Proc. Natl. Acad. Sci. USA
Vol. 86, pp. 9727–9731, December 1989
Biochemistry

A computational approach to the mechanism of self-cleavage of hammerhead RNA

(CHARMm program/energy minimization/viroids/virusoids/satellite RNA)

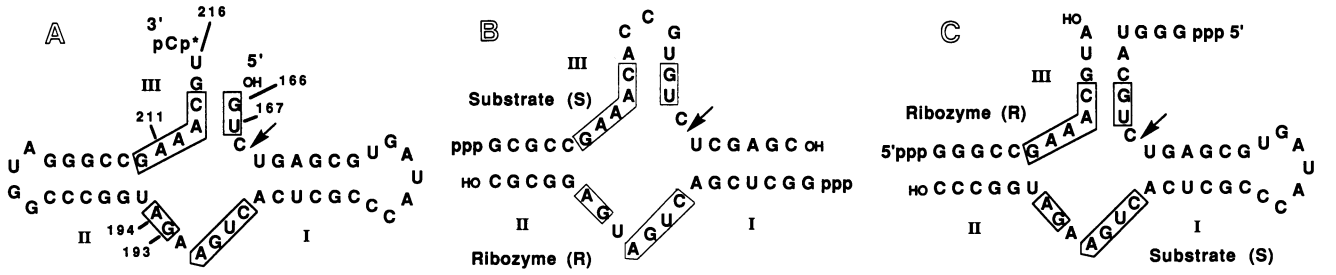
HOUNG-YAU MEI, THOMAS W. KAARET, AND THOMAS C. BRUICE

Department of Chemistry, University of California at Santa Barbara, Santa Barbara, CA 93106

Contributed by Thomas C. Bruice, September 14, 1989

ABSTRACT **Extensive minimization and dynamics computational studies of the hammerhead structural domain of the virusoid of lucerne transient streak virus have been carried out. The following observations at the self-cleavage position are derived from the resulting three-dimensional structure: (*i*) the cytosine base is at the surface and does not interact with another base (it is free to move), and (*ii*) the ribose-phosphate backbone is forced to take an abrupt turn since it bridges stems I and III, and this turn points the pro-S and pro-R oxygens of the phosphate to the inward side of the hammerhead. These structural features are independent of the hammerhead being open or closed and allow an unencumbered 3′- to 2′-endo conformational change of the ribose with the resultant creation of an unusual stereochemistry that allows a direct in-line self-cleavage reaction. In the closed hammerhead structure, interactions of stems I and II create a vacancy into which the catalytic hydrated Mg(II) may be docked on labile phosphate. This opening is not present if stems I and II are shortened.**

Certain single-stranded, circular RNAs undergo a site-specific self-cleavage reaction during replication. The self-cleavage reactions occur in a common structural domain known as a hammerhead (1, 2). The hammerhead structure is sufficient for self-cleavage in avocado sunblotch viroid and in virusoid of lucerne transient streak virus (structure 𝔸) and most likely is sufficient in newt satellite RNA (3). In the self-cleaving hammerhead structure there are 13 conserved nucleotides (boxed) and cleavage is always at the same phosphodiester linkage (arrow). Self-cleavage involves nucleophilic displacement by a 2′-hydroxyl upon the adjacent 3′-phosphate to generate a 2′,3′-cyclic phosphate ester with the release of a 5′-hydroxyl group. In this respect the self-cleavage reaction resembles that of pancreatic RNase A (4). Uhlenbeck (5), Symons (6), and coworkers have shown that the self-cleavage reaction can be achieved by catalytic RNA strands operating upon a complementary RNA strand as substrate (structures 𝔹 and ℂ). 𝔹 and ℂ represent the smallest known self-cleaving hammerhead structures.



