Articles

Subsite Binding in an RNase: Structure of a Barnase–Tetranucleotide Complex at 1.76-Å Resolution‡

Ashley M. Buckle and Alan R. Fersht*

Cambridge Centre for Protein Engineering, Medical Research Council Centre, Hills Road, Cambridge CB2 2QH, U.K.

Received October 4, 1993; Revised Manuscript Received November 15, 1993®

ABSTRACT: A set of subsites in barnase has been proposed from kinetic studies. A specific substrate analog, the tetradeoxynucleotide, CGAC, has been designed from this information. We report the crystal structure of its complex with barnase at 1.76-Å resolution. The structure was solved by molecular replacement from a model of free barnase and refined to a crystallographic *R* factor of 19.0%. The stoichiometry of the asymmetric unit dimeric complex is [barnase:d(CGAC)]$_2$, with 2-fold noncrystallographic symmetry. Each barnase molecule binds one oligonucleotide, whereby the recognition site is occupied by guanine, and all three phosphate groups of the nucleotide make electrostatic interactions with basic residues in a strongly electropositive region at the bottom of the active site. The active-site His102 packs against the adenine base of the nucleotide in an almost identical manner to the guanine base in the barnase–d(GpC) complex and defines a possible subsite in the Michaelis complex. The overall protein structure is unchanged on forming the complex with d(CGAC), but there are small differences in the active site and in crystal packing regions. The protein coordinates will be useful for theoretical calculations since some disorder induced by packing constraints in the crystals of the free enzyme are absent in the crystals of the complex. The interface of the dimer is formed by a His102-adenine–adenine-His102 face-to-face ring stack directly on the 2-fold axis. The edge of the adenine–adenine stack packs closely onto the face of a 3′-cytosine–3′-cytosine interaction, which has a "base-pair"-like conformation but too great a separation of the bases to form hydrogen bonds. This unusual arrangement is the major stabilizing interaction within the dimeric complex, since there are no direct protein–protein interactions. Using the structure of the complex as a starting point for model building, the nature of the enzyme–substrate and enzyme–transition state complexes is investigated.

Barnase, a member of the family of homologous microbial ribonucleases, catalyzes the cleavage of single-stranded RNA via a two-step mechanism thought to be similar to that of pancreatic ribonuclease. The mechanism involves a transesterification reaction to give a 2′,3′-cyclic phosphate intermediate, followed by hydrolysis to yield a 3′-nucleotide (Figure 1). The active-site residues His102 and Glu73 act as general-acid–base groups during catalysis, while Arg83, Arg87, and Lys27 are important in binding the reactive phosphate, the latter probably binding the phosphate in the transition state (Mossakowska et al., 1989; Meiering et al., 1991). Rates of catalysis are increased by the occupation of subsites by the substrate as follows[1] (Day et al., 1992; Figure 2). For substrates of the type $Zp_0Gp_1Xp_2Y$, where X, Y, and Z are nucleotides, p is phosphate, and G is guanosine, (1) G occupies the primary specificity site; (2) the most important subsite is for p_2 (occupation gives rise to a 1000-fold increase in k_{cat}/K_m, composed of a 100-fold increase in k_{cat} and a 10-fold decrease in K_m); (3) the next important subsite is for Y; and (4) there is no indication of subsites for Z or p_0. Most of the additional binding energy gained on going from dinucleotide monophosphates to tetranucleotide substrates is associated with binding the transition state rather than the substrate. Values of k_{cat}/K_m for tetranucleotide substrates approach the values associated with diffusion control (Day et al., 1992).

Direct study of ribonuclease–nucleotide interactions is limited to RNase[2] complexes with mono- and dinucleotides [Hill et al., 1983; Sevcik et al. (1990) and references therein]. The structure of a barnase–3′-GMP complex has been determined by X-ray crystallography (Guillet et al., 1993) and NMR methods (Meiering et al., 1993), in which guanine is located in the recognition site and the mode of binding is similar to that in mononucleotide phosphate complexes with microbial RNases Sa (Sevcik et al., 1991) and binase (which is identical to barnase in all but 17 residues) (Pavlovsky et al., 1988). Further structural information on barnase–nucleotide interactions comes from the crystal structure of a barnase–d(GpC) complex (Baudet & Janin, 1991), in which the nucleotide does not bind in a productive manner, but is located outside the active site in a putative subsite. By far the largest body of structural data originates from RNase T1–nucleotide complexes (Heinemann & Saenger, 1982; Kostrewa et al., 1989; Ding et al., 1991; Lenz et al., 1991). A study using longer nucleotides is that of the crystal structure of a complex between mammalian RNase A and $d(pA)_4$ (McPherson et al., 1986). This revealed extensive protein–nucleotide interactions that are mostly between phosphate groups and basic residues and allowed the visualization of a putative complex between the enzyme and a nucleotide strand 12 bases in length. There is no sequence or structural homology between the mammalian and microbial RNases. Whereas RNase T1 catalyzes the hydrolysis of dinucleotide monophosphates with

‡ Structural coordinates have been deposited in the Brookhaven Protein Data Bank under filename 1BRN.

® Abstract published in *Advance ACS Abstracts,* January 15, 1994.

[1] A subsite is referred to as P + 1 if it is occupied by phosphate p_1.

[2] Abbreviations: RNase, ribonuclease; rms, root-mean-square.

0006-2960/94/0433-1644$04.50/0 © 1994 American Chemical Society

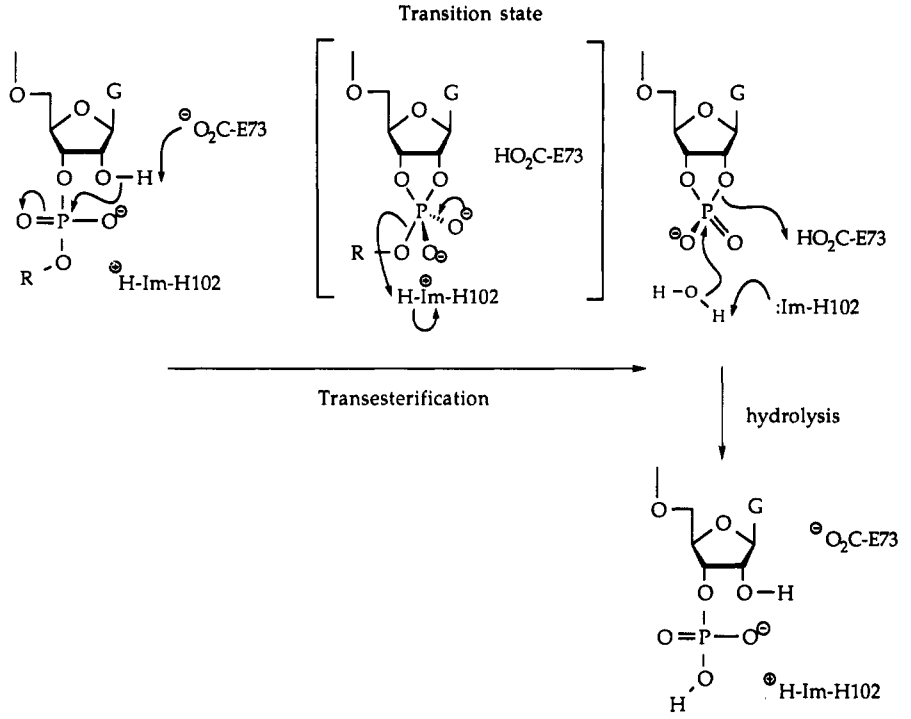

FIGURE 1: Catalytic mechanism of the ribonuclease, barnase. Im represents the histidine imidazole group.

