DISCUSSION
DISCUSSION

Regulation of cellular functions by extracellular matrix (ECM), a dynamic assembly of a variety of interacting biomolecules, is a fundamental mechanism influencing cellular behaviour and phenotypic expression. Interaction of the individual components of ECM, either with specific cell surface molecules, with integrin receptors or with proteoglycans; initiates a cascade of signal transduction pathways leading to numerous transient or persistent cellular responses. An intimate involvement of ECM in all aspects of these cellular responses have established the role of ECM as a key programmer right up to gene expression, far from the longstanding notion of it being an inert scaffold around the cell and an inactive bystander.

Compositional analysis of ECM isolated from different sources has revealed enormous structural and functional heterogeneity. Notwithstanding these differences, most of the ECMs contain a backbone of fibrillar proteins e.g. collagens or elastin into which a variety of glycoproteins, carbohydrates, and proteoglycans are enmeshed. The few best studied adhesive glycoprotein component of ECM are fibronectin, laminin, thrombospondin, tenascin and entactin. Proteoglycans represent the most abundant, heterogeneous and perhaps functionally the most versatile nonfibrillar component of the ECM. These complex macromolecules are made up of a core polypeptide to which linear heteropolysaccharides called glycosaminoglycans (GAGs) are covalently attached. GAG chains are polymers of repeating disaccharide unit of D-glucosamine/galactosamine and D-glucuronic acid. Except hyaluronic acid (HA), all the other GAGs found in the ECM, namely, heparan sulfate, chondroitin sulfate, dermatan sulfate and keratan sulfate are sulfated.

Hyaluronan or hyaluronic acid (HA), the nonsulfated GAG in ECM, is ubiquitously distributed in the body and facilitates diversified cellular functions like cell migration, proliferation during embryonic development (Toole, 1981) and tumor invasion, (Knudson et al., 1989). One of the important functional properties of HA is its ability to bind specifically to HA-binding proteins present both, in the ECM and on the cell surface (Toole, 1990; Underhill, 1989; Turley, 1989) in order to perform its biological functions.
This binding is predominantly non-covalent in nature, except a recent report showing the covalent binding of HA with C1q, the initiator protein in complement mediated lysis pathway (Prehm, 1995). A number of HA-binding proteins have been identified from various normal and transformed cells and their distinctness has been confirmed by molecular characterization. The present study describes molecular characterization of one such 68 kDa HA-binding protein (34 kDa subunit) and elucidates the possible involvement of this protein in cellular signalling processes. We have been able to clone a partial cDNA encoding this protein which revealed many informations regarding the multifunctional activities of the protein including the cellular signal transduction. These informations have been partly confirmed by in vivo and in vitro findings.

Our initial investigation focussed on the purification of the HA-binding protein. The protein was purified from rat kidney tissue by HA-Sepharose-4B affinity chromatography, a conventional way for the purification of most of the HA-binding proteins. Under SDS-PAGE, the purified protein moves with equal electrophoretic mobility both under reducing and non-reducing conditions and shows up as 34 kDa band, thus implying the absence of intrachain disulfide bond in the primary sequence of the protein. This observation was supported by the fact that there is only one cysteine residue present in the primary sequence of the mature protein (described in detail in the latter part of the discussion). The formation of immunoglobulin (Ig) fold which is normally a common phenomenon in matrix HA-binding proteins like cartilage link protein (Goetinck et al., 1987), aggrecan (Doege et al., 1991), versican (Zimmermann and Ruoslahti, 1989), glial HA-binding protein (Perdies et al., 1989) and cell surface HA-binding protein CD44 (Stamenkovic et al., 1989),is actually absent in this protein. Infact , 34 kDa HA-binding protein resembles RHAMM (Hoare et al., 1993),a protein which also does not possess any intrachain disulfide bond. The native molecular mass of the purified HA-binding protein as determined by gradient native PAGE was found to be 68 kDa implying that the native HA-binding protein is a homodimer of 34 kDa subunits. In all subsequent discussions, the protein has been referred to by its subunit molecular mass i.e. 34 kDa.
DISCUSSION

HA-binding protein.

As the molecular nature of 34 kDa HA-binding protein was unknown, our next objective was to isolate a cDNA encoding this protein by screening a λgt11 human skin fibroblast cDNA expression library using affinity purified specific anti-HA-binding protein antibodies. For this purpose, the polyclonal antibodies against this protein were developed which were later used both for screening cDNA library and as a probe for the subsequent biochemical studies.

