Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 5666–5670, September 1985
Biochemistry

Amino acid sequence of the UP1 calf thymus helix-destabilizing protein and its homology to an analogous protein from mouse myeloma

(DNA unwinding protein/single-stranded DNA binding proteins/amino acid sequence/N^G,N^G-dimethylarginine)

KENNETH R. WILLIAMS*, KATHRYN L. STONE*, MARY B. LOPRESTI*, BARBARA M. MERRILL*, AND STEPHEN R. PLANCK†

*Department of Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, CT 06510; and the †Department of Cell Biology and Anatomy, Oregon Health Sciences University, Portland, OR 97201

Communicated by Bruce M. Alberts, May 6, 1985

ABSTRACT **A complete amino acid sequence has been determined for the UP1 single-stranded DNA binding protein from calf thymus that was first described by G. Herrick and B. M. Alberts [(1976) *J. Biol. Chem.* 251, 2124–2132]. Peptides required to establish the UP1 sequence were isolated by reversed-phase HPLC of digests produced by endoproteinase Lys-C, trypsin, chymotrypsin, *Staphylococcus aureus* V8 protease, and cyanogen bromide cleavage of UP1. The purified peptides were coupled to aminopolystyrene prior to solid-phase sequencing. UP1 contains 195 amino acids and has a molecular weight of 22,162. UP1 has a blocked NH_2 terminus and contains a single N^G,N^G-dimethylarginine residue near its COOH terminus. Gas-phase sequencing of tryptic peptides derived from an analogous protein from mouse myeloma cells [Planck, S. R. & Wilson, S. H. (1980) *J. Biol. Chem.* 255, 11547–11556] revealed that this mouse helix-destabilizing protein shares a high degree of sequence homology with UP1. Of the 59 amino acids in the mouse protein that have so far been found to be homologous with UP1, 48 correspond exactly to sequences found in UP1. Most of the 11 differences that have been found between the sequences of these two proteins are conservative in nature, involving primarily the interchange of chemically similar amino acids. One 9-residue mouse sequence that is not obviously homologous to UP1 may be a result of the larger size of the mouse myeloma protein as compared to UP1. Since none of the UP1 or mouse myeloma helix-destabilizing protein sequence appears to be homologous to that of any previously sequenced protein, we presume that these two proteins represent a distinct class of single-stranded nucleic acid binding proteins that probably play a role in metabolism of single-stranded RNA or DNA *in vivo*.**

The single-stranded DNA (ss DNA) binding proteins encoded by gene *32* of bacteriophage T4, gene *2.5* of bacteriophage T7, and *ssb* of *Escherichia coli* have served as prototypes for a class of proteins that are probably ubiquitous in nature (see refs. 1–3 for recent reviews). Two common features of these proteins *in vitro* are their ability to bind cooperatively and preferentially to ss DNA, thus destabilizing partially double-stranded polynucleotides such as poly[d(A-T)], and to stimulate the activity of their homologous DNA polymerase. Even though the ss DNA binding proteins from bacteriophages T4 and T7 and from *E. coli* lack extensive sequence homology, they all nonetheless contain a very acidic region near their COOH terminus that appears to be functionally homologous (4) and is also found in the *E. coli* F-plasmid ss DNA binding protein (5). By the use of conditional-lethal mutants, it has been possible to demonstrate that the bacteriophage T4 and *E. coli* ss DNA binding proteins fulfill functionally homologous and essential roles in DNA replication, repair, and recombination (see refs. 1–3 for reviews).

Although proteins analogous to the bacteriophage T4 and *E. coli* ss DNA binding proteins are probably present in eukaryotes, the absence of a well-defined and specific *in vitro* assay and the difficulty in using a genetic approach to define functions of individual proteins in eukaryotes has hampered the search for these proteins. A further complication in the identification of functionally homologous proteins in eukaryotes is that numerous proteins, such as dehydrogenases (6, 7), a protocollagen precursor (8), serum proteins involved in complement activation (9, 10), α_1-antichymotrypsin (11), and several other plasma proteins (12), that apparently have no *in vivo* role in DNA metabolism nonetheless bind reasonably well to ss DNA-cellulose. This affinity support is usually the first chromatographic step used for purifying ss DNA binding proteins.

Despite the obvious need for caution, several eukaryotic ss DNA binding proteins have been found that probably do play important roles in the metabolism of single-stranded nucleic acids *in vivo*. Two of these proteins are the calf thymus unwinding protein 1 (UP1, ref. 13) and the mouse myeloma helix-destabilizing protein (HDP-1, ref. 14). The mouse HDP-1 protein is heterogeneous with regard to apparent molecular weight, with individual species ranging from 24,000 to 33,000 (14). Although UP1 has a reported molecular weight of 24,000 (13), there is evidence for higher molecular weight (24,000–31,000) forms of this protein as well (15). Both the mouse HDP-1 and calf UP1 protein bind specifically but noncooperatively to ss DNA and in so doing they destabilize duplex DNA (14, 16). In addition, both proteins stimulate the activity of their homologous DNA polymerase (17, 18). To determine whether these two ss DNA binding proteins are structurally homologous and to assist in identifying functionally homologous proteins from other organisms, we have determined the complete primary structure for the calf thymus UP1 protein and some partial sequences for the mouse HDP-1.

