilyl)propionic-2,2,3,3-$d_4$ acid, sodium salt, in $D_2O$. Imino proton spectra were collected in $H_2O$-$D_2O$ (4:1, v/v), that contained 0.1 M NaCl and 10 mM phosphate buffer (pH 7.0), and non-exchangeable proton spectra were collected in 99.98% $D_2O$ in the same buffer after H-D exchange. The final concentration of the oligoribonucleotide was 5 mM.

Exchangeable proton spectra were obtained with a 1–1 pulse sequence for $H_2O$ signal suppression. NOE experiments were recorded at 5 °C to reduce the rate of the exchange with the water protons.

All 2D NMR spectra were recorded in the phase-sensitive mode by the TPPI method at 15°C. NOESY spectra were recorded with 1,024 points in $t_2$ and 256 points in $t_1$ at different mixing times (60, 90, 120, and 250 ms). The sweep width was 5,000 Hz in both dimensions, and 64 scans were acquired per $t_1$, value. The data sets were zero-filled to 512 data points along the $t_1$, axis and multiplied by Lorentz–Gaussian or exponential window functions before Fourier transformation. Noise ridges in the $t_1$, dimension were attenuated by multiplying the first row by one-half prior to transformation. The DQF-COSY spectrum was obtained from data sets of the same size. HOHAHA spectra were obtained from data sets of the same size with an MLEV-17 spin locking pulse sequence (90, 120 and 150 ms mixing times). A natural-abundance $^{13}C$–$^1H$ HSQC spectrum was collected using the standard pulse sequence. $^{13}C$ decoupling was achieved by a GARP decoupling sequence. The spectral widths were 24,000 Hz in the $^{13}C$ dimension ($t_1$,) and 4,000 Hz in the $^1H$ dimension ($t_2$). The data sets were recorded with 512 points in $t_2$, and 128 points in $t_1$. A total of 96 scans were averaged for each FID. Digital resolution in $t_2$ was 7.81 Hz/point and in $t_1$, it was 47.30 Hz/point after zero-filling to 512 points. The experiment was optimized for a 135 Hz $^1H$–$^{13}C$ coupling. A proton-detected $^{31}P$–$^1H$ heteronuclear COSY experiment (21, 22) was performed in the inverse mode. The spectral widths were 1,000 Hz in the $^{31}P$ dimension ($t_1$,) and 1,500 Hz in the $^1H$ dimension ($t_2$). The data sets were recorded with 1,024 points in $t_2$ and 128 points in $t_1$. 128 $t_1$, FID's of 96 acquisitions each were collected. Digital resolution in $t_2$, was 1.46 Hz/point and in $t_1$, it was 1.95 Hz/point after zero-filling to 512 points.

### Determination of structure

A three-dimensional structure, based on NMR data, was created with NMRGraf (Version 3.0, Molecular Simulation ) on Silicon Graphics IRIS 4D/35TG. Cross-peak volumes were measured using the integration routine in the NMR2 Model (Version 3.98, NMRI) on Sun 4/370. Inter-proton distances between nonexchangeable protons were estimated from the buildup rates of cross-peak volumes in short mixing time NOESY's (60, 90, 120, 250 msec). Pyrimidine H5−H6 distances were used as a reference (2.41 Å). Then distances between nonexchangeable protons were classified as short (1.8−3 Å), medium (2−4 Å), and long (2.5−5 Å). The total number of NOE-derived distances was 179 (106 intranucleotide interproton distances and 73 internucleotide interproton distances).

Backbone torsion angles were derived from coupling constants and NOES. Sugar pucker and $\delta(C_{4'}\text{-}C_{3'})$ were determined from the Hl′ and H2′ coupling constants. $\alpha(P\text{-}O_{5'})$ and $\zeta(O_{3'}\text{-}P)$ values were estimated from phosphorus chemical shifts *(gauche* or *trans)* (23). $\beta(O_{5'}\text{-}C_{5'})$ values were obtained from P-H5′ and P-H5' coupling constants. $\gamma C_{5'}\text{-}C_{4'})$ values were estimated from H8/H6/H5-H5′/H5″ NOE connectivities. $\epsilon(C_{3'}\text{-}O_{3'})$ Values were derived from H3′-P coupling constants. However, several ranges of angles had to be considered because any value for a coupling constant corresponds to several angles in the Karplus equation. The protocol used for refinement of the structure was as follows: 1) initial energy minimization to remove any highly strained components; 2) restrained molecular dynamics (RMD) at 600 °K during which the NOE force constant was gradually increased; 3) RMD to cool the system to 300 °K; 4) RMD and restrained energy minimization at 300°K; and 5) final energy minimization with no restraints.