Using these polyclonal antibodies, we were able to detect the 34 kDa HA-binding protein in the cell lysate of J774, a transformed macrophage cell line. In addition, another immunoreactive band at 68 kDa was detected not only in the same macrophage cell lysate but also in the medium of the growing J774 cell, thereby suggesting its secretory nature. This is not a surprising observation, since several isoforms of the cell surface HA-binding proteins like CD44 and RHAMM have been identified by other laboratories. On the basis of unique sequences encoded in the genomic RHAMM, RHAMM isoforms (secreted extracellular isoforms sRHAMM, plasma membrane isoforms pRHAMM) appear to be generated by alternate splicing. Besides these, the differential glycosylation and phosphorylation also contribute to the heterogeneity in RHAMM, leading to its various isoforms (Pilarski et al., 1994). The same is true for CD44, the well-studied cell surface and transmembrane HA-binding protein which is also reported to have many splice variants and glycosylated isoforms (reviewed by Herrlich et al., 1993; Lesley et al., 1993). It is important to mention here that one isoform of CD44 causes the poorly metastatic tumor cells to become highly metastatic (Gunthert et al., 1991) and an overexpression of RHAMM isoforms transforms normal fibroblast which exhibit low rates of spontaneous transformation to highly malignant tumor cells (Hall et al., 1995). As regards the 34 kDa protein, it is probably too premature to say from our present data whether immunoreactive 68 kDa protein is a glycosylation product of the same 34 kDa HA-binding protein backbone or is a result of alternate splicing of the mRNA of genomic fragment coding 34 kDa HA-binding protein. However, it is
important to mention that the N-terminal sequence of both immunoreactive proteins of 34 kDa and 68 kDa are identical. The precise function performed by this 68 kDa protein in tumorigenesis is also a subject of our future study.

As we opted for immunoscreening of the cDNA encoding 34 kDa HA-binding protein from cDNA expression library, the authenticity of the immunoscreened clone was of prime importance to us. The only way to check the authenticity of the clone was to match the few internal peptide sequences of the purified HA-binding protein with the deduced amino acid sequences of the cDNA encoding 34 kDa HA-binding protein. For this purpose, 14 peptides were separated from the purified HA-binding protein after partial protease digestion and microsequenced. Subsequently, these sequences were utilized to confirm the clone obtained.

To progress further, a commercial human skin fibroblast λgt11 cDNA expression library was immunoscreened by affinity purified rabbit anti-HA-binding protein antibodies (IgG) and two positive signals were picked up after three rounds of screening. One of the immunoscreened clone was used to absorb the plating bacteria LE 392 and plate lysates were prepared. λgt11 DNA was then purified from the plate lysates. As the commercial λgt11 cDNA expression library made use of the EcoRI site as the cDNA cloning site, the cDNA could well be amplified using the commercial λgt11 forward and reverse primers specific for the EcoRI site. Using the same strategy, the cDNA insert was amplified by polymerase chain reaction (PCR). The 2.0 kb cDNA insert was subcloned as two fragments of 1.0 kb and 0.9 kb separately in plasmid pBS KS(+). The 0.9 kb fragment of the cDNA was completely sequenced. The internal polypeptide sequence (83 residues) of the purified HA-binding protein was found to be identical to the predicted protein sequence derived from HA-binding protein partial cDNA suggesting the authenticity of the clone.

Interestingly, the partial cDNA sequence of this HA-binding protein showed complete homology with the cDNA sequence of a protein P-32, co-purified with the human pre-mRNA splicing factor SF2 (Krainer et al., 1991). Furthermore, the data on
DISCUSSION

the N-terminal sequence of HA-binding protein and the predicted polypeptide of P-32 revealed the identical coding sequence of 209 amino acids for both the proteins, which was later on shown to be the mature protein, as discussed later. Though P-32 protein was co-purified with splicing factor SF2, its functional characteristics were unknown. Thus, our result on identity of HA-binding protein with P-32 led us to carry out the experiments to determine the HA-binding activity of P-32 protein. P-32 cDNA encoding the mature protein already cloned in an expression plasmid (pT7A.A-32) by Krainer et al (1991) was expressed in E. coli BL21(DE3) by IPTG induction and the recombinant P-32 protein was purified by HA-Sepharose-4B affinity chromatography, used for purification of HA-binding protein from tissues. Both the recombinant P-32 and HA-binding protein migrate with same electrophoretic mobility corresponding to 34 kDa, although the theoretical molecular mass of 209 amino acids is only 24 kDa. This ambiguity can be explained by the highly acidic nature of this protein, since an anomalous movement and overestimation of molecular mass of highly acidic/charged proteins in SDS-PAGE is known (Kaufmann et al., 1984; Wada et al., 1991). The recombinant P-32 protein showed immunocrossreactivity with the polyclonal antibodies raised against 34 kDa HA-binding protein. The anti-P-32 polyclonal antibodies which was developed against the recombinant P-32 also immunodetected purified HA-binding protein thus once again confirming the antigenic similarity of both the proteins.

