The EMBO Journal vol.5 no.9 pp.2267–2273, 1986

Mammalian single-stranded DNA binding protein UP I is derived from the hnRNP core protein A1

S.Riva[1], C.Morandi[2], P.Tsoulfas[1], M.Pandolfo[1], G.Biamonti[1], B.Merrill[3], K.R.Williams[3], G.Multhaup[4], K.Beyreuther[4], H.Werr[5], B.Henrich[5] and K.P.Schäfer[5]

[1]Istituto di Genetica Biochimica ed Evoluzionistica CNR Via Abbiategrasso 207, 27100 Pavia, [2]Dipartimento di Genetica e Microbiologia, Universitá di Pavia, Italy, [3]Department of Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, CT 06510, USA, [4]Institute for Genetics, University of Köln, Weyertal 121, D-5000 Köln, and [5]Ruhr-Universität Bochum, Lehrstuhl für Biochemie, D-4630 Bochum 1, FRG

Communicated by K.Beyreuther

Antibodies induced against mammalian single-stranded DNA binding protein (ssDBP) UP I were shown to be cross-reactive with most of the basic hnRNP core proteins, the main constituents of 40S hnRNP particles. This suggested a structural relationship between both groups of proteins. Using the anti-ssDBP antibodies, a cDNA clone (pRP10) was isolated from a human liver cDNA library in plasmid expression vector pEX1. By DNA sequencing this clone was shown to encode in its 949 bp insert the last 72 carboxy terminal amino acids of the ssDBP UP I. Thereafter, an open reading frame continued for another 124 amino acids followed by a UAA (ochre) stop codon. Direct amino acid sequencing of a V8 protease peptide from hnRNP core protein A1 showed that this peptide contained at its amino terminus the last 11 amino acids of UP I followed by 19 amino acids which are encoded by the open reading frame of cDNA clone pRP10 immediately following the UP I sequence. This proves that ssDBP UP I arises by proteolysis from hnRNP core protein A1. This finding must lead to a re-evaluation of the possible physiological role of UP I and related ssDBPs. The formerly assumed function in DNA replication, although not completely ruled out, should be reconsidered in the light of a possible alternative or complementary function in hnRNA processing where UP I could either be a simple degradation product of core protein A1 (as a consequence of controlling the levels of active A1) or may continue to function as an RNA binding protein which has lost the ability to interact with the other core proteins. This conclusion is supported by the location of the proteolytic cut which leads to UP I. Northern blotting experiments reveal two major species of mRNA at 1.5 and 1.8 kb. They are expressed at levels several hundred-fold above that of dihydrofolate reductase mRNA.

Key words: hnRNP core proteins/hnRNA processing/proteolysis/single-stranded DNA binding proteins

Introduction

Mammalian single-stranded DNA binding proteins (ssDBPs) have been purified from calf thymus, HeLa, mouse myeloma and other eukaryotic cells (Herrick and Alberts, 1976; Planck and Wilson, 1980; Riva *et al.*, 1980; Valentini *et al.*, 1984; Chase and Williams, 1986). Calf thymus UP I (Herrick and Alberts, 1976) and mouse myeloma HDP-1 (Planck and Wilson, 1980) are typical examples. SsDBPs bind preferentially to ssDNA and depress the melting temperature of double-stranded DNA (dsDNA), hence their alternative denomination as helix-destabilizing proteins (HDP). However, since these proteins bind equally well to ssRNA and ssDNA and promote conformational changes in RNA with secondary structure, their correct designation would be single-stranded nucleic acid binding proteins or nucleic acid helix-destabilizing proteins.

Under certain conditions ssDBPs can stimulate DNA polymerase α activity *in vitro* (Herrick and Alberts, 1976; Riva *et al.*, 1980; Henner and Furth, 1977). Whether these proteins participate in chromosome replication as DNA polymerase accessory proteins is still obscure. In most mammalian cells, the properties of ssDBPs are consistent with the assumption that these proteins are structurally related despite their differences in size and function (Riva *et al.*, 1980; Chase and Williams, 1986). This and other observations led some workers to propose that ssDBPs derive from one or more higher molecular weight precursors (Valentini *et al.*, 1984).

In a previous paper (Valentini *et al.*, 1985) it was shown that antibodies raised against bovine ssDBP cross-react with the main proteins (core proteins) constituting the 40S hnRNP particles (M_r 31–40 kd) in both calf thymus and HeLa cells. An antigenic and structural correlation between these two families of proteins was thus established. HnRNP proteins are a family of about nine different but related polypeptides (Wilk *et al.*, 1985). SDS–polyacrylamide gel electrophoresis of 40S complexes reveals three closely spaced doublets corresponding to the A, B and C proteins, respectively (Beyer *et al.*, 1977). These proteins have a high selective affinity for single-stranded nucleic acids and are suggested to participate in post-transcriptional hnRNA processing and transport (Samarina *et al.*, 1968; Pederson, 1976; LeStourgeon *et al.*, 1978; Pederson, 1983). More recently, it has been shown that hnRNP core proteins can undergo specific proteolysis in crude extracts resulting in the production of the described ssDBP. In addition, a trypsin-like protease which cleaves hnRNP core proteins to yield ssDBPs was purified from HeLa cells (Pandolfo *et al.*, 1985).

Here, the correlation between ssDBPs and hnRNP core proteins is studied in more detail by means of cDNA cloning, h.p.l.c. comparative peptide mapping and amino acid sequencing of peptides and fragments obtained after limited proteolysis. The results of these experiments unequivocally prove that the 24 kd ssDBP (UP I) is derived from the 31 kd core protein A1 (M_r on SDS gels: Wilk *et al.*, 1985) through specific proteolysis C-terminal of amino acid 195 of A1. The discovery of the hnRNP protein A1/ssDBP UP I relationship raises several interesting questions on the possible biological significance of this conversion in relation to both DNA replication and hnRNP metabolism.

