UC Irvine

UC Irvine Previously Published Works

Title

Crystal structure disposition of thymidylic acid tetramer in complex with ribonuclease A.

Permalink

<https://escholarship.org/uc/item/74g9m6ns>

Journal

The Journal of biological chemistry, 267(31)

ISSN

0021-9258

Authors

Birdsall, DL
McPherson, A

Publication Date

1992-11-01

DOI

10.1016/s0021-9258(18)41659-6

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, availalbe at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Crystal Structure Disposition of Thymidylic Acid Tetramer in Complex with Ribonuclease A*

(Received for publication, April 20, 1992)

David L. Birdsall‡§ and Alexander McPherson¶

From the ‡Division of Biomedical Sciences, Lawrence Livermore National Laboratory, University of California, Livermore, California 94550 and the ¶Department of Biochemistry, University of California, Riverside, California 92521

The crystal structure of ribonuclease A with bound thymidylic acid tetramer is reported at 2.5-Å resolution. The diffusion of the tetramer into native orthorhombic crystals of the ribonuclease allows for the formation of a structurally stable complex where the single-stranded nucleic acid enters and leaves the enzyme's catalytic region in a persistent 5'-3' direction. The binding of the tetramer to the enzyme's surface is facilitated and mediated by electrostatic interactions between basic protein residues and nucleotide phosphates. Two pyrimidine nucleotides are bound to the enzyme's active site in a manner similar to that observed for other complexes between ribonuclease A and nucleic acid oligomers.

The x-ray structure of a complex of ribonuclease A (RNase A)¹ with the single-stranded DNA oligomer d(pT)₄ is described in this paper. This study extends the list of crystallographically studied complexes between RNase and nucleic acids by the inclusion of a complex with a pyrimidine oligomer larger than a dinucleotide and further illustrates the role of a pyrimidine nucleotide bound at the active site for RNase catalysis.

Bovine pancreatic ribonuclease A is a protein of molecular mass 13,800 (daltons) and sequence length of 124 amino acids. Its enzymatic activities, primarily associated with the degradation of RNA polymers, have been previously reviewed (Anfinsen and White, 1961; Smyth *et al.*, 1963; Richards and Wyckoff, 1971). It is a member of a class of enzymes that catalyze the hydrolysis of the 3',5'-phosphodiester linkage in polynucleotide RNA (Brown and Todd, 1955). The enzyme preferentially cleaves RNA following pyrimidine residues, which are resolved into terminal pyrimidines with 3'-phosphate groups (Usher *et al.*, 1972).

Catalytically important residues at the RNase active site are Lys-41, His-12, and His-119 as inferred from studies of

the stereochemistry of catalysis (Usher *et al.*, 1970, 1972) and from complexes of the enzyme with inhibitors (Richards and Wyckoff, 1971; Wodak *et al.*, 1977; Pavlovsky *et al.*, 1978). The side chains of His-12 and His-119 mediate proton transfers with the 2'-oxygen of the pyrimidine ribose to form and resolve a 2',3'-cyclic nucleotide intermediate, and Lys-41 is thought to stabilize a pentacoordinate phosphorus intermediate involving the 2', 3', and (vicinal) 5'-oxygen with the two phosphate oxygens.

The RNase sites for binding of the pyrimidine ribose and base and the positioning of the phosphate (involved in the hydrolysis) 3' to the pyrimidine nucleotide have been well characterized within the enzyme's active site cavity (Richards and Wyckoff, 1971; Borkakoti, 1983). A purine nucleotide, in particular adenosine, preferably contributes its 5'-phosphate for hydrolysis. The chemical, physical, and structural properties of RNase A and RNase S (formed by enzymatic cleavage of the peptide bond between residues Ala-20 and Ser-21 of RNase A) have also been extensively reviewed (Anfinsen and White, 1961; Richards and Wyckoff, 1971; Blackburn and Moore, 1982; Wlodawer, 1984).

Numerous studies of three-dimensional structures from x-ray analyses of crystals of the native protein and of complexes between the protein and single-stranded mono-, di-, and oligonucleotides and various nucleotide inhibitors have been reported (see review by Wlodawer (1985) and references cited therein; McPherson *et al.*, 1986a, 1986b; Campbell and Petsko, 1987; Wlodawer *et al.*, 1988; Nachman *et al.*, 1990). The enzyme cannot cleave DNA due to the absence of 2'-hydroxyl groups on the nucleotides. However, RNase can bind to DNA with an affinity comparable to that for RNA (Sekine *et al.*, 1969; Walz, 1971), and denatured DNA acts as a competitive inhibitor for the enzyme's hydrolysis of substrate analogues and RNA itself (Sekine *et al.*, 1969). Therefore, it has been concluded that single-stranded DNA forms a stable complex with RNase and probably binds to the enzyme at the active site by mechanisms similar to RNA binding (McPherson *et al.*, 1986b).

It is of interest to determine the structural characteristics of the binding of polymeric and oligomeric single-stranded DNAs to RNase as a means of inferring methods of RNase interactions with, and catalysis on, polymeric RNA substrates. Such studies help to elucidate the nature (efficiency, stability, stereospecificities, etc.) of protein and enzyme interactions with nucleic acids of nonspecific base sequence. The present study involving the association of RNase A with a pyrimidine oligomer offers further insights into nonspecific protein/nucleic acid interactions and enzyme catalysis.