Recent studies have demonstrated that binding of GAGs such as HA and chondroitin sulfate to proteins, depends upon the interaction between negatively charged carboxyl and/or sulfate groups in the GAG and positively charged cluster of basic amino acids in the protein (Jackson et al., 1991). Using site directed mutagenesis, Yang et al (1994) have further defined the critical and minimally required amino acid structure of the HA-binding motif to be B(X7)B, where B is any basic amino acid other than histidine and X is an non-acidic amino acid. The 34 kDa HA-binding protein also possesses one such site 119KLVRKVAGEK128, which contains one extra glutamic acid (127E) residue. The one step purification of recombinant P-32 protein from the bacterial lysate and the
presence of only this site throughout the composite sequence obtained together from N-terminal sequence and partial cDNA sequence (total 209 amino acids) made us to believe that this is the only and correct HA-binding motif in conformity with the other reported HA-binding motifs. Peptide mimicry experiments (Yang et al., 1994) suggest that internal residues can vary by one amino acid without having a detrimental effect on the ability of the protein to bind HA. Presence of glutamic acid is also reported in the HA-binding motif of TSG-6, a known HA-binding protein (Lee et al., 1992). Experimentally, the binding of recombinant P-32 was observed in a concentration dependent manner with $^{125}$I-tyramine-HA suggesting the HA-binding activity of recombinant P-32. This observation was further strengthened showing the specific affinity of recombinant P-32 to HA by its binding to biotinylated HA in a concentration dependent manner which can be competed only with excess unlabeled HA but not by RNA or DNA, ruling out its non-specific anionic interactions. It may be mentioned here that though P-32 is co-purified with pre-mRNA splicing factor, Krainer et al have already demonstrated that P-32 does not bind to RNA (Krainer et al., 1991). Furthermore, general RNA-binding motifs RNP1 and RNP2 are also absent in P-32 though present in the associated protein SF2. The observation that P-32 binds to HA in both reducing and non-reducing conditions suggests the lack of conserved cysteine residues in this protein in contrast to CD44 and cartilage link protein (Goetinck et al., 1987; Wolffe et al., 1990; Goldstein et al., 1989). Rather P-32 behaves like RHAMM, which also lacks conserved cysteine residues and binds HA equally well under reducing conditions (Hoare et al., 1993).

So far we have described here the cloning and sequencing of the partial cDNA clone for 34 kDa HA-binding protein and confirmed this protein as P-32, a protein that co-purified with pre-mRNA splicing factor SF2 from human HeLa cell. The conclusion is based on the following facts; firstly, both the proteins, namely, P-32 and HA-binding protein, are completely identical at cDNA levels; secondly, the predicted amino acid sequence from cDNA sequence of P-32 protein (Krainer et al., 1991) is again almost similar with partial peptide sequences of purified HA-binding protein; and finally, the
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>N-linked glycosylation sites</td>
</tr>
<tr>
<td>2.</td>
<td>HA-binding motif</td>
</tr>
<tr>
<td>3.</td>
<td>Tyrosine sulfation site</td>
</tr>
<tr>
<td>4.</td>
<td>Protein kinase C phosphorylation site</td>
</tr>
<tr>
<td>5.</td>
<td>Casein kinase II phosphorylation sites</td>
</tr>
<tr>
<td>6.</td>
<td>Extracellular signal regulated kinase (ERK)/cdc2 kinase phosphorylation site</td>
</tr>
<tr>
<td>7.</td>
<td>Integrin binding site</td>
</tr>
<tr>
<td>8.</td>
<td>Nuclear localization sequence (NLS)</td>
</tr>
</tbody>
</table>

Table 2: Prosite analysis of the deduced amino acid sequence of mature 34 kDa HA-binding protein. Asterisk denotes the potential residue of post-translational modification.
## Table 3: Potential HA-binding motifs \([B(X)\_B]\) of HA-binding proteins (reproduced from Yang et al. EMBO J. 13, 286-296, 1994)