Results

Comparative h.p.l.c. peptide mapping on ssDBP UP I and hnRNP proteins A1/2

A possible relation between ssDBP and some of the hnRNP core proteins has been surmised previously (Valentini *et al.*, 1985;

© IRL Press Limited, Oxford, England

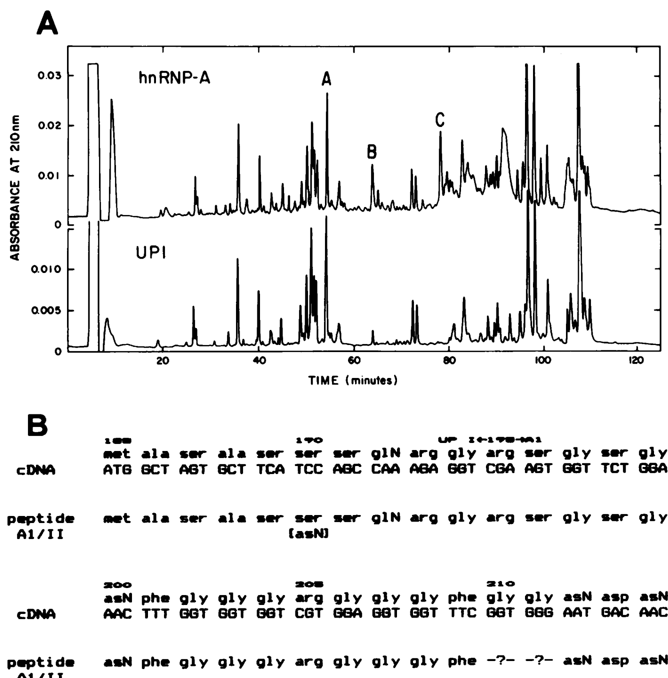

Fig. 1. **(A)** Comparative h.p.l.c. tryptic peptide mapping of the hnRNP A1/2 peptides (top) and UP I (bottom). Peptides were eluted as described. Aliquots of peaks labeled A, B and C were taken for gas phase sequencing (for the sequence of the three peptides, see text). **(B)** Comparison of amino acid sequence of core protein peptide A1/II with DNA sequence of putative cDNA clone of A1 (see Figure 3). Core protein A1 from HeLa cells was excised from two-dimensional gels and then digested by V8 protease (Glu-specific) as described. Peptide A1/II was electroeluted from the polyacrylamide gel (see Figure 4) and sequenced by gas phase sequencing. At position 6 a serine and an asparagine derivative were found. Positions 26 and 27 could not be determined.

Pandolfo *et al.*, 1985) on the basis of indirect evidence. Here, we describe a more direct approach to the problem based on the use of h.p.l.c. comparative peptide mapping and amino acid sequencing of peptides and large fragments.

The comparative h.p.l.c. tryptic peptide maps shown in Figure 1A demonstrate a close relationship between the 24 kd UP I ssDNA binding protein from calf thymus and the 31/32.5 kd A1/2 hnRNP proteins from HeLa cells. A1 and A2 are two closely related proteins which constitute the major components of the 40S hnRNP complex (Beyer *et al.*, 1977; Wilk *et al.*, 1983). As shown in Figure 1, in every instance where there is a major peak present in the UP I chromatogram there is a corresponding peak in the A1/2 profile. As expected from their larger size, there are several additional peaks present in the A1/2 digest that are not found in UP I, suggesting that UP I is a proteolytic product derived from these two proteins. Further confirmation of this idea comes from the amino acid sequences of the three A1/2 peptides labeled A–C in Figure 1. Peptide 'A' (with the sequence Glu-Asp-Ser-Gln-Arg-Pro-Gly-Ala-His-Leu-Thr-Val-Lys) corresponds exactly to residues 92–104 of UP I (Williams *et al.*, 1985). Peptide 'B' has the sequence Asp-Tyr-Phe-Glu-Gln-(Tyr/Phe)-Gly-Lys which corresponds to residues 122–129 of UP I with the exception of Phe at a position corresponding to residue 127. The ratio of Tyr/Phe at this cycle in the sequence of peptide 'B' was almost exactly 2:1 which suggests that this peak from the h.p.l.c. actually contains two peptides in a 2:1 ratio. One of these peptides could be derived from A1 and the other from A2 or — if A1 and A2 are themselves mixtures of homologs — may represent a subpeptide of one of such homologs.

Note that there are peaks in the UP I chromatogram that correspond to the 'A' and 'B' peptides as expected from their sequences. The yield of the 'B' peptide appears somewhat low in the UP I chromatogram but this is not unexpected since an Arg-Asp cleavage is required to produce this peptide. In general, nearby acidic residues significantly decrease the rate of trypsin cleavage. Peptide 'C' (Figure 1) clearly does not correspond to a peak in the UP I chromatogram, therefore it was not surprising to find that its sequence (Ser-Gly-Ser-Gly-X-Phe-Gly-Gly-Gly-X-Gly-Gly-Gly-Phe-Gly-Gly-Asn-Asp-Asn-Phe) could not be matched to UP I. Based on the high glycine content of this peptide we would predict that it arises from the glycine-rich C-terminal domain of A1/2 (see Discussion).

A further experiment established even more firmly a structural relationship between UP I and core protein A1. V8 protease-created fragments of protein A1 were fractionated on one-dimensional gels and electroeluted. One of these peptides (A1/II; 13 kd; see Figure 4B) proved to have an open amino terminus and yielded a sequence of 30 amino acids (Figure 1B): Met-Ala-Ser-Ala-Ser-(Ser/Asn)-Ser-Gln-Arg-Gly-Arg-Ser-Gly-Ser-Gly-Asn-Phe-Gly-Gly-Gly-Arg-Gly-Gly-Gly-Phe-X-X-Asn-Asp-Asn. As can be seen, the first 11 amino acids of this sequence are identical to the last 11 C-terminal residues of UP I (Williams *et al.*, 1985) and the last 19 residues of this peptide overlap the sequence of tryptic peptide 'C' of the A1/2 digest. This result strongly supports our original hypothesis that ssDBP UP I is a proteolytic fragment of a higher M_r precursor that can now be identified as hnRNP protein A1.