EXPERIMENTAL PROCEDURES

Crystallizations and Complex Formation—Native RNase A crystals were grown in unbuffered polyethylene glycol 4000 (PEG 4000) so-

* This work was performed under the auspices of the United States Department of Energy by Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48 and supported in part by National Institute of General Medical Science National Research Service Award Fellowship 5 F32 GM11235 (to D. L. B.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors have been deposited in the Protein Data Bank, Brookhaven National Laboratory, Upton, NY.

§ To whom correspondence should be addressed. Tel.: 510-422-3097; Fax: 510-422-2282.

¹ The abbreviations used are: RNase, ribonuclease A; d(pA)₆, hexadeoxyphosphoadenosine; d(pA)₄, tetradexyphosphoadenosine; d(pT)₄, tetradexyphosphothymidine; PEG, polyethylene glycol; rms, root-mean-square.

lutions using vapor diffusion methods (McPherson, 1982) as follows. Bovine RNase A (Sigma) was dissolved in distilled water to a concentration of 60 mg/ml with no further purification. Crystallizations were carried out using siliconized 9-well depression plates sealed in plastic boxes and stored for equilibration at 4 °C. Reservoir solutions (25 ml) consisted of 28% PEG 4000 and 0.02% azide. Each well sample contained 7 μ l of RNase A solution combined with 3 μ l of 5% β -octyl glucoside (Sigma), a nonionic detergent, and 7 μ l of reservoir solution. Native protein crystals appeared within 2 weeks and were allowed to grow for about 4 months. They were judged to be typical of the orthorhombic prismatic habit seen for co-crystals of the RNase A·d(pA)4 complex (Brayer and McPherson, 1982). X-ray precession photographs determined that they belonged to the space group P2₁2₁2₁. Reflections from three native crystals were used to solve the protein structure. The crystal yielding the most extensive set of reflections had unit cell lengths of $a = 44.9$ Å, $b = 74.9$ Å, and $c = 44.0$ Å with 1 molecule/asymmetric unit.

Particular care was made during nucleic acid diffusion attempts so as not to perturb the physical structure of the native protein crystals. A low mole ratio of d(pT)4 tetramer to RNase A (2:9) was eventually employed. To form the RNase A·d(pT)4 complex within the crystals the tetramer (McPherson *et al.*, 1986c) was dissolved in 27% PEG 4000 and adjusted to a pH of 7.0 at a concentration of 1.0 μ g/ μ l. The d(pT)4 solution was added to wells containing native crystals in four steps of 2 μ l each over 5 days. The complex was allowed to form for over 1 month at 4 °C before crystals were selected for data collection. A single crystal with unit cell lengths of $a = 44.9$ Å, $b = 75.0$ Å, and $c = 43.6$ Å yielded the most extensive data set and was used for the structure determination and refinement of the RNase A·d(pT)4 complex.

Data Collection—An Enraf Nonius CAD-4 automated diffractometer with an extended counter arm and helium-filled path was used for all data collection. The x-ray generator was operated at 40 kV and 40 mA with a fine focus x-ray tube and omega scan rate of 1.33°/min. Data were collected at 16 °C. For each data set 32 ϕ independent standard reflections were used to generate an empirical absorption correction curve applied to all reflections. Two standard reflections, measured at intervals of about 2 h of x-ray exposure, were used to generate a time-dependent crystal decay curve which was applied to all data. Net intensities were determined by subtracting from the measured peak intensity twice the sum of the 16 left and 16 right nearest neighbors in reciprocal space.

Crystals of the native protein yielded a collective data set of 4466 independent reflections to 2.5-Å resolution at better than 2 σ . A crystal for the RNase A·d(pT)4 complex was used to collect 4347 independent reflections to 2.5 Å at better than 2 σ . The integrated intensities for both data sets were corrected for Lorentz-polarization effects and absorption; backgrounds were subtracted. Reflection intensities were converted to structure factors for use in refinement programs.

Structure Solutions and Refinements—A model for the RNase A structure, based on a combination of x-ray diffraction and neutron diffraction data (Wlodawer and Hendrickson, 1982; Wlodawer and Sjolin, 1983) from crystals not grown in the presence of β -octyl glucoside, had previously been used as a molecular replacement model to position the RNase A molecule in the P2₁2₁2₁ unit cells of native protein crystals with the typical prismatic habit seen of RNase A crystals grown in PEG 4000 (Brayer and McPherson, 1982). This model, with its positioning in the orthorhombic unit cell, was initially used to solve the structure of the enzyme using reflection data from native crystals of the present study. A refinement summary is given in Table I. The general strategy involved initial least-squares rigid body and atomic refinements under CORELS (Sussman, 1985) followed by energy minimizations and simulated annealing refinements under X-PLOR (Brunger, 1988). X-PLOR refinements for the complex typically resulted in a loss of perfect thymine base planarity, so they were followed by replacement of bases with planar ones and further CORELS positional refinements leading to the final structure. Refinements were interspersed with calculations of difference Fourier maps using the FFT algorithm (TenEyck *et al.*, 1976), and model fitting under the molecular modeling programs Frodo (Jones, 1982) and Quanta (Polygen Corp., Waltham, MA), as implemented on Evans and Sutherland PS330 and PS390 and Silicon Graphics Iris graphics stations. A refined structure of the native protein (in β -octyl glucoside solvent) was then used to determine the location of nucleic acid for the RNase A·d(pT)4 complex, followed by tetramer fitting in difference Fourier maps and complex refinement as outlined above. Rotation function and translation function searches under X-PLOR (Brunger, 1990), using the refined structure for RNase A (in β -octyl