To understand the mechanism of self-cleavage the three-dimensional structure of the hammerhead (as 𝔸) must be known. The use of x-ray and NMR techniques is plausible with stable hammerhead structures; however, these RNA molecules are in a noncleaving conformation (7). For those hammerhead molecules that do cleave, the cleavage rate is too great for the use of spectroscopic methods. In any event, with the exception of $tRNA^{Phe}$ (8) and $tRNA^{Asp}$ (9), crystallographic methods have not been applicable for the determination of RNA structure. Chemical modifications in the vicinity of the cleavage site may allow NMR structural determinations, but the catalytic structural tuning of the self-cleavage site may be lost. Herein we present a computational approach to the wild-type hammerhead structure 𝔸, using the molecular mechanics programs of CHARMm (10). Our goal is to determine whether the three-dimensional structure obtained by this means will allow the formulation of a rational mechanism of self-cleavage that might be used as a working hypothesis. We assure the reader that we are aware of the limitations of this computational approach to mechanism.

Nilsson and Karplus (11) have employed CHARMm in molecular dynamics simulations of the anticodon loop of yeast $tRNA^{Phe}$, and Doudna *et al.* (12) have employed CHARMm in a study of the docking of the nucleophilic guanosine in the *Tetrahymena* intron self-splicing reaction. Structural modeling of the three-dimensional structure of the catalytic center of *Tetrahymena* intron RNA has been reported by Kim and Cech (13).

MATERIALS AND METHODS

All computational analyses and graphics were carried out on a Silicon Graphics Iris 4D/70 workstation using CHARMm (versions 21.1.5 and 21.1.7) and QUANTA (version 2.4) programs (Polygen Corporation). Energy minimizations of RNA structures were performed in vacuum using "steepest descents" until energy changes for the entire structures were smaller than 10 kcal/mol per 50 steps. This was followed by application of the adopted basis Newton–Raphson (ABNR) minimization procedure until the energy-change tolerance 10^{-9} was satisfied. In general this required 40,000–75,000

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

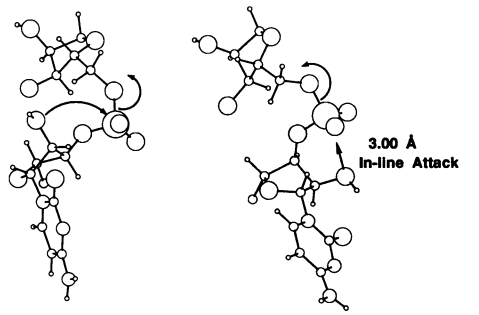
ABNR energy-minimization steps. The cutoff for nonbonded distance was 11.5 Å and the cutoff for hydrogen-bonding distance was 7.5 Å (cutoff angle, 90 degrees). During minimization both the nonbonded and hydrogen-bonded lists were updated at frequencies varied from every 1 step to every 400 steps. Updating was less frequent as the minimizations proceeded. A distance-dependent dielectric constant and the CHARMm shift functions were used in all minimizations. Molecular dynamics were performed using the atomic coordinates of the extensively minimized structures. The structures were heated to 300 K over 1.0 ps and velocities were assigned from a Gaussian distribution. Verlet integration was invoked with a time step of 0.001 ps. The nonbonded and hydrogen-bonded parameters were the same as used in minimization. Following the heating procedures, 60 ps of equilibration and production dynamics were performed. Velocities were scaled every 0.05 ps for temperatures out of the range 290–310 K. Vacuum dynamics trajectories were calculated and saved every 0.05 ps during the production dynamics procedure.

RESULTS AND DISCUSSION

We have *explicitly* assumed that the secondary structure for the self-cleaving hammerhead RNA is as proposed by Forster and Symons (2) and as supported by mutational studies (14, 15). Structure 𝔸 consists of three A-form RNA helices (stems I, II, and III) and an interior loop. Stems I (5′-UGAGC-3′) and II (5′-GGCCC-3′) were generated using the atomic coordinates of the 5-base-paired A helix of the T-arm from the x-ray crystal structure of yeast $tRNA^{Phe}$ (Brookhaven database code name 4TNA). Coordinates for the bases were deleted and regenerated according to the sequences of stems I and II of 𝔸. The coordinates for the new bases of stems I and II, the residues in stem III, and the residues in the single-stranded loop were constructed using the topology and parameter files in CHARMm. The same procedure was used for all the hydrogen atoms. Nuclear Overhauser effect (NOE) constraints with a scale factor of 50 were used to maintain the A-form RNA helices of stems I, II, and III during energy minimization. Examination of 𝔸 shows base-pairing of U^{195} and G^{210}. Similar base-pairing of U and G is seen in ℂ but not in 𝔹. Thus, this particular base-pairing (U·G) in 𝔸 and ℂ should not be required for self-cleavage. For this reason, no constraints were applied to base-pair U^{195} to G^{210}.