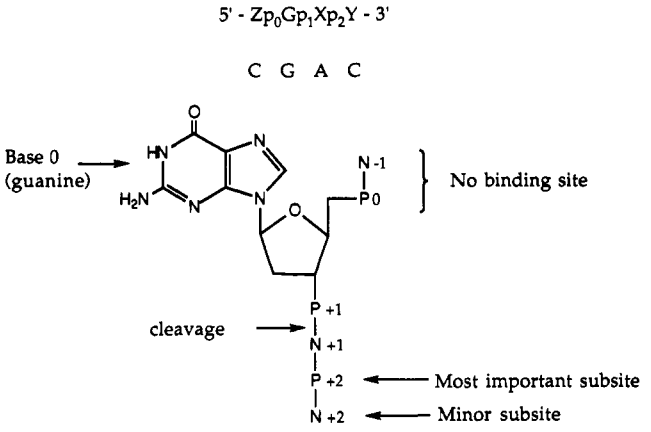

FIGURE 2: Schematic diagram of the ribonucleotide substrate showing the relative importance of subsites investigated in this article.

a value of k_{cat} 200 times that of barnase, the k_{cat} for catalysis of RNA hydrolysis is one-half that of barnase (Mossakowska et al., 1989). To address the above observations on a structural level, we have crystallized, solved, and refined the structure of a complex between barnase and the inhibitor, d(CGAC). This inhibitor is not cleaved by the enzyme since it lacks the 2'-hydroxyl group present in RNA that is necessary for the formation of the cyclic intermediate during the reaction pathway (Figure 1).

EXPERIMENTAL PROCEDURES

The active site in the crystal structure of wild-type barnase is blocked by neighboring symmetry-related molecules, so that inhibitors or substrates cannot be soaked into crystals. Therefore, in this study, cocrystallization of barnase with the inhibitor d(CGAC) was attempted. The stages of structure solution are described below. Unless stated otherwise, all data processing, data reduction, molecular replacement, electron density syntheses, and structural analyses were carried out using CCP4 software (CCP4, 1979).

Crystallization. Expression and purification of barnase have been described (Serrano et al., 1990). The DNA tetranucleotide, d(CGAC), was chemically synthesized on an Applied Biosystems 380B DNA synthesizer, using a 10-μmol-scale column. The tetranucleotide was deprotected by heating in 30% ammonia at 56 °C for 6 h, diluted 5-fold with H_2O, and loaded onto a small Mono Q FPLC column (Pharmacia) that had been preequilibrated in H_2O. d(CGAC) was eluted using a linear gradient of 0–1 M triethylamine carbonate at pH 7.5. After six cycles of adding H_2O and lyophilizing the sample, most of the triethylamine was removed. Any remaining triethylamine was removed by passing the sample over a Na^+ cation-exchange resin (Dowex) several times.

Crystals were grown by vapor diffusion using the hanging drop method (McPherson, 1982). An incomplete factorial search was used to screen for initial crystallization conditions (Jancarik & Kim, 1991). Using this procedure, 48 different experiments were set up to explore a wide range of salts, pH, and precipitants. Five of these conditions resulted in crystals of varying size and quality. All of these had similar concentrations of PEG and salt but varied in pH. On going from pH 9 to 5, the size and quality of the crystals increase dramatically. The largest crystals that diffract the best were grown in the presence of 30% PEG 4000, 0.1 M sodium acetate buffer, and 0.2 M ammonium acetate at 23 °C, pH 5.5. Typically, 5 μL of an equimolar mixture of barnase and oligonucleotide (10 μM each) was mixed with 5 μL of well solution containing the above chemicals. Large prismatic crystals grew in 2 days and were characterized with a FAST area detector, using in-house monochromatic Cu Kα radiation generated from a rotating anode. The crystals belong to space group $P1$ with cell dimensions a = 39.87 Å, b = 54.29 Å, c = 29.61 Å, α = 105.77°, β = 89.05°, and γ = 96.37°.

Data Collection and Processing. X-ray diffraction data were collected at 4.5 °C on a Mar Research Image Plate at the Synchrotron Radiation Source at Daresbury (station px 9.5), using radiation of wavelength 0.92 Å. Starting with an unknown orientation of the crystal in the X-ray beam, a 123° wedge of data to 1.76-Å resolution was collected from one crystal, using an oscillation angle of 2.5° and an exposure time of 6.7 min per image. Indexing and intensity measure-

Table 1: X-ray Data Collection and Refinement Statistics[a]

Data Collection	
unit cell dimensions	a = 39.87 Å, b = 54.29 Å, c = 29.61 Å, α = 105.77°, β = 89.05°, γ = 96.37°
space group	$P1$
maximum resolution (Å)	1.76
total reflections collected	34544
unique reflections	18358
completeness of data (%)	78.0 (83.8)
R_{merge} (%)[b]	3.0 (5.0)
$\langle F/\sigma F \rangle$	28.2 (23.2)
reflections with $\|F\| > 3\sigma_F$ (%)	98.3 (95.3)
solvent content of unit cell (%)[c]	47

Structure Refinement	
R_{cryst} (%), 6.0–1.76 Å, $F > \sigma F$[d]	19.0
Δ_{bond} (Å)	0.010
Δ_{angle} (deg)	1.92
no. atoms in refinement	2048
no. solvent molecules	230
mean B factor (protein) (Å^2)	15.0
mean B factor (nucleotide) (Å^2)	15.4

[a] Values given in parentheses are for the highest resolution shell. [b] R_{merge} gives the agreement between intensities of repeated measurements of the same reflections and can be defined as $\sum(I_{h,i} - \langle I_h \rangle)/\sum I_{h,i}$, where $I_{h,i}$ are individual values and $\langle I_h \rangle$ is the mean value of the intensity of reflection h. [c] Calculated by CCP4 program TRUNCATE. [d] The crystallographic R factor, R_{cryst}, is defined as $\sum|F_o - F_c|/\sum F_o$.

ments of diffraction data were carried out with the MOSFLM program suite (Leslie, 1990). Table 1 summarizes the data collection and processing results. The presence of low-resolution satellite spots on each diffraction image indicated that the crystal was twinned. This was not apparent from the morphology of the crystal, but was found to be the case for all of the crystals tested. However, optimization of the size and shape of the measurement box in MOSFLM decreased the number of reflections that were rejected because of neighboring satellite spots. The low-resolution spots that were rejected mean that the completeness of the data gradually decreases on going from 2.8- to 10.0-Å resolution. Since the medium- to high-resolution data is over 80% complete, no attempt was made to collect more low-resolution data. The radiation damage to the crystal was negligible.

Molecular Replacement. The volume of the unit cell is consistent with only two molecules of barnase in the asymmetric unit (corresponding to ~47% solvent content). Each molecule of barnase was oriented separately from a cross-rotation function, calculated by the program ALMN. Molecule C in the barnase wild-type structure refined at pH 6 (K. Henrick and A. Cameron, unpublished results) was used as a search model. Structure factors were calculated for this model in a $P1$ cell. A sphere of integration of radius 20 Å was used in order to minimize the contribution of cross-vectors from neighboring symmetry molecules. All data in the resolution range 8.0–2.0 Å were used in the calculations. The cross-rotation function was calculated in 5-deg steps in the Eulerian angle β, covering a total range of 0–180°. Two outstanding peaks at 5.5σ and 5.4σ were found (the third and fourth peaks were both 4σ high, where σ is the rms deviation in the rotation function map). After the position of one molecule was fixed, the second molecule was correctly placed in the correct unit cell by calculating a translation function with the program TFPART. Using all data to 1.76 Å, an outstanding peak 22σ high was found; the next peak was at 5σ. After this solution was applied, the resulting dimer (with individual molecules given chain identifiers L and M, respectively) has noncrystallographic 2-fold symmetry.