<table>
<thead>
<tr>
<th>Proteins</th>
<th>Binding region</th>
<th>Amino acid sequence</th>
<th>Binding confirmed</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>RHAMM</td>
<td>401-411</td>
<td>401-EKIKKHKVHKLK</td>
<td>+</td>
<td>Yang et al., 1993</td>
</tr>
<tr>
<td>RHAMM</td>
<td>423-432</td>
<td>423-KLRSQLVRKK</td>
<td>+</td>
<td>Yang et al., 1994</td>
</tr>
<tr>
<td>Link protein</td>
<td>316-325</td>
<td>316-KYPISRPPFR</td>
<td>+</td>
<td>Goetinck et al., 1986;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Yang et al., 1994</td>
</tr>
<tr>
<td>CD44</td>
<td>38-46</td>
<td>38-ENGRYSISR</td>
<td>+</td>
<td>Yang et al., 1994</td>
</tr>
<tr>
<td>CD44</td>
<td>150-162</td>
<td>150-RDGTRYQKGEYR</td>
<td>+</td>
<td>Yang et al., 1994</td>
</tr>
<tr>
<td>CD44</td>
<td>292-300</td>
<td>292-RRRCGQKKK</td>
<td>+</td>
<td>Yang et al., 1994</td>
</tr>
<tr>
<td>HAse</td>
<td>96-104</td>
<td>96-RGTRSGSTR</td>
<td>+</td>
<td>Hope et al., 1993</td>
</tr>
<tr>
<td>HAse</td>
<td>106-117</td>
<td>106-RRKMQGRSRK</td>
<td>+</td>
<td>Hope et al., 1993</td>
</tr>
<tr>
<td>Link protein</td>
<td>103-112</td>
<td>103-RKSYKFQR</td>
<td>+</td>
<td>Goetinck et al., 1986</td>
</tr>
<tr>
<td>Versican</td>
<td>2319-2327</td>
<td>2319-KTFOGRMKRPR</td>
<td>+</td>
<td>Zimmermann and Ruoslahti, 1989</td>
</tr>
<tr>
<td>Aggrecan</td>
<td>71-79</td>
<td>71-KIKMSRSVK</td>
<td>+</td>
<td>Doege et al., 1987</td>
</tr>
<tr>
<td>Aggrecan</td>
<td>2109-2117</td>
<td>2109-KRTMRPTKR</td>
<td>+</td>
<td>Doege et al., 1987</td>
</tr>
<tr>
<td>Human GHAP</td>
<td>54</td>
<td>54-KVGKSPPPVR</td>
<td>+</td>
<td>Perides et al., 1989</td>
</tr>
<tr>
<td>TSG-6</td>
<td>39-48</td>
<td>39-HREARGKDYK</td>
<td>+</td>
<td>Lee et al., 1992</td>
</tr>
<tr>
<td>34 kDa HA-binding protein</td>
<td>119-128</td>
<td>119-KLRKVKAGEK</td>
<td>+</td>
<td>Deb and Datta, 1996</td>
</tr>
</tbody>
</table>
DISCUSSION

identity of P-32 protein as specific HA-binding protein is confirmed by the presence of known HA-binding motif, purification of recombinant P-32 protein using HA-Sepharose-4B affinity chromatography, their identical electrophoretic mobility, immunocrossreactivity of P-32 protein with anti-HA-binding protein antibodies or vice-versa and its specific affinity toward HA.

Functional characteristics of P-32 were not known earlier. Krainer et al (1991) reported the sequence of P-32 cDNA which does not contain any conventional ATG (Met) start codon but initiates with a CTG (Leu) codon. From the matching of predicted polypeptide sequence with the N-terminus of HeLa cell purified P-32 and the absence of upstream ATG codon or of consensus 3' splice sites, perhaps had led them to conclude that the mature N-terminus does not arise by proteolytic cleavage of a precursor. Later on, Honore et al (1993) purified a protein by 2D-gel electrophoresis from transformed fibroblasts, microsequenced the protein and cloned the cDNA using oligos made from the microsequenced protein (Honore et al., 1993). The cDNA analysis revealed that it also codes for P-32, the pre-mRNA splicing factor SF2 co-purified protein. The cDNA sequence extends beyond the 5' end of previously reported cDNA by Krainer et al, showing that ATG is the start codon at nt 79-81, which is 12 nt upstream of the 5' end of the sequence published by Krainer et al. However, N-terminus sequence of P-32 protein synthesized by cells infected with vaccinia virus construct of full length cDNA (Honore et al., 1993) including the conventional ATG start codon gave the N-terminus amino acid sequence of P-32 as reported by Krainer et al, clearly defining the synthesis of P-32 as a pro-protein of 282 amino acid, which is post-translationally processed by removal of initial 73 amino acids to a mature protein of 209 amino acids. From our study, it is clear that the composite amino acid sequence of HA-binding protein containing 209 residues is completely identical to that of mature P-32 protein reported by both the above groups. Sequence analysis further supports our earlier finding on the role of this protein in transformation. Our previous report has shown a higher expression of this protein in AK-5, a histiocytic tumor cell line (Gupta et al., 1991). Recently, Celis
et al (1990, 1991) also reported that the synthesis of P-32 protein is approximately two fold upregulated in SV40 transformed human keratinocytes and MRC-5 fibroblasts as compared to their normal counterpart establishing some transformation dependent role of HA-binding protein/P-32. In support, the computer analysis reveals the presence of phosphorylation sites of ERK/cdc2 kinase known to have definite role in cell cycle progression thus, regulating the transformation.

