^{1}H NMR (500 MHz) of Gene 32 Protein–Oligonucleotide Complexes†

Richard V. Prigodich, Jose Casas-Finet, Kenneth R. Williams, William Konigsberg, and Joseph E. Coleman*

ABSTRACT: In concentrated solutions, gene 32 single-stranded DNA binding protein from bacteriophage T4 (gene 32P) forms oligomers with long rotational correlation times, rendering ^{1}H NMR signals from most of the protons too broad to be detected. Small flexible N- and C-terminal domains are present, however, the protons of which give rise to sharp resonances. If the C-terminal A domain (48 residues) and the N-terminal B domain (21 residues) are removed, the resultant core protein of 232 residues (gene 32P*) retains high affinity for ssDNA and remains a monomer in concentrated solution, and most of the proton resonances of the core protein can now be observed. Proton NMR spectra (500 MHz) of gene 32P* and its complexes with ApA, $d(pA)_n$ (n = 2, 4, 6, 8, and 10), and $d(pT)_8$ show that the resonances of a group of aromatic protons shift upfield upon oligonucleotide binding. Proton difference spectra show that the ^{1}H resonances of at least one Phe, one Trp, and five Tyr residues are involved in the chemical shift changes observed with nucleotide binding. The number of aromatic protons involved and the magnitude of the shifts change with the length of the oligonucleotide until the shifts are only slightly different between the complexes with $d(pA)_8$ and $d(pA)_{10}$, suggesting that the binding groove accommodates approximately eight nucleotide bases. Many of the aromatic proton NMR shifts observed on oligonucleotide complex formation are similar to those observed for oligonucleotide complex formation with gene 5P of bacteriophage fd, although more aromatic residues are involved in the case of gene 32P*. For gene 5P, complex formation is known to involve intercalation with one Phe and three Tyr residues [O'Connor, T. P., & Coleman, J. E. (1983) *Biochemistry 22*, 3375–3381]. The NMR results of the present study are consistent with the evidence from other techniques which have implicated the tyrosine-rich region of the gene 32 protein, residues 72–116, in DNA binding. Hydrophobic interactions between the side chains of aromatic amino acids and the bases of a polynucleotide which confer a regular "ladder" pattern on the single-stranded nucleotide may be a general structural feature of single-stranded DNA binding proteins.

There is a class of DNA binding proteins found in both prokaryotic and eukaryotic cells which show high binding affinity for single-stranded DNA (ssDNA). The result of the preferential interaction of these proteins with ssDNA is the destabilization (melting) of the double helix, hence the name helix-destabilizing proteins (Alberts & Frey, 1970; Alberts et al., 1972; Coleman & Oakley, 1980). It is the binding of these proteins to ssDNA, however, which is responsible for the functions of these proteins and the various ways by which they regulate genome replication (Kornberg, 1980). The mechanism of interaction of such proteins with ssDNA is, therefore, of considerable importance.

Gene 32 of bacteriophage T4 codes for a 33 487-dalton ssDNA binding protein known as gene 32P. Gene 32P is required for the replication (Epstein et al., 1963), recombination (Tomizawa et al., 1966), and repair (Wu & Yeh, 1973) of T4 DNA. Gene 32P coats the single-stranded regions of T4 DNA just ahead of the T4 DNA polymerase. The multiplicity of functions of this protein is made possible by the ability of gene 32P to interact not only with ssDNA (Alberts & Frey, 1970) but also with other T4-coded proteins (Mosig et al., 1979) and adjacent molecules of gene 32P (Huberman et al., 1971). Furthermore, particular structural domains of gene 32P have been identified with these specific functions (Williams & Konigsberg, 1978). Amino acid residues 1–21 comprise the B region of gene 32P which is responsible for gene 32P self-aggregation. Thus, the B region mediates the cooperativity observed for the binding of multiple molecules of gene 32P to ssDNA. Residues 254–301 constitute a C-terminal or A domain. This domain appears to provide a "kinetic block" to the denaturation of native double-stranded DNA (dsDNA), since a conformational change involving displacement of the A region from the ssDNA binding locus on gene 32P is postulated to occur prior to ssDNA binding. The A region is then proposed to be available for interactions between gene 32P and other proteins involved in T4 DNA replication (Hosoda & Moise, 1978).

The best characterized of the ssDNA binding proteins is the 9689-dalton protein encoded by gene 5 of the filamentous bacteriophage fd (M13) [for a review, see Coleman & Oakley (1980)]. Gene 5 protein binds the ssDNA product synthesized from the double-stranded replicative form of the phage DNA. This prevents subsequent synthesis of the complementary strand, and hence gene 5P is responsible for the shift of M13 DNA synthesis from the double-stranded replicative form to the single-stranded viral genome [see Kornberg (1980, 1982) for a review].

Proton and ^{31}P NMR studies (Coleman & Armitage, 1978; Alma et al., 1981; O'Connor & Coleman, 1982, 1983) and chemical modification data (Anderson et al., 1975) have defined the gene 5P binding locus for ssDNA. In addition to a number of Arg and Lys side chains, the DNA binding groove is composed of three Tyr residues and one Phe residue which can be assigned to Tyr-26, Phe-73, Tyr-34, and Tyr-41 on the basis of the crystal structure (Brayer & McPherson, 1983). Phe-73 is a residue from the opposite monomer of the symmetrical gene 5P dimer which is the basic unit in solution (Brayer & McPherson, 1983). The energy of nucleotide complex formation is provided by the association of the phosphate moieties with Lys and Arg side chains, coupled with partial intercalation of the nucleotide heterocycles with the aromatic side chains of the amino acid residues cited above (Coleman & Armitage, 1978; Alma et al., 1981).

As in the case of gene 5P, nitration of gene 32P has identified five Tyr side chains as being involved in ssDNA binding

† From the Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06510. *Received August 1, 1983.* This work was supported by NIH Grants GM21919 (to J.E.C.) and GM12607 (to W.K.) and NSF Grant PCM8104118 (to K.R.W.). R.V.P. is the recipient of a National Institutes of Health Postdoctoral Fellowship (AMO7100). The 500-MHz NMR facility is supported by NSF Grant CHE-7916210.

0006-2960/84/0423-0522$01.50/0

(Anderson & Coleman, 1975). In contrast to gene 5P, gene 32P, at concentrations required for NMR studies, forms long oligomers even in the absence of ssDNA. The long rotational correlation time of the oligomers broadens most of the proton resonances beyond detection, which has prevented significant NMR studies of oligonucleotide complexes of gene 32P (Coleman & Oakley, 1980). This difficulty can be overcome by preparing a species of gene ^{32}P from which the A and B domains have been removed by limited proteolysis, gene 32P* (Williams & Konigsberg, 1978). The modified protein (M_r 26 024) no longer oligomerizes and has lost the cooperativity of binding, but the monomer retains its high binding affinity for ssDNA and oligonucleotides. Both the proteolytically modified protein and its oligonucleotide complexes show well-resolved ^{1}H NMR spectra. This paper presents the ^{1}H spectra for gene 32P* and its complexes with oligonucleotides from $d(pA)_2$ to $d(pA)_{10}$. Changes in the chemical shifts of aromatic protons of the protein on formation of the oligonucleotide complexes identify a series of aromatic side chains involved in nucleotide binding to gene 32P*.