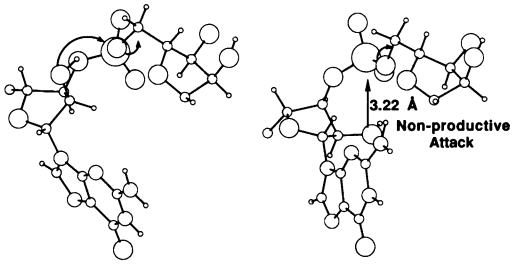
Twenty thousand steps of energy minimization provided the "open" hammerhead structure (Fig. 1 *Left*) and extensive minimization the "closed" hammerhead structure (Fig. 1 *Right*). For clarity, the structures of Fig. 1 are shown with only the first three nucleotides of stems I and II. The bases of the open and closed hammerheads are shown in Fig. 2 *Left* and *Right*, respectively. Examination of Fig. 1 shows that the phosphate linkage at the cleavage site (arrow) takes an abrupt turn since it bridges stems I and III. Fig. 2 shows that the cytosine base at the cleavage site does not interact with any other purine or pyrimidine base. *The freedom of movement of the base and the sharp change in the direction of the ribose-phosphate backbone differentiates the ribonucleotide at the cleavage site from all others in the structure. The conformation of this nucleotide is not perceptibly dependent upon whether the hammerhead structure is open or closed.*

The facility and mechanism [in-line or adjacent attack followed by pseudorotation (16)] of displacement of the 5′-hydroxyl leaving group by nucleophilic attack of a 2′-hydroxyl group upon the neighboring 3′-phosphate depend upon: (*i*) the conformation of the ribose ring (3′- or 2′-endo) and (*ii*) the direction of the ribose-phosphate backbone. Structures 𝔻 and 𝔼 represent the 3′-endo and 2′-endo sugar conformations of the nucleotide at the self-cleavage site. These structures have been taken from the coordinates for minimized 𝔸, where constraints were applied to restrict the ribose to be 3′-endo and 2′-endo, respectively. In 𝔼 the attacking 2′-oxygen is in line with the leaving 5′-oxygen. The energy barrier for the 3′- to 2′-endo change is small (17) and the equilibrium is rapid in solution [occurs in the nanosecond time scale for DNA (18, 19)]. Consistently, the *total* energies of the two computer-generated hammerhead structures are comparable, with the 2′-endo structure favored by about 10 kcal/mol.



𝔻 3′-endo ribose 2′-endo ribose 𝔼

The usual 3′-endo sugar stereochemistry of a nucleotide (from x-ray coordinates) in a RNA A-form helix is shown as structure 𝔽. The helical conformation disfavors a 3′- to 2′-endo sugar transformation. A 3′-endo to 2′-endo conformational change in 𝔽 would provide 𝔾. In 𝔾, displacement of the 5′-oxygen can only be by adjacent attack of the 2′-hydroxyl upon the phosphorus due to the direction of the ribose-phosphate backbone leading from the 5′-oxygen (20). The direction of the ribose-phosphate backbone prevents an in-line displacement of O5′ by O2′. Observations of Newman projections through the C2′—C3′ bond of the ribose structural units of the hammerhead RNA were employed to judge whether 3′- to 2′-endo ribose conformational changes, at positions other than the self-cleavage site, could align the 2′ and 5′ oxygens for an in-line displacement. All nucleotides of the minimized hammerhead structure, with the exception of the nucleotide at the self-cleavage site (𝔻 and 𝔼), have the salient features of structures 𝔽 and 𝔾.