TABLE I
Refinement summary

	Native	Complex
Resolution	10.0–2.5 Å	10.0–2.5 Å
No. reflections	4352	4218
No. atoms (non-hydrogen)	951	1031
<i>r</i> factor ^a	0.228	0.235
Average deviations from ideality		
All distances (Å)	0.071	0.075
rms bond lengths (Å)	0.022	0.024 (protein), 0.043 (DNA)
rms bond angles	4.2	4.5 (protein), 5.4 (DNA)
Protein atomic b-factors (Å ²)		
Average backbone	7.52	6.02
Average side chain	8.42	6.54
Highest	40.0	35.7
Molecular replacement ^b parameters		
Cross-rotation		
PC value ^c	0.227	0.222
<i>R</i> function, σ	2.26, 0.107	2.26, 0.106
Translation		
<i>p</i> value ^c	0.390	0.391
<i>T</i> function, σ	0.776, 0.037	0.776, 0.041

$$^a r \text{ factor} = \frac{\sum |F(\text{obs})_{hkl}| - |F(\text{calc})_{hkl}|}{\sum |F(\text{obs})_{hkl}|}$$

^b 10.0–3.5 Å.

^c Refer to Brunger, 1990.

TABLE II
Stages of refinement for RNase A·d(pT)4 complex

- 1 CORELS rigid body and atomic positional refinements of native RNase A structure with diffraction data from the RNase A·d(pT)4 complex. *R* factor is 0.251 at 3.0 Å resolution.
- 2 Generation of 2-Fo-Fc map and first attempt to fit in DNA by locating possible phosphate positions. At this point the difference density map is very noisy. *R* factor is 0.265 at 3.0-Å resolution.
- 3 Refinement of RNase A·d(pT)4 complex and generation of Fo-Fc omit map using protein phases only, followed by refitting of the DNA and further refinement. The Fo-Fc omit map yielded the first clear evidence of active site density from the tetramer. *R* factor is 0.258 at 3.0-Å resolution.
- 4 Generation of Fo-Fc omit map and tetramer refitting, followed by further refinement at an extended resolution range. *R* factor is 0.249 at 2.7-Å resolution.
- 5 After more tetramer refitting, the resolution range of CORELS refinements is extended to 2.5 Å. The refined model is used to initiate rounds of X-PLOR energy minimization and simulated annealing refinements followed by replacement of bases for perfectly planar ones and further CORELS refinements, resulting in much cleaner difference density maps. *R* factor is 0.239 at 2.5-Å resolution.
- 6 Generation of Fo-Fc omit map, tetramer refitting, and further CORELS refinements to the final structure. The final Fo-Fc map is very sparse and clean with tetramer density clearly seen. 2Fo-Fc maps for the final model reveal density from a β -octyl glucoside molecule associated with the protein. *R* factor is 0.235 at 2.5-Å resolution.

glucoside solvent) as the molecular replacement model, were done to confirm the correct positioning of the protein in the native and complex unit cells (Table I). Table II gives a summary of the refinement stages for the RNase A·d(pT)4 complex. Throughout the text the individual nucleotides of the d(pT)4 tetramer are referred to by using the following notation convention: 5'-T1-T2-T3-T4-3'.

RESULTS AND DISCUSSION

Electron Density Fitting—The final Fo-Fc difference Fourier “omit maps” (based on phases from the protein refined alone) for the RNase A · d(pT)4 complex were very clean and sparse and showed densities which could be associated with only 1 molecule of d(pT)4 (Fig. 1). No attempts were made to fit nucleic acid moieties from possible oligomers with low occupancies. Electron densities were particularly clear near and within the enzyme’s active site cavity and near catalytically important protein residues. The nucleic acid moieties of the 3′-nucleotide, including the phosphate at the P1 site, were well fitted to difference densities for the final structure. Positive electron densities from $F_o - F_c$ maps for most of the d(pT)4 tetramer were persistent features. The T2 nucleotide may be partially disordered since some deoxyribose and base atoms show unclear or no $2F_o - F_c$ density. The T3 phosphate is associated with weak $F_o - F_c$ “omit map” density and a high temperature factor (Table III), but its position is supported by the stereochemistry of the tetramer and its density is evident in $2F_o - F_c$ maps from phases for the final model (Fig. 2).