Materials and Methods

Gene 32 Proteins. Homogeneous gene 32 protein was prepared as previously described (Bittner et al., 1979). The gene 32P fragment, gene 32P*, was isolated from a tryptic digest and prepared according to a modified version of the procedure of Hosoda & Moise (1978) (Williams & Konigsberg, 1978). The protein concentration was determined by using $\epsilon_{280} = 3.7 \times 10^4$ M^{-1} cm^{-1} (Jensen et al., 1976) and amino acid analysis. The protein samples were deuterium exchanged by repeated dialysis against 10 mM $NaDCO_3$–0.3 M NaCl, pH 8. Protein concentrations of the NMR samples were ~0.5 mM for native gene 32P and ~0.8 mM for gene 32P*.

Nucleotides. $d(pA)_n$'s (n = 4, 6, 8, and 10) and $d(pT)_8$ were purchased from P-L Biochemicals, Inc. $d(pA)_2$ was purchased from Collaborative Research, Inc., and ApA was purchased from Sigma. The nucleotides were dissolved in D_2O, and the appropriate amount of nucleotide was determined from known extinction coefficients (Kelly & von Hippel, 1976). The deuterium-exchanged nucleotides were lyophilized, and then protein solution was added. All complexes were 1:1 with protein, which yields >95% complex formation with the concentrations employed (Kelly et al., 1976).

^{1}H NMR Spectra. The spectra were obtained at 500 MHz on a Bruker WM-500 spectrometer at 303 K. The solvent, D_2O, served as a field-frequency lock. The samples were 0.4 mL in 5-mm tubes. Chemical shifts were measured in parts per million (ppm) downfield from a standard reference frequency (6684 Hz). Nuclear Overhauser effect (NOE) experiments were performed by allowing a 3.0-s relaxation delay to precede a 1.0-s preirradiation pulse which generated the NOE. The preirradiation pulse was followed by an observation pulse. Sets of eight transients with on-resonance preirradiation pulses were interleaved with sets of eight transients with off-resonance preirradiation pulses. The NOE difference spectra were made by subtracting the free induction decay (FID) of the on-resonance experiments from the FID of the off-resonance FID. Thus, the saturated on-resonance absorptions and negative NOE's appear as positive peaks.

Results

*^{1}H NMR Spectra of Gene 32P and Gene 32P**. The ^{1}H NMR spectrum of native gene 32P shows a number of sharp resonances in the aliphatic region and a relatively small

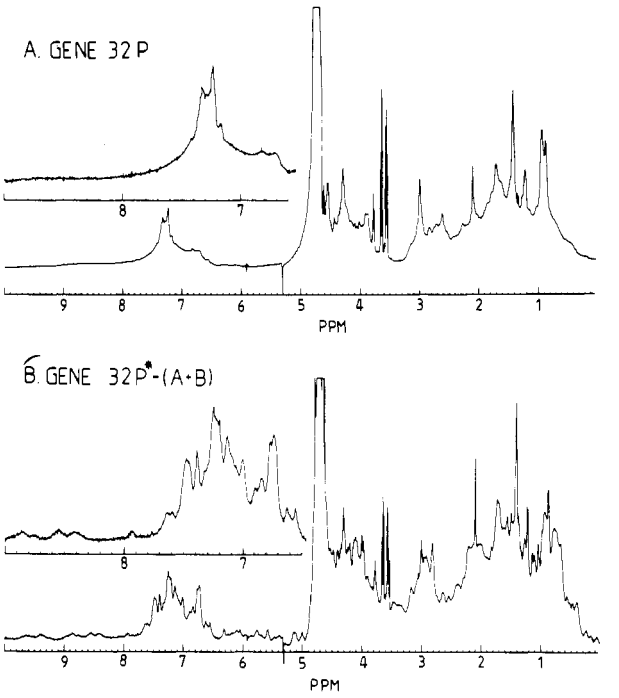

FIGURE 1: 500-MHz ^{1}H NMR spectra of the aromatic and aliphatic regions of (A) gene 32P and (B) gene 32P*. Insets are expansions of the respective aromatic regions.

number of overlapping aromatic proton resonances grouped near 7.3 ppm (Figure 1A). The latter are predominantly Phe protons (see below). Tyr protons are not well represented. A prominent signal for Lys ϵ-CH_2 protons appears at ~3 ppm. In addition to the residual HDO peak, two sets of sharp quartets at 3.56 and 3.66 ppm are present in all spectra from the β-mercaptoethanol in the buffer.

Removal of the A and B domains by limited proteolysis to produce gene 32P* results in a dramatic change in the ^{1}H NMR spectrum (Figure 1B). Many more aromatic resonances appear, especially Tyr protons from 5 to 7 ppm. New aliphatic resonances appear, and many of the narrow aliphatic resonances originally present seem to have disappeared. Most of the original Phe resonances may also have disappeared. As we previously suggested in an exploratory ^{1}H NMR spectrum at 270 MHz (Coleman & Oakley, 1980), the sharp resonances in Figure 1A appear to be accounted for by the amino acid side chains in the flexible A and B domains of the protein which have a combined composition of 12 Asp, 7 Asn, 4 Thr, 9 Ser, 2 Glu, 1 Gln, 2 Gly, 8 Ala, 2 Val, 3 Met, 7 Leu, 4 Phe, 7 Lys, and 1 Arg. The set of resonances expected for this composition accounts for the major resonances observed for native gene 32P, suggesting that the majority of the proton resonances are broadened beyond detection. There are in fact indications of broad humps under the narrow resonances of gene 32P (Figure 1A).

^{1}H NMR Spectra of Gene 32P–$d(pA)_6$ and Gene 32P–$d(pA)_6$ Complexes.* The absence from the spectra of a large fraction of the proton resonances expected from gene 32P suggested that ^{1}H NMR of the oligonucleotide complexes would not provide much information. Indeed, addition of $d(pA)_6$ to the native protein causes few changes in the spectrum other than the appearance of appropriately broadened H(2), H(8), and H(1′) resonances from the bases and sugars of the bound nucleotide (Figure 2A). On the other hand, addition of $d(pA)_6$ to gene 32P* causes significant alterations in the envelope of the aromatic protons (compare Figures 2B and 1B) while relatively minor changes occur in the upfield aliphatic region on the addition of $d(pA)_6$ (Figure 2B).