MATERIALS AND METHODS

The calf thymus UP1 protein was purified by the procedure used by Herrick and Alberts (13) except that all buffers contained the following protease inhibitors: 1.0 mM aminoacetonitrile, 1.0 mM phenylmethylsulfonyl fluoride, and egg-white trypsin inhibitor (0.1 mg/ml). In the absence of these inhibitors, the resulting UP1 preparations contained several proteins which had molecular weights between 22,000 and 24,500 as judged by $NaDodSO_4$/PAGE. In contrast, in

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

Abbreviations: ss DNA, single-stranded DNA; HDP, DNA helix-destabilizing protein; UP, unwinding protein.

the presence of these inhibitors the resulting UP1 migrated as a single M_r 24,500 protein. Two other alterations that were made in the original isolation procedure (13) were that elution from the ss DNA-cellulose column was achieved with a linear 0.25–1.2 M NaCl gradient and that an additional anion-exchange step on PBE-94 resin (Pharmacia) was added after the ss DNA-cellulose chromatography. The PBE-94 column (1.5 × 15 cm) was equilibrated with 20 mM Tris Cl, pH 8.8/25 mM NaCl/0.1 mM dithiothreitol/10% (vol/vol) glycerol and the UP1 was eluted with a linear 0.025–0.2 M NaCl gradient. All of the various enzymatic digestions of UP1 were done on 25-nmol aliquots of trichloroacetic acid-precipitated protein that had been resolubilized in 0.25 ml of 8 M urea. After dilution to a final concentration of 2 M urea with 50 mM NH_4HCO_3 at pH 8.0 (for trypsin, chymotrypsin, or *Staphylococcus aureus* V8 protease), the enzyme was added at a UP1/enzyme ratio (wt/wt) of 30:1 and digestion was allowed to proceed at 37°C for 24 hr. In the case of endoproteinase Lys-C (Boehringer Mannheim), 25 mM Tris Cl at pH 7.6 was used in place of NH_4HCO_3 and digestion was at a UP1/enzyme (wt/wt) ratio of 400:1 for 4 hr at 37°C. After each digestion, the cysteine residues were carboxamidomethylated by addition of dithiothreitol to a final concentration of 10 mM followed by 50 μCi (1 Ci = 37 GBq) of iodo[1-^{14}C]acetamide. After 10 min at room temperature, unlabeled iodoacetamide was added to a final concentration of 20 mM. After an additional 20 min, the carboxamidomethylated digests were injected directly onto a Waters Associates C_{18} μBondapak column equilibrated with 0.1% trifluoroacetic acid.

The mouse myeloma HDP-1 was purified through the Sephacryl S-200 column step as described (14) except that the second ss DNA-cellulose column was omitted and preparative $NaDodSO_4$/PAGE was added. The HDP-1 preparation contained two major bands with apparent molecular weights of 24,000 and 27,000, respectively (see electrophoresis of fraction IV, ref. 14). The mouse HDP-1 was precipitated with trichloroacetic acid prior to digestion with trypsin as described above.

All of the HPLC-purified UP1 peptides were coupled to aminopolystyrene and then subjected to solid-phase sequencing as described (19). The mouse HDP-1 tryptic peptides (0.5–2 nmol) were dissolved in 0.1 ml of 100% trifluoroacetic acid and were subjected to gas-phase sequencing using the 02NRUN program supplied with the Applied Biosystems instrument. Procedures used for amino acid analysis and for identification of the phenylthiohydantoin derivatives that resulted from gas- and solid-phase sequencing have been described (20).

RESULTS AND DISCUSSION

Both the mouse HDP-1 and the calf thymus UP1 protein appear to have blocked NH_2 termini. Attempts to sequence 32 nmol of UP1 and 2.8 nmol of the mouse HDP-1 on a Beckman 890C liquid-phase sequencer failed to produce any discernible sequence. The UP1 sequence shown in Fig. 1 was therefore derived entirely from solid-phase sequencing of peptides that resulted from cleavage with endoproteinase Lys-C, trypsin, chymotrypsin, *S. aureus* V8 protease, and cyanogen bromide. In each instance the resulting peptides were separated by HPLC on C_{18} μBondapak columns and then coupled to aminopolystyrene prior to solid-phase sequencing. By far, the most fruitful digest was that obtained with endoproteinase Lys-C. Every major peak that resulted from reversed-phase chromatography of this digest (Fig. 2 *Upper*) was subjected to solid-phase sequencing from which more than 91% of the proposed UP1 sequence (Fig. 1) was derived. No cleavage occurred at any of the 15 arginines in UP1 and the only lysine residue that was resistant to cleavage was that contained within the sequence -proline-lysine-