Structure Refinement. Rigid-body refinement of each barnase molecule in the dimer was performed with X-PLOR (Brunger, 1992). Using all 10–5.0-Å data, the crystallographic R factor dropped from 50.0% to 46.6% after 20 cycles. The structure was then refined against all 6.0–2.0-Å data by simulated annealing with X-PLOR, as follows. Each atom was given a constant temperature factor of 15 Å^2, calculated from a Wilson (1949) plot within the program TRUNCATE. A "prep stage" consisting of 120 cycles of restrained energy refinement was followed by slow cooling, in which the structure was cooled from 4000 to 300 K. A third stage consisted of an additional 120 cycles of positional refinement. Finally, 20 cycles of atomic temperature factor refinement resulted in a model with an R factor of 27.9% for all 6.0–1.76-Å data.

Strong peaks in the $F_o - F_c$ electron density map at the dimer interface and near the active sites of both barnase molecules were easily interpreted as two inhibitor molecules, giving one molecule of inhibitor per molecule of barnase in the dimer (tetrahedral density into which phosphate groups were placed was 10σ high). An initial model of the d(CGAC) inhibitor was taken from the crystal structure of a double-helical oligonucleotide in the Brookhaven database (filename 1BNA; Dickerson & Drew, 1981). One molecule of d(CGAC) was built into electron density at the active site of barnase molecule L. Since the density for the 5′-sugar and base was uninterpretable, these atoms were not included. The second molecule of inhibitor was generated from the noncrystallographic 2-fold symmetry operation. Nucleotide molecules in the asymmetric unit were given chain identifiers A and B, respectively, such that barnase chain L is associated with nucleotide A, and barnase chain M is associated with nucleotide B. Model rebuilding was carried out on an Evans & Sutherland ESV10 graphics workstation using the program O (Jones et al., 1991). Each monomer in the asymmetric unit was inspected separately, and rebuilding was greatly aided by the use of a real-space R factor calculation in the program O. The fit between the protein atoms and the $3F_o - 2F_c$ electron density was calculated and used to color code the protein chain, as a means of identifying poorly fitting regions. Several side chains could be modeled in more than one discrete conformation. The occupancy of discretely disordered residues was not refined.

Peaks above 3.5σ in the $F_o - F_c$ map were found, using the program PEAKMAX, and examined for contacts with protein or solvent atoms with WATPEAK. Solvent molecules were added only if they had acceptable hydrogen-bonding geometry contacts of 2.5–3.5 Å with protein atoms or with existing solvent. Further refinement consisted of rounds of positional and atomic temperature factor refinement with XPLOR followed by model building. Electron density for the 5′-sugar and base of both nucleotides in the asymmetric unit did not appear during refinement. Solvent molecules with B values greater than 50 Å^2 were kept only if they were in good electron density and satisfied the above hydrogen-bonding criteria. The occupancy of solvent molecules was set at unity and was not refined.

Structural Analysis and Model Building. Structural comparisons between the barnase–d(CGAC) complex and unligated barnase were performed using protein chain L in the d(CGAC) complex and chain C of the unligated barnase wild-type crystal structure solved at pH 6.0 (K. Henrick and A. Cameron, unpublished results). Least-squares superposition of structures was carried out with the program LSQKAB (Kabsch, 1978). Structural analysis of protein–nucleotide interactions in the barnase–d(CGAC) complex described in this paper was carried out using barnase chain L and nucleotide chain A in the asymmetric unit. Potential hydrogen-bonding

FIGURE 3: Dimeric complex in the asymmetric unit, viewed approximately down the noncrystallographic 2-fold axis. Each nucleotide is represented by a ball-and-stick model; bonds from one barnase–d(CGAC) complex are drawn filled, whereas those from the other complex that are related by the noncrystallographic 2-fold axis are drawn open [drawn with MolScript (Kraulis, 1991)].

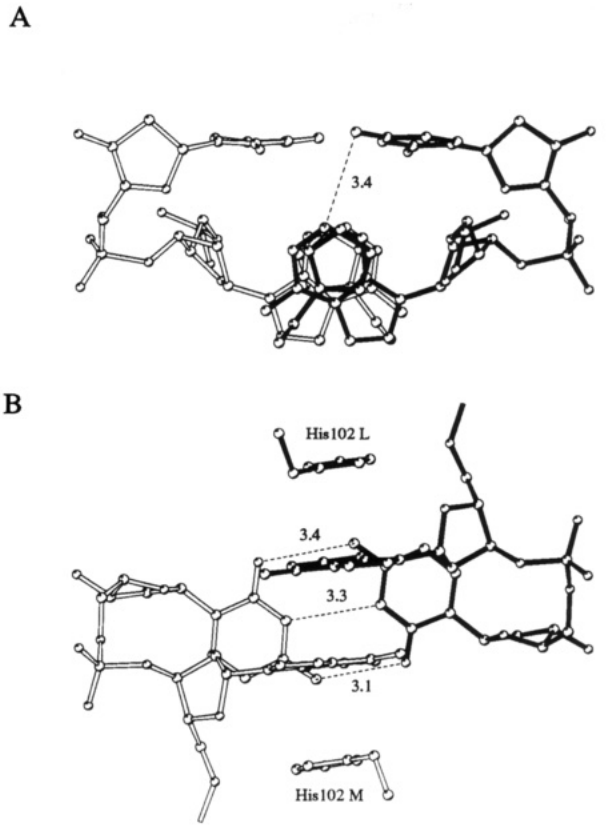

FIGURE 4: His-adenine–adenine-His stack at the dimer interface. Bonds from one barnase–d(CGAC) complex are drawn filled, whereas those from the other complex that is related by the noncrystallographic 2-fold axis are drawn open. (A) Viewed perpendicular to the 2-fold axis. (B) Looking down the 2-fold axis. Distances between potential hydrogen-bonding partners are given in angstroms [drawn with MolScript (Kraulis, 1991)].

interactions were analyzed using the CCP4 program CONTACT. The coordinates of the structure have been deposited in the Brookhaven Protein Data Bank.

RESULTS

Overall Structure of the Asymmetric Unit. The dimeric complex has 2-fold noncrystallographic symmetry, with the active site of each monomer located at the dimer interface (Figure 3). Each barnase molecule has its active site occupied by one molecule of inhibitor. That is, the stoichiometry of the asymmetric unit complex is [barnase:d(CGAC)]$_2$. Since the two molecules of inhibitor are located at the active sites, the noncrystallographic 2-fold symmetry results in a His102-adenine–adenine-His102 stack and a base-pair-like interaction between the two 3′-cytosine bases (Figure 4). Although there is probably one weak hydrogen bond between the closest approaching atoms in the pair, a closer, fully hydrogen-bonded base-pair interaction would require the imine nitrogen (N3) of one of the bases to be protonated (the proton would probably switch rapidly between the two bases). Repeating 150 cycles of the structure refinement (using XPLOR) incorporating a charged, protonated form of one 3′-cytosine base did not result in a closer interaction. Since the electron density does not suggest a closer interaction between the bases, this leads us to believe that, in the crystal, the pH is not sufficiently close enough to the pK_a of the cytosine N3 (estimated at approximately 4.6; Blackburn & Gait, 1990) to allow protonation to occur, resulting in repulsion between the lone-pair electrons of both N3 atoms. Protonated CC base pairs have been observed in crystal structures and NMR solution structures of oligonucleotides in acidic conditions (Saenger, 1983; Gehring et al., 1993). Since there are no direct protein–protein contacts at the dimeric interface, the stack of aromatic groups "links" the barnase monomers together. The CC base interaction packs tightly onto the His-adenine–adenine-His stack in an orthogonal face-to-edge fashion (Figure 4).