At the start, a structure for the single-stranded r-CUCAGAG-

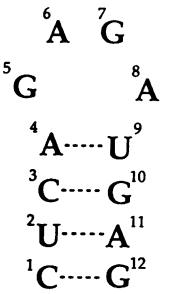


**Figure 1.** Secondary structure of r-CUCAGAGAUGAG.

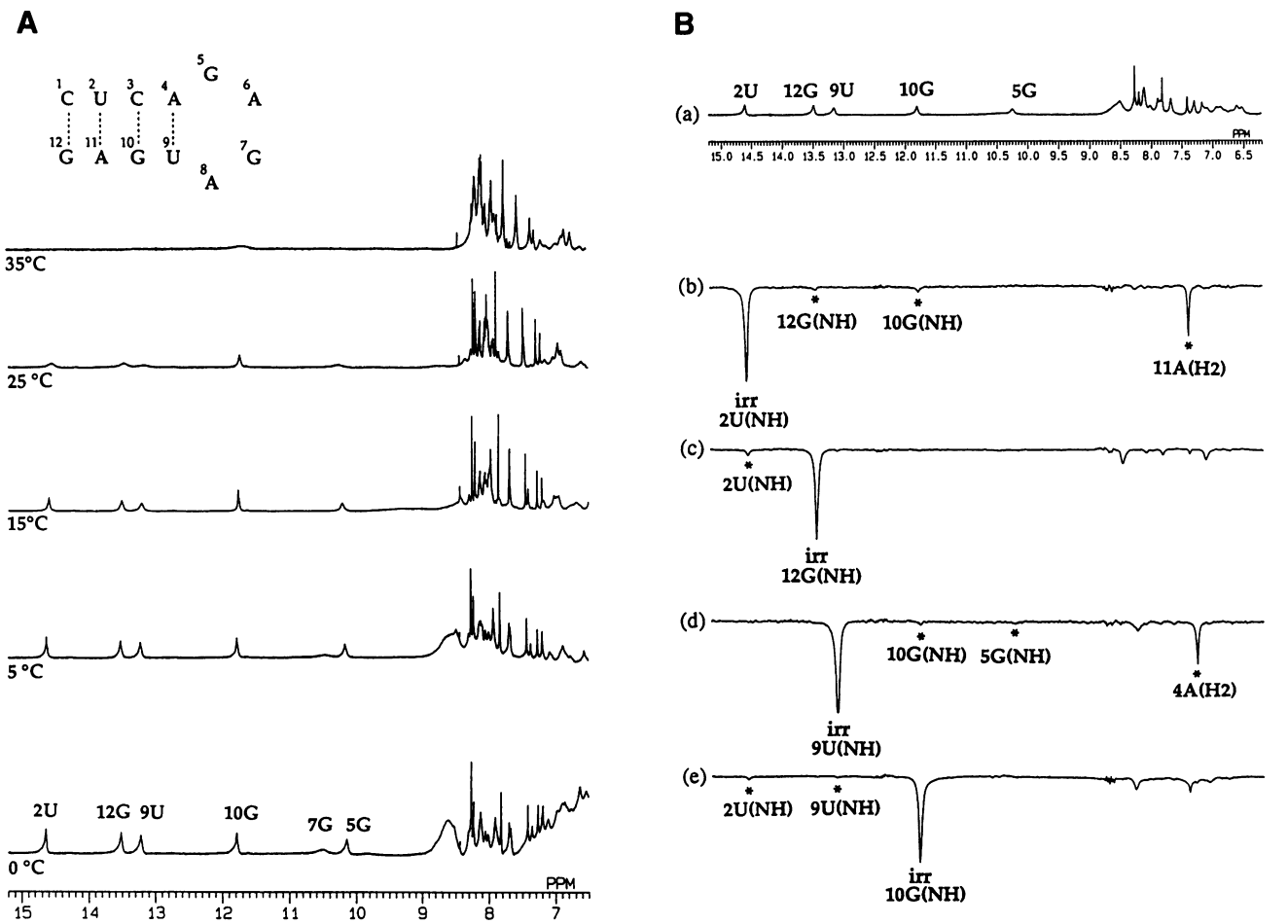


**Figure 2.** **(A)** $^1$H NMR spectra of r-CUCAGAGAUGAG in $H_2O$-$D_2O$ (4:1, v/v) that contained 0.1 M NaCl, 10 mM sodium phosphate buffer (pH 7.0) at 0, *5*, 15, 25, and 35 °C. **(B)** NOE experiments with r-CUCAGAGAUGAG in $H_2O$-$D_2O$(4:1, v/v) that contained 0.1 M NaCl, 10 mM sodium phosphate buffer (pH 7.0) at 5 °C. (a): Normal spectrum. (b)−(e): NOE difference spectra. Observed NOEs are indicated by asterisks.

AUGAG in which torsion angles were randomized was used. Because the stem region of the oligomer adopted an A-form duplex, we used, in addition to NOE-derived distance constraints, base-pairing distance constraints (H-N, H-O, N-N, N-O, N1-N9, and C1′-C1′ distance constraints), and dihedral angle constraints derived from canonical A-form helical parameters ($\alpha-\zeta$) for the IC:12G, 2U:11A, and 3C:IOG base pairs at the first annealing step. This procedure resulted in a more standard A-form appearance of the stem region, without any major effect on the structure of the loop and the loop-closing base pairs. This initial annealing was followed by refinement annealing during which the sugar pucker and the torsion angle restraints were introduced.

## RESULTS AND DISCUSSION

### Thermodynamic study

Recent studies have indicated that Watson−Crick and non-Watson−Crick base pairs contribute to the stability of RNA folding. In addition to a wobble type G:U base pair, a G:A base pair has also been identified in DNA and RNA by X-ray crystal-structure and NMR analysis (24, 25), and Santa Lucia *et al.* have shown that neighboring G:A mismatches are unusually stable (26). Since r-CUCAGAGAUGAG (Figure 1) has the possibility of forming a self-complementary duplex that contains a GAGA/GAGA internal loop, we first analyzed the dependence on concentration of the melting transition temperature (Tm) of the oligomer. The profile of UV absorption at 260 nm *versus* temperature at different concentrations (from 2.54 $\mu$M to 71.5 $\mu$M) revealed the same Tm value (50.0 °C) in each case and, furthermore, the Tm determined from the $^1$H NMR chemical shifts of base protons was also 50 °C at 5 mM. Since the oligomer exhibited the same Tm value at concentrations from 2.54 $\mu$M to 5 mM, we concluded that r-CUCAGAGAUGAG adopts a monomolecular hairpin stem-loop structure in solution, even at concentrations used for NMR.