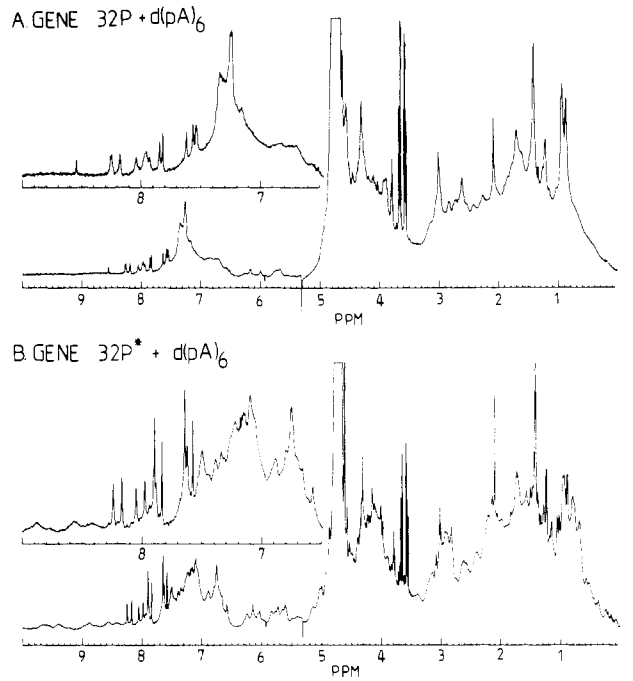

FIGURE 2: 500-MHz ^{1}H NMR spectra of the aromatic and aliphatic regions of (A) the 1:1 gene 32P–d(pA)$_6$ complex and (B) the 1:1 gene 32P*–d(pA)$_6$ complex. Insets are expansions of the respective aromatic regions.

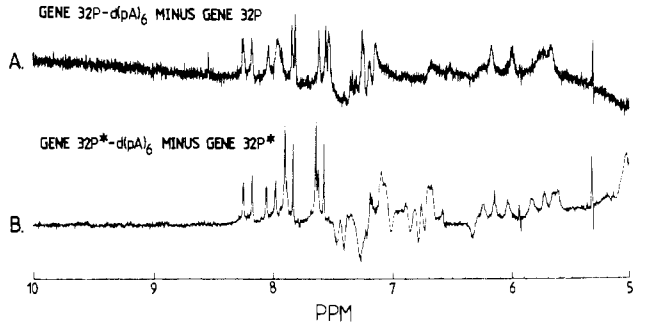

FIGURE 3: ^{1}H NMR difference spectra of the aromatic regions of (A) the gene 32P–d(pA)$_6$ complex spectrum minus that of gene 32P and (B) the gene 32P*–d(pA)$_6$ complex spectrum minus that of gene 32P*.

The changes (or lack of changes) can be better displayed by the difference spectra for these two examples, native gene 32–d(pA)$_6$ minus gene 32 and gene 32P*–d(pA)$_6$ minus gene 32P*, as shown in Figure 3. The difference in Figure 3A was made by nulling the signals from β-mercaptoethanol and the difference in Figure 3B by nulling the nonexchangeable amide proton resonances between 8.3 and 10.0 ppm. In the case of the complex of d(pA)$_6$ with gene 32P*, a series of well-resolved negative and positive peaks between 7.6 and 6.2 ppm of the difference spectrum document the shifts of the aromatic protons of several amino acid side chains upon nucleotide complex formation (Figure 3B). The corresponding difference spectrum of the native protein shows mainly the adenine and sugar protons with just a hint of shifts in the region of 7.6 ppm. The latter were not obvious in the direct spectrum (Figure 2A), but the difference spectrum shows that a few of the "core" resonances are contributing to the gene 32P spectrum, largely obscured, however, by the relatively sharp resonances from the A and B domains.

Difference Spectra: ^{1}H NMR Spectra of Gene 32P–ApA and –d(pA)$_n$ Complexes minus ^{1}H NMR Spectra of Gene 32P**. To elucidate the number and nature of the protein aromatic side chains involved in ssDNA binding, the binding site was probed with adenine oligonucleotides of varying lengths (2–10 residues). The difference spectra were made by subtracting the gene 32P* spectrum from the gene 32P*–oligonucleotide complex spectrum and nulling the several nonexchangeable amide proton resonances downfield of 8.3 ppm. Figure 4 displays the difference spectra for the ApA and d(pA)$_n$ (n = 2, 4, 6, 8, and 10) complexes. The difference spectrum for the d(pA)$_{10}$ complex is shown in Figure 5B and compared to the similar difference spectrum for the d(pA)$_4$ complex of gene 5P. The latter nucleotide just fills the nucleotide binding groove of gene 5P.

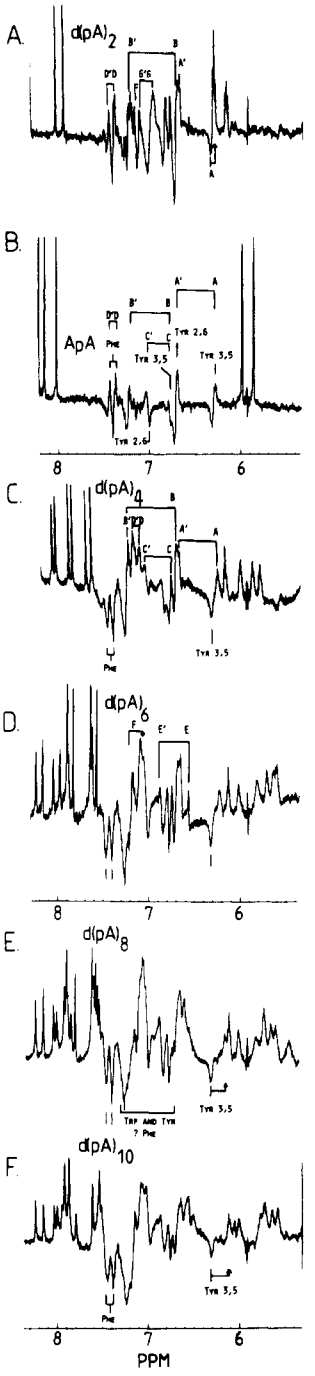

FIGURE 4: ^{1}H NMR difference spectra of the aromatic region of gene 32P*–adenine oligonucleotide complex spectra minus that of gene 32P*. The oligonucleotides in the complexes are (A) d(pA)$_2$, (B) ApA, (C) d(pA)$_4$, (D) d(pA)$_6$, (E) d(pA)$_8$, and (F) d(pA)$_{10}$. No difference in binding gene 32P* was noted between the ribose and deoxyribose dinucleotides.