$2F_o - F_c$ maps based on phases from the refined complex reveal density resulting from a β -octyl glucoside molecule within the asymmetric unit near the side chains of protein residues Gln-69, Asn-71, and Glu-111. Although the detergent molecule was not used in refinements nor for R factor and electron density map calculations, density for the D-glucopyranoside moiety clearly shows not only the general contours of the pyranoside but also the location of the terminal C6-O6 group. No density, however, can be discerned for the anomeric linkage and the alkyl group, and the density for the sugar is much attenuated in the more stringent Fo-Fc omit maps. A glucopyranoside molecule, taken from the x-ray crystal structure of 1-decyl- α -D-glucopyranoside (Moews and Knox, 1976) with the O1 (anomeric) oxygen placed in a β orientation, can be convincingly fitted (as a rigid body) into the density (Fig. 3). The original α linkage would direct the alkyl group away from the protein; however, a β linkage directs the alkyl group along the side of the protein surface to facilitate intermolecular hydrophobic interactions. The proximity of the detergent molecule may affect the orientations of nearby protein residues and short contacts seen between these residues and the 3′-terminal thymine base, since a manual fitting of the glucopyranoside places its O6 atom around 3 Å from both the NE2 nitrogen of Gln-69 and the OD1 oxygen of Asn-71. Similar difference density is not seen for the native RNase A structure (in β -octyl glucoside solvent), suggesting that the diffusion of d(pT)4 into native crystals may have aided the

trapping of the detergent molecule in the crystalline complex.

Protein Structure—The inclusion of β -octyl glucoside in the crystallization medium for RNase A allows for the diffusion of single-stranded DNA oligomers larger than dinucleotides into native crystals without conformational disruption or denaturation of the protein molecule. Major conformational changes involving the polypeptide backbone for the RNase A molecule in the protein-nucleic acid complex were avoided and none were necessary during the refinements. Backbone conformations for the refined native and complex protein structures were virtually identical and only slight changes in the positioning of side chain atoms were found. The average rms differences between the proteins of the refined native and complex structures were for the backbone and side chain atoms 0.48 and 1.2 Å, respectively. Noticeable changes in the orientations of some amino acid side chains, especially those involved in tetramer binding, were observed. Very few protein-protein intermolecular contacts were found for the native and complex crystal structures.

The fractional unit cell volume occupied by protein in the native RNase A crystals of the present study, 0.433, results in an overall volume to mass ratio, V_m (Matthews, 1968), of 2.69 Å³/dalton. For crystals of the RNase A · d(pT)4 complex this ratio (protein + DNA) decreases to 2.48 Å³/dalton, assuming that only 1 d(pT)4 molecule associates with each RNase A molecule. The corresponding fractional unit cell volumes are 0.434 for protein and 0.030 for DNA. The calculated V_m values for the native and complex crystals are higher than average (2.15 Å³/dalton) but are within normal limits for small proteins (Matthews, 1968). Therefore, β -octyl glucoside allows for the growth of RNase A crystals with a greater fractional volume of solvent than otherwise.

Protein residues Lys-41, His-12, and His-119 appear most crucial for the configuration of the nucleotides within the active site region of the complex. The side chains of Lys-41 and His-12 moved little during refinement of their positions in the native structure, a lack of repositioning also reported for the RNase A · d(pA)4 complex (McPherson *et al.*, 1986b). A noticeable torsional adjustment of the imidazole ring of residue His-119 from the present study was observed. Its side chain, however, remains in the minor position for His-119 (Borkakoti, 1983).

Nucleic Acid Structure—There are some structural similarities between the nucleic acid of the present RNase A · d(pT)4 complex and that for the co-crystalline complex of RNase A with four d(pA)4 oligomers (McPherson *et al.*, 1986a). Each of the four deoxyribose sugars refine to within the C3′-endo conformational family of sugar puckers (see Table III) ac-

FIG. 1. Stereoview of the Fo-Fc difference electron density map at the d(pT)4 region of the complex. Bias in the map was minimized by 50 cycles of positional refinement for the protein without phases from the tetranucleotide. The omit map is contoured to the 3 σ level at 3.0 Å. Some amino acid residues involved with nucleic acid binding are also shown. They are (running clockwise from the top) Arg-39, Lys-66, His-119, His-12, and Lys-41.

cording to calculated pseudorotation angles. The C2'-exo sugar pucker observed for two of the nucleotides is not unexpected. A C2'-exo deoxyribose conformation was found in the crystal structure of a deoxythymidine covalently bound to RNase A at residue His-12 (Nachman *et al.*, 1990).

Group temperature factors for the DNA moieties (phosphate, deoxyribose, base) of the d(pT)4 tetramer compare favorably to those of the most firmly bound tetramers in the co-crystalline RNase A·d(pA)4 complex (McPherson *et al.*, 1986b). In particular, the thermal parameters of the 3'-D1-C4-C3-C2-5'-phosphates of the co-crystalline complex are similar to the B factors for the 3'-T4-T3-T2-T1-5'-phosphates, with the T3 phosphate B factor (like that for the C4 phosphate) being the highest of the four. As suggested for the RNase A·d(pA)4 complex, the thermal parameters associated with the d(pT)4 tetramer probably reflect statistical disorder for the nucleotides. Although the crystalline complex was formed under physiological conditions, many of the protein-nucleic acid interactions appear electrostatic in nature and are also similar to those observed in crystals of low ionic strength for complexes between RNase and single-stranded nucleic acid oligomers (Wodak *et al.*, 1977; McPherson *et al.*, 1986b). Each of the thymine bases of the d(pT)4 tetramer have different plane normals and there is no evidence of base stacking, in contrast to the adenosine bases of a d(pA)4 oligomer bound at the active site for the RNase A·d(pA)4 complex (McPherson *et al.*, 1986b). The thymidine tetramer is associated with the RNase A molecule in a partially extended manner as similarly defined by Webb *et al.* (1973). Nearest neighbor distances between C2 base atoms range from

8.9 Å to 10.1 Å. There is no stacking of the T3 base to the phenyl moiety of the Phe-120 side chain. The tetramer can be divided into essentially two regions, an extended dinucleotide bound at the active site (T3-T4) and a more contracted dinucleotide (T1-T2).