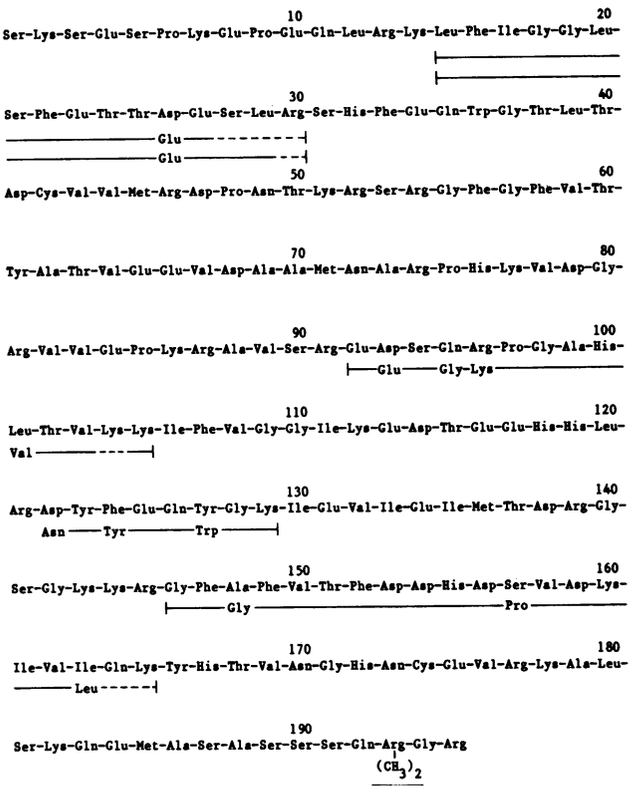

FIG. 1. Amino acid sequence of the UP1 ss DNA binding protein from calf thymus. The N^G,N^G-dimethylarginine at position 193 is shown. Five of the six tryptic peptides from the mouse myeloma HDP-1 that have been sequenced are indicated below the UP1 sequence; sequences shown to be identical in these two proteins on the basis of gas-phase sequencing or amino acid composition are represented by solid lines or dashed lines, respectively.

glutamic acid-proline- (residues 6–9). Acidic amino acids apparently decrease the rate of endoproteinase Lys-C attack, since only about 50–60% cleavage was observed at lysine-112 (Fig. 1), which is followed by two acidic residues. In the case of a lysine-lysine sequence, cleavage occurs primarily in between the two lysines and to a lesser extent after the second lysine (see yields for cleavages after lysines 104–105 and 143–144 in Table 1). The sequence of the first two amino acids in UP1 was deduced from the amino acid compositions of a tryptic, chymotryptic, and *S. aureus* protease peptide that we were unable to sequence. The serine-lysine dipeptide that is predicted to be at the NH_2 terminus of UP1 (Fig. 1) was in fact isolated from the endoproteinase Lys-C digest (Fig. 2 *Upper* and Table 1). The COOH terminus of UP1 was confirmed by the finding that the peptide corresponding to residues 183–195 was the only endoproteinase Lys-C peptide isolated that did not contain lysine. Similarly, the peptide corresponding to residues 185–195 was the only *S. aureus* protease peptide isolated that did not contain an aspartic or glutamic acid. The amino acid compositions for UP1 (Table 2) and its endoproteinase Lys-C peptides (Table 1) are in excellent agreement with the compositions expected from the UP1 sequence shown in Fig. 1. UP1 has a calculated molecular weight of 22,162, which is within 10% of the value of 24,500 that we estimated from the mobility of UP1 in $NaDodSO_4$/PAGE. UP1 contains a single N^G,N^G-dimethylarginine (unsymmetric dimethylarginine) residue near its COOH terminus (Fig. 1). This unusual amino acid has been found in histones (21) and high mobility group proteins (22) as well as in ribosomal proteins (23) and in proteins associated with heterogeneous RNA (hnRNP particles, refs. 24–26). A more detailed account of the sequencing studies on UP1 is in preparation.