There is excellent agreement between the conformations of protein monomers and between oligonucleotides in the asymmetric unit. Barnase main chains L and M, when superimposed, give an rms deviation of 0.14 Å. The rms deviation in atomic positions for nucleotides A and B is 0.47 Å.

Several side chains on the surface of the protein adopt more than one discrete conformation. The following residues have two well-defined side-chain conformations: Ser28, Val36, Asp54, Ser57, Ser80, Ile96, and Thr99 in chain L and Lys19 and Ser28 in chain M. In both chains, Ser85 has three side-chain conformations. None of these side chains are disordered in the unligated enzyme structures at pH 6.0, 7.5, or 9.0 (Cameron, 1992). All of these residues lie on the surface of the protein, with the interesting exception of Ile96, which is completely buried in the hydrophobic core. This residue is not disordered in the 2.0-Å crystal structure of the unligated enzyme, suggesting that the observed disorder is probably a result of increased resolution in the barnase–d(CGAC) structure. Discrete disorder of buried hydrophobic side chains has been observed previously in the structure of a barnase Ile→Val88 mutant, in which the side chain of Val88 adopts three discrete conformations (Buckle et al., 1993). The structurally equivalent residue in RNase T1, Val78, is disordered in the apo structure, but becomes ordered when the active site of the enzyme is occupied by a nucleotide inhibitor (Martinez-Oyanedel et al., 1991).

Nucleotide Binding Does Not Cause Large Changes in Protein Structure. Overall, the structure of the protein in the

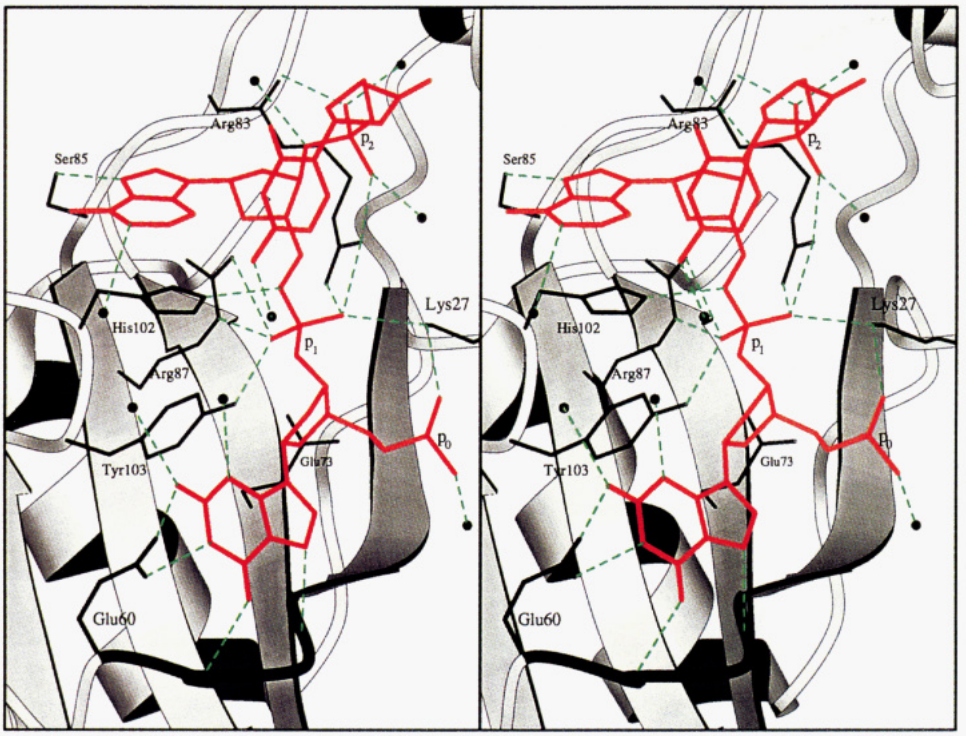

FIGURE 5: Stereoview of the barnase–d(CGAC) complex showing the hydrogen bonding (green broken lines) of nucleotide (red) with protein (side chains in black) and solvent (black spheres). The guanine recognition loop is shown in black. Hydrogen bonds between the guanine base and the main chain of the recognition loop involve the amide groups of the protein [drawn with MolScript (Kraulis, 1991)].

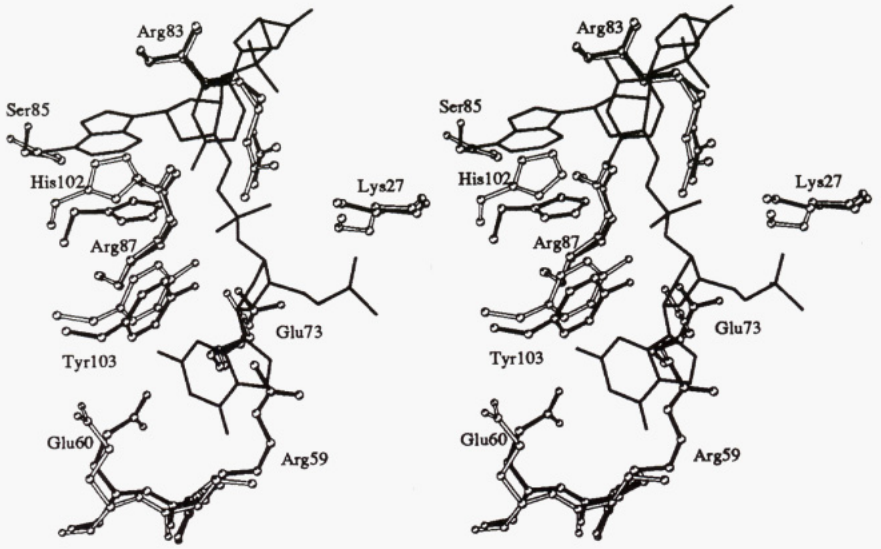

FIGURE 6: Comparison, in stereo, of unligated barnase structure (pH 6.0, chain C) and barnase–d(CGAC) complex (chain L). Active-site residues are shown after a least-squares superposition of the main-chain atoms of both structures. Filled bonds, ligated structure; open bonds, unligated structure. The nucleotide is drawn as solid lines [drawn with MolScript (Kraulis, 1991)].

barnase–d(CGAC) complex is almost identical to that found in the unligated structure. The rms deviation between free and unligated protein structures is 0.33 and 0.95 Å for main-chain and side-chain atoms, respectively. Most structural differences result from a combination of crystal packing and inhibitor binding, which are mainly in the flexible G-binding loop, around His102 and at the C-terminus of the first α-helix (6–18). In all structures of barnase wild-type and mutant proteins, there is a pH-dependent disorder around histidines −18 and −102 that is responsible for conformational variation in these regions (Buckle et al., 1993). Since this disorder is absent from the complex structure, the protein coordinates will be useful for theoretical calculations. Differences around the second α-helix (26–34) between the free and complexed structures are due to interactions with the 3′-cytosine base of the inhibitor. Along with main-chain and side-chain shifts at His102 and Tyr103, the remaining significant movements in the active site are mostly at residues that make electrostatic interactions with phosphate groups of the inhibitor (Glu73, Arg83, and Lys27; Figure 6).