Below we describe the isolation and characterization of a cDNA clone encoding the UP I protein. The nucleotide sequence of this cDNA confirms a precursor/product relationship between A1 and UP I.

Isolation and characterization of cDNA clone pRP10

Three immunoreactive clones were isolated from a human liver cDNA library in expression vector pEX1 as described under Materials and methods. For further characterization, the fusion proteins (β-galactosidase/antigen) produced by these clones were individually blotted onto nitrocellulose filters. Following incubation of the filters with total rabbit antiserum and repeated washings, specific IgGs which reacted with the antigens were eluted and tested against purified hnRNP core proteins and ssDBP UP I in a Western blot assay. In the case of one clone (pRP10) it was found that the IgGs so obtained reacted equally well with UP I and with the most abundant polypeptides (A1 and A2) of the 40S hnRNP particles. This clone was subsequently chosen for DNA sequencing. The restriction map of pRP10, together with the sequencing strategy, is shown in Figure 2.

The nucleotide sequence of the cDNA insert and the deduced amino acid sequence of the corresponding protein are shown in Figure 3. The pRP10 insert is 949 bp in length, it contains an open reading frame (ORF) of 591 bp (encoding about 20 kd of protein and ending in a TAA ochre stop codon) followed by a 358 bp stretch of presumed 3′ untranslated region and by a residual poly(A) sequence of 24 adenines. As will be discussed below, the N-terminal part of the ORF-encoded protein is missing probably due to premature termination of the cDNA transcript.

Identification of the ORF-encoded protein in pRP10

As shown above, comparative h.p.l.c. peptide mapping and peptide sequencing clearly indicate a close correlation between hnRNP core protein A1 and the ssDBP of 24 kd (UP I). The analysis of the pRP10 cDNA sequence not only supports this con-

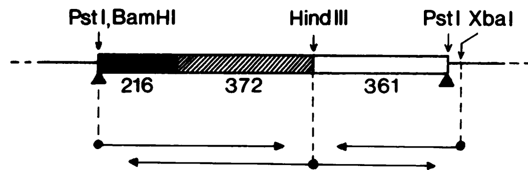

Fig. 2. Restriction map and sequencing strategy for pRP10 DNA. The 949 bp cDNA insert is boxed: the solid part represents the coding region, corresponding to protein UP I, the hatched part corresponds to the carboxy terminal region of core protein A1 and the open part represents the 3′ trailer. Solid triangles mark the flanking polylinker sites of the vector pEX1. The vector sequences are shown as flanking thin lines. Fragments obtained by restriction nuclease digestion were labeled at their respective 5′ ends and used for sequencing as represented by the arrows.

Table I. Amino acid compositions[a] of the hnRNP proteins A1/2 and UP I

Amino acid	A1/2 proteins	A1/2 proteins[b]	UP I[c]	A1[d]	pRP10 cDNA[e]
Asp	10.2	9.0	9.1	10.9	10.9
Thr	3.9	3.1	5.5	3.9	3.8
Ser	9.5	11.7	6.5	9.8	10.6
Glu	12.7	11.7	13.7	9.5	9.4
Pro	2.9	3.8	2.8	2.7	2.8
Gly	19.0	24.0	8.4	19.9	20.0
Ala	5.3	7.0	5.5	4.2	3.4
Val	4.4	2.9	9.0	5.6	5.3
Met	0.7	1.1	1.9	1.5	1.9
Ile	3.5	3.0	4.2	2.7	2.5
Leu	6.5	4.6	4.5	3.4	2.5
Tyr	3.4	2.0	2.1	4.6	3.8
Phe	4.0	3.5	5.3	6.4	6.9
His	1.9	1.6	4.4	2.5	2.5
Lys	5.8	6.9	8.5	5.5	5.6
Arg	6.3	4.3	8.8	6.1	7.2
Dma[f]	1.3	0.3	0.5	1.1	–

[a]Given in mol % of each amino acid.
[b]Average of A1 and A2 values from Beyer *et al.* (1977).
[c]From Merrill *et al.* (1986).
[d]From Wilk *et al.* (1985).
[e]Deduced amino acid sequence (extrapolated to complete protein; see text).
[f]N^G-N^G-dimethylarginine.

clusion but shows that the UP I protein is part of a larger ORF which encodes hnRNP protein A1. In fact (see Figure 3), the first 72 amino acids encoded by pRP10 cDNA are identical to the last 72 amino acids of the known UP I sequence of 195 amino acids (Williams *et al.*, 1985). This indicates that the cDNA clone isolated by us corresponds to the UP I gene starting at nucleotide 372 (amino acid 124). pRP10 cDNA, however, is 949 nucleotides long and, assuming that the first 372 nucleotides are missing, it can code for a larger protein of 34 kd. The identity of this larger protein with hnRNP core protein A1 [M_r (on SDS gels) 31 kd; (on urea gels) 38 kd; Wilk *et al.*, 1985] is inferred on the basis of the following arguments.

(i) If the N-terminal part (123 amino acids) of the large protein is assumed to have the same sequence as that of UP I, the overall amino acid composition is identical to that of hnRNP core protein A1 (see Table I).

(ii) The sequence of several A1 peptides (separated by h.p.l.c.) matches exactly that predicted from the pRP10 DNA sequence (Figures 1 and 3). In particular the peptide A1/II which overlaps the C-terminal region of UP I and continues into A1 has a sequence which is found identically in pRP10 (Figures 2 and 3).