*Assignments of Proton Resonances in Gene 32P**. Detailed assignment of individual aromatic proton resonances in the

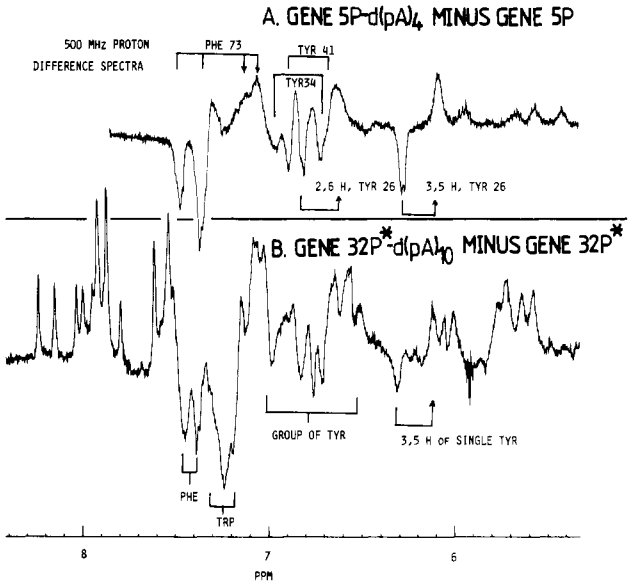

FIGURE 5: ^{1}H NMR difference spectra of the aromatic region of (A) the gene 5P–d(pA)$_4$ complex minus gene 5P and (B) the gene 32P*–d(pA)$_{10}$ complex minus gene 32P*.

^{1}H NMR spectrum of gene 32P* is much more difficult than in the case of gene 5P which has no Trp and only eight aromatic residues. The aromatic spectrum of gene 32P* is not resolved into individual ^{1}H resonances (Figure 1B), and the difference spectra give nonoverlapping peaks in only two cases (Figure 4). One is the most upfield resonance, the two 3,5-protons of a single Tyr residue at 6.25 ppm (resonance A). This is the only case where the position of the resonance in the complex, ~0.2 ppm upfield in the d(pA)$_{10}$ complex (Figure 5B), can be unequivocally assigned. An NOE induced by irradiating the 3,5-protons of this Tyr in the d(pA)$_4$ complex shows the 2,6-protons of this Tyr to resonate at 6.67 ppm (resonance A′, Figures 4 and 6A). An additional small negative NOE peak at ~7.03 ppm is a secondary NOE with the protons of another aromatic residue.

The downfield Phe protons at 7.4–7.5 ppm are also clearly isolated, although the new positions of these resonances in the complex overlap with Trp and even Tyr 2,6-protons (Figures 4 and 5). In the intervening region where the Trp and the rest of the Tyr resonances are overlapping, only general comments on the Tyr and Trp resonances as a group can be made. The overlapping resonances have made identification of protons of the same aromatic side chain by the NOE method difficult, since the irradiation spills on to adjacent signals. Such an example is shown in Figure 6B where irradiation of the 2,6-protons of Tyr-1 (a position established by the NOE in Figure 6A) excites the 3,5-protons, but also adjacent protons. Irradiating just downfield from the position of the 2,6-protons of Tyr-1 excites another resonance at 7.23 ppm (Figure 6C). This pair (B and B′) probably represents the 3,5- and 2,6-protons of a second Tyr. A third Tyr residue may be assigned from the NOE at 6.74 ppm generated by irradiation at 7.02 ppm (C and C′, Figure 6D).

We chose to do the NOE's on the d(pA)$_4$ complex so we could select resonance positions that were represented in the difference spectra of the complex and hence represented aromatic residues involved in oligonucleotide complex formation. The gene 32P*–d(pA)$_4$ minus gene 32P* difference spectrum is given in Figure 6E with indications of the proton resonances that have been connected by observable NOE's. These general assignments give a picture of the aromatic residues of the protein involved in nucleotide complex formation. The best way to make these assignments more precise (which is currently under way) is to biosynthesize separate protein samples containing perdeuterated Phe, Tyr, and Trp by T4 infection of aromatic auxotrophs of *E. coli*.

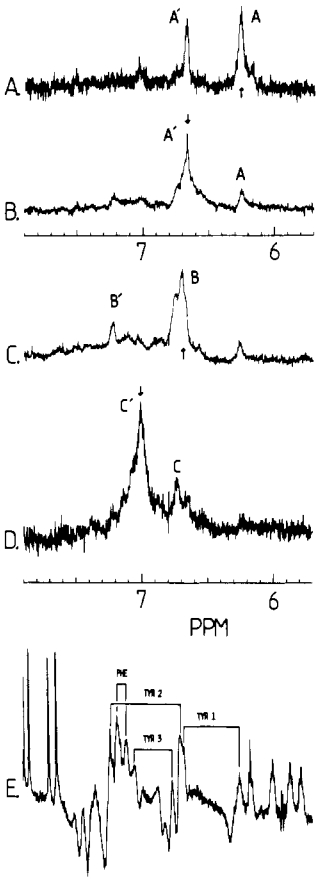

FIGURE 6: NOE difference spectra for the gene 32P*–d(pA)$_4$ complex. The arrows indicate the preirradiated frequency in spectra A–D. Spectrum E is the difference spectrum for the gene 32P*–d(pA)$_4$ complex minus gene 32P* with tentative assignments for aromatic side chain resonances based on the NOE data.

Figure 4C presents the difference spectrum for the d(pA)$_4$ complex. We have labeled the protein–oligonucleotide complex (positive peaks) of this difference spectra in as self-consistent a manner as possible, coupling resonances that arise from the same residue according to the NOE data (e.g., AA′, BB′, CC′, etc.). Resonances like D and D′ clearly belong to Phe, while A and A′ are the 3,5- and 2,6-protons of a Tyr residue, respectively. The difference spectrum of the d(pA)$_4$ complex (Figure 4C) shows further upfield perturbations of the same aromatic protons that were affected by ApA (Figure 4B). The D and D′ resonances assigned to Phe appear to shift further upfield in the d(pA)$_4$ complex by about 0.3 ppm (Table I). The effect of oligonucleotide binding on the B and C resonances is much more developed in the d(pA)$_4$ complex than in the ApA complex. As the difference spectra become more complex with the longer oligonucleotides, we have attempted to deconvolute the spectra by simply labeling the new peaks which appear and retaining the peaks already labeled in the d(pA)$_4$ complex. This procedure gives rise to the assignments and progressive upfield shifts as the nucleotide is lengthened as shown in Table I.