### Exchangeable imino proton NMR spectra

Figure 2-A shows NMR spectra of r-CUCAGAGAUGAG in the low-field region. There are four hydrogen-bonded imino proton signals in the region between 11.5 ppm and 15.0 ppm, and two unpaired imino protons in the region between 10.0 and 11.0 ppm. The hydrogen-bonded imino protons were assigned by nuclear Overhauser effect (NOE) experiments, as shown in Figure 2-B. The two resonances at 14.60 and 13.17 ppm were easily identified as being due to A:U imino protons from their strong NOEs to adenine H2 resonances between 7.0 and 8.0 ppm, and NOEs between neighboring bases can reasonably be expected. Generally, a hydrogen-bonded uridine imino proton resonates in a lower field than a hydrogen-bonded guanosine imino proton. However, the 9U imino proton showed a specific upfield shift and resonated in a higher field than the 12G imino proton.

The signal at 10.65 ppm did not give any NOEs to the others signals but the 9U imino proton gave a weak NOE to the signal at 10.15 ppm (Figure 2-B d). Therefore, we assigned the signal at 10.15 ppm to the 5G imino proton and the signal at 10.65 ppm to the 7G imino proton. The temperature-dependence of the imino proton signals (Figure 2-A) revealed that the 7G imino proton signal was lost at 10 °C while the 5G imino proton signal was lost at a higher temperature at which signals for the stem region were lost. These results indicate that 7G NH exchanges rapidly with water protons and 5G NH exchanges relatively slowly. Because the 5G imino proton is thought to be buried more deeply inside the loop than the 7G imino proton, the rate of exchange of the 5G imino proton with a water proton should be lower than that of the 7G imino proton. In GAAA and GCAA loops, the imino proton of the first guanosine exchanges relatively slowly with water protons (17).

### Nonexchangeable proton NMR spectra

In order to obtain detailed information about the conformation of the oligomer, we assigned the nonexchangeable proton resonances by two-dimensional NOESY, DOF-COSY, HOHAHA, and heteronuclear $^{31}$P−$^1$H COSY spectroscopy. Assignments were performed as follows (27, 28). 1) Adenine H2 could be assigned from the NOE experiments in $H_2O/D_2O$ (Figure 2-A). H5 and H6 of pyrimidine bases were identified from the strong H5 to H6 cross-peaks in the DQF-COSY spectrum. 2) H8 of purine, H6 of pyrimidine, and sugar HI′ were assigned by the usual sequential-assignment method. 3) H2′ could he assigned by the strong NOEs to HI′ in the short-mixing-time NOESY (60 msec) and the cross-peaks in the DQF-COSY. We also confirmed the H2′ assignments by sequential NOES. 4) H3′ could be assigned from the heteronuclear $^{31}$P−$^1$H COSY spectrum and sequential NOEs. 5) The assignment of H4′ used the HI′-H4′ NOEs and H3′-H4′ cross-peaks in the DQF-COSY

**Table 1.** Chemical shifts of proton[a] and phosphorus[b] resonances at 15 °C

| | Imino | H8/6 | H2/5 | HI′ | H2′ | H3′ | H4′ | HS′/5″ | P |
|---|---|---|---|---|---|---|---|---|---|
| 1C | na[c] | 7.96 | 5.94 | 5.18 | 4.41 | 4.41 | 3.95 | 3.91/3.84 | na |
| 2U | 14.60 | 8.14 | 5.42 | 5.59 | 4.52 | 4.65 | 4.42 | 4.46/4.26 | −4.37 |
| 3C | na | 7.98 | 5.75 | 5.59 | 4.30 | 4.67 | 4.56 | 4.58/4.47 | −3.93 |
| 4A | na | 8.04 | 7.28 | 5.97 | 4.57 | 4.84 | 4.47 | | −3.77 |
| 5G | 10.15 | 6.93 | na | 5.57 | 4.14 | 4.48 | 4.27 | 4.48/4.29 | −2.63 |
| 6A | na | 8.22 | 7.85 | 5.67 | 4.72 | 4.32 | 4.18 | 4.27/4.19 | −1.17 |
| 7G | 10.65 | 7.40 | na | 5.12 | 4.18 | 4.42 | 4.26 | 4.33/4.18 | −2.27 |
| 8A | na | 7.46 | 8.26 | 6.13 | 4.38 | 4.80 | 4.20 | 4.28/4.28 | −4.72 |
| 9U | 13.17 | 8.08 | 5.93 | 3.92 | 4.52 | 4.17 | | 4.24/4.16 | −1.73 |
| l0G | 11.80 | 7.69 | na | 5.79 | 4.57 | 4.68 | 4.08 | | −3.70 |
| 11A | na | 7.69 | 7.44 | 5.79 | 4.46 | 4.57 | | 4.45/4.42 | −3.68 |
| 12G | 13.48 | 7.20 | na | 5.71 | 3.95 | 4.16 | | | −3.77 |

[a]Relative to *t*BuOD (1.23 ppm)
[b]Relative to TMP (0 ppm)
[c]not applicable

spectrum. 6) H5′ and H5″ were assigned from the heteronuclear $^{31}P-^{1}H$ COSY spectrum, but stereospecific assignment was not possible. The results of assignments are shown in Table 1. Figure 3 shows the expanded NOESY spectrum from the aromatic to the HI′ region. The H8/H6-HI′ cross-peaks can be traced sequentially along 1C−5G and 6A−12G but there are no cross-peaks between 5G HI′ and 6A H8. We assumed that, in the loop region, the bases of 6A, 7G, and 8A form a continuous stack while 5G is isolated from them. In other words, the phosphodiester backbone seems to have a turn between 5G and 6A.

In the stem region, H6 and H8s NOEs to 5′-neighboring H2′ were stronger than those to their own H2′ (data not shown). This result indicates that the stem region of the oligomer forms an A-form helix. The observation of the pyrimidine H5 cross-peaks with 5′-neighboring H2′ and H3′ also supports the proposed A-form duplex conformation forms (28).