General Protein-DNA Interactions—The path of the d(pT)4 tetramer across the protein surface is roughly similar to that followed by nucleotides C2 to C4 and continued by the 5'-phosphate of tetramer D in the RNase A·d(pA)4 complex. In both cases the majority of intermolecular contacts involve binding of single-stranded DNA oligomers to a single protein molecule in the asymmetric unit. The d(pT)4 tetramer runs from the 5' terminus into the RNase A active site cleft with the next to last nucleotide (T3) and the 3'-terminal phosphate adjacent to catalytically important amino acid residues (Fig. 4). The 3'-end of the d(pT)4 tetramer is anchored to the protein with the T4 phosphate occupying the P1-binding site (Borkakoti, 1983). The T4 thymine base and deoxyribose occupy the respective B2 and R2 sites of the enzyme (Wyckoff *et al.*, 1970). The ribose and base of the T3 nucleotide are found at the corresponding pyrimidine R1- and B1-binding sites (Borkakoti, 1983). Both nucleotides from the 3'-end are involved in intermolecular contacts with side chain from amino acid residues at the enzyme's active site.

The disposition of the d(pT)4 tetramer probably reflects the complementarity of positive charges and nucleoside-binding sites on the surface of the protein with negatively charged phosphate groups and bases, as was concluded for the RNase A·d(pA)4 complex (McPherson *et al.*, 1986a). These electrostatic interactions are of primary importance for complex formation since similar complexes form between RNase and heteronucleotides devoid of heterocyclic bases (Karpel *et al.*, 1981). The persistent 5'-3' orientation of the tetramer is common for many dinucleoside monophosphate crystalline complexes involving RNase A and RNase S molecules (Richards and Wyckoff, 1973; Wodak *et al.*, 1977; Wlodawer, 1984), and would be as required for catalysis on RNA substrates (Richards and Wyckoff, 1971; Blackburn and Moore, 1982; Wlodawer, 1984).

Many direct interactions between d(pT)4 and RNase A involve the nucleic acid backbone. The location of the d(pT)4 tetramer supports, in general, the positioning of a virtually continuous strand of eight nucleotides suggested by two

TABLE III
Nucleotide stereochemistry

Nucleotide	T1	T2	T3	T4
Sugar pucker	C3'-endo	C2'-exo	C2'-exo	C3'-endo
Pseudorotation Angle ^a	0.813	327.0	340.0	12.90
Base ^b	Anti	Syn	Anti	Anti
Group b-factors				
Phosphate	26.1	27.1	68.7	13.5
Deoxyribose	38.9	37.0	21.9	10.9
Base	51.8	47.2	38.2	44.2

^a Refer to Altona and Sundaralingam, 1972.

^b Determined from torsion angle about glycosyl bond, defined as $\chi:O1'-C1'-N1-C2$.

FIG. 2. Stereoview of the complex looking out from the RNase A active site. The T3 nucleotide and the T4 phosphate and deoxyribose can be seen. The 2Fo-Fc map is contoured to the 1 σ level at 3.0 Å.

FIG. 3. Stereoview of the positioning of a β -D-glucopyranoside moiety within the crystalline complex. The sugar unit was fitted as a rigid body into difference electron density lying near the side chains of protein residues Gln-69, Asn-71, and Glu-111. The 2Fo-Fc map is derived from protein and nucleic acid phases of the final model and is contoured to the 1σ level at 2.5 Å.

FIG. 4. Stereoview of the disposition of d(pT)4 tetramer in complex with RNase A. The α -carbon backbone of the protein is shown with the nucleic acid tetramer. The 5'-end of the tetramer is on top.

d(pA)4 oligomers bound at the active site cleft and emerging from near the pyrimidine-binding site, as postulated by McPherson *et al.* (1986a). In this way a "virtual DNA strand" can be intramolecularly associated with the protein as the general nature of single-stranded nucleic acid binding to RNase A. The influence of the enzyme in the binding of the nucleic acid extends throughout the length of the tetramer, particularly at the active site region, and directly for at least three of the four nucleotides of the tetramer. The T2 phosphate is most distant from basic protein residues within the complex, and the T2 nucleotide is the most awkwardly situated unit of the tetramer with its thymine base adopting a slightly syn conformation ($\chi = -35^\circ$) to its deoxyribose.