These data, together with our previous observations on hnRNP

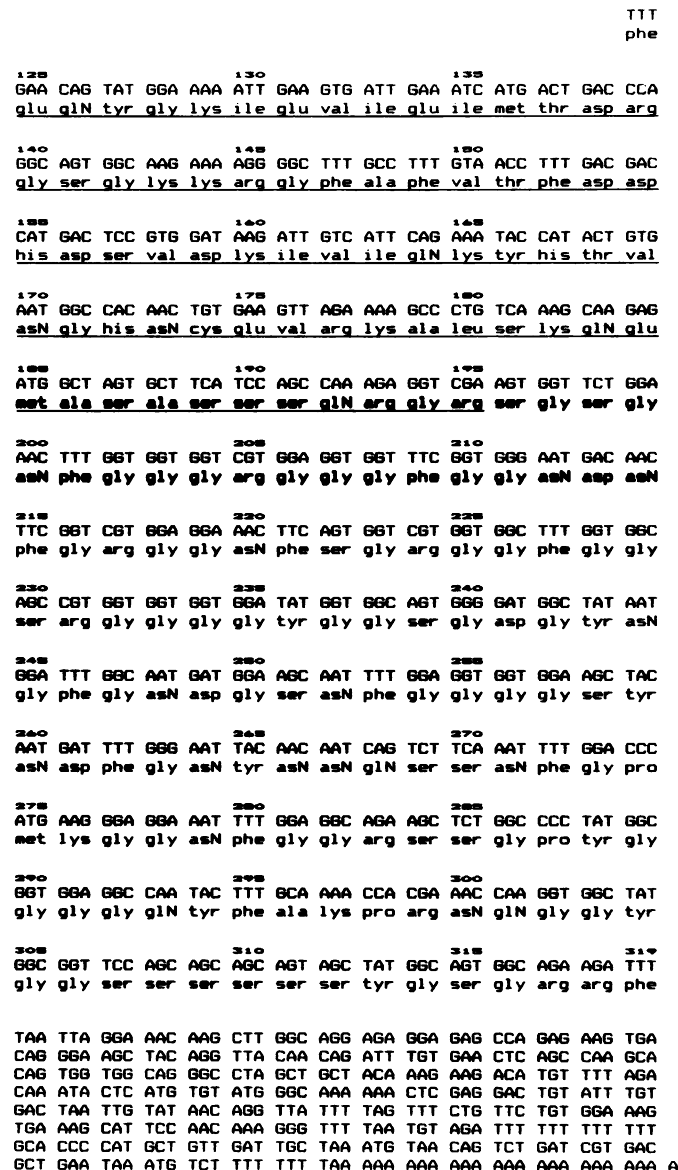

Fig. 3. Nucleotide sequence of pRP10 insert and deduced amino acid sequence of ORF protein. The nucleotide sequence of the insert DNA in clone pRP10 and the corresponding amino acid sequence of the ORF protein are given. Number of amino acids is according to Williams *et al.* (1985). The underlined section in the amino acid sequence represents that part of the sequence which is identical to UP I (amino acids 124–195; the arginine at position 193 is dimethylated). The V8 protease peptide A1/II which transcends the C terminus of UP I and which was sequenced is printed in bold face. A putative polyadenylation signal in the 3′ non-translated region is underlined.

→ ssDBP conversions by proteolysis (Pandolfo *et al.*, 1985), strongly point to a precursor/product relationship between A1 and UP I.

Core protein A1-specific mRNAs

Poly(A)$^+$ RNA from HeLa cells was analyzed in a Northern type assay by means of a ^{32}P-labeled pRP10 probe. Two RNA species of about 1.8 and 1.5 kb specifically hybridize to the pRP10 probe (data not shown). The relative abundance of the two molecular species was determined by comparison of the hybridization signal with that obtained with a human dihydrofolate reductase (DHFR) probe on HeLa poly(A)$^+$ RNA (Morandi *et al.*, 1982). Densitometry of the autoradiograms reveals that the A1-specific

Table II. Stimulation of DNA polymerase α by ssDBP UP I and hnRNP proteins

	Ratio of protein/ DNA	dNTP incorporation pmol/assay	Stimulation index
DNA polymerase α	–	1.2	1
+ calf thymus ssDBP (24–26 kd)	5/1	9.5	7.9
+ HeLa ssDBP (25 kd)	5/1	8.3	6.9
+ HeLa hnRNP proteins A1/2	5/1	1.8	1.5
+ ditto cleaved with trypsin	5/1	6.0	5.0
+ ditto cleaved with endogenous protease	5/1	7.5	6.2

Purification of DNA polymerase α and assay conditions were as previously described (Valentini *et al.*, 1984). Proteins (1.25 μg) were incubated with poly[d(AT)] (0.25 μg) in 50 μl of reaction mixture for 5 min at 0°C. The reaction was started by adding 0.25 units of DNA polymerase α and incubation was continued for 60 min at 37°C. 25 μl samples were taken at 30 and 60 min and the radioactivity in acid-precipitable material was determined.

mRNAs are 600-fold (1.5 kb species) and 300-fold (1.8 kb species) as abundant as the DHFR mRNA (data not shown). Consequently, A1-specific mRNA may constitute up to >1% of the poly(A)$^+$ RNA population in HeLa cells. This finding, which further confirms the specificity of our cDNA clone, is in agreement with the fact that the hnRNP core proteins are the most abundant non-histone nuclear proteins in mammalian cells.

In addition, *in vitro* translation of HeLa poly(A)$^+$ RNA with [^{35}S]methionine and two-dimensional autoradiography of the translation products also shows a spot which co-migrates with authentic core protein A1 (C.C.Kuhn and K.P.Schäfer, unpublished results). The difference in length of the two mRNAs may be due to the use of differential polyadenylation sites on the same hnRNP precursor (Tosi *et al.*, 1981; Setzer *et al.*, 1982) or may be caused by transcription of two independent genes.