9U H1′ shows an unusual upfield shift (more than I ppm from the normal position). The ring-current effects of aromatic rings are assumed to be the cause of this shift. We confirmed the assignment of 9U HI′ by a natural-abundance $^{13}C-^{1}H$ HSQC (Figure 4). As a rule, a $^{13}C$ nucleus does not show as much of a ring-current shift as a $^{1}H$ nucleus. Figure 4 shows that the $^{13}C$ that is directly bonded to the upfield-shifted proton resonates in the C1′ region. Thus, it is clear that 9U H1′ resonates at 3.92 ppm. In the $^{1}H$ NMR spectra at around the Tm (from 45°C to 55°C), we can also detect the signal that moves from the H2′-H5′ region to the H1′ region.

Because the stem region adopts a typical A-form duplex structure, the upfield shift of 9U H1′ is thought to come from the ring-current effect of 8A. In other words, the adenosine base of 8A is located directly over the C1′ proton on 9U. Thus, the adenine ring of 8A slides into the minor groove. NOE data revealed that 5G and 8A adopt an anti glycosidic conformation and that 5G is stacked on 4A and 8A is stacked on 9U (data not shown). These data suggest indirectly the presence of an unusual 5G:8A base pair with hydrogen bonds formed between the 5G amino proton and the 8A N7, and between the 8A amino proton and 5G N3 (Figure 5). The three-dimensional structure derived from NMR data also supports the existence of this G:A base pair (Figure 7).

## $^{31}$P-NMR studies

The backbone conformation of the oligomer was examined by $^{31}P$-NMR. The phosphorus signals were assigned from a heteronuclear $^{31}P-1H$ COSY spectrum (Figure 6). When an oligonucleotide forms a normal, right-handed double helix, which makes a smooth sugar-phosphate backbone, $^{31}P$ signals are observed in the region of −3.0 to −4.5 ppm. However, the 4Ap5G, 5Gp6A, 6Ap7G, and 8Ap9U phosphorus resonances were shifted downfield and the 7Gp8A phosphorus resonance was shifted upfield from this region. This dispersion of chemical shifts has been seen in other hairpin RNAs (20) and in Z-RNA (29). Phosphorus chemical shifts give information about torsion angles of $\zeta$-($O_{3'}$-P) and $\alpha$(P-$O_{5'}$) (23). It is well-known that the downfield shift is caused by a change in conformation about

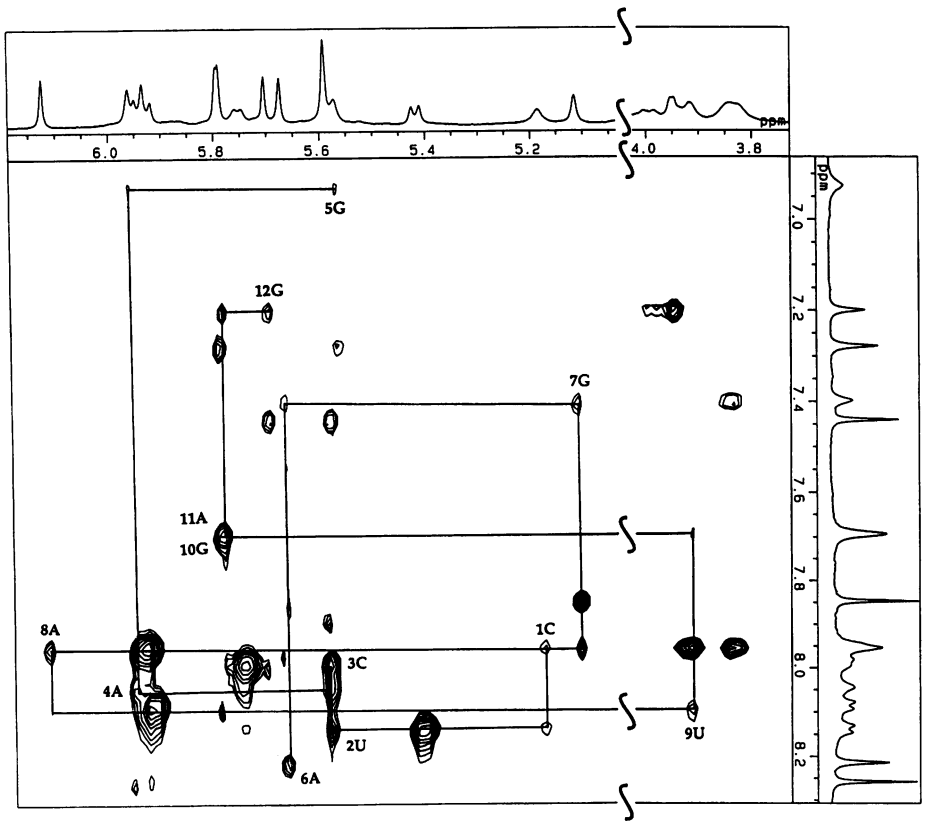


**Figure 3.** Expanded contour plots of the NOESY spectrum (250-ms mixing time) of r-CUCAGAGAUGAG in $D_2O$ that contained 0.1 M NaCl, 10 mM phosphate buffer (pH 7.0) at 15 °C. The sequential connectivities from 1C to 5G and from 6A to 12G through base proton-H1′ cross-peaks are shown by continuous lines. The intraresidue cross-peaks are labeled.