𝔽 3′-endo ribose 2′-endo ribose 𝔾

The difference between the 2′-endo structures 𝔼 and 𝔾 resides in the direction of the ribose-phosphate chain. Structure 𝔼 is brought about by the change in the direction of the ribose-phosphate chain at the cleavage site since this chain spans stems I and III. Examination of 𝔻 and 𝔼 shows that the cytosine base must move during the conversion of the 3′-endo conformation to 2′-endo. The cytosine base at the cleavage site is not associated with other bases (*loc. cit.*) and, since it is on the surface of the hammerhead, it is free to move.

Dynamic computations have been carried out at 300 K to 60 ps, using the extensively minimized hammerhead structure (Fig. 1 *Right*) with the ribose at the cleavage site in the 3′-endo conformation. These calculations show that the movements of the atoms reconstitute structure 𝔻 about every

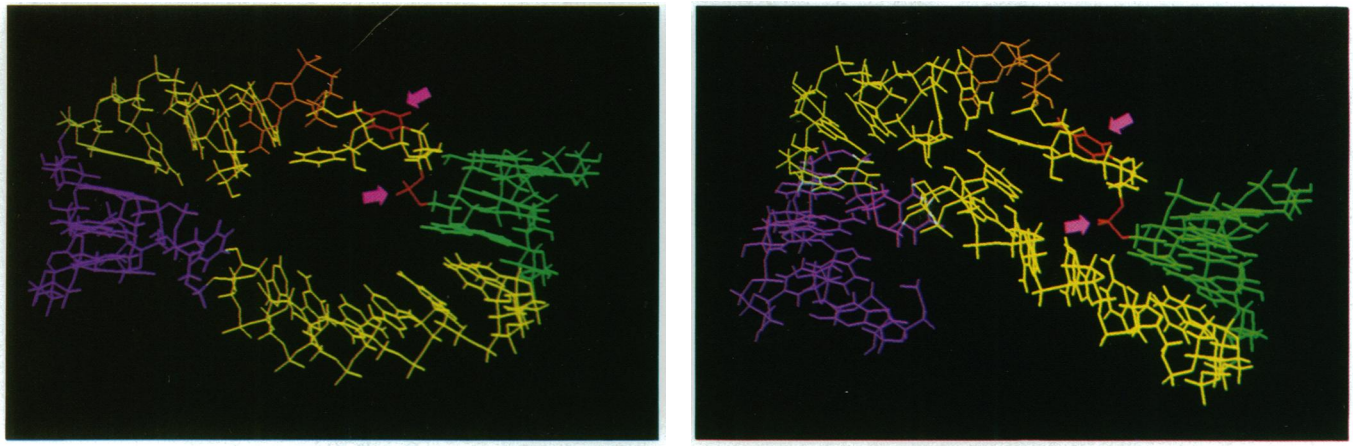


FIG. 1. Stick models of open (*Left*) and closed (*Right*) hammerhead RNAs. Stem I (UGA/UCA) is colored green, stem II (GGC/GCC) purple, stem III (GU/AC) orange, and the interior loop residues yellow. Cytosine base and phosphate (arrows) at each self-cleavage site are highlighted in red.