Stimulation of DNA polymerase α by ssDBP UP I and hnRNP proteins

Stimulation of DNA polymerase α activity by ssDBP UP I and

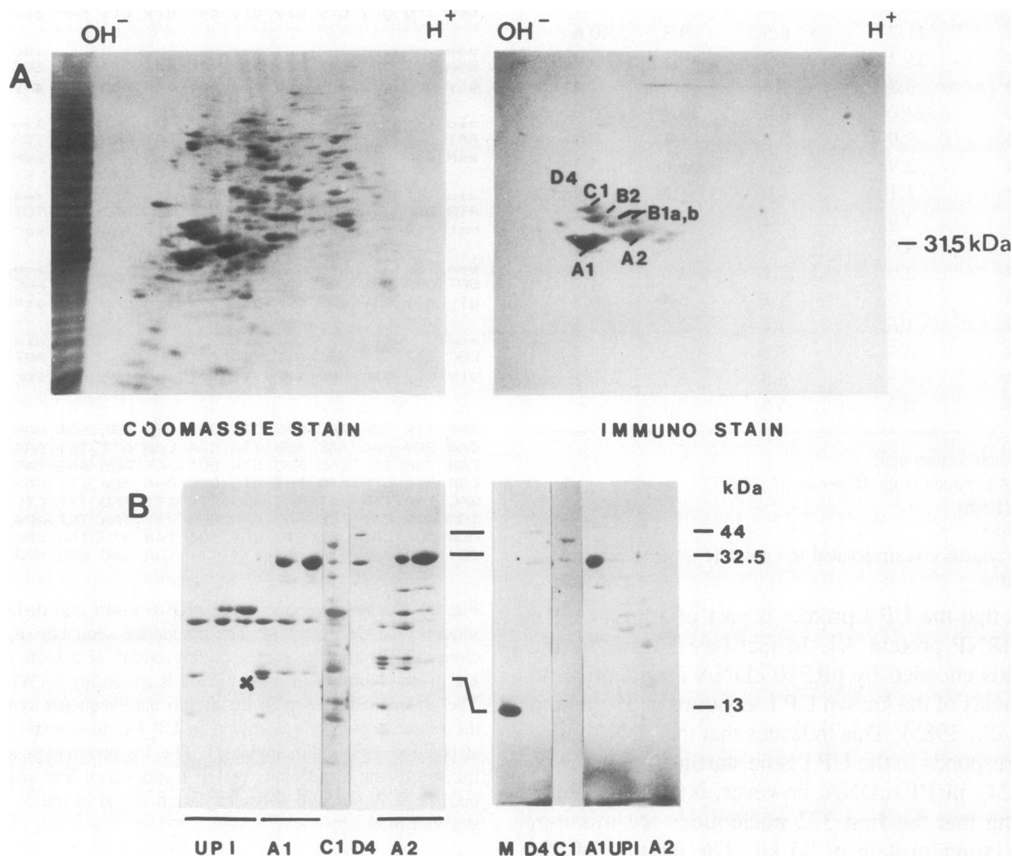

Fig. 4. Western type analysis of core proteins and related V8 protease peptides. **(A)** Blot of two-dimensional gel of hnRNP core proteins/nuclear extract. Nuclear extract from HeLa cells was analyzed by two-dimensional NEPHGE (Wilk *et al.*, 1983). Immunoblotting was conducted as described under Materials and methods. **(B)** Comparative V8 protease mapping of UP I and core proteins A1, C1 and D4. Individual core proteins were excised from two-dimensional NEPHGE gels and subsequently digested with V8 protease while still in the gel. The resulting peptides were separated by one-dimensional SDS–PAGE and stained with Coomassie Blue. The peptide A1/II which was sequenced is marked by a cross (x) (**left panel**). Immunoreactive peptides were detected under Western blotting on different gels as described under Materials and methods (**right panel**). Lysozyme was used as a marker and becomes visible due to unspecific binding of glycoslyated immunoglobulins. Assay conditions were as follows (40 min V8 protease digestion in all experiments). **Left panel** from left to right: **lanes 1–3:** 4 μg of UP I was digested with 150, 30 and 5 ng of V8 protease, respectively; **lanes 4–6:** 8 μg of A1 was treated with 150, 30 and 5 ng of V8 protease; **lane 7:** 8 μg of C1 was treated with 60 ng of of V8 protease; **lane 8:** 3 μg of D4 was treated with 60 ng of V8 protease; **lanes 9–11:** 8 μg of A2 was treated with 150, 30 and 5 ng of V8 protease. **Right panel** from left to right: **lane 1:** 6 μg of lysozyme marker; **lane 2:** 3 μg of D4 was treated with 40 ng of V8 protease; **lane 3:** 5 μg of C1 was treated with 60 ng of V8 protease; **lane 4:** 8 μg of A1 was treated with 80 ng of V8 protease; **lane 5:** 5 μg of undigested, crude UP I was used (a gift of Martin Sapp, Universität Konstanz; the higher M_r bands are due to the presence of other hnRNP proteins); **lane 6:** 8 μg of A2 was treated with 80 ng of V8 protease.

hnRNP proteins was tested on poly[d(AT)] co-polymer as previously described (Valentini *et al.*, 1984). As shown in Table II, while ssDBP UP I from calf thymus and the homologous 25 kd ssDBP from HeLa cells show substantial stimulation of DNA polymerase activity, hnRNP proteins A1/2 tested at the same protein/DNA ratio show no effect. Interestingly, however, if A1/2 are first submitted to limited digestion with trypsin or with a trypsin-like protease described in a previous paper (Pandolfo *et al.*, 1985) stimulation is resumed. We had previously shown (Valentini *et al.*, 1985; Pandolfo *et al.*, 1985) that this proteolytic treatment quantitatively converts hnRNP proteins A1/2 into smaller polypeptides corresponding to ssDBP.

These results indicate that the removal of the glycine-rich C-terminal portion of these proteins probably affects their DNA binding (unwinding?) properties. This point is now being further followed up.

Immunoblotting indicates other possible precursor/product protein pairs

When a two-dimensional gel was probed with anti-ssDBP antibody most of the basic core proteins (A1, A2, Bla/b, B2 and C1) gave a positive signal (Figure 4A). In addition, D4 also reacted positively. When the intensity of the Western blot staining is used as gauge of relatedness among this group of proteins, A1, B2 and C1 form a group of strong reactants whereas A2, Bla and Blb react less intensely. To probe this relatedness further, we used V8 protease digestion of proteins A1, A2, C1 and D4 followed by Western type analysis of V8 peptides (Figure 4B).