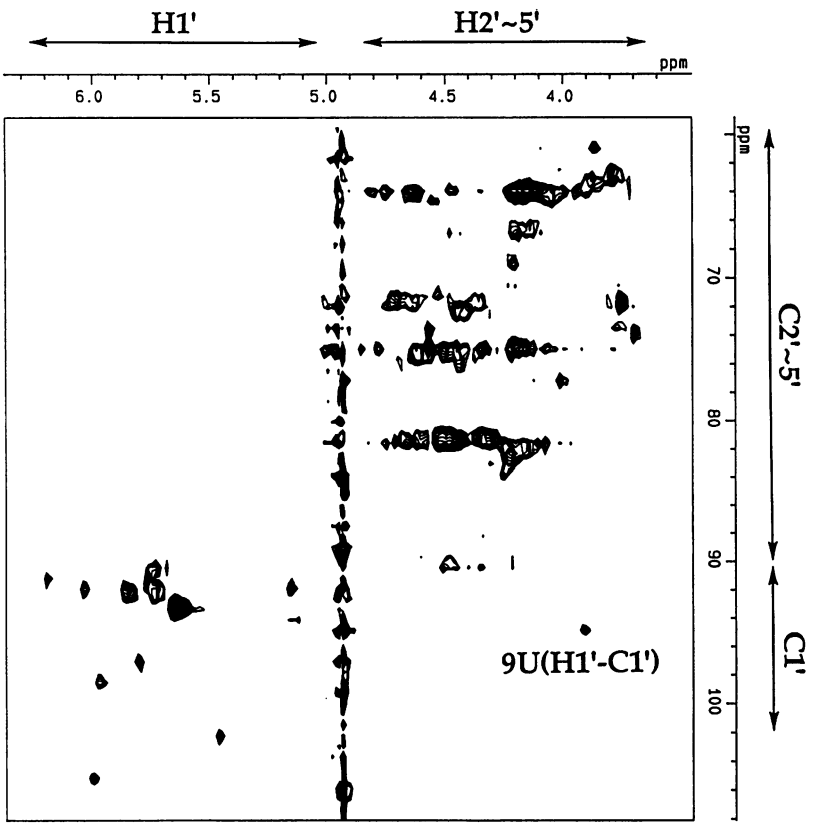


Figure 4. Natural-abundance $^{1}H$-$^{13}C$ HSQC spectrum of r-CUCAGAGAUGAG at 15 °C. T'he connectivity between 9U H1′ and 9U C1′ is indicated.

$O_{3'}$-P and P-$O_{5'}$, bonds, from normal *gauche-gauche* to *gauche-trans*. We assumed that the $\zeta$-$\alpha$ conformation of the most downfield-shifted 5Gp6A phosphorus (−1.17 ppm) was *gauche-trans or trans-gauche*. The downfield-shifted 5Gp6A phosphorus signal and the absence of sequential NOE between 5G HI′ and 6A H8 suggested that the phosphodiester backbone in the loop region has a turn between 5G and 6A. No assumptions were made about other shifted phosphorus signals. Several conformational effects might cause such shifts.

All P-H5′ and P-H5″ coupling constants, except for those of 9U, were less than 4 Hz or were not detected in the heteronuclear $^{31}P-^{1}H$ COSY spectrum. The conformation about $\beta(O_{5'},C_{5'})$ was interpreted as *trans*, as can be seen in a normal, right-handed double helix. The P-H5′ (or P-H5″) coupling constants of 9U were 16 ± 2Hz and 9 ± 2Hz, so the 9U $\beta$ torsion angle seemed to be *gauche*.

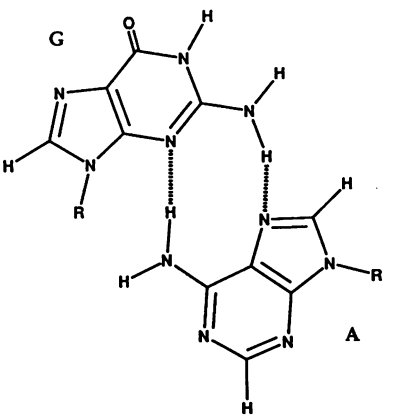


Figure 5. Schematic representation of the G:A base pair between 5G and 6A in the GAGA loop.

## The solution structure

The solution structure of r-CUCAGAGAUGAG was elucidated by restrained molecular-dynamic calculations. In the first step of the restrained molecular-dynamic calculations, as well as 179 NOE-derived distance constraints, we used base-pairing distance constraints (hydrogen bonds, NI-N9 and C1′-C1′ distance constraints) for the stem region and dihedral angle constraints, derived from canonical A-form helical parameters ($\alpha$-$\zeta$), for the IC:12G, 2U:11A, and 3C:10G base pairs since the stem region of r-CUCAGAGAUGAG seems to exist as an A-form duplex. This strategy resulted in a more standard A-form appearance of the stem region, yet there was little effect on the structure of the loop and the loop-closing base pairs. A total of 179 interproton distances (106 intranucleotide interproton distances and 73 internucleotide interproton distances) were estimated from the build-up rates of cross-peak volumes in short-mixing-time NOESY's (60, 90, 120, 250 msec) to reduce the effect of spin diffusion. Most internucleotide NOEs were for adjacent residues but there were also some NOEs between nonadjacent residues in the loop region. At the next step of the refinement, we introduced the sugar pucker and torsion angle restraints that were derived from coupling constants and NOES. Figure 7 shows the