Stability of the Dimeric Complex. Apart from the stabilization conferred upon the dimer by crystal packing, the asymmetric unit complex is stabilized by the unusual arrangement of nucleotide bases at the dimeric interface. Previous NMR experiments in this laboratory investigating barnase–d(CGAC) binding (E. M. Meiering and A. R. Fersht, unpublished results) were hindered by the crystallization of the complex in the NMR tube. Crystallographic characterization of these highly twinned crystals showed them to be isomorphous with the form studied here. At the high (>40

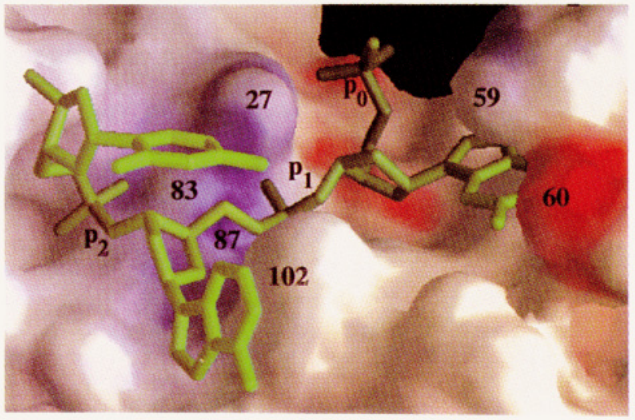

FIGURE 7: Barnase–d(CGAC) complex with the protein surface color-coded according to the electrostatic potential (calculated by the Poisson–Boltzmann solver within GRASP, and displayed within GRASP; Nicholls, 1992). The calculation was carried out in the absence of d(CGAC). Lys and Arg residues were assigned a single positive charge, and Asp and Glu residues were assigned a single negative charge. The side chain of His102 was treated as neutral, in accordance with its measured pK_a of 6.3 (Sali et al., 1988). All other residues were considered neutral. The calculation was performed by assuming a uniform dielectric constant of 80 for the solvent and 2 for the protein interior. The ionic strength was set to zero. The color of the surface represents the electrostatic potential at the protein surface, going from blue (potential of +4.7 kT/e) to red (potential of −2.4 kT/e). T is temperature, e is the charge of the electron, and k is the Boltzmann constant. The probe radius used was 1.4 Å.

mg mL^{-1}) concentrations required for NMR studies and in the absence of precipitants, the potentially stable nature of this arrangement is a considerable driving force for crystallization. It is noteworthy that these crystals, in contrast to those grown from all barnase wild-type and mutant proteins to date, have unprecedentedly high stability in the X-ray beam and intense high-resolution diffraction.

Barnase–Nucleotide Interactions. The nucleotide spans most of the active-site cleft, covering a surface area of 305 Å^2, that is, 7% of the total solvent-accessible area of barnase (Figures 5 and 7). The guanine base stacks onto the aromatic rings of Phe56 and Tyr103 in the recognition site and makes an almost identical set of hydrogen bonds with the specificity loop of barnase previously seen in the NMR solution structure of the barnase–3′-GMP complex (Meiering et al., 1993), in the barnase–3′-GMP crystal structure (Guillet et al., 1993), in the structure of the 3′-GMP complex with the structural homolog of barnase, binase (Pavlovsky et al., 1988), and in a microbial ribonuclease Sa–3′-GMP complex (Sevcik et al., 1991). In the unligated barnase structure and in the barnase–d(GpC) complex (Baudet & Janin, 1991), the side chain of Arg59 is too disordered to be located in the electron density, but when the guanine recognition site is occupied, it twists back on itself and stacks directly onto the face of the guanine ring. This is also the case in the binase–3′-GMP structure. NMR studies in this laboratory have shown that the mobility of this residue, and indeed of all active-site residues, decreases when barnase binds 3′-GMP (Meiering et al., 1993).

The dominant barnase–nucleotide interactions are electrostatic in nature between phosphate groups p_0, p_1, and p_2 and the side chains of Lys27, Arg83, Arg87, His102, and Tyr103. In total, there are 15 hydrogen bonds between inhibitor and protein atoms (Figure 5 and Table 2). Seven of these are between protein and the reactive phosphate p_1. There are nine hydrogen bonds between nucleotide and water atoms. Upon inhibitor binding, two solvent molecules are displaced from the active site and several solvent positions remain conserved, but there is more ordered solvent resolved

Table 2: Hydrogen-Bonding Interactions between Barnase and d(CGAC)[a]

	residue	D···A (Å)	D–H···A (deg)
	5′-Cytidine		
deoxyribose			
O3′	Wat175	2.6	
phosphate p_0			
O1P	NZ Lys27	2.8	
	Guanosine		
base			
N7	N Ser57	3.0	163
O6	N Arg59	2.8	177
N2	OE2 Glu60	2.8	110
	Wat60	2.9	
N1	OE1 Glu60	2.9	111
N3	Wat154	2.9	
phosphate p_1			
O1P	NZ Lys27	2.6	
	NH1 Arg83	3.0	148
	NH2 Arg83	3.2	137
O2P	NE Arg87	2.9	161
	NH2 Arg87	3.2	138
	OH Tyr103	2.6	117
	Adenosine		
base			
N7	OH Ser85	2.9	113
N1	Wat143	3.0	
N6	Wat134	2.7	
deoxyribose			
O5′	NE2 His102	2.8	168
O3′	Wat107	2.7	
phosphate p_2			
O1P	NH1 Arg83	2.8	148
	Wat85	2.8	
O2P	N Arg83	2.9	166
	Wat19	2.5	
	3′-Cytidine		
base			
N4	Wat100	3.0	

[a] All contacts shorter than 3.25 Å with a D–H···A angle greater than 120° and an X–O···H angle greater than 90° are shown (D, hydrogen bond donor; A, hydrogen bond acceptor). Bond angles are not given for hydrogen bonds involving water and lysine NZ groups since the hydrogen position is ambiguous. Values were calculated with CCP4 program CONTACT.

in the electron density in the complex structure. This is a consequence of the increased resolution and higher quality of data and model, the presence of new hydrophilic inhibitor atoms in the active site, and a change of crystal packing. The latter difference is probably the most important since, in the free structure, residues in the active site pack onto a symmetry-related molecule, leaving little room for solvent. Consequently, it is impossible to determine whether an entropically favorable release of bound solvent in the active site upon inhibitor binding contributes to the stability of the barnase–inhibitor interaction.