Extracellular signal regulated kinases (ERKs), also known as mitogen activated protein kinases (MAPKs) are important intermediates in signal transduction pathways that are initiated by many types of eukaryotic cell surface receptors. A detailed analysis of substrate specificity using synthetic peptides indicated that Pro-Xaa-Ser/Thr-Pro represents the optimal primary sequence of MAP kinase phosphorylation (Gonzalez et al., 1991), Clark et al., 1991). However, the investigation of phosphorylation of proteins indicated that while many substrates conform to the Pro-Xaa-Ser/Thr-Pro-motif, the minimal consensus sequence is Ser/Thr-Pro (Alvarez et al., 1991). Together these data indicated that the recognition of protein substrates by MAP kinase may require structural determinants (conformation) in addition to primary sequence requirements (Ser/Thr-Pro).

The substrate specificity of MAP kinase, therefore, overlaps with other proline directed protein kinases (e.g. cdc2 kinase) that are present within the cell. The 34 kDa HA-binding protein has a proline directed 160PELTSTP166 sequence, which may act as the substrate phosphorylation site of protein kinases like ERK and cdc2 family. It is already shown that Ser/Thr-Pro motif is sufficient for phosphorylation by ERK and the presence of N-terminal proline residue atleast one amino acid distant to the phosphorylation site in the motif increases the efficiency of substrate recognition (Gonzalez et al., 1994).

The substrate specificity of the ERK and cdc2 kinase being the same, the 34 kDa should behave as the substrate of either ERK or cdc2 kinase depending upon the regulation on these phosphorylations. The phosphorylation by cdc2 kinase implies the role of this protein in cell division control. Recently, Grammatikakis et al (1995) have isolated a cDNA which encodes a 29.3 kDa protein homologous to avian cdc37, an
DISCUSSION

essential cell cycle regulatory factor. The cdc37 also binds HA and contain HA- binding
motif. The biochemical function of cdc37 is unknown but genetic evidences strongly
suggest that cdc37 influences the activity of cdc2 kinase and consequently cell cycle
progression (Reed, 1992; Cutforth and Rubin, 1994). In our knowledge, cdc37 and 34
kDa HA-binding protein are the only two HA-binding proteins till date to perform some
role in cell cycle progression.

Besides the extracellular localization, GAGs have been shown to be present in the
cytoplasm and also in the nucleus, at least transiently. The types of GAGs present in the
nucleus include hyaluronic acid, chondroitin sulfate, and heparan sulfate (Furukawa and
Terayama, 1977; Fedarko and Conrad, 1986; Ishihara et al., 1986; Ripellino et al., 1988,
1989; Hiscock et al., 1994). Of particular interest is the observation that targeting of
heparan sulfate to the nucleus of rat hepatoma cells increases markedly under reduced
growth rate (Ishihara and Conrad, 1989; Fedarko et al., 1989). It has also been shown that
heparin and related polysaccharides inhibit the action of Fos and Jun on transcription
events involved in cell cycle progression (Busch et al., 1992). It seems likely that HA
being the ligand of both cdc37 (cell cycle regulatory protein) and the cdc2 kinase
substrate (34 kDa HA-binding protein) are actively involved in cell cycle regulation.