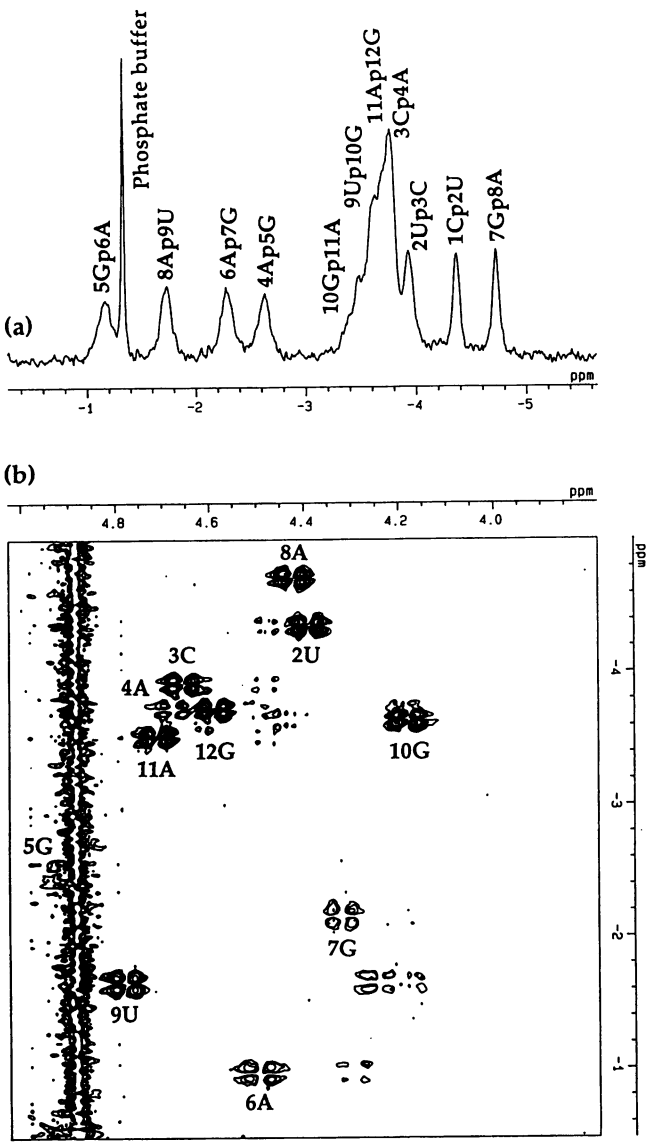


**Figure 6.** (a) Proton-decoupled phosphorus NMR spectrum of r-CUCAGAG-AUGAG at 15 °C. (b) Heteronuclear $^{1}H$-$^{31}P$ COSY spectrum of rCUCAGAG-AUGAG at 15 °C. Peaks corresponding to the internucleotide H3′($i$)−P($i$+1) coupling are labeled.

stereoview of r-CUCAGAGAUGAG. A schematic representation is shown in Figure 8.

*The stem region.* The stem region seems to exist as a canonical A-form duplex, as indicated by the strong NOEs observed between H2′ and the 3′-neighboring H6 or H8 protons, and the NOEs between the pyrimidine H5 and 5′-neighboring H2′ or H3′. All sugar puckers in the stem region, except for 1C, are of the N-type (C3′-endo), which is generally observed in an A-form duplex. The sugar pucker of 1C is a mixture of the N- and the S-types. A similar phenomenon is usually observed for the terminal base pair in a stem region. A helical stem with canonical Watson−Crick base pairs is essential for recognition of RNA by ricin. The activity is independent of the sequence (11). The stem region of r-CUCAGAGAUGAG meets this requirement.

*The loop region.* In the three-dimensional structure derived from our NMR data, 5G $NH_2$ is within hydrogen-bonding distance of 8A N7 (< 2 Å), and 8A $NH_2$ is also near 5G N3. 5G and 8A seem to form an unusual G:A base pair (Figure 5). The 5G imino proton in the G:A base pair seems to be exchanged relatively slowly with water protons. 9U H1′ exhibits an upfield shift of more than 1 ppm from the normal position because of the ring-current effect of 8A. The 9U imino proton is also associated with an upfield shift. The base-stacking pattern of the GAGA loop is shown in Figure 8. In our model of the GAGA loop, amino proton of 5G is located within 2 Å of the phosphate oxygen of 7Gp8A, and 7G N7 is also located near 5G 2′OH. In the NOESY spectrum, a weak 7G H8 NOE to 5G H2′ and the slight upfield shift of 5G H2′ are observed. These data suggest the hydrogen bonding between 5G $NH_2$ and the phosphate oxygen of 7Gp8A, and the hydrogen bonding between 2′-OH of 5G and N7 of 7G. The H1′-H2′ coupling constants indicate that the sugar puckers of 5G, 6A, and 8A are N-type (C3′-endo) while that of 7G is a mixture of the N- and the S-types. Saenger *et al.* suggested that a C2′-endo sugar pucker extends the backbone by about 2 Å and helps bridge the gap between opposite strands of a stem (30). In contrast to the GAGA loop, in GAAA and GCAA loops the sugars in the middle, at positions 2 and 3, have a mixture of N-and S-type puckers and sugars at positions 1 and 4 have N-type (C3′-endo) puckers (17). A difference in the loop sequence seems to cause a small conformational change but the geometry of GAGA, GAAA, and GCAA loops as a whole seems to be similar.

## CONCLUSION

In the GAAA and GCAA hairpin structures (17), Heus and Pardi reported the presence of a G:A mismatch and the hydrogen bonds that we observed in the GAGA hairpin. The overall structure of the GAGA loop seems to be similar to those of GAAA and GCAA loops (but in the GCAA loop the second C does not stack on the third A). However ricin specifically depurinates only the second adenosine in the GAGA tetraloop. As well as the adenosine at the depurination site, the guanosine at the third position in the GAGA loop also seems to participate in the recognition by ricin.

Monzingo and Robertus studied the structures of complexes of formycin monophosphate (FMP) and adenyl(3′−5′)guanosine (ApG) with the ricin A-chain by X-ray crystallography and they proposed three hypothetical models of the binding between the GAGA loop and the ricin A-chain (31). In addition to the hydrogen bonds between the adenosine at the depurination site and the ricin A-chain, some hydrogen bonds between the 3′-neighboring guanosine and the protein were also proposed in these models. These hydrogen bonds that involve the 3′-neighboring guanosine may be important for recognition by ricin.