The difference spectrum of the d(pA)$_6$ complex (Figure 4D) shows that the two new peaks, E and E′, become more distinct. They are barely discernible in the other spectra (Figure 4A–C) if one uses their apparent chemical shifts in Figure 4D to locate them. Additional aromatic protein protons which had chemical shifts of ~7.2 ppm in gene 32P* seem to be shifted upfield

Table I: Changes in the Chemical Shifts (ppm) of Aromatic ^{1}H Resonances of Gene 32P* upon Oligonucleotide Binding[a]

	Tyr-1		Tyr-2		Tyr-3		Phe		Tyr-4		Tyr-5	
nucleotide	A	A′	B	B′	C[b]	C′	D	D′	E	E′	G	G′
$d(pA)_2$	0.07	0.07	0.10	0.04	0.8/0.3	?	0.03	0.03			0.09	0.04
ApA	0.05	0.05	~0.05	0.04	~0.05	−0.04	0.03	0.03				
$d(pA)_4$	0.10	0.08	0.10	0.06	0.10/0.02	−0.04	0.35	0.35				
$d(pA)_6$	0.10	0.08	0.10	0.09	0.10/0.04	−0.06	~0.4	~0.4	−0.01	−0.01		
$d(pA)_8$	0.18	0.13	0.13	~0.2	0.13/0.04	−0.06	~0.4	~0.4	−0.01	−0.01		
$d(pA)_{10}$	0.25	0.20	0.19	~0.2	0.18/0.04	−0.06	~0.4	~0.4	0.01	−0.01		
$d(pT)_8$	~0.10	0.15	0.18	0.10	0.10	−0.06	0.21	0.22			0.13	0.05

[a] Assignments are based on chemical shift and NOE data. Upfield changes are positive; downfield changes are negative. [b] The values under column C indicated with a slant refer to the split C difference resonance that appears with the longer oligonucleotides, giving two possible calculations of the upfield shift.

of their original resonance by 0.05–0.15 ppm in the complex. This new difference is labeled F. Furthermore, the original C resonance appears to be split by $d(pA)_6$ binding.

In the $d(pA)_8$ complex (Figure 4E), the differences at A and A′ are even more developed, as are the upfield shifts of most of the other resonances. One of the split C resonances has shifted upfield sufficiently to attain a frequency equivalent to the A′ resonance of the unbound protein such that these resonances cancel. The difference spectrum of the $d(pA)_{10}$ complex is very similar to that of the $d(pA)_8$ complex (Figure 4). The only variations are additional small increases in the magnitude of the upfield shifts of some of the resonances (Table I).

Examination of Figure 4A reveals that the $d(pA)_2$ complex is not entirely consistent with the trends observed in Figure 4B–E. While resonances at A, A′, D, and D′ are perturbed in a very similar fashion to the ApA complex, for example, B and B′ resonances resemble the B and B′ resonances of the $d(pA)_4$ complex, and the C, C′, E, and E′ resonances are like those of the $d(pA)_6$ complex. It is possible that $d(pA)_2$ can occupy several positions in the binding lattice, since there are potentially four different occupancies for $d(pA)_2$ possible with a binding locus accommodating eight bases. Indeed, two $d(pA)_2$ molecules bind to the four-base binding locus in gene 5P (O'Connor & Coleman, 1983). Clarification of this will require spectra on perdeuterated samples.

Gene 32P–d(pT)$_8$ Complex.* The difference spectrum from 5 to 10 ppm for the gene 32P*–$d(pT)_8$ complex minus gene 32P* is given in Figure 7A. It is clear that many of the same resonances that are perturbed in the $d(pA)_8$ complex are also perturbed when $d(pT)_8$ is bound. A number of the upfield shifts are larger for $d(pA)_8$ (e.g., the A, C, D, D′, and B′ resonances) than for $d(pT)_8$. This pattern is not, however, observed for all the perturbed aromatic resonances. For example, the G and G′ perturbations do not appear to be present in the $d(pA)_8$ complex, are relatively much smaller than in the $d(pT)_8$ complex, or are masked by other superimposing transitions in the $d(pA)_8$ complex. It also appears that the A′ perturbation gives a slightly greater upfield shift in the $d(pT)_8$ complex than in the $d(pA)_8$ complex (Table I). These variations are more clearly seen in the difference spectrum of gene 32P*–$d(pA)_8$ minus gene 32P*–$d(pT)_8$ (Figure 7B).

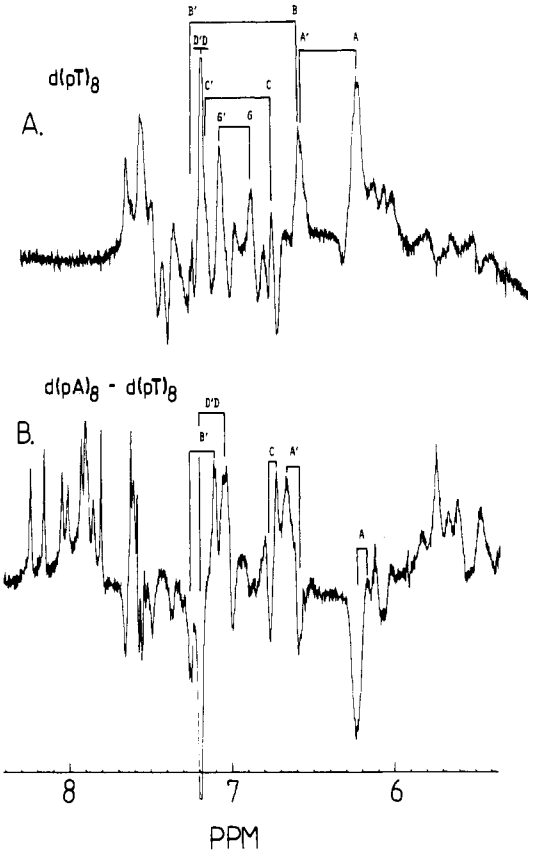

FIGURE 7: ^{1}H NMR difference spectra of the aromatic region of (A) the gene 32P*–$d(pT)_8$ complex minus gene 32P* and (B) the gene 32P*–$d(pA)_8$ complex minus that of the gene 32P*–$d(pT)_8$ complex.

Discussion

The primary structure of gene 32P consists of 301 amino acid residues (Williams et al., 1981). There are indications that gene 32P shares some structural characteristics with the much smaller helix-destabilizing protein gene 5, containing 87 amino acid residues, from bacteriophage fd (Nakashima et al., 1974a,b). Nitration of Tyr residues in both proteins destroys DNA binding, and nitration is prevented by prior DNA binding (Anderson et al., 1975; Anderson & Coleman, 1975). The Tyr residues involved in the binding of oligonucleotides to the symmetric gene 5P dimer have been clearly identified by a combination of chemical modification, ^{1}H NMR, and crystallographic studies (Anderson et al., 1975; O'Connor & Coleman, 1983; Brayer & McPherson, 1983). The residues are Tyr-26, -34, and -41 which together with Phe-73 from the opposite monomer are in positions along the DNA binding groove to form an alternating stack with the bases of a tetranucleotide. Gene 5P lacks tryptophan, and its secondary structure is ~90% β-pleated sheet (Anderson et al., 1975; McPherson et al., 1980). In contrast, a secondary structure prediction from the sequence of gene 32P suggests a mixed structure of approximately 26% β-sheet, 22% α-helix, and 52% random coil (Williams et al., 1981). The sequence from residues 72–116 contains Tyr-73, -84, -92, -106, and -115 as well as two Trp, both with Tyr in adjacent positions, Trp^{72}-Tyr^{73} and Tyr^{115}-Trp^{116}. This sequence is located in the

relatively positively charged N-terminal half of gene 32P and is a candidate for at least part of the region involved in nucleotide binding in view of the evidence for participation of Tyr and Trp residues in binding (see below). There is also some sequence homology between this region of gene 32P and a Tyr-containing sequence from residues 26–39 of Rauscher murine leukemia virus protein p10 which binds single-stranded DNA and RNA (Henderson et al., 1981).