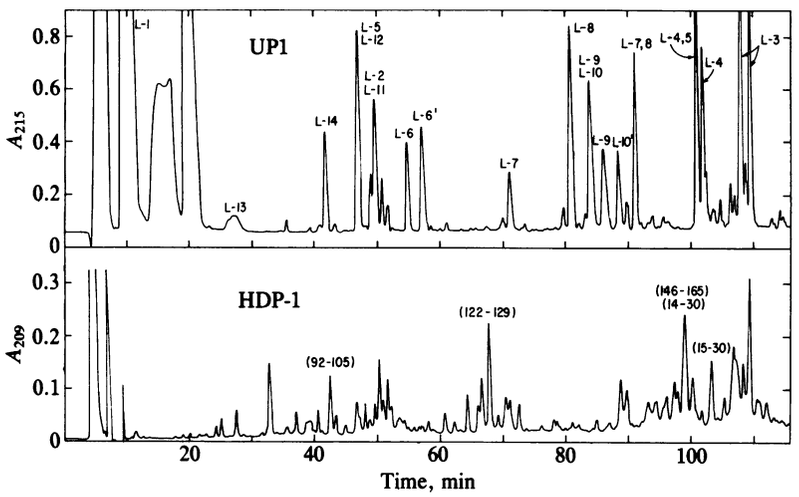

FIG. 2. Reversed-phase HPLC of peptides derived from the calf thymus UP1 (*Upper*) and mouse myeloma HDP-1 protein (*Lower*). The C_{18} μBondapak column was equilibrated with 0.1% (*Upper*) or 0.05% (*Lower*) trifluoroacetic acid. Elution (flow rate 0.7 ml/min) was with increasing concentrations of buffer B (0.05% trifluoroacetic acid/80% acetonitrile): 0–30% B (0–86 min); 30–60% B (86–129 min). (*Upper*) Endoproteinase Lys-C digest of 31.2 nmol of calf thymus UP1. (*Lower*) Tryptic digest of 7.5 nmol of mouse myeloma HDP-1. The UP1 digest was carboxamidomethylated before HPLC.

Our contention that the calf thymus ss DNA binding protein whose sequence is shown in Fig. 1 actually is the UP1 protein reported by Herrick and Alberts (16) is supported by the excellent agreement between both the molecular weight and the amino acid composition predicted from Fig. 1 and that observed previously (16). In addition, our UP1 protein is eluted from ss DNA cellulose with about 0.6 M NaCl, which is just slightly higher than the value of about 0.5 M NaCl reported for UP1 by Herrick and Alberts (16).

Both the UP1 protein from calf thymus and the HDP-1 protein from mouse myeloma cells stimulate the activity of their homologous DNA polymerase α and are heterogeneous with respect to molecular weight (14, 15, 18, 27). Their similar properties suggest that these two proteins are analogous (14). Indeed, both the mouse HDP-1 and calf thymus UP1 proteins have similar amino acid compositions (Table 2). Comparative HPLC tryptic peptide maps for these two proteins are also similar. As shown in Fig. 3, tryptic peptide maps for both proteins contained three major clusters of peptides that were eluted between 35–55, 70–75, and 90–110 min, respectively. Although their relative intensities varied, several peaks appeared to be eluted at the same position in the two chromatograms (Fig. 3). Definitive proof that these two helix-destabilizing proteins are structurally homologous was obtained by gas-phase sequencing of several mouse tryptic peptides that were isolated by reversed-phase HPLC (Fig. 2 *Lower*). Of the 59 mouse HDP-1 amino acids so far sequenced, 48 correspond exactly to sequences found in the calf thymus UP1 protein (Fig. 1). In addition, most of the 11 substitutions that have been found are relatively conservative, involving the interchange of chemically similar amino acids such as aspartic and glutamic acid, lysine and arginine, isoleucine and leucine, aromatic amino acids, and glycine and alanine. In addition to the mouse HDP-1 sequence shown in Fig. 1, one other tryptic peptide has been partially sequenced (glycine-glycine-asparagine-phenylalanine-glycine-phenylalanine-glycine-aspartic acid-serine) that does not appear to be homologous to UP1. This mouse HDP-1 sequence ei- ther results from the larger size of mouse HDP-1 (M_r 24,000–27,000 as compared to the estimate of 24,500 derived for UP1 from NaDodSO$_4$/PAGE) or perhaps arises from a region in the UP1 sequence that is not as highly conserved as those shown in Fig. 1. Nonetheless, our results clearly show that the UP1 protein from calf thymus and the HDP-1 from mouse myeloma cells are structurally homologous and, therefore, probably functionally homologous as well.