Our data can be summarized as follows. 1) The G:A base pair in the loop is very important because it seems to help formation of a turn between 5A and 6A. 2) The G:A base pair seems to be less stable than a Watson−Crick base pair so, in order to form a G:A base pair, G:A must be close to a region of a stable base pairs. We suggest that this observation explains why GAGA-containing hexaloops are not recognized by ricin. 3) Because of the turn between 5G and 6A, 6A is flipped out of the loop and ricin is able to gain access to the depurination site 6A. 4) A GAAA loop, whose overall structure is similar to that of the

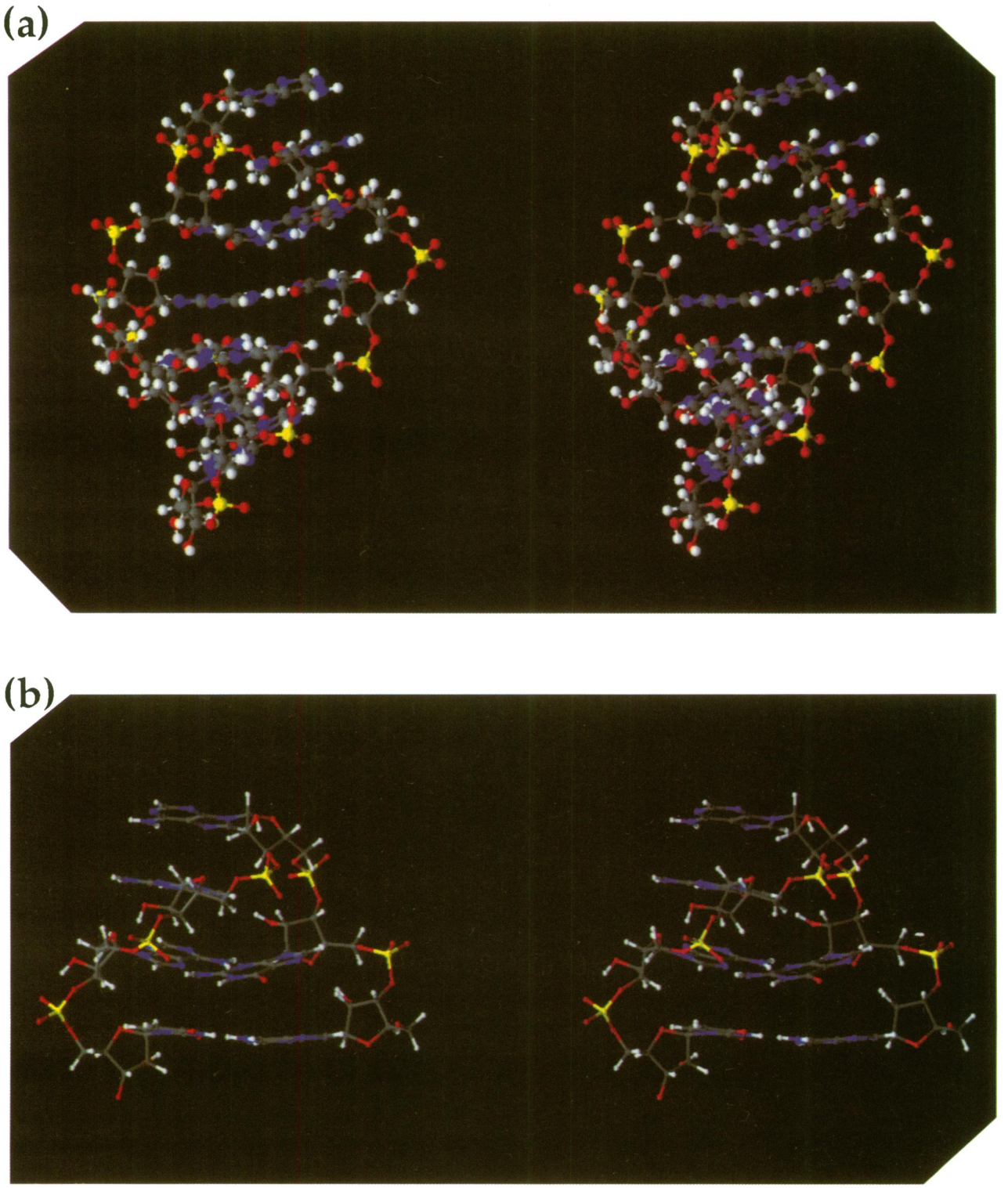


**Figure 7.** Stereoview of the structure of r-CUCAGAGAUGAG that was obtained from restrained molecular dynamics. **(a)**: Overall view. **(b)**: Loop region. In (b) the structure has been rotated by about 180° with respect to that in (a).

GAGA hairpin, is not recognized by ricin. Therefore, as well as the adenosine at the depurination site, the guanosine at the third position in the GAGA loop seems to be important for recognition by ricin.

The sequence of the ricin-substrate loop in 28S rRNA is UCAGUACGAGAGGAACC. This 17-member loop is assumed to form a GAGA-tetraloop conformation although the stem region contains mismatched base pairs (32). It is likely that, in 28S rRNA, ricin recognizes the alternate structure (GAGA tetraloop) as the sensitive form and does not recognize the 17-base loop.