Comparison of the ^{1}H NMR aromatic difference spectra produced from gene 5P and a nucleotide complex which fills the nucleotide binding surface, $d(pA)_4$, with that produced from gene 32P* and a nucleotide which fills its binding surface, $d(pA)_{10}$, shows some remarkable similarities in the aromatic proton shifts of the two proteins induced by nucleotide binding (Figure 5). The clearly identifiable shifts of Phe and Tyr protons in both proteins suggest similar nucleotide interactions with aromatic side chains. The extensive chemical modification and crystallographic work on gene 5P combined with extensive NMR analysis of a relatively simple proton spectrum has allowed a rather precise identification of the residues involved in the binding groove (O'Connor & Coleman, 1983; Brayer & McPherson, 1983). Part of the additional complexity of the gene 32P* difference spectrum (Figure 5B) could be explained by the involvement of five rather than three Tyr in the longer nucleotide binding groove as suggested by the nitration experiments (Anderson & Coleman, 1975). In addition, one or more Trp side chains may be involved in nucleotide complex formation, and part or most of the difference spectrum between 7.3 and 7.0 ppm (Figure 5B) may involve indole protons. This assignment is compatible with fluorescence studies suggesting that at least one tryptophan side chain is near the nucleotide binding site (Kelly & von Hippel, 1976; Hélène et al., 1976).

The existence in gene 32P of N- and C-terminal domains which modulate separate functions of gene 32P also distinguishes it from the simple gene 5P prototype. The ^{1}H NMR spectra of gene 32P and its proteolytic derivative, gene 32P*, offer good support for the three-domain model of gene 32P structure which has been based on limited proteolysis (Williams & Konigsberg, 1978), amino acid sequence (Williams et al., 1981), and ultracentrifugation (Carroll et al., 1975) studies. The native protein does appear to oligomerize sufficiently to give extreme broadening of the ^{1}H resonances from the core protein; yet domains corresponding to the A and B amino acid compositions have additional rotational motion relative to the core protein, giving rise to relatively sharp resonances (Figure 1). That the oligomerization process depends critically on the B domain is dramatically illustrated by the appearance of the well-resolved ^{1}H NMR spectrum expected of a 26 024-dalton monomer when the A and B domains are removed (Figure 2). The well-resolved ^{1}H NMR spectrum of the fragment monomers and the capability of the fragments to interact with nucleotides make mapping the DNA binding groove with NMR data possible.

There is a progressive increase in the number of aromatic protons shifted and the magnitude of the observed upfield shifts as the bound oligonucleotide increases in length (Figure 4A–E, Table I). These changes are at a maximum when a nucleotide between 8 and 10 residues is bound (Table I). This suggests a binding groove accommodating at least eight nucleotide residues per monomer. This is close to but not identical with estimates based on other methods, including the gene 32 induced hyperchromicity of poly(dA) [6.7:1 (Jensen et al., 1976)], the melting of poly[d(A-T)] [7.5:1 (Jensen et al., 1976)], the change in the CD of poly[d(A-T)] or fd DNA [10:1 (Anderson & Coleman, 1975)], or the quenching of the protein fluorescence [~5:1 (Kelly et al., 1976)].

The upfield nature of most of the proton shifts of the aromatic side chains induced by nucleotide binding to gene 32P* is compatible with the hypothesis that these proton shifts represent ring currents from the adenine rings inserted into the base-binding pockets. The maximum values of these shifts, 0.06–0.4 ppm, are all smaller than would be expected for complete overlap of the aromatic rings at the minimum stacking distance of 3.4 Å. The latter would place the aromatic protons within 0.5 Å of the center of the five- or six-membered base rings (Giessner-Prettre et al., 1976). This suggests that while the bases of the bound nucleotides in gene 32P* approach the aromatic rings in the base-binding pockets, the stacking distances are larger than 3.4 Å and/or the ring overlap is less than maximal, i.e., proton to center of the base rings >1 Å. Other than providing complementary pockets for a "ladder-type" orientation of the bases, the contribution of the stacking interactions to the thermodynamic stability of the complexes is probably not great. As mentioned above, the electrostatic interactions between the phosphate backbone and the positive charges on the Lys and Arg residues appear to account for the largest fraction of the binding energy, although the cooperative protein–protein interactions when polynucleotide complexes are formed may make a significant contribution to the very low dissociation constants, $<10^{-8}$ M, estimated for the latter complexes.

In the oligonucleotide complexes studied here, the tyrosyl and phenylalanyl rings in the complexes may continue to have significant flipping rates relative to the NMR time scale, making the 3,5- and 2,6-protons equivalent. In addition, location of the aromatic protons of the tyrosyl and phenylalanyl rings at more than the minimal distance from the magnetic field of the base rings would tend to decrease the differential proton shifts expected across a given ring, a differential that might apply to a rigidly held aromatic ring intercalated at 3.4 Å. These factors both probably contribute to the similar magnitudes observed for the upfield shifts of all the aromatic protons involved (Figure 4). The aromatic proton shifts induced in gene 32P* by oligonucleotide binding do not allow construction of a rigid or specific model of the complex. The distances between the base rings and the protons on the aromatic rings found in the gene 32P*–oligonucleotide complexes are, however, probably considerably greater than the minimum possible in a tightly intercalated structure. These same considerations apply to the previously documented interaction of the aromatic residues of gene 5P with the bases of oligonucleotides [see O'Connor & Coleman (1983) for a discussion].

The upfield shifts are in most cases greater for A-containing nucleotides than for T-containing nucleotides, supporting the interpretation that they represent ring-current shifts. While A-containing oligonucleotides bind more tightly to gene 5P than T-containing oligonucleotides (Coleman et al., 1976), poly(dT) has been reported to bind more tightly to gene 32P than poly(dA) (Williams & Konigsberg, 1978). While some of this difference may involve unstacking of poly(dA), different conformations of bound A and T bases may be involved which could account for some of the NMR differences in Figure 7.