Although we had described D4 as a constituent of oligomeric hnRNP complexes (Wilk *et al.*, 1983) the data in Figure 4 show that D4 may be more closely related to the basic core proteins than the difference indices [calculated on the basis of amino acid compositions (Wilk *et al.*, 1985)] originally allowed us to assume. The V8 peptide pattern of UP I and A1 demonstrates the matching peptides (Figure 4). The peptide A1/II which was sequenced is indicated. Proteins D4, C1 and A2 share some bands at identical positions. For C1 and A2 the difference indices are remarkably low (C1/A2 = 3.7) indicating a close relationship, whereas the M_r of D4 differs too much from that of both the former proteins to yield meaningful information via the difference indices. We have consistently observed that the C1 protein spot on two-dimensional gels has an identical substructure to that of A1, exhibiting two series of satellite protein spots analogous to the conformeric forms A1 and A1x (Wilk *et al.*, 1985). All these data, though indirect, suggest that there may be additional precursor/product pairs among the core proteins.

Discussion

The data presented here together with our previous results (Valentini *et al.*, 1985; Pandolfo *et al.*, 1985) demonstrate that the most widely studied single-stranded DNA binding protein from mammalian cells (typified by calf thymus UP I, 24 kd) is a proteolytic fragment corresponding to the first 195 amino acid residues of protein A1 of the hnRNP complex.

Calf thymus UP I is heterogeneous in size; while the most abundant polypeptide has a M_r of 24 kd; other less represented species range in size from 22.5 to 25.5 kd (Herrick and Alberts, 1976). Furthermore, with a slightly different purification technique we obtained polypeptides immunologically and structurally related to UP I, ranging in size from 26 to 28 kd (Riva *et al.*, 1980; Valentini *et al.*, 1984). It is likely that all those species represent limited digestion products of protein A1. However, we cannot rule out the possibility that proteolysis of A2, closely related to A1 (Beyer *et al.*, 1977; Wilk *et al.*, 1985), may give rise to some ssDBP species. More detailed studies are required to clarify this point. Furthermore, some preliminary data (Figure 4) indicate that some type of specific proteolysis may be a general attribute of hnRNP core protein turnover.

This work raises two different but related questions on the physiological significance of the described proteolysis and on the *in vivo* existence of the UP I-related single-stranded DNA binding proteins. The first problem has already been addressed in a previous paper (Pandolfo *et al.*, 1985) where a trypsin-like protease that converts hnRNP proteins into ssDBP was discovered in nuclear extracts and shown to co-purify with the bulk of the hnRNP proteins. The results reported here seem to add more significance to the initial observations since the proteolytic cleavage that generates UP I from A1 occurs at the border between two structurally different domains: the α helix-rich UP I domain (Williams *et al.*, 1985) and the random coil glycine-rich C-terminal domain. These different domains may play different roles in protein/nucleic acid and protein/protein interactions. Multiple domains binding different ligands could be a general property of eukaryotic single-stranded nucleic acid binding proteins as reported for other types of proteins (Chase and Williams, 1986).

An intriguing, although highly speculative, idea is that limited *in vivo* proteolysis may be a method of changing the function of a protein that normally binds RNA into one that participates in DNA metabolism. Alternatively, it is also possible that UP I may function in hnRNA processing either by being a simple degradation product of A1 (as a consequence of controlling the level of A1) or by playing some unknown role as an RNA binding protein which has lost the ability to interact with other core proteins.

On the other hand, as reported in previous papers (Valentini *et al.*, 1985; Pandolfo *et al.*, 1985), in crude cellular extracts very little, if any, UP I can be detected with polyclonal antibodies against UP I. No increase in UP I was observed with cell age or state of differentiation (Riva *et al.*, unpublished observations). It therefore seems clear that most of the UP I purified from cells or tissues results from *in vitro* proteolysis of the A1 protein. At the moment, other possibilities such as the existence of an independent UP I-specific gene or creation of UP I by differential splicing of a common hnRNP precursor (together with A1) cannot be ruled out with certainty.

A piece of evidence that argues against UP I being simply an *in vitro* artifact is that a specific proteolytic cleavage endows the protein with a unique functional property, namely the ability to stimulate DNA polymerase α (Table II).

It is worth mentioning that routine computer analysis of the pRP10 sequence has shown a partial homology of the C-terminal region of A1 with an N-terminal region (also glycine-rich) of human keratin type II (Marchuk *et al.*, 1984), the constituent of intermediate filaments (Steinert *et al.*, 1985). The interpretation of this finding is not straightforward, although it is tempting to propose that the glycines in both types of proteins could play a role in protein/protein interactions. In this regard it is known that under suitable conditions, hnRNP proteins can form filaments (Lothstein *et al.*, 1985). Furthermore, nascent RNP complexes show some filamentous features (Skoglund *et al.*, 1983). It is also worth mentioning that hnRNP particles are found associated with the nuclear matrix (van Eekelen and van Venrooij, 1981) and may even be one of its constituents. Finally, recent studies, independently carried out, on the structure of a rodent helix-destabilizing protein (HDP: 27 kd, homologous to UP I) as revealed by cDNA cloning gave results very similar to some of those reported here (Cobianchi *et al.*, 1986).

Materials and methods

Purification of ssDBP

The 24 kd ssDBP from calf thymus (UP I) was prepared by the procedure of Herrick and Alberts (1976). A 25 kd ssDBP from HeLa cells (highly homologous to calf thymus UP I) was purified to near homogeneity by a modification of the procedure of Herrick and Alberts (Riva *et al.*, 1980; Valentini *et al.*, 1984). Polyclonal antibodies against bovine ssDBP were raised in rabbit and purified as described (Valentini *et al.*, 1985).