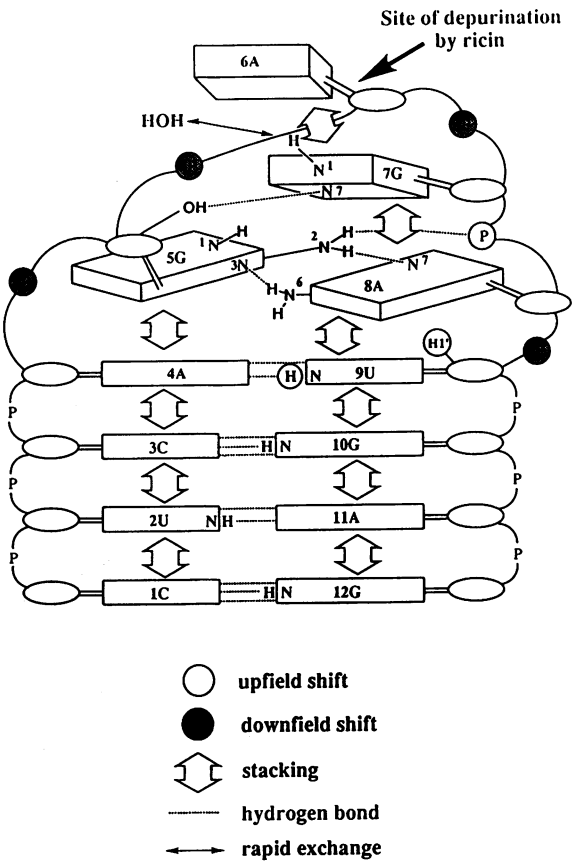


**Figure 8.** Schematic representation of the structure of r-CUCAGAGAUGAG. Open circles: upfield-shifted $^{1}H$ and $^{31}P$. Stippled circles: downfield-shifted $^{31}P$ Open arrows: stacking. Dotted lines: hydrogen bond. Arrow with two heads: rapid exchange with water protons.

## ACKNOWLEDGEMENTS

We thank Dr. Seichi Uesugi for a critical reading of the manuscript, Dr. Masato Katahira for helpful discussions, and Dr. Kenji Ogura for application of the technique of the heteronuclear COSY that we employed.

## REFERENCES

1. Olsnes,S. and Pihl,A. (1982) in *Molecular Action of Toxins and Viruses*, eds Cohen,P. and van Heyningen,S. Elsevier, New York, pp. 51–105.
2. Nicolson, G. L. and Blaustein, J. (1972) *Biochim. Biophys. 4ctct*, **266**, 543–547.
3. van Deurs, B., Tønnessen, T. I., Petersen, O. W., Sandvig, K. and Olsnes, S. (1986) *J. Cell Biol.* 102, 37–47.
4. Endo, Y. and Tsurugi, K. (1987) J. *Biol. Chem.* **262**, 8128–8130.
5. Endo, Y. and Tsurugi, K. (1988) J. *Biol. Chem.* **263**, 8735–8739.
6. Hausner, T. P., Atmadja, J. and Nierhaus, K. H. (1987) *Biochimie* 69, 911–923.
7. Moazed, D., Robertson, J. M. and Nolier, H. F. (1988) *Nature (London) 334*, 362–364.
8. Morris, K. N. and Wool, I. G. (1992) *Proc. Natl. Acad. Sci. USA* 489, 4869–4873.
9. Ready, M. P., Kim, Y. S., Robertus, J. D. (1991) *Proteins 10*, 270–178.
10. Endo, Y., Glück, A. and Wool, I. G. (1991).*J Mol. Biol.* 221, 193–207.
11. Glück, A., Endo, Y., Wool, 1. G. (1992).*J Mol. Biol.* 226, 411–424.
12. Uhlenbeck, O. C. (1990) *Nature (London)* 346, 613–614.
13. Pace, N. R. Smith, D. K., Olsen, G. J. and James, B. D. (1989) *Gene* 82, 65–75.
14. Waring, R. B. and Davis, R. W. (1984) *Gene* 28, 277–291.
15. Jaquier, A. and Michel, F. (1987) *Cell 50*, 17–29.
16. Antao, V. P., Lai, S. Y. and Tinoco, I., Jr (1991) *Nucleic.Acids Res.* 19, 5901–5905.
17. Heus, H. A. and Pardi, A. (1991) *Science* 253, 191–194.
19. Cheong, C., Varani, G. and Tinoco, I., Jr (1990) *Nature (London)* 346, 680–682.
19. Varani, G., Cheong, C. and Tinoco, I., Jr (1991) *Biochemistry 30*, 3280–3289.
20. Sakata, T., Hiroaki, H., Oda, Y., Tanaka, T., Ikehara, M. and Uesugi, S. (1990) *Nucleic Acids Res.* 18, 3831–3839.
21. Sklenar, V., Miyashiro, H., Zon, G., Miles, H. T., and Bax, A. (1986) *FEBS Lett.* 208, 94–98.
22. Sklenar, V., and Bax, A. (1987).*J. Am. Chem.* 109, 7525–7526.
23. Gorenstein, D. G. (1984) *Phosphorus–31 NMR: Principles and.Applications* (Gorenstein,D. G., ed.) Academic Press, New York.
24. Li, Y., Zon, G., and Wilson, W. D. (1991) *Proc. NcitL, cad. Sci.* U.,5.,4. 88, 26–30.
25. Lane, A., Martin, S. R., Ebel, S. and Brown, T. (1992) *Biochemistry* 31, 12087–1,–7095.
26. Santa Lucia, J., Jr., Kierzek, R., and Turner, D. H. (1990) *Biochemistry* 29 8813–8819.
27. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids.* John Wiley & Sons, New York.
28. Chou, S. H., Flynn, P. and Reid, B. (1989) *Biochemistry* 28, 2422–2435.
29. Gorenstein, D. G., Luxon, B. A., Goldfield, E. M., Lai, K., Negeais, D. (1982) *Biochemistry* 21, 580–589.
30. Saenger, W. (1984) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.
31. Monzingo, A. F. and Robertus, J. D. (1992).*J. Mol. Biol.* 227, 1136–1145.
32. Wimberly, B., Varani, G. and Tinoco, I., Jr (1993) *Biochemistry* 32, 1078–1087.