The assignments of the perturbed resonances in the difference spectra (Figures 4 and 5) were made in pairs (i.e., A, A′ etc.) and in keeping with a tentative identification of the types of residues involved based on the original chemical shifts, NOE results, and the assignments made for gene 5P–oligonucleotide complexes. The tentative identification scheme links

resonances A–A′, B–B′, C–C′, E–E′, and G–G′ with tyrosine side chains and links D-D′ and F with phenylalanine and tryptophan residues, respectively. This assignment sets the binding site as seven to eight nucleotides long. This is consistent with the great similarity of the ^{1}H NMR difference spectra of the $p(dA)_8$– and $d(pA)_{10}$–gene 32P* complexes. Although many of the same aromatic residues are perturbed by both $d(pA)_8$ and $d(pT)_8$, there are differences in the degree to which the residues are shifted upfield. The A, C, D, D′, and B′ resonances are all subject to greater upfield transitions in the $d(pA)_8$ complex, while only the A′ resonance is shifted to a greater degree by $d(pT)_8$. This is consistent with the greater ring-current shifts expected for purine rings (Giessner-Prettre & Pullman, 1976). The effect of $d(pT)_8$ on the A′ resonance could be due to a unique geometry of binding of one thymidine ring.

Gene 32P provides the second example of a helix-destabilizing protein which uses hydrophobic interactions between the bases of the nucleotide residues and aromatic amino acid side chains to stabilize its DNA complexes in addition to electrostatic interactions between the phosphate backbone and positively charged protein side chains. Whether the observed ring-current shifts can be interpreted to represent intercalation in the usual sense of the word requires a more precise limit on the geometry of the complex than is possible with NMR data alone. Stacking of a base ring on a Tyr residue has been found in ribonuclease T_1. In the crystal structure of a ribonuclease T_1–guanylic acid complex, Tyr-45, located in the active center of the enzyme, is stacked 3.5 Å above the six-membered ring of the guanine ring of the bound nucleotide (Heineman & Saenger, 1982).

The presence of a set of sequentially placed aromatic side chains which confer a regular ladder pattern on the bases in a single-stranded polynucleotide may be a general theme for ssDNA binding proteins. This structural feature would appear to differentiate ssDNA binding proteins from proteins which recognize dsDNA, e.g., *lac* repressor, CAP protein, the headpiece of the λ repressor, and cro protein from λ. Both NMR and X-ray crystal structures suggest that these proteins do not have aromatic side chains as a prominent feature of their DNA binding loci (Anderson et al., 1981; McKay & Steitz, 1981; Steitz et al., 1982; Pabo et al., 1982; Kirpichnikov et al., 1982; Kurochin & Kirpichnikov, 1982). On the other hand, nitration of the monomeric T7 RNA polymerase nitrates five Tyr residues and destroys template binding (Oakley et al., 1975). Hence, polymerases which melt dsDNA and establish an open complex with catalytic functions centered on a single template strand may also involve binding interactions with aromatic side chains.

Registry No. ApA, 2391-46-0; $d(pA)_2$, 16240-63-4; $d(pA)_4$, 15279-64-8; $d(pA)_6$, 24512-53-6; $d(pA)_8$, 12105-21-4; $d(pA)_{10}$, 60117-28-4; $d(pT)_8$, 54284-61-6; Phe, 63-91-2; Trp, 73-22-3; Tyr, 60-18-4.

References

Alberts, B., & Frey, L. (1970) *Nature (London) 227*, 1313–1318.

Alberts, B., Frey, L., & Delius, H. (1972) *J. Mol. Biol. 68*, 139–152.

Alma, N. C. M., Harmsen, B. J. M., Hilbers, C. W., Van der Marel, G., & van Boom, J. H. (1981) *FEBS Lett. 135*, 15–20.

Anderson, R., & Coleman, J. E. (1975) *Biochemistry 14*, 5485–5491.

Anderson, R. A., Nakashima, Y., & Coleman, J. E. (1975) *Biochemistry 14*, 907–917.

Anderson, W. F., Ohlendorf, D. H., Takeda, Y., & Matthews, B. W. (1981) *Nature (London) 290*, 754–758.

Bittner, M., Burke, R. L., & Alberts, B. M. (1979) *J. Biol. Chem. 254*, 9565–9572.

Brayer, G. D., & McPherson, A. (1983) *J. Mol. Biol. 169*, 565–596.

Carroll, R. B., Neet, K., & Goldthwait, D. A. (1975) *J. Mol. Biol. 91*, 275–291.

Coleman, J. E., & Armitage, I. M. (1978) *Biochemistry 17*, 5038–5045.

Coleman, J. E., & Oakley, J. L. (1980) *CRC Crit. Rev. Biochem. 7*, 247–289.

Coleman, J. E., Anderson, R. A., Ratcliffe, R. G., & Armitage, I. M. (1976) *Biochemistry 15*, 5419–5430.

Epstein, R. H., Bolle, A., Steinberg, C. M., Kellenberger, E., Boy de la Tour, E., Chevalley, R., Edgar, R. S., Susman, M., Denhardt, G. B., & Lielausis, A. (1963) *Cold Spring Harbor Symp. Quant. Biol. 28*, 375–392.

Giessner-Prettre, C., & Pullman, B. (1976) *Biochem. Biophys. Res. Commun. 70*, 578–581.

Giessner-Prettre, C., Pullman, B., Borer, P. N., Kan, L.-S., & Ts'o, P. O. P. (1976) *Biopolymers 15*, 2277–2286.

Heineman, U., & Saenger, W. (1982) *Nature (London) 299*, 27–31.

Hélène, C., Toulmé, F., Charlier, M., & Yaniv, M. (1976) *Biochem. Biophys. Res. Commun. 71*, 91–98.

Henderson, L. E., Copeland, T. D., Sowder, R. C., Smythers, G. W., & Oroszlan, S. (1981) *J. Biol. Chem. 256*, 8400–8406.

Hosada, J., & Moise, H. (1978) *J. Biol. Chem. 253*, 7547–7555.

Huberman, J. A., Kornberg, A., & Alberts, B. (1971) *J. Mol. Biol. 62*, 39–52.

Jensen, D. E., Kelly, R. C., & von Hippel, P. H. (1976) *J. Biol. Chem. 251*, 7215–7228.

Kelly, R. C., & von Hippel, P. H. (1976) *J. Biol. Chem. 251*, 7229–7239.

Kelly, R., Jensen, D., & von Hippel, P. (1976) *J. Biol. Chem. 251*, 7240–7250.

Kirpichnikov, M. P., Kurochin, A. V., & Skryabin, K. G. (1982) *FEBS Lett. 150*, 407–410.

Kornberg, A. (1980) in *DNA Replication*, pp 479–496, W. H. Freeman, San Francisco, CA.

Kornberg, A. (1982) in *Supplement to DNA Replication*, p 571, W. H. Freeman, San Francisco, CA.

Kurochin, A. V., & Kirpichnikov, M. P. (1982) *FEBS Lett. 150*, 411–415.