Purification of hnRNP core proteins from HeLa cells, isolation of V8 protease peptides and amino acid sequencing

HnRNP core proteins were purified from HeLa cell nuclei by one of the following methods: (i) hnRNP core protein doublet A1/2 was quantitatively purified from HeLa cells by a chromatographic method described elsewhere (Pandolfo *et al.*, 1985); (ii) hnRNP core proteins were also purified from HeLa cell nuclei by the pH 8.0/diffusion technique (Samarina *et al.*, 1968; Wilk *et al.*, 1983). Individual proteins were electroeluted from two-dimensional polyacrylamide gels as described by Wilk *et al.* (1985). For V8 peptide isolation, stained protein spots excised from two-dimensional gels were squeezed on top of a second one-dimensional gel and incubated in the gel matrix with V8 protease (Cleveland *et al.*, 1977). After electrophoresis the peptide bands were stained, excised again and electroeluted (Wilk *et al.*, 1985).

The lyophilized eluate was re-dissolved in formic acid and 30 μl aliquots containing 200−600 pmol of peptide were dried onto the preconditioned glass filter disc (containing 1.5 mg polybrene and 3 μmol NaCl) of the protein microsequencer (Hewick *et al.*, 1981). To assess the immobilization of the sample on the glass filter in the presence of SDS the filter was treated with 30 μl of 25% aqueous trifluoroacetic acid (TFA) (v/v) after sample loading and immediately dried again. The latter treatment restores an even polybrene film from crystalline polybrene due to the presence of SDS in the sample. Automated Edman degradation was done using a protein microsequencer model 470A (Applied Biosystems, Foster City, CA) equipped with on-line detection for phenylhydantoin amino acids (PTH-Analyser, Model 120A, Applied Biosystems). Separation of the phenylhydantoin amino acids was performed on cyanopropyl columns (Zorbax CN, 0.47 × 25 cm) as described previously (Beyreuther *et al.*, 1983) or on C-18 columns (Brownlee C-18, 0.21 × 22 cm) as specified by the manufacturer (Applied Biosystems).

H.p.l.c. peptide mapping and amino acid sequencing

For analytical peptide mapping, aliquots of UP I (1 nmol) or of the A1/2 mixture (460 pmol) that had been precipitated with trichloroacetic acid were re-dissolved in 50 μl of 8 M urea. After dilution to 2 M urea with 25 mM NH_4HCO_3, proteins were digested with a 1/25 (enzyme/protein, w/w) ratio of trypsin (Cooper Medical) for 24 h at 37°C. Peptides for amino acid sequencing were obtained from a similar digest done on 4.8 nmol of the A1/2 proteins. In both cases the digests were injected directly onto a reverse-phase column (Vydac C-18, 5 μm) equilibrated in 0.005% TFA at a flow rate of 0.7 ml/min.

Peptides were eluted with increasing concentrations of solvent B [0.05% TFA, 80% (v/v) acetonitrile] as follows: 0−90 min (0−37.5% B), 90−135 min (37.5−75% B) and 135−150 min (75−100% B). The Waters Associates h.p.l.c. system used for these studies consisted of a model 721 system controller, a WISP automated sample injector, two model 510 pumps, and an ISCO model V^4 absorbance detector that was connected to a Nelson analytical model 4416 Chromatography Data Station. The absorbance profile for a buffer blank was subtracted from the analytical h.p.l.c. runs which were then replotted using the Nelson analytical chromatography software. Aliquots of three A1/2 peptides (labeled A−C in Figure 1) were dried *in vacuo*, re-dissolved in 0.1 ml 100% TFA and loaded onto the protein microsequencer (Model 430, Applied Biosystems) as described by Williams *et al.* (1985).

Isolation of cDNA clones

A human liver cDNA library in plasmid expression vector pEX1 was screened by the immunological procedure of Stanley and Luzio (1984) by utilizing the rabbit anti-bovine ssDBP antiserum described above. Three positive clones producing β-galactosidase−antigen fusion proteins were selected for further analysis.

Affinity purification of antibodies

Nitrocellulose filters were saturated with the fusion proteins (β-galactosidase−antigen) produced by the three clones and used to affinity purify antibodies from the whole rabbit serum by a modification of the procedure of Olmsted (1981). Briefly, filters were incubated overnight with a 1:500 dilution of the serum and washed extensively with Tris-buffered saline solution (pH 7.4). Bound antibodies were then eluted by incubating filters in 0.2 M glycine buffer (pH 2.8) for 5 min on ice, immediately neutralized with 2 M NaOH, concentrated and tested against purified hnRNP core proteins in a Western blot assay.

DNA sequencing

Sequencing was performed by the chemical modification method of Maxam and Gilbert (1980) except that diphenylamine formate was used instead of piperidine formate in the (A+G) reaction.

Preparation and hybridization of nucleic acids

Total RNA from HeLa cells and from calf thymus was prepared by the guanidinium thiocyanate method (Chirgwin *et al.*, 1979). Poly(A)$^+$ RNA was isolated by passage through an oligo(dT) cellulose column (Aviv and Leder, 1972) and utilized for Northern blot analysis (Thomas, 1983) after fractionation on agarose/formaldehyde gels (Rave *et al.*, 1979). Hybridization was performed under standard high stringency conditions (0.2 × SSC), 68°C.

Immunological characterization of hnRNP core proteins

Proteins or peptides from one- or two-dimensional gels were transferred to nitrocellulose and specifically visualized by incubation with rabbit anti-bovine ssDBP antiserum (1:1000 dilution) followed by a peroxidase-coupled goat anti-rabbit antibody (Yeda) (Towbin *et al.*, 1979). 3,3′-Diaminobenzidine was used as peroxidase substrate.

Stimulation of DNA polymerase

HeLa cell DNA polymerase α purification and stimulation assay on poly[d(AT)] template were done as described (Valentini *et al.*, 1984). Conditions for limited proteolytic digestion of hnRNP proteins were as reported (Valentini *et al.*, 1985; Pandolfo *et al.*, 1985).