Nucleotide Conformation. The conformational parameters of the nucleotide are summarized in Table 3. The sugar puckers of the deoxyribose rings are clearly defined as C2′-*endo*, C2′-*endo*, and C4′-*exo* for nucleotides G, A, and C, respectively. Figure 8 shows a comparison of the nucleotide conformation in the d(CGAC) complex with those in the binase–3′-GMP and barnase–d(GpC) complexes. From the comparison with 3′-GMP, it can be seen that there is a small difference in deoxyribose and guanine base position. In the binase–3′-GMP complex, the exact position of the nucleotide is dictated by relatively few protein–nucleotide interactions, whereas in barnase–d(CGAC), more extensive interactions with the extra two nucleotides have to be accommodated. The conformation of the ApC portion of the d(CGAC) nucleotide and its interactions with active-site residues are almost identical

Table 3: Torsion Angles of d(CGAC) Nucleotide[a]

dihedral angle	atoms involved	G	A	C
backbone				
α[b]	(-)O3′–P–O5′–C5′	–97	97	–73
β	P–O5′–C5′–C4′	–123	123	138
γ	O5′–C5′–C4′–C3′	48	171	48
δ	C5′–C4′–C3′–O3′	152	149	80
ϵ	C4′–C3′–O3′–P	–59	–97	
ξ[c]	C3′–O3′–P–O5′(+)	–148	–70	
χ	O4′–C1′–N1–C2[b]	–145 *anti*	77 *syn*	–105 *anti*
sugar pucker		C2′-*endo*	C2′-*endo*	C4′-*exo*

[a] The 5′-cytidine is not included since density was uninterpretable and atoms were not included in the final model. [b] (–) and (+) indicate that the atom belongs to nucleotides –1 and +1, respectively. [c] This applies to pyrimidines; for purines the angle is defined as O4′–C1′–N9–C4.

to those in the barnase–d(GpC) structure. In both structures, the purine ring stacks face-to-face onto the side chain of His102. The small difference in the position of the 3′-cytosine base is most likely the result of crystal packing interactions in the barnase–d(GpC) structure and the cytosine–cytosine interaction in the barnase–d(CGAC) structure.

Inhibitor Binding Decreases the Mobility of Active-Site Residues and the G-Binding Loop. The mean atomic temperature factor for the complexed barnase is very similar to that of free barnase solved at pH 6.0 (15.0 and 11.2 Å^2, respectively). Whereas the overall atomic order and continuity of the electron density maps between the two structures are very similar, residues in the flexible G-binding loop and in the active site become much more ordered upon inhibitor binding. The crystal packing related disorder in the regions around His18 and His102 in the unligated structure is absent from the complexed structure.

DISCUSSION

Relationship of Crystal Structure to Kinetics and NMR Data. Structural features of the model can be correlated with the kinetics data (Day et al., 1992) and results from NMR studies of a barnase–3′-GMP complex (Meiering et al., 1991). The validity of this comparison is strengthened by the close similarity between the buffer, ionic strength, pH, and temperature of the crystallization medium and those in kinetic experiments. There is no evidence to contradict the accepted mechanism of catalysis and the role of His102 and Glu73 as general-acid–base groups (Mossakowska et al., 1989).

The lack of density for the 5′-cytosine base and sugar is consistent with kinetic data that show that the presence of these groups has no detectable effect on rates of catalysis. The 5′-phosphate group p_0, however, makes one hydrogen bond with the Lys27 NE1 atom. This is a direct consequence of the nonproductive, *anti* conformation of the guanine nucleotide. When modeled in the catalytically productive *syn* conformation, as observed in both the crystal structure (Guillet et al., 1993) and the NMR solution structure of the barnase–3′-GMP complex (Meiering et al., 1991), p_0 points into the solvent and makes no interactions with the protein (see below).

The binding mode of guanine in the specificity site is very similar to that in other microbial ribonucleases (Sevcik et al., 1990). The primary specificity of binase and barnase for guanine has been attributed to the donor and acceptor pattern of the four hydrogen bonds between the guanine base and residues 57–60 in the large recognition loop (Sevcik et al., 1990). The specificity for guanine is absolute in dinucleotide monophosphates but relaxed in longer substrates (Mossakowska et al., 1989), presumably because of additional binding energy from subsite interactions. An important stabilizing factor at the recognition site is the "lid" over the guanine base that is formed by the side chain of Arg59. Mutation of this residue to Ala leaves only 15% of enzyme activity (Meiering et al., 1992). The NMR solution structure of a barnase–3′-GMP complex shows two main conformational isomers of the nucleotide, varying in ribose pucker and glycosidic torsion angle (*syn* and *anti* conformations). There is no indication of disorder of the ribose in the crystal structure, however.

As it is accepted that Glu73 acts as the general base in the catalysis mechanism by extracting a proton from the 2′-OH group (Mossakowska et al., 1989), it is surprising that in this structure (with modeled 2′-OH), in a binase–3′-GMP complex (Pavlovsky et al., 1988), and in a RNase Sa–3′-GMP complex (Sevcik et al., 1991) Glu73 (or its homologous counterpart) is pointing away from the 2′-OH (~7-Å separation) and clearly not in a position to fulfill this role. For a catalytically productive situation to occur, the glycosidic torsion angle must be *syn*, as opposed to the *anti* conformation seen here. In the crystal structure of a barnase–3′-GMP complex, the guanine adopts a *syn* conformation, allowing a Glu73–O2′ hydrogen bond. In solution, the guanine nucleotide is probably in equilibrium between *syn* and *anti* (Meiering et al., 1993), but in the d(CGAC) inhibitor the possible interaction between the 5′-phosphate and Lys27 would favor the *anti* conformation. In addition, the lack of the 2′-OH of the guanine nucleotide group in d(CGAC), and thus a potential hydrogen bond with Glu73, might destabilize the active *syn* conformation. In the RNase Sa–3′-GMP complex, the ribose *anti* conformation is stabilized by a hydrogen bond between the O2′ atom and an Asp residue belonging to a neighboring molecule in the crystal. Taken together, these structures demonstrate that, for mononucleotide phosphates, several binding modes are possible, depending on the particular environment. It is possible that these enzymes exploit this conformational flexibility in some way during catalysis [for example, in order to promote product dissociation (Meiering et al., 1993)].

Structural Basis for Increases in Catalysis Rate Due to the Presence of Phosphate p_2. The increases in the rate of catalysis due to the presence of p_2 in an RNA substrate are manifested to a small extent in K_m, but mostly in k_{cat}, and are thus a result of increased stabilization of the transition state (TS). It is crucial, therefore, to determine whether the conformation of active-site residues and tetranucleotide in the crystal structure resembles an enzyme–substrate complex (ES), an enzyme–transition state complex (ES*), or some other, possibly unrelated, state. It has been shown that the precise arrangement of residues in the active site actually destabilizes the enzyme (Meiering et al., 1992), presumably because of the unfavorable clustering of positively charged residues Arg83, Arg87, and Lys27. This compromise between stability and function is the result of a requirement for optimal complementarity with the TS in order to achieve maximum rates of catalysis. Thus, we can assume that the crystal structure gives us an approximate picture of the active-site residue conformation in a TS scenario. Extension of this assumption to the nucleotide conformation is a more complex matter, however. The first step in the transesterification reaction is the extraction of a proton from the 2′-OH group of the guanine nucleotide by the general base, Glu73. In the barnase–d(CGAC) complex, the *anti* conformation of the guanine nucleotide sequesters the Glu73 from a modeled 2′-OH group. As already mentioned, a catalytically productive *syn* conformation is predominant in solution, but in our structure the *anti* conformer is stabilized by extra interactions. When the guanine in d(CGAC) is modeled in a *syn* confor-