Specific Protein-DNA Interactions—The extension within the d(pT)4 tetramer allows the 5'-phosphate to form a salt bridge with the guanidinium group of Arg-39, similar to that formed by the C1 nucleotide's phosphate in the RNase A·d(pA)4 complex (McPherson *et al.*, 1986b), even though the d(pT)4 tetramer runs into and out of the active site. The C5 methyl group of the 5' base lies less than 3.5 Å from the NH2 nitrogen of Arg-39. All 3 enzyme residues associated with catalytic activity (Richards and Wyckoff, 1971) are involved in the binding of the thymidine tetramer (Table IV). The oxygens of the phosphate bound at the P1 site have intermolecular contacts to the NE2 nitrogen of protein residue His-12 and the NZ nitrogen of residue Lys-41. The His-119 residue influences the positioning of the P1 site phosphate by forming electrostatic contacts with the O3'- and O5'-oxygens

TABLE IV
Possible protein/DNA contacts within the RNase A·d(pT)4 complex

Protein residue	Atom	DNA nucleotide	Atom	Distance Å
His-12	NE2	T3	O2	3.0
Arg-39	NE	T1	C5M	3.1
Arg-39	NH2	T1	O1P	2.9
Arg-39	NH2	T1	O3P	3.1
Cys-40	N	T1	O2P	2.8
Lys-41	NZ	T3	O1'	3.0
Asn-71	ND2	T4	O4	2.9
Asn-71	ND2	T4	C5M	2.4
Glu-111	OE1	T4	C5M	2.3
His-119	ND1	T3	O3'	2.7
His-119	ND1	T4	O2	3.0
His-119	NE2	T4	O2	3.3
Phe-120	N	T4	O2P	3.1

of nucleotides T3 and T4, respectively.

The O1'-atom of the deoxyribose bound at the R1 site (T3) is near the NZ nitrogen of Lys-41, and the O5'-oxygen from this nucleotide may also be involved in a contact with the Lys-41 side chain. The O2 atom from the thymine base bound at the B1 site (T3) is close to the NE2 nitrogen of His-12. The NZ nitrogen of Lys-66 lies 5.4 Å from the O4 oxygen of the T3 nucleotide's base, and less than 4 Å from the C5 methyl group of the same base. The O3'-oxygen of nucleotide T3 is associated with the (T4) phosphate bound at the P1 site

and is in a position similar to the O8'-oxygen from the covalently bound thymidine nucleotide in the T-H12·RNase A crystalline complex (Nachman *et al.*, 1990). This O3'-atom (present complex) is under 3 Å from the ND1 nitrogen of His-119 within the complex. Apparently, the T3 nucleoside has sufficient mobility to allow positioning of the (T3) thymine base into the B1 site, as opposed to the T-H12·RNase A complex.

The NE2 nitrogen of Glu-111 is located within hydrogen bonding distance from one of the phosphate oxygens of the T4 nucleotide, similar to that observed for some RNase A-nucleotide inhibitor complexes (Borkakoti, 1983). Contacts of under 3.6 Å between the phosphate oxygens of the 3'-nucleotide, the OE1 atom from the side chain of Gln-11, and the backbone nitrogen of residue Phe-120 are observed, similar to the crystalline complex of RNase S with cytidyl (2', 5') adenosine (Wodak *et al.*, 1977).

The side chain of protein residue Asn-67 is involved in contacts with the 3'-deoxythymidine (T4) similar to those found for the covalently attached thymidine of the T-H12·RNase A complex. The distances between the O4 oxygen of the T4 thymine base to both the ND2 nitrogen of Asn-67 and the NE3 nitrogen of Gln-69 are about 5 Å. The C5 methyl group of the T4 base is very close to the ND2 atom of Asn-67 and is less than 4.5 Å from the OE2 oxygen of the same protein residue. Also similar to the T-H12·RNase A complex, the (T4) O4 oxygen may be hydrogen bonded to the side chain of residue Asn-71. The side chain of His-119 may stabilize the location of the T4 base, the O2 oxygen of the this base lying relatively close to the ND1 and NE2 nitrogens of His-119.

The syn conformation for a pyrimidine nucleotide is not commonly observed (Saenger, 1984); the crystallization of 4-thiouridine from aqueous solution has lead to a reported syn conformation of the base due to the inclusion of water of hydration in the crystal lattice (Saenger and Scheit, 1970; Saenger, 1984), but most pyrimidine nucleotides are found in the anti conformation. For the present RNase A·d(pT)4 complex the T2 nucleotide is found within the syn conformational range (Table III). This may have been influenced by a number of interactions of the nucleotide with nearby charged amino acid side chains within the complex. The NZ nitrogen of Lys-66 may form a contact (under 5 Å) with the O1'-atom from the T2 deoxyribose. The O2 oxygen of the T2 thymine base lies under 6 Å from the NZ nitrogen of Lys-66. The O4 oxygen and C5 methyl group of this base are under 7 Å from the ND1 and NE2 nitrogens of the His-105 residue from a neighboring protein molecule. The O4 atom also lies less than 8 Å from the OD1 oxygen of the Asn-113 residue from a different symmetry related protein molecule and is about 9 Å from the ND2 nitrogen of the same protein residue. A noticeable peak of difference Fourier electron density, which may be attributable to a solvent molecule, lies between the Lys-66 and Asn-67 side chains and the T2 deoxyribose within the complex, possibly influencing the syn conformation. Attempts to reposition the T2 and T3 nucleotides so that the T3 phosphate could occupy this density resulted in structurally inferior models with higher *R* factors. The syn orientation of the T2 base was a persistent feature of the model during refinements and reoccurred even after adjusting the glycosyl bond's torsion angle (χ) to well within the anti range.