The entire sequence of UP1 (in 25-residue segments starting at every 15th residue) as well as all of the available

Table 1. Amino acid compositions of endoproteinase Lys-C peptides of UP1

	L-1	L-2[a]	L-3	L-4	L-5[b]	L-4,5	L-6	L-6′	L-7	L-8	L-7,8	L-9[c]	L-10[c]	L-10′	L-11[a]	L-12[b]	L-13	L-14
Cys			ND (1)													ND (1)		
Asp			4.6(4)	2.2(2)	1.2(1)	3.1(3)	1.2(1)	1.1(1)		2.0(2)	2.6(2)	1.3(1)	4.0(4)	3.8(4)		2.4(2)		
Thr			4.5(5)	1.6(2)		1.8(2)	0.9(1)	0.9(1)		1.1(1)	1.1(1)	0.9(1)	0.9(1)	0.9(1)		0.8(1)		
Ser	0.9(1)	1.5(2)	2.8(3)	1.7(1)		1.0(1)	1.5(2)	1.5(2)				0.8(1)	0.8(1)	0.9(1)			0.8(1)	3.0(4)
Glu		4.3(4)	4.6(4)	2.2(2)	1.1(1)	3.3(3)	2.1(2)	2.1(2)		4.9(5)	4.3(5)	1.7(2)		0.4	1.1(1)	1.1(1)		3.3(3)
Pro		1.9(2)	1.0(1)	0.9(1)	0.8(1)	1.8(2)	0.9(1)	0.9(1)										
Gly	0.4		3.7(4)	2.7(2)	0.7(1)	3.4(3)	1.3(1)	1.2(1)	2.2(2)	1.3(1)	3.5(3)	2.1(2)	1.1(1)	1.5(1)		1.4(2)		1.2(1)
Ala	0.3			3.6(4)		4.0(4)	1.8(2)	2.0(2)					1.0(1)	1.0(1)			1.0(1)	2.2(2)
Val			1.9(2)	2.9(3)	2.6(3)	5.6(6)	2.3(2)	2.0(2)	1.1(1)		1.4(1)	1.1(1)	2.1(2)	2.0(2)	1.1(1)	1.8(2)		
Met			0.8(1)	0.6(1)		0.8(1)						0.7(1)						1.0(1)
Ile			1.0(1)						1.7(2)		1.5(2)	2.2(3)			1.8(2)			
Leu		1.1(1)	4.3(4)				1.0(1)	1.0(1)		1.0(1)	0.8(1)						1.1(1)	
Tyr				0.9(1)		1.0(1)				1.8(2)	1.6(2)					1.1(1)		
Phe			3.0(3)	1.7(2)		2.0(2)			1.0(1)	0.9(1)	2.3(2)	0.4	2.9(3)	2.8(3)				
His			1.1(1)	1.1(1)		1.2(1)	1.2(1)	1.0(1)		1.8(2)	1.8(2)		1.1(1)	1.1(1)		2.2(2)		
Lys	1.1(1)	2.1(2)	1.2(1)	1.1(1)	1.0(1)	2.0(2)	2.0(2)	1.1(1)	1.0(1)	1.0(1)	1.8(2)	1.0(1)	2.1(2)	1.1(1)	1.1(1)	1.0(1)	1.1(1)	
Arg		1.2(1)	2.3(2)	2.9(3)	1.0(1)	4.2(4)	3.1(3)	3.1(3)		1.0(1)	1.2(1)	1.0(1)	1.0(1)	1.0(1)		1.0(1)		1.2(1)
Trp			ND (1)															
DNA[d]																		1.2(1)
Residue Numbers	1–2	3–14	15–51	52–77	78–86	52–86	87–105	87–104	106–112	113–129	106–129	130–143	144–160	145–160	161–165	166–178	179–182	183–195
Yield (%)	18	70	66	42	44	38	34	51	47	49	32	91	51	19	85	58	86	62

[a]This peptide was isolated as a mixture containing 21.8 nmol L-2 and 14.5 nmol L-11.
[b]This peptide was isolated as a mixture containing 13.8 nmol L-5 and 18.0 nmol L-12.
[c]This peptide was isolated as a mixture containing 7.6 nmol L-9 and 16.0 nmol L-10.
[d]N^G,N^G-dimethylarginine (unsymmetrical).

Table 2. Amino acid compositions of mouse HDP-1 and calf UP1 and of mouse HDP-1 tryptic peptides