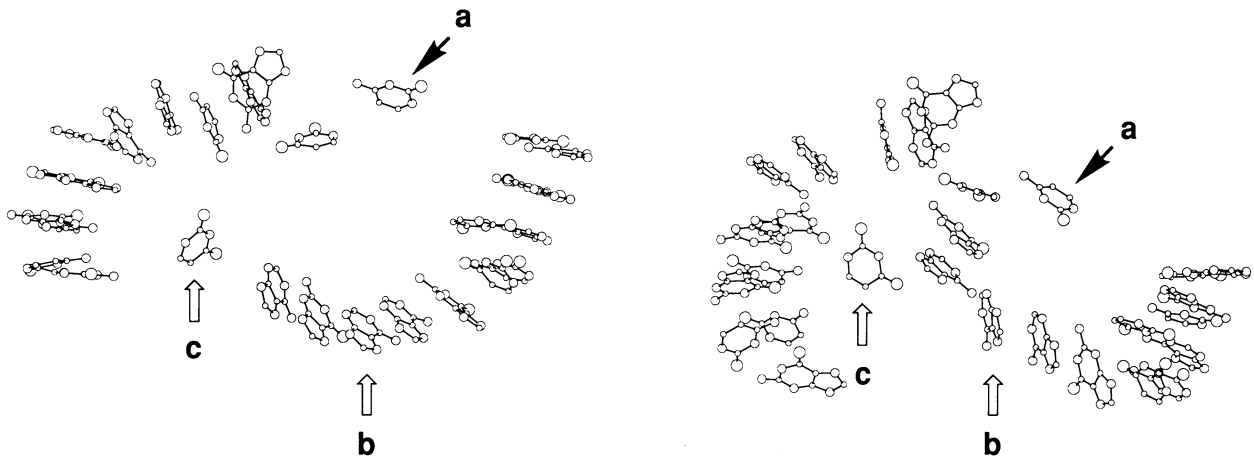


FIG. 2. Bases of the open (*Left*) and closed (*Right*) hammerhead RNAs. Cytosine base at the self-cleavage site is indicated by arrow a, A^{192} by arrow b, and U^{195} by arrow c. Hydrogen atoms are not included.

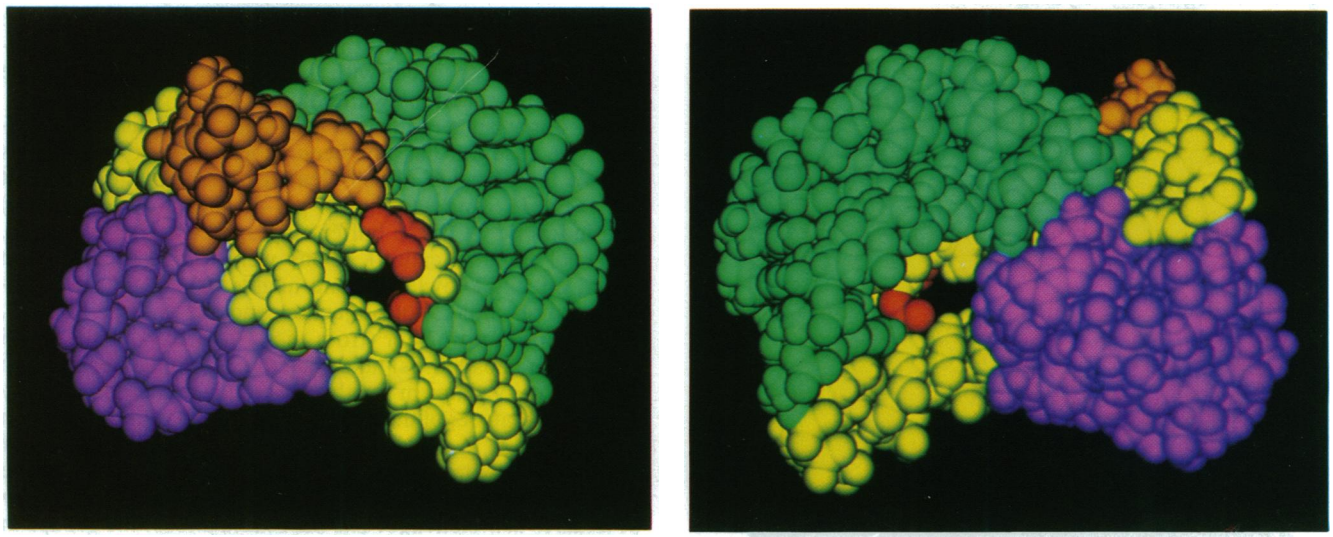
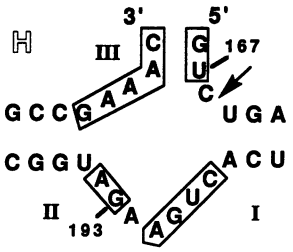


FIG. 3. Space-filling models of closed hammerhead RNA. (*Left*) The "front" view of Fig. 1 *Right*, showing a hole next to the self-cleaving phosphate (red). (*Right*) The "back" view, showing the interactions of stems I (green) and II (purple).