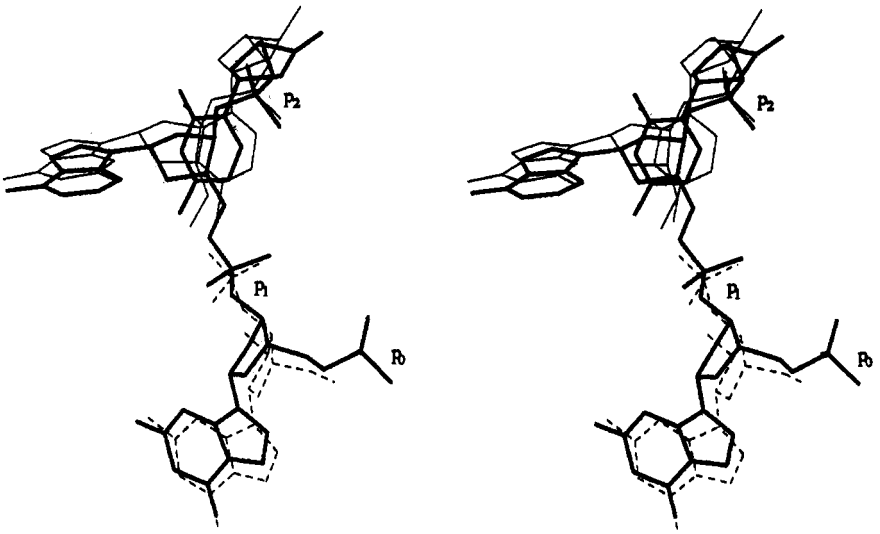

FIGURE 8: Stereoview showing the comparison of nucleotide conformations of barnase–d(CGAC) (thick lines), barnase–d(GpC) (thin lines), and a binase–3′-GMP complex (broken lines). The comparison was done after a superposition of the barnase–d(GpC) and binase–3′-GMP coordinates onto the barnase–d(CGAC) structure using protein main-chain atoms [drawn with MolScript (Kraulis, 1991)].

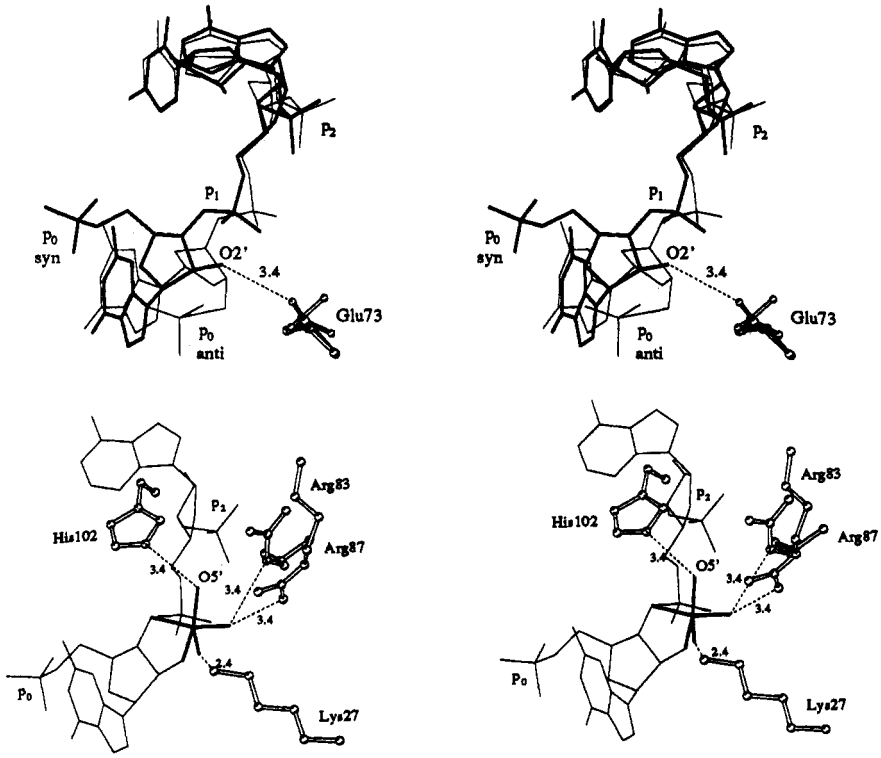

FIGURE 9: Stereo diagrams showing modeled enzyme–substrate and enzyme–transition state interactions. (a, top) Enzyme–substrate model with the guanine modeled in a catalytically productive *syn* conformation (thick bonds), indicating the Glu73–O2′ interaction. The *anti* conformation is shown with thin lines for comparison. Glu73 is represented as a ball-and-stick model (unligated as filled bonds, ligated as open bonds). (b, bottom) Modeled enzyme–substrate complex with guanine in the *syn* conformation (thin bonds), also showing p_1 modeled in a transition-state geometry having trigonal bipyramidal geometry (thick bonds). The trigonal bipyramidal geometry was adapted from the structure of the transition-state analog, uridine vanadate (Wlodawer et al., 1983). Active-site residues that could interact with the transition-state phosphate are shown as a ball-and-stick model. For reasons of clarity, the 3′-cytidine atoms are omitted. Distances are in angstroms [drawn with MolScript (Kraulis, 1991)].

mation (by rotating around the glycosidic bond) and a 2′-OH group modeled into the deoxyribose sugar, a Glu73–O2′ hydrogen bond is made possible by a small change of the Glu73 side-chain torsion angles, as in the barnase–3′-GMP crystal structure (Guillet et al., 1993). As a check we carried out energy minimization of the resulting nucleotide while keeping the surrounding protein atoms fixed (150 cycles using the Powell minimizer in XPLOR). This results in little change in the nucleotide conformation (Figure 9a).

Having modeled an enzyme–substrate complex, we can examine the effect of phosphate p_2 by comparing the binding of substrates GpN and GpNp (N is any nucleoside). Owing to the large area of positive charge at the bottom of the active site, there are two sites in which a phosphate group can bind. A GpN substrate can bind in a catalytically productive fashion as in the barnase–3′-GMP complex (Guillet et al., 1993; Meiering et al., 1993), but it can also bind nonproductively (Baudet & Janin, 1991) where p_1 occupies the P + 2 subsite

of the d(CGAC) complex. In the case of GpNp binding, both phosphates can make electrostatic interactions with the protein only when the substrate is productively bound. In other words, the presence of p_2 is crucial to ensuring that the substrate enters the active site in a productive manner and will effectively eliminate nonproductive binding.

Subsequent formation of the ES* leads to a change in geometry around p_1 to that of a trigonal bipyramid. The in-line mechanism of nucleophilic attack on p_1 by the 2′-oxygen requires this group and the leaving group, O5′, to be at the apices of the bipyramid (Uscher et al., 1970). If the TS geometry is modeled into the ES complex (Figure 9b), we can see that, to remain consistent with the catalytic mechanism, this must be accompanied by a small conformational movement of protein and/or nucleotide, so that the His102–O5′ distance is reduced by 0.5–1.0 Å to be within hydrogen-bonding distance. From this simple modeling, we can see that the enzyme might take advantage of the change in geometry around p_1 as it goes from the tetracoordinate substrate to the pentacoordinate transition state, in order to achieve stronger protein–phosphate interactions. This phenomenon has been proposed before in the study of the catalytic mechanism of tyrosyl-tRNA synthetase, in which the transition state is stabilized by improved binding of the γ-phosphoryl group of ATP by interactions with an enzyme subsite (Leatherbarrow et al., 1985). The active-site Lys27, a residue that is crucial to stabilizing the transition state (Mossakowska et al., 1989), makes one hydrogen bond to p_1 in the d(CGAC) complex. In the modeled TS geometry, however, Lys27 could make interactions with both charged equatorial oxygen atoms (Figure 9b). Its flexibility would allow it to maximize this stabilization in a manner similar to the movement and ordering of Lys41 in RNase A upon binding a transition-state analog (Alber et al., 1983).