	Mole %		Residues per mole of protein or peptide						
					HDP-1 Peptides				
	HDP-1	UP1	HDP-1[a]	UP1[b]	14–30[c]	15–30[d]	92–105	122–129	146–165
Asp	9.0	9.1	21.0	17.3 (17)		1.2		1.1 (1)	4.9 (4)
Thr	4.4	5.5	10.2	10.5 (12)	1.9 (2)	1.9 (2)	1.0 (1)		1.0 (1)
Ser	6.0	6.5	14.0	12.3 (16)	2.3 (2)	2.2 (2)	1.2 (1)		
Glu	14.5	13.7	33.8	26.1 (25)	3.0 (3)	3.2 (3)	2.4 (2)	2.0 (2)	1.1 (1)
Pro	3.4	2.8	7.8	5.4 (6)			0.8 (1)		0.9 (1)
Gly	10.3	8.4	24.1	16.0 (15)	2.5 (2)	2.6 (2)	2.3 (2)	1.6 (1)	2.4 (2)
Ala	11.4	5.5	26.6	10.4 (10)		0.9	1.0 (1)		
Val	6.4	9.0	15.0	17.1 (17)			1.7 (2)		2.4 (3)
Met	2.3	1.9	5.3	3.6 (4)					
Ile	4.3	4.2	10.1	7.9 (8)	0.7 (1)	1.0 (1)			0.7 (1)
Leu	5.2	4.5	12.1	8.6 (8)	2.7 (3)	2.3 (3)			0.8 (1)
Tyr	2.6	2.1	6.1	4.0 (4)				1.6 (2)	
Phe	5.0	5.3	11.6	10.0 (10)	1.8 (2)	2.1 (2)			2.8 (3)
His	2.8	4.4	6.5	8.3 (8)		0.6	0.9 (1)		1.1 (1)
Lys	6.5	8.5	15.2	16.1 (16)	1.0 (1)		1.7 (2)	0.7 (1)	2.1 (2)
Arg	6.1	8.8	14.2	16.7 (15)	0.9 (1)	0.6 (1)			
Me_2Arg[e]	—	0.5	—	1.0 (1)					
Yield (%)					13	13	20	29	13

[a]Based on molecular weight of 25,500.
[b]Based on a molecular weight of 22,162 as calculated from the sequence shown in Fig. 3. Numbers in parentheses were also based on this sequence.
[c]Based on amino acid sequencing, this peptide was isolated as an equimolar mixture with that corresponding to residues 146–165 (see Fig. 2). Amino acid compositions shown were based on this assumption.
[d]Although its composition indicates that this peptide is not pure, no secondary sequence was detected after gas-phase sequencing.
[e]N^G,N^G-dimethylarginine.

sequence from the mouse HDP-1 has now been searched for in the National Biomedical Research Foundation's Protein Data Base. This analysis failed to detect any significant homology with any other protein sequence in that data base. The mouse HDP-1 and calf UP1 proteins therefore seem to belong to a previously unidentified family of eukaryotic single-stranded nucleic acid binding proteins. The completion of the amino acid sequence of UP1 and the apparently high degree of sequence homology among UP1-like proteins, as evidenced by our studies on the mouse myeloma HDP-1, should make it possible to rapidly identify functionally homologous proteins in other organisms. Based on our studies on mouse HDP-1, even relatively short stretches of sequence should be sufficient for this purpose. In the absence of a specific *in vitro* assay for this class of proteins, it would seem advisable to obtain some amino acid sequence for any otherwise unidentified eukaryotic, single-stranded nucleic acid binding protein. Such an analysis succeeded in identifying one rat liver HDP as lactate dehydrogenase (7) and, because of their known primary structures, should easily identify high mobility group (HMG) proteins. This latter group of proteins share many properties in common with prokaryotic ss DNA binding proteins. In fact, another rat liver ss DNA binding protein, this one isolated by Duguet and de Recondo (28), was subsequently shown to be HMG-1 (29). This particular HMG protein [at least when isolated from regenerating rat liver (28)] behaves much like the *E. coli* and bacteriophage T4 ss DNA binding proteins in that it binds preferentially to single-stranded nucleic acids, destabilizes double-stranded DNA, and stimulates *in vitro* the activity of its homologous DNA polymerase (28).

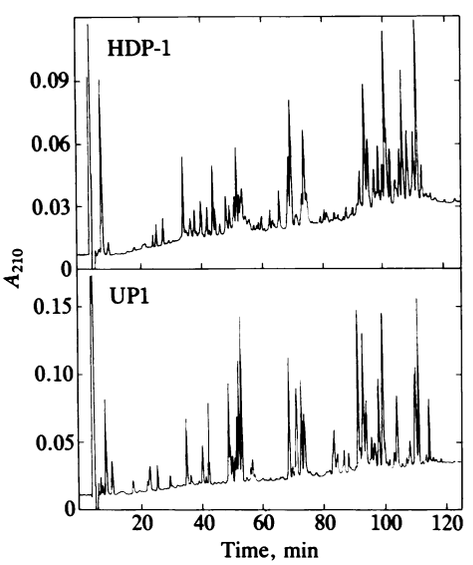

FIG. 3. Reversed-phase HPLC of tryptic peptides from 0.9 nmol of the mouse myeloma HDP-1 (*Upper*) or from 1.0 nmol of calf thymus UP1 (*Lower*). Elution from the C_{18} μBondapak column was as described in the legend to Fig. 2.