10–15 ps. A much longer time frame would be required to observe any conversion of $\mathbb{D}$ to $\mathbb{E}$ (nanosecond range). Calculations to the nanosecond range were judged to be prohibitive in computer time.

Fig. 3 *Left* and *Right* are "front" and "back" views of the extensively minimized structure $\mathbb{A}$ showing the phosphate and the mobile cytosine of the cleavage site at the surface and a hole next to the cleavage site. This opening is brought about

by interactions of stems I and II (green and purple regions). Interactions of the D- and T- arms of $tRNA^{Phe}$ are important in the assembly of its three-dimensional structure. The influence of stem interactions on hammerhead structure is shown from studies with the "chopped-hammerhead" structure H, where much of stems I and II of A has been removed.



Extensive minimization of H provides a structure that is completely closed but still has the same stereochemistry around the cleavage site and a freely movable cytosine as seen with the complete hammerhead (D and E). Thus, the interactions of stems I and II in A prevent complete closure of the hammerhead structure but do not create the stereochemistry favorable to self-cleavage at the self-cleavage site. The importance of the interactions of stems I and II requires further exploration. All self-cleaving hammerhead structures have at least four or five base pairs in stems I and II (15, 16).

Self-cleavage of the hammerhead is quite slow in the absence of Mg(II) or Mn(II), and when the pro-R oxygen of the self-cleaving phosphate is replaced by sulfur, a 180-fold decrease in the rate of cleavage is observed (O. C. Uhlenbeck, personal communication); Mg(II) is a catalyst only at high concentrations, whereas Mn(II) is still a very effective catalyst. These observations may be explained by assuming that (*i*) the rate-limiting step is self-cleavage rather than a structural rearrangement to form an active species, and (*ii*) Mg(II) is bound to the pro-R oxygen of the phosphate group that is undergoing cleavage. Docking of a $Mg(H_2O)_5^{2+}$ on the pro-R oxygen anion (of the reactive phosphate) in the space provided by the hole adjacent to this phosphate is shown in Fig. 4. The $Mg(H_2O)_5^{2+}$ hydrogen-bonds with conserved G^{190}

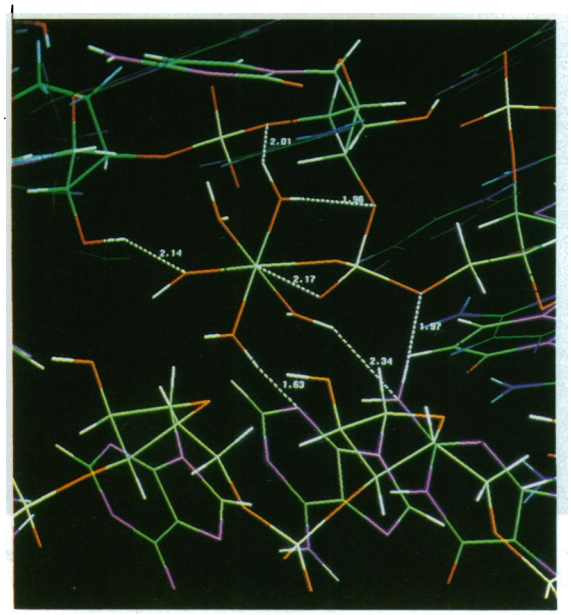
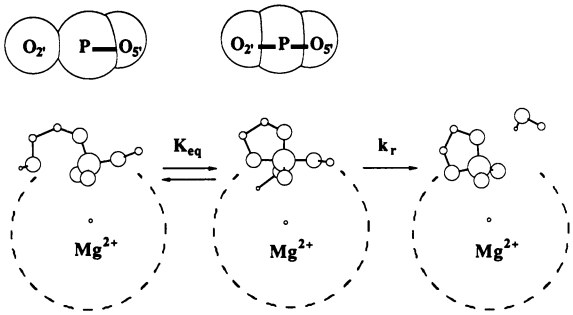