Comparisons to Other Complexes—No association of nucleic acid to the region of a cluster of basic amino acids, the "deep anion trap" (McPherson *et al.*, 1986b), was detected in contrast to the RNase A·d(pA)4 complex. This may be due to

the relatively neutral pH of the d(pT)4 solution used for diffusion into native crystals. Such a pH may inhibit electrostatic interactions with residues from this electropositive cavity (Lys-31, Lys-37, Lys-91, Arg-33). The binding of the d(pA)4 oligomers to this region may also have been enhanced by the co-crystallization of the complex. Binding to the deep anion trap region of RNase A has not been reported for other complexes formed by diffusion of single-stranded nucleic acid monomers and oligomers into native crystals (Borkakoti, 1983; Wlodawer *et al.*, 1983), even when the pH of the crystallization medium is 5.5 (Wodak *et al.*, 1977).

The 3.1-Å structure of a complex formed by the diffusion of d(pT)4 into native crystals of the Klenow fragment of DNA polymerase I suggests that the pyrimidine oligomer binds to the 3'-5'-exonuclease site with splayed thymine bases, anti-conformations, and C3'-endo sugars (Beese and Steitz, 1991). Unlike the polymerase I complex, where binding of the d(pT)4 tetramer to the protein seems uninfluenced by the base type, for the RNase A·d(pT)4 complex there are specific interactions between protein residues and base atoms. For example, probable van der Waals interactions involving the 5'- and 3'-C5 methyl groups allow for short contacts which may stabilize tetramer binding. The 3'-thymine base of the RNase A·d(pT)4 complex lies within the B2 or purine base-binding site. Its C5 methyl group appears to function in a role similar to that of the N1 and N6 atoms of an adenine base bound at the B2 site. The crystal structure of cytidyl (2', 5') adenosine complexed with RNase S, for example, suggests that the N1 atom is protonated and forms a hydrogen bond to the side chain of Glu-111 and the N6 atom may hydrogen bond to the side chain of Asn-71 (Wodak *et al.*, 1977). Short contacts which cannot be relieved by further refinements without loss of model integrity occur between the C5 methyl group of the 3'-nucleotide (T4) and both the OE1 oxygen of Glu-111 and the ND2 nitrogen of Asn-71 in the RNase A·d(pT)4 complex. While interactions involving the thymine C5 methyl group are generally expected to be hydrophobic in nature, Kissinger *et al.* (1990) report an important sequence-specific van der Waals contact within the major groove of DNA in the 2.8-Å crystal structure of an engrailed homeodomain-DNA complex. Here the interaction is between the side chain of residue Gln-50 and the C5 methyl group from the thymine of base pair 14. It is formed in preference to making a hydrogen bond with the O4 oxygen of the thymine base. The theoretical calculation of a partial atomic charge of -0.42 for the C5 methyl carbon of deoxythymidine (Pearlman and Kim, 1990) implies that under some circumstances the C5 methyl group may be available for such van der Waals interactions.

Low resolution (4 Å) x-ray diffraction studies on crystalline complexes between RNase S with thymidine 3',5'-diphosphate and adenylyl-3',5'-cytidine suggest the presence of a nonspecific phosphate-binding site near protein residue Lys-66 (Sawada and Irie, 1969; Mitsui *et al.*, 1978; Iwahashi *et al.*, 1981). A phosphate for the nucleotide 5' to the p1 site could occupy this position, termed site p0. The location of the p0 site was determined in part from coordinates of the Lys-66 residue elucidated crystallographically for uncomplexed RNase S (Wyckoff *et al.*, 1970; Iwahashi *et al.*, 1981). It was concluded that one of the phosphate oxygens at the p0 site may form a hydrogen bond with the NZ nitrogen of Lys-66 (Iwahashi *et al.*, 1981). In the present RNase A·d(pT)4 complex the T3 phosphate oxygens are not close enough to form such a hydrogen bond. The T3 phosphate lies closer to the side chain of residue Lys-41, with one of its oxygens being under 5 Å from the NZ nitrogen of Lys-41. The p0 site appears to be occupied by the most 3'-phosphate of one of the d(pA)4

oligomers bound near the active site in the RNase A · d(pA)4 co-crystallization complex (McPherson *et al.*, 1986b). However, for that complex the p0 and p1 sites are not occupied by phosphates from the same tetramer. The p0 site itself cannot fix the target phosphate group and does not attract free anions (such as sulfate ions) for the uncomplexed RNase molecule (Iwahashi *et al.*, 1981). Therefore, occupation of the p0 site by nucleic acid moieties may be limited to phosphates from the 5' or 3' terminus of single-stranded DNA/RNA oligomers, where there is greater conformational flexibility.

Acknowledgment—We thank the San Diego Supercomputer Center for allowing us computing time on the Cray Y-MP8/864.