Although progress has been made in determining the primary structures of mouse myeloma HDP-1 and calf thymus UP1 proteins, our knowledge of the roles played by these proteins *in vivo* is still largely speculative and is based primarily on inferences drawn from studies on prokaryotic DNA binding proteins. Both the mouse and calf thymus helix-destabilizing proteins bind noncooperatively to single-stranded nucleic acids (14, 16), stimulate the activity of DNA polymerase α (17, 18), and are present in large amounts *in vivo* ($\approx 10^6$ copies per cell, refs. 13 and 14). Both proteins have been found in the cytoplasm (14, 30) but at least 75% of the mouse protein is found in the nucleus. There is no evidence that either protein is associated with chromatin. The ability of these two proteins to stimulate the activity of their homologous DNA polymerase might argue for a role in DNA replication, whereas their ability to bind single-stranded RNA and their appearance in both the nucleus and the cytoplasm might favor a role in ribonucleoprotein metabo-

lism. In support of the latter notion, both UP1 and mouse HDP-1 lack two of the distinguishing and probably essential (1, 4, 31) characteristics of those prokaryotic DNA binding proteins that have been shown to play an active role in DNA replication, repair, and recombination. That is, UP1 and the mouse HDP-1 bind ss DNA noncooperatively, as contrasted with the highly cooperative binding of the *E. coli* and bacteriophage T4 ss DNA binding proteins and, based on the available sequence, neither UP1 nor mouse HDP-1 contain an acidic COOH terminus. The acidic COOH terminus of the prokaryotic ss DNA binding proteins is thought to be essential for modulating the helix-destabilizing activity of these proteins and perhaps also for interacting with other proteins in the replication complex (4, 31). At least with the bacteriophage T4 gene *32* protein, these latter protein–protein interactions may help to restrict the helix-destabilizing activity of gene *32* protein to just that region in front of an advancing DNA replication complex (31, 32). This line of reasoning is weakened by the presence of higher molecular weight ss DNA binding proteins during the purification of both UP1 (15) and HDP-1 (14). Since both proteins have blocked NH_2 termini, it is likely that if these higher molecular weight forms are in fact related to UP1 and HDP-1, then they must have additional sequences at their COOH termini. They may therefore contain acidic COOH termini after all. It is also possible that these higher molecular weight species might differ in other properties such as their cooperativity and strength of binding to ss DNA. Valentini *et al.* (27) have found that the various M_r 24,000–28,000 ss DNA binding proteins that they isolate from calf thymus do, in fact, differ in their affinities for ss DNA and in their abilities to stimulate DNA polymerase α. It is not known whether these higher molecular weight forms are encoded by different genes. If on the other hand, UP1 represents a limited-proteolysis product, it is not known whether this cleavage occurs *in vivo* or *in vitro* during the purification.

It is becoming increasingly clear that eukaryotes contain a wide variety of proteins that have a preferential affinity for single-stranded nucleic acids (33). Some of these proteins probably are similar to the prokaryotic ss DNA binding proteins and will be found to play essential roles in DNA replication, recombination, and repair. It is, however, quite possible that some of these eukaryotic proteins may be more highly specialized than their prokaryotic counterparts. For instance, although the ss DNA binding protein from the meiotic cells of lily plants seems to play a role in chromosome recombination, it is unlikely that this protein is also required for DNA replication (34–36). Eukaryotes probably contain several ss RNA binding proteins that could be envisioned to play essential roles in transcription, heterogeneous RNA processing, RNA transport, and protein synthesis. Since most single-stranded nucleic acid binding proteins bind both ss DNA and ss RNA, additional criteria will probably be required before a decision can be made as to whether a particular eukaryotic, single-stranded nucleic acid binding protein functions primarily in RNA or DNA metabolism *in vivo*. Such is certainly the case with the calf thymus UP1 and mouse HDP-1 proteins. Although these two proteins seem to belong to a distinct class of single-stranded nucleic acid binding proteins, additional studies are required to establish their *in vivo* functions.

We thank William Konigsberg for his advice and encouragement and Bruce Alberts for his assistance in selecting suitable protease inhibitors. This work was supported by National Institutes of Health Grants GM31539 (K.R.W.), GM12607 (William Konigsberg), and CA30466 and GM34543 (S.R.P.).