Acknowledgements

The Pavia authors thank G.Astaldi-Ricotti and A.Falaschi for helpful discussions, O.Valentini, M.T.Chiesa and F.Nuzzo for collaboration. S.R. and M.P. also wish to acknowledge the support of NATO Collaborative Research Grant No. R.G. 85/0511 with R.Karpel (University of Maryland). We are grateful to K.Stanley and R.Cortese (EMBL, Heidelberg) for kindly providing the cDNA library. We thank F.Cobianchi for making available unpublished results. K.R.W. was supported by N.I.H. grant G.M. 31539. K.B. received support from the Bundesminister für Wissenschaft und Technologie and the Deutsche Forschungsgemeinschaft (SFB 74). K.P.S. was supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie. K.P.S. also thanks Prof. Olaf Pongs (Bochum) for financial support. This work was supported by 'Progetto Finalizzato Ingegneria Genetica e Basi Molecolari delle Malattie Ereditarie' and by 'Progetto Finalizzato Oncologia' of CNR, Rome. M.P. was supported by a grant of the Istituto Neurologico C.Besta, Milan.

References

Aviv,H. and Leder,P. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 1408−1412.

Beyer,A., Christensen,M.E., Walker,B.W. and LeStourgeon,W.M. (1977) *Cell*, **11**, 127−138.

Beyreuther,K., Bieseler,B., Bovens,J., Dildrop,R., Neifer,K., Stüber,K., Zaiss,S., Ehring,R. and Zabel,P. (1983) In Tschesche,H. (ed.), *Modern Methods in Protein Chemistry.* Walter de Gruyter, Berlin/NY, pp. 303−355.

Chase,J.W. and Williams,K.R. (1986) *Annu. Rev. Biochem.*, in press.

Chirgwin,J.M., Przybyla,A.E., MacDonald,R.J. and Rutter,W.J. (1979) *Biochemistry*, **24**, 5294−5299.

Cleveland,D.W., Fischer,S.G., Kirschner,M.W. and Laemmli,U.K. (1977) *J. Biol. Chem.*, **252**, 1102−1106.

Cobianchi,F., SenGupta,D.N., Zmudzka,B.Z. and Wilson,S.H. (1986) *J. Biol. Chem.*, **261**, 3536−3543.

Henner,D. and Furth,J.J. (1977) *J. Biol. Chem.*, **252**, 1932−1937.

Herrick,G. and Alberts,B. (1976) *J. Biol. Chem.*, **251**, 2124−2132.

Hewick,R.M., Hunkapiller,M.W., Hood,L.E. and Dryer,W.J. (1981) *J. Biol. Chem.*, **256**, 7990−7997.

LeStourgeon,W.M., Beyer,A.L., Christensen,M.E., Walker,B.W., Poupore,S.M. and Daniels,L.P. (1977) *Cold Spring Harbor Symp. Quant. Biol.*, **42**, 885−898.

Lothstein,L., Arenstorf,H.P., Chung,S.-Y., Walker,B.W., Wooley,J.C. and LeStourgeon,W.M. (1985) *J. Cell Biol.*, **100**, 1570−1581.

Marchuk,D., McCrohon,S. and Fuchs,E. (1984) *Cell*, **39**, 491−498.

Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499−560.

Merrill,B.M., LoPresti,M.B., Stone,K.L. and Williams,K.R. (1986) *J. Biol. Chem.*, **261**, 878−883.

Morandi,C., Masters,J.N., Mottes,M. and Attardi,G. (1982) *J. Mol. Biol.*, **156**, 583−607.

Olmsted,B.J. (1981) *J. Biol. Chem.*, **256**, 11955−11957.

Pandolfo,M., Valentini,O., Biamonti,G., Morandi,C. and Riva,S. (1985) *Nucleic Acids Res.*, **13**, 6577−6590.

Pederson,T. (1976) *J. Cell Biol.*, **70**, 308a.

Pederson,T. (1983) *J. Cell Biol.*, **97**, 1321−1326.

Planck,S.R. and Wilson,S.H. (1980) *J. Biol. Chem.*, **255**, 11547−11556.

Rave,N., Crkvenjakov,R. and Boedtker,H. (1979) *Nucleic Acids Res.*, **11**, 3559−3567.

Riva,S., Clivio,A., Valentini,O. and Cobianchi,F. (1980) *Biochem. Biophys. Res. Commun.*, **96**, 1053−1062.

Samarina,O.P., Lukanidin,E.M., Molnar,J. and Geogiev,G.P. (1968) *J. Mol. Biol.*, **33**, 251−263.

Setzer,D.R., McCrogan,M. and Schimke,R.T. (1982) *J. Biol. Chem.*, **257**, 5143–5147.
Sjoglund,V., Andersson,K., Bjorkroth,B., Lamb,M.M. and Daneholt,B. (1983) *Cell*, **34**, 847–855.
Stanley,K.K. and Luzio,J.P. (1984) *EMBO J.*, **3**, 1424–1434.
Steinert,P.M., Steven,A.C. and Roop,D.R. (1985) *Cell*, **42**, 411–419.
Thomas,P. (1983) *Methods Enzymol.*, **100**, 255–266.
Tosi,M., Young,R.A., Hagenbuechle,O. and Schibler,U. (1981) *Nucleic Acids Res.*, **9**, 2313–2323.
Towbin,H., Staehelin,T. and Gordon,J. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 4350–4354.
Valentini,O., Biamonti,G., Mastromei,G. and Riva,S. (1984) *Biochim. Biophys. Acta*, **782**, 147–153.
Valentini,O., Biamonti,G., Pandolfo,M., Morandi,C. and Riva,S. (1985) *Nucleic Acids Res.*, **13**, 337–346.
van Eekelen,C.A.G. and van Venrooij,W.J. (1981) *J. Cell Biol.*, **88**, 554–563.
Wilk,H.-E., Angeli,G. and Schäfer,K.P. (1983) *Biochemistry*, **22**, 4592–4600.
Wilk,H.-E., Werr,H., Friedrich,D., Kiltz,H.-H. and Schäfer,K.P. (1985) *Eur. J. Biochem.*, **146**, 71–81.
Williams,K.R., Stone,K., LoPresti,B.M., Merrill,M.B. and Planck,S.R.L (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 5666–5670.

Received on 28 April 1986; revised on 6 June 1986