FIG. 4. A model for the Mg(II) binding site in hammerhead RNA. $Mg(H_2O)_5^{2+}$ (octahedral complex at the center) is docked on the pro-R oxygen of the self-cleaving phosphate. Possible hydrogen bonds are indicated by dashed lines with distances in angstroms.

and A^{191} bases of the loop, with a second phosphate, and with an adjacent 2'-hydroxyl substituent. The x-ray structure of $tRNA^{Phe}$ possesses four bound magnesium hydrate species (21). Two are at the corner position, where one is coordinated to a single phosphate and hydrogen-bonded to three bases and the other is coordinated to two phosphates. At the anticodon loop a hydrated magnesium is coordinated to one phosphate and hydrogen-bonded to another phosphate group and neighboring bases, while the hydrated magnesium in the central core region is held in position by hydrogen-bonding to four phosphate groups. The interactions of the hydrated magnesium in Fig. 4 appear eminently reasonable.

We propose, as our working hypothesis, the mechanism of Eq. 1.



[1]

The structures of Eq. 1 are from the coordinates of the extensively minimized A with the ribose conformation being 2'-endo, *Left* (only the 2'- and 3'-carbons of the ribose ring are included); the cleavage intermediate with pentacoordinated phosphate constraint, *Center*; and the cleavage product with 2',3'-cyclic phosphate and 5'-hydroxyl constraints, *Right*. The relative positions of the 2',3'-cyclic phosphate and 5'-hydroxyl substituent were obtained by creating the product structure from the hammerhead (45,000 minimizations) followed by 20,000 additional minimizations. (The product structure was found to be of lower total energy than the hammerhead.) The attacking 2'-hydroxyl oxygen is in-line with the leaving 5'-oxygen and, from the space-filling structures included in Eq. 1, it is almost in bonding distance of the phosphorus. The movement of the 5'-hydroxyl away from the cyclic phosphate prevents the reversal of the in-line displacement. This is consistent with the irreversibility found in the self-cleavage reactions with hammerhead RNAs (7).

The importance of general acid–base catalysis in the self-cleavage reaction has not been thoroughly explored. It is known that the rate of self-cleavage is essentially constant between pH 6 and 9 (at higher pH the structure is unstable) and independent of the buffer employed (O. C. Uhlenbeck, personal communication). If one were forced to a decision based on this knowledge, one might assume that the self-cleavage reaction is not subject to general catalysis. In Eq. 1, the first proton transfer (from the 2'-hydroxyl to the pro-S oxygen on phosphorus) takes place in a pre-equilibrium step so that the mechanism for this transfer would be kinetically invisible. Transfer of the proton in the rate-limiting step (from pro-S oxygen to the leaving 5'-oxygen) could then occur through the hydration shell (dashed circle in Eq. 1) of the Mg(II). *We propose that the structure of the hammerhead per se leads to self-cleavage. The bases and phosphate groups in the structure do not participate as catalysts [though they may be of importance in the coordination of the Mg(II) hydrate that serves as a catalyst by stabilizing the pentacoordinate intermediate and facilitating hydrogen transfer].*