1. Kowalczykowski, S. C., Bear, D. G. & von Hippel, P. H. (1981) in *The Enzymes*, ed. Boyer, P. (Academic, New York), Vol. 14a, pp. 373–444.
2. Williams, K. R. & Konigsberg, W. H. (1981) in *Gene Amplification and Analysis*, eds. Chirikjian, J. G. & Papas, T. S. (Elsevier/North-Holland, New York), Vol. 2, pp. 475–508.
3. Williams, K. R. & Konigsberg, W. H. (1983) in *Bacteriophage T4*, eds. Mathews, C. K., Kutter, E. M., Mosig, G. & Berget, P. B. (Am. Soc. Microbiol., Washington, DC), pp. 82–89.
4. Williams, K. R., Spicer, E. K., LoPresti, M. B., Guggenheimer, R. A. & Chase, J. W. (1983) *J. Biol. Chem.* **258,** 3346–3355.
5. Chase, J. W., Merrill, B. M. & Williams, K. R. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 5480–5484.
6. Perucho, M., Salas, J. & Salas, M. L. (1977) *Eur. J. Biochem.* **81,** 557–562.
7. Patel, G. L., Reddigari, S., Williams, K. R., Baptist, E., Thompson, P. E. & Sisodia, S. (1985) in *Proceedings of the Tenth International Cell Cycle Conference* (Humana, Clifton, NJ), in press.
8. Tsai, R. L. & Green, H. (1972) *Nature (London) New Biol.* **237,** 171–173.
9. Gardner, W. D., Haselby, J. A. & Hoch, S. O. (1980) *J. Immunol.* **124,** 2800–2806.
10. Gardner, W. D., White, P. J. & Hoch, S. O. (1980) *Biochem. Biophys. Res. Commun.* **94,** 61–67.
11. Siddiqui, A. A., Hughes, A. E., Davies, R. J. H. & Hill, J. A. (1980) *Biochem. Biophys. Res. Commun.* **95,** 1737–1742.
12. Kroll, J., Larsen, J. K., Loft, H., Ezban, M., Wallevik, K. & Faber, M. (1976) *Biochim. Biophys. Acta* **434,** 490–501.
13. Herrick, G. & Alberts, B. (1976) *J. Biol. Chem.* **251,** 2124–2132.
14. Planck, S. R. & Wilson, S. H. (1980) *J. Biol. Chem.* **255,** 11547–11556.
15. Riva, S., Clivio, A., Valentini, O. & Cobianchi, F. (1980) *Biochem. Biophys. Res. Commun.* **96,** 1053–1062.
16. Herrick, G. & Alberts, B. (1976) *J. Biol. Chem.* **251,** 2133–2141.
17. Herrick, G., Delius, H. & Alberts, B. (1976) *J. Biol. Chem.* **251,** 2142–2146.
18. Detera, S. D., Becerra, S. P., Swack, J. A. & Wilson, S. H. (1981) *J. Biol. Chem.* **256,** 6933–6943.
19. Hemmings, H. C., Williams, K. R., Konigsberg, W. H. & Greengard, P. (1984) *J. Biol. Chem.* **259,** 14486–14490.
20. Merrill, B. M., Williams, K. R., Chase, J. W. & Konigsberg, W. H. (1984) *J. Biol. Chem.* **259,** 10850–10856.
21. Paik, W. K. & Kim, S. (1967) *Biochem. Biophys. Res. Commun.* **29,** 14–20.
22. Boffa, L. C., Sterner, R., Vidali, G. & Allfrey, V. G. (1979) *Biochem. Biophys. Res. Commun.* **89,** 1322–1327.
23. Chang, F. N., Navickas, I. J., Chang, C. N. & Dancis, B. M. (1976) *Arch. Biochem. Biophys.* **172,** 627–633.
24. Boffa, L. C., Karn, J., Vidali, G. & Allfrey, V. G. (1977) *Biochem. Biophys. Res. Commun.* **74,** 969–976.
25. Christensen, M. E., Beyer, A. L., Walker, B. & LeStourgeon, W. M. (1977) *Biochem. Biophys. Res. Commun.* **74,** 621–629.
26. Marvil, D. K., Nowak, L. & Szer, W. (1980) *J. Biol. Chem.* **255,** 6466–6472.
27. Valentini, O., Biamonti, G., Mastromei, G. & Riva, S. (1984) *Biochim. Biophys. Acta* **782,** 147–155.
28. Duguet, M. & de Recondo, A. M. (1978) *J. Biol. Chem.* **253,** 1660–1666.
29. Bonne, C., Sautiere, P., Duguet, M. & de Recondo, A. M. (1982) *J. Biol. Chem.* **257,** 2722–2725.
30. Karpel, R. L., Miller, N. S. & Fresco, J. R. (1976) in *Molecular Mechanisms in the Control of Gene Expression*, eds. Nierlich, D. P., Rutter, W. J. & Fox, C. F. (Academic, New York), pp. 411–419.
31. Burke, R. L., Alberts, B. M. & Hosoda, J. (1980) *J. Biol. Chem.* **255,** 11484–11493.
32. Moise, H. & Hosoda, J. (1976) *Nature (London)* **259,** 455–458.
33. Falaschi, A., Cobianchi, F. & Riva, S. (1980) *Trends Biochem. Sci.* **5,** 154–157.
34. Hotta, Y. & Stern, H. (1971) *Dev. Biol.* **26,** 87–99.
35. Hotta, Y. & Stern, H. (1971) *Nature (London) New Biol.* **234,** 83–86.
36. Hotta, Y. & Stern, H. (1979) *Eur. J. Biochem.* **95,** 31–38.