McKay, D. B., & Steitz, T. A. (1981) *Nature (London) 290*, 744–749.

McPherson, A., Jurnak, F. A., Wang, A., Kolpak, F., Rich, A., Molineux, I., & Fitzgerald, P. (1980) *Biophys. J. 32*, 155–173.

Mosig, G., Luder, A., Garcia, G., Dannenberg, R., & Bock, S. (1979) *Cold Spring Harbor Symp. Quant. Biol. 43*, 501–515.

Nakashima, Y., Dunker, A., Marvin, D., & Konigsberg, W. (1974a) *FEBS Lett. 40*, 290–292.

Nakashima, Y., Dunker, A., Marvin, D., & Konigsberg, W. (1974b) *FEBS Lett. 43*, 125.

Oakley, J. L., Pascale, J. A., & Coleman, J. E. (1975) *Biochemistry 14*, 4684–4691.

O'Connor, T. P., & Coleman, J. E. (1982) *Biochemistry 21*, 848–854.
O'Connor, T. P., & Coleman, J. E. (1983) *Biochemistry 22*, 3375–3381.
Pabo, C. O., & Lewis, M. (1982) *Nature (London) 298*, 443–447.
Pabo, C. O., Krovatin, W., Jeffrey, A., & Sauer, R. T. (1982) *Nature (London) 298*, 441–443.
Spicer, E., Williams, K. R., & Konigsberg, W. (1979) *J. Biol. Chem. 254*, 6433–6436.
Steitz, T. A., Ohlendorf, D. H., McKay, D. B., Anderson, W. F., & Matthews, B. W. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 3097–3100.
Tomizawa, J., Anraku, N., & Iwama, Y. (1966) *J. Mol. Biol. 21*, 247–253.
Toulmê, J. J., & Hêlène, C. (1980) *Biochim. Biophys. Acta 606*, 95–104.
Williams, K. R., & Konigsberg, W. (1981) in *Gene Amplification and Analysis* (Chirirjian, J., & Pappas, T., Eds.) Vol. II, pp 476–508, Elsevier, Amsterdam.
Williams, K. R., Lo Presti, M. B., & Setoguchi, M. (1981) *J. Biol. Chem. 256*, 1754–1762.
Wu, J.-R., & Yeh, Y.-C. (1973) *J. Virol. 12*, 758–765.

Multiple Deoxyribonucleic Acid Dependent Adenosinetriphosphatases in FM3A Cells. Characterization of an Adenosinetriphosphatase That Prefers Poly[d(A-T)] as Cofactor†

Yasunori Tawaragi, Takemi Enomoto,* Yoshinari Watanabe,‡ Fumio Hanaoka, and Masa-atsu Yamada

ABSTRACT: Four chromatographically distinct DNA-dependent ATPases, B, C_1, C_2, and C_3, have been partially purified from mouse FM3A cell extracts. These ATPases are distinguished from each other by their physical and enzymological properties. DNA-dependent ATPases B, C_1, C_2, and C_3 have sedimentation coefficients in 250 mM KCl of 5.5, 5.3, 7.3, and 3.4 S, respectively. ATPases B, C_2, and C_3 hydrolyze dATP as efficiently as ATP, whereas C_1 does not. ATPase B hydrolyzes other ribonucleoside triphosphates with relatively high efficiency as compared to the other three enzymes. ATPase C_3 prefers poly[d(A-T)] to poly(dT) as cofactor, whereas the other three enzymes prefer poly(dT) to poly[d(A-T)]. Among the four ATPases, ATPase C_3 has been highly purified and characterized in detail. Sodium dodecyl sulfate–polyacrylamide gel electrophoresis of the most purified fraction of ATPase C_3 showed two major bands corresponding to molecular weights of 66 000 and 63 000. The K_m values of the enzyme for ATP and dATP are 0.53 and 0.86 mM, respectively. As cofactor, poly[d(A-T)] is the most effective among the DNAs tested. Heat-denatured DNA and native DNA are also effective but used with less efficiency. Almost no or very little activity has been detected with ribohomopolymers and oligonucleotides. The activity attained with poly(dT) and poly(dA) is 11 and 6% of that with heat-denatured DNA, respectively. When both polymers were added at a molar ratio 1 to 1, very high activity was obtained with these polymers. On the other hand, little activity was observed by the combination of noncomplementary homopolymers such as poly(dT) and poly(dG).

It is well known that ATP is required for DNA replication in various prokaryotic systems. Some of the proteins that have DNA-dependent ATPase activity such as *dnaB* protein, protein *n′*, and *rep* protein have been proved to play important roles in the replication of ϕX174 DNA by a combination of genetic and biochemical approaches (Kornberg, 1980). The requirement of ATP for DNA replication of eukaryotic cells has been also shown with various in vitro systems in permeabilized cells, cell lysate, and isolated nuclei. We have demonstrated an absolute requirement for high levels of ATP for the DNA synthesis in isolated nuclear systems (Tanuma et al., 1980; Nagata et al., 1981), especially for the synthesis of Okazaki fragments (Enomoto et al., 1981, 1983). The molecular basis of the requirement for ATP still remains unclear. To clarify the roles of ATP in the processes of DNA replication, we have paid attention to DNA-dependent ATPases and attempted to purify and characterize them.

In recent years, DNA-dependent ATPases have been isolated from various eukaryotic cells and tissues (Hachmann & Lezius, 1976; Otto, 1977; Hotta & Stern, 1978; Cobianchi et al., 1979; Assairi & Johnston, 1979; Boxer & Korn 1980; Plevani et al., 1980; DeJong et al., 1981; Yaginuma & Koike, 1981; Thomas & Mayer, 1982). We have also isolated three forms of DNA-dependent ATPases from calf thymus (Watanabe et al., 1981). Recently, we have sought analogous enzymes in FM3A cells, which are a good source for the isolation of DNA replication enzymes, because it is possible to obtain a large amount of S phase accumulated cells grown in mice by the method reported by us (Hanaoka et al., 1981). In addition, several mutants related to DNA replication have been isolated from the cells in our laboratory (Nakano et al., 1978; Nishimura et al., 1979; Tsai et al., 1979; Yasuda et al., 1981).

We described here the existence of multiple DNA-dependent ATPases in FM3A cell extracts and the purification and characterization of one form of the ATPases, ATPase C_3, which prefers poly[d(A-T)] as cofactor. Partial purification of another form of the ATPases, ATPase B, has been recently

† From the Department of Physiological Chemistry, Faculty of Pharmaceutical Sciences, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan. *Received June 30, 1983.* This work was supported in part by a Grant-in-Aid for Scientific Research and a Grant-in-Aid for Cancer Research from the Ministry of Education, Science and Culture, Japan.

‡ Present address: Department of Biochemistry I, Fukui Medical School, Matsuoka, Fukui 910-11, Japan.

0006-2960/84/0423-0529$01.50/0