One might inquire as to just how far this computational approach might be taken—in particular, can the requirements for base identities be explained. This will be the topic of a future communication. A few observations will be

mentioned here. In the assessment of the importance of conserved bases the following considerations must be taken into account. The means that have been employed to follow the self-cleavage reaction limit detection of self-cleavage reactions. Only those self-cleavage reactions whose $\Delta G^{\ddagger}$ values are less than 2 kcal/mol greater than the $\Delta G^{\ddagger}$ for self-cleavage of the wild-type hammerheads are considered to self-cleave (ref. 14; O. C. Uhlenbeck, personal communication). Structures that cleave at a slower rate are considered nonreactive. Modeling of a structural activity relationship is constrained by this imposed limitation on the range of rate constants that may be considered. C^{168} at the cleavage site has freedom of motion (*loc. cit.*). This explains the observations that this base may be replaced by A, U, or G without loss of self-cleavage. We find that replacement of C^{168} by A does not affect the required stereochemistry for the self-cleavage reaction (data not shown). Inspection of Fig. 2 shows structurally distinctive features of the two nonconserved bases, A^{192} (arrow b) and U^{195} (arrow c), in the interior loop. Both A^{192} and U^{195} are at the hinge where base-stacking changes its direction. The discontinuity in base-stacking and the changes in the direction of the phosphate backbone may explain why these bases are not conserved. U^{195}, like C^{168}, does not interact with any other bases of the hammerhead structure. We have provided a rational working hypothesis for the mechanism of the self-cleavage reaction of hammerhead RNAs. We are aware of the limitations of this computational approach to the description of detailed tertiary structure. Further computational studies are needed.

We thank Drs. O. C. Uhlenbeck and D. H. Gregory for helpful discussions. This study was supported by grants from the National Science Foundation and the Office of Naval Research.

1. Hutchins, C. J., Rathjen, P. D., Forster, A. C. & Symons, R. H. (1986) *Nucleic Acids Res.* **14,** 3627–3640.
2. Forster, A. C. & Symons, R. H. (1987) *Cell* **49,** 211–220.
3. Epstein, L. M. & Gail, J. G. (1987) *Cell* **48,** 535–543.
4. Usher, D. A., Richardson, D. I., Jr., & Eckstein, F. (1970) *Nature (London)* **228,** 663–665.
5. Uhlenbeck, O. C. (1987) *Nature (London)* **328,** 596–600.
6. Jeffries, A. C. & Symons, R. H. (1989) *Nucleic Acids Res.* **17,** 1371–1377.
7. Forster, A. C., Jeffries, A. C., Sheldon, C. C. & Symons, R. H. (1987) *Cold Spring Harbor Symp. Quant. Biol.* **52,** 249–259.
8. Sussman, J. L., Holbrook, S. R., Warrant, R. W., Church, G. M. & Kim, S.-H. (1978) *J. Mol. Biol.* **123,** 607–630.
9. Westhof, E., Dumas, P. & Moras, D. (1985) *J. Mol. Biol.* **184,** 119–145.
10. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983) *J. Comp. Chem.* **4,** 187–217.
11. Nilsson, L. & Karplus, M. (1987) in *Structure and Dynamics of RNA*, ed. van Knippenberg, P. H. & Hilbers, C. W. (Plenum, New York), pp. 151–159.
12. Doudna, J. A., Gerber, A. S., Cherry, J. M. & Szostak, J. W. (1987) *Cold Spring Harbor Symp. Quant. Biol.* **52,** 173–180.
13. Kim, S.-H. & Cech, T. R. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 8788–8792.
14. Koizumi, M., Iwai, S. & Ohtsuka, E. (1988) *FEBS Lett.* **228,** 228–230.
15. Koizumi, M., Iwai, S. & Ohtsuka, E. (1988) *FEBS Lett.* **239,** 285–288.
16. Westheimer, F. H. (1968) *Acc. Chem. Res.* **1,** 70–78.
17. Saenger, W. (1984) *Principles of Nucleic Acid Structure* (Springer, New York).
18. Levy, G. C., Kraik, D. J., Kumar, A. & London, R. E. (1983) *Biopolymers* **22,** 2703–2726.
19. Kearns, D. (1984) *CRC Crit. Rev. Biochem.* **15,** 237–290.
20. Usher, D. A. (1972) *Nature (London) New Biol.* **235,** 207–208.
21. Teeter, M. M., Quigley, G. J. & Rich, A. (1980) in *Nucleic Acid–Metal Ion Interactions*, ed. Spiro, T. G. (Wiley, New York), pp. 145–177.