REFERENCES

- Altona, C. & Sundaralingam, M. (1972) *J. Am. Chem. Soc.* **94**, 8205–8212
- Anfinsen, C. B. & White, F. H., Jr. (1961) *The Enzymes* (Boyer, P. D., ed) Vol. 5, 2nd ed., Academic Press, New York
- Beese, L. S. & Steitz, T. A. (1991) *EMBO J.* **10**, 25–33
- Blackburn, P. & Moore, S. (1982) *The Enzymes* (Boyer, P. D., ed) Vol. 15, 3rd ed., Academic Press, New York
- Borkakoti, N. (1983) *Eur. J. Biochem.* **132**, 89–94
- Brayer, G. & McPherson, A. (1982) *J. Biol. Chem.* **257**, 3359–3361
- Brown, D. M. & Todd, A. R. (1955) in *The Nucleic Acids* (Chargaff, E. & Davidson, G. N., eds) pp. 409–445, Academic Press, New York
- Brunger, A. T. (1988) *J. Mol. Biol.* **203**, 803–816
- Brunger, A. T. (1990) *Acta Crystallogr. Sect. A* **46**, 46–57
- Campbell, R. L. & Petsko, G. A. (1987) *Biochemistry* **26**, 8579–8584
- Iwahashi, K., Nakamura, K., Mitsui, Y., Ohgi, K. & Irie, M. (1981) *J. Biochem. (Tokyo)* **90**, 1685–1690
- Jones, T. A. (1982) in *Computational Crystallography* (Sayer, D., ed) pp. 307–317, Oxford University Press, New York
- Karpel, R. L., Yrttimaa, V. A. & Patel, G. L. (1981) *Biochem. Biophys. Res. Commun.* **100**, 760–768
- Kissinger, C. R., Liu, B., Martin-Bianco, E., Kornberg, T. B. & Pabo, C. O. (1990) *Cell* **63**, 579–590
- Matthews, B. W. (1968) *J. Mol. Biol.* **33**, 491–497
- McPherson, A. (1982) *The Preparation and Analysis of Protein Crystals*, pp. 94–97, John Wiley & Sons, New York
- McPherson, A., Brayer, G., Cascio, D. & Williams, R. (1986a) *Science* **232**, 765–768
- McPherson, A., Brayer, G. D. & Morrison, R. D. (1986b) *J. Mol. Biol.* **189**, 305–327
- McPherson, A., Brayer, G. & Morrison, R. (1986c) *Biophys. J.* **49**, 209–219
- Mitsui, Y., Urata, Y., Torii, K. & Irie, M. (1978) *Biochim. Biophys. Acta* **535**, 299–308
- Moews, P. C. & Knox, J. R. (1976) *J. Am. Chem. Soc.* **98**, 6628–6633
- Nachman, J., Miller, M., Gilliland, G. L., Carty, R., Pincus, M. & Wlodawer, A. (1990) *Biochemistry* **29**, 928–937
- Pavlovsky, A. G., Borisova, S. N., Broisov, V. V., Antonov, I. V. & Karpeisky, M. Y. (1978) *FEBS Lett.* **92**, 258–262
- Pearlman, D. A. & Kim, S.-H. (1990) *J. Mol. Biol.* **211**, 171–187
- Richards, F. M. & Wyckoff, H. W. (1971) in *The Enzymes* (Boyer, P. D., ed) Vol. 4, 3rd ed., Academic Press, New York
- Richards, F. M. & Wyckoff, H. W. (1973) in *Atlas of Molecular Structure in Biology* (Phillips, D. C. & Richards, F. M., eds) Vol. I, Clarendon, Oxford
- Saenger, W. (1984) in *Principles of Nucleic Acid Structure* (Cantor, C. R., ed) pp. 69–73, 132–140, Springer-Verlag, New York
- Saenger, W. & Scheit, K. H. (1970) *J. Mol. Biol.* **50**, 153–169
- Sawada, F. & Irie, M. (1969) *J. Biochem. (Tokyo)* **66**, 415–418
- Sekine, H., Nakano, E. & Sakaguchi, K. (1969) *Biochim. Biophys. Acta* **174**, 202–210
- Smyth, D. G., Stein, W. H. & Moore, S. (1963) *J. Biol. Chem.* **238**, 227
- Sussman, J. L. (1985) *Methods Enzymol.* **115**, 271–303
- TenEyck, L., Weaver, L. H. & Matthews, B. W. (1976) *Acta Crystallogr. Sect. A* **32**, 349
- Usher, D. A., Richardson, D. I. & Eckstein, F. (1970) *Nature* **228**, 663–665
- Usher, D. A., Erenrich, E. S. & Eckstein, F. (1972) *Proc. Natl. Acad. Sci. U. S. A.* **69**, 115–118
- Walz, F. G. (1971) *Biochemistry* **10**, 2156–2162
- Webb, L. E., Hill, E. J. & Banaszak, L. J. (1973) *Biochemistry* **12**, 5101–5109
- Wlodawer, A. (1984) in *Biological Macromolecules and Assemblies, Nucleic Acids and Interactive Proteins* (Jurnak, F. A. & McPherson, A., eds) pp. 393–440, John Wiley Co., New York
- Wlodawer, A. (1985) in *Biological Macromolecules and Assemblies* (Jurnak, F. A. & McPherson, A., eds) Vol. II, pp. 395–439, John Wiley Co., New York
- Wlodawer, A. & Hendrickson, W. A. (1982) *Acta Crystallogr. Sect. A* **38**, 239–247
- Wlodawer, A. & Sjolín, L. (1983) *Biochemistry* **22**, 2720–2728
- Wlodawer, A., Miller, M. & Sjolín, L. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80**, 3628–3631
- Wlodawer, A., Svensson, L. A., Sjolín, L. & Gilliland, G. L. (1988) *Biochemistry* **27**, 2705–2717
- Wodak, S. Y., Liu, M. Y. & Wyckoff, H. W. (1977) *J. Mol. Biol.* **116**, 855–875
- Wyckoff, H. W., Tsernoglou, D., Hanson, A. W., Knox, J. R., Lee, B. & Richards, F. M. (1970) *J. Biol. Chem.* **245**, 305–328