Do Alternative Modes of Binding and Additional Subsites Exist? The conformation of the ApC portion is almost identical to that in the barnase–d(GpC) structure (Baudet & Janin, 1991), in which the 3′-cytosine is involved in crystal contacts with a symmetry-related molecule. Purine–His stacking has been observed in an RNase T1–2′-AMP complex (Ding et al., 1991), a T1–guanylyl(2′–5′)guanosine complex (Koepke et al., 1989), and a T1–guanylyl(3′–5′)guanosine interaction modeled from the structure of a complex between T1 and two molecules of 3′-GMP (Lenz et al., 1991). This strongly suggests that this His base packing constitutes a stable subsite interaction and is not a product of crystal packing.

However, there is evidence from kinetics that the formation of an enzyme–transition state complex is probably accompanied by a conformational change 3′ of the scissile bond. Day et al. (1992) found that there is a favorable interaction between bases B + 1 and B + 2 in the ground state that becomes unfavorable in the transition state. It is possible that this favorable interaction is seen here as a face-to-edge stack of adenine and cytosine bases that is probably absent or severely weakened in the transition state.

We cannot rule out the possibility that the presence of 2′-hydroxyl groups in RNA (in addition to that involved in the catalysis mechanism discussed above) has an effect on the rates of catalysis, which we are unable to account for in structural studies using 2′-deoxy analogs of RNA substrates. These groups make no interactions with the protein when modeled into the d(CGAC) structure; therefore, we expect that the only effect could be through their influence on the nucleotide conformation (Saenger, 1983).

Conclusions. High-resolution structural information for a barnase–deoxynucleotide interaction has allowed us to rationalize much of the existing kinetic data on the effect of subsites on ribonucleotide hydrolysis. We have been able to postulate how the occupation of a P + 2 subsite has a dramatic effect on the rates of catalysis, achieved through the stabilization of the transition state. It is clear that barnase has evolved to maximize subsite interactions with RNA in a transition-state conformation, involving nucleotides at least two positions 3′ of the scissile phosphodiester bond. The possibility of the presence of subsite interactions involving nucleotides farther downstream awaits kinetic and structural studies with longer substrates and inhibitors. The precise nature of the barnase–RNA interaction will be understood in greater detail by undertaking structural studies with transition-state analogs combined with further kinetics studies with site-directed mutants.

ACKNOWLEDGMENT

We are grateful to Dr. Kim Henrick for his help and tuition and to Tony Day for helpful discussions.

REFERENCES

Alber, T., Gilbert, W. A., Ringe Ponzi, D., & Petsko, G. A. (1983) in *Mobility and function in proteins and nucleic acids*, pp 4–24, Pitman, London.

Baudet, S., & Janin, J. (1991) *J. Mol. Biol. 219*, 123–132.

Blackburn, G. M., & Gait, M. J., Eds. (1990) *Nucleic Acids in Chemistry and Biology*, IRL Press/Oxford University Press, Oxford, U.K.

Brunger, A. T. (1992) *XPLOR Manual*, Version 3.0., Yale University, New Haven, CT.

Buckle, A. M., Henrick, K., & Fersht, A. R. (1993) *J. Mol. Biol. 234*, 847–860.

Cameron, A. D. (1992) Ph.D. Thesis, York University, Heslington, York, U.K.

CCP4 (1979) *The SERC (UK) Collaborative Computing Project No. 4, a suite of programs for protein crystallography*, distributed from Daresbury Laboratory, Warrington WA4 4AD, U.K.

Day, A. G., Parsonage, D., Ebel, S., Brown, T., & Fersht, A. R. (1992) *Biochemistry 31*, 6390–6395.

Dickerson, R. E., & Drew, H. R. (1981) *J. Mol. Biol. 149*, 761–786.

Ding, J., Koellner, G., Grunert, H. P., & Saenger, W. (1991) *J. Biol. Chem. 266*, 15128–15134.

Gehring, K., Leroy, J.-L., & Guéron, M. (1993) *Nature (London) 363*, 561–565.

Guillet, V., Lapthorn, A., & Mauguen, Y. (1993) *FEBS Lett. 330*, 137–330.

Heinemann, U., & Saenger, W. (1982) *Nature (London) 299*, 28–32.

Hill, C., Dodson, G., Heinemann, U., Saenger, W., Mitsui, Y., Nakamura, Y. M. K., Borisov, S., Tischenko, G., Polyakov, K., & Pavlovsky, S. (1983) *Trends Biochem. Sci. 8*, 364–369.

Jancarik, J., & Kim, S. H. (1991) *J. Appl. Crystallogr. 24*, 409–411.

Jones, T. A., Zou, J.-Y., Cowan, S. W., & Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110–119.

Kabsch, W. (1978) *Acta Crystallogr. A42*, 922–923.

Koepke, J., Maslowska, M., Heinemann, U., & Saenger, W. (1989) *J. Mol. Biol. 206*, 475–488.

Kostrewa, D., Choe, H. W., Heinemann, U., & Saenger, W. (1989) *Biochemistry 28*, 7592–600.

Kraulis, P. (1991) *J. Appl. Crystallogr. 24*, 946–950.

Leatherbarrow, R. J., Fersht, A. R., & Winter, G. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 7840–7844.

Lenz, A., Cordes, F., Heinemann, U., & Saenger, W. (1991) *J. Biol. Chem. 266*, 7661–7667.

Leslie, A. G. W. (1990) in *Crystallographic Computing*, Oxford University Press, Oxford, U.K.
Martinez-Oyanedel, J., Choe, H.-W., Heinemann, U., & Saenger, W. (1991) *J. Mol. Biol. 222*, 335–352.
McPherson, A. (1982) *The Preparation and Analysis of Protein Crystals*, John Wiley & Sons, New York.
McPherson, A., Brayer, G., Cascio, D., & Williams, R. (1986) *Science 232*, 765–768.
Meiering, E. M., Bycroft, M., & Fersht, A. R. (1991) *Biochemistry 30*, 11348–11356.
Meiering, E. M., Serrano, L., & Fersht, A. R. (1992) *J. Mol. Biol. 225*, 585–589.
Meiering, M., Bycroft, M., Lubienski, M. J., & Fersht, A. R. (1993) *Biochemistry 32*, 10975–10987.
Mossakowska, D. E., Nyberg, K., & Fersht, A. R. (1989) *Biochemistry 28*, 3843–50.
Nicholls, A. (1992) *GRASP, Graphical Representation and Analysis of Surface Properties*, Columbia University, New York.
Pavlovsky, A. G., Borisova, S. N., Strokopytov, B. V., Sanishvili, R. G., Vagin, A. A., & Chepurnova, N. K. (1988) *Metabolism and Enzymology of Nucleic Acids Including Gene Manipulations*, Plenum Press, New York.
Saenger, W. (1983) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.
Sali, D., Bycroft, M., & Fersht, A. R. (1988) *Nature (London) 335*, 496–500.
Serrano, L., Horovitz, A., Avron, B., Bycroft, M., & Fersht, A. R. (1990) *Biochemistry 29*, 9343–9352.
Sevcik, J., Sanishvili, R. G., Pavlovsky, A. G., & Polyakov, K. M. (1990) *Trends Biochem. Sci. 15*, 158–162.
Sevcik, J., Dodson, E. J., & Dodson, G. G. (1991) *Acta Crystallogr. B47*, 240–253.
Uscher, D. A., Richardson, D. I., & Eckstein, F. (1970) *Nature (London) 228*, 663–665.
Wilson, A. J. C. (1949) *Acta Crystallogr. 2*, 318.
Wlodawer, A., Miller, M., & Sjölin, L. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3628–3631.