Extracellular signal regulated kinase (ERK) or MAP kinase has been reported to
phosphorylate a major cell surface substrate, EGF receptor, but the functional
significance of this phosphorylation is unclear (Northwood et al., 1991; Takishima et
al., 1991). On the other hand, MAP kinase phosphorylate a host of DNA binding proteins
like c-Myc, ATF-2, c-Jun which predominantly act as transcription factors. At this stage
, we cannot say anything about the specific DNA binding activity of this protein as
transcription factor except presenting the evidence of nuclear localization sequence (NLS)
in the protein.

The presence of the bipartite nuclear localization sequence (NLS) motif
94RKIQHK100 and 118AKLVRK123 in the deduced amino acid sequence of the mature
HA-binding protein was derived by search analysis. Though this HA-binding protein was

94
DISCUSSION

reported to be present on cell surface (Gupta et al., 1991), the presence of NLS raised the possibility of the nuclear import of this protein. This view is strengthened by the presence of cdc2 phosphorylation site and five CKII phosphorylation sites on HA-binding protein since several protein substrates of CKII have already been shown to be phosphorylated by cdc2 kinase which regulate their nuclear localization and activity. The nucleocytoplasmic distribution of nucleoplasmin is influenced by its overall degree of phosphorylation as well as NLS (Sealy et al., 1986; Paine et al., 1995). Peter and co-workers demonstrated that two specific phosphorylation sites of SV40 large T antigen, the 112S for CKII and 124T for the cdc2 kinase in the amino flank of NLS regulate its nuclear transport/accumulation (Rihs and Peters, 1989; Jans and Jans, 1994). Vancurova et al. (1995) further extended this work demonstrating that the cdc2 kinase greatly enhances the facilitated transport of SV40 large T antigen through nuclear pore complex and additional presence of CKII enhances subsequent intranuclear binding (Vancurova et al., 1995). Nuclear localization of p53 is also likely to be regulated by cdc2 kinase.

The other examples of phosphorylation regulated nuclear translocation is provided by the Rel-related family of transcription factors which includes NF-κB p50, NF-κB p65, c-Rel, (Gilmore 1991; Schmitz et al., 1991) where the non-phosphorylated transcription factor is predominantly cytoplasmic and signal stimulated phosphorylation accelerate the rate of nuclear transport by affecting the activity of NLS. Recently, the regulation on helicase activity of nucleolin, a ubiquitous eukaryotic protein essential for pre-ribosome assembly is reported (Tuteja et al., 1995) by cdc2 kinase and CKII phosphorylation. The effect of the two phosphorylation seems to be additive, hinting that both the kinases are positive regulator of the helicase activity and act on different sites of the helicase. In addition, it may be mentioned that NLS 118AKLVRK123 present in the HA-binding protein is overlapping with HA-binding motif suggesting that the availability of HA to bind this protein may regulate the compartmentalization of this protein.

Keeping this in view, the role of this HA-binding protein in cellular signalling was further explored. As depicted in the sequence, the protein lacks a conserved kinase
DISCUSSION

domain, ruling out the possibility of the protein as a receptor tyrosine kinase or serine/threonine kinase having defined kinase and GS domain (Heldin, 1995). The other possibility left to be a phosphoprotein is by specific enzymatic phosphorylation of a protein by either tyrosine kinase or serine/threonine kinase. Most of the tyrosine kinase substrates contain defined Src homology domain 2 (SH2) which recognize phosphorylated tyrosine residues in specific amino acid sequences (Paterson and Gish, 1995; Schlessinger, 1994) and SH3 domains binding to proline-rich region in several protein (Cicchetti et al., 1992). SH2 and SH3 domains are involved in complex networks of protein-protein interactions. Several SH2 and/or SH3 domain containing proteins like PLC-γ, Ras-GAP, Grb7 and spectrin also contain another conserved domain called PH-domain (Mayer et al., 1995) which is involved in signal transduction pathways. However, all these conserved domains are absent in the 34 kDa HA-binding protein.

The computer search revealed the presence of multiple phosphorylation sites like one protein kinase C (PKC) and five casein kinase (CKII) phosphorylation sites besides the already discussed ERK/cdc2 kinase site. All of these sites are phosphorylation sites of serine/threonine kinase. These observations were supplemented by an in vitro phosphorylation of the recombinant P-32/HA-binding protein by the crude liver extract known to contain a lot of kinases including CKII (Marin et al., 1990). Specifically, the recombinant P-32 was shown to be phosphorylated by CKII implying that HA-binding protein/P-32 may be an endogenous substrate of CKII. The heparin inhibition of the CKII and P-32/HA-binding protein phosphorylation and alkaline phosphatase treatment abolishing this phosphorylation prove the specificity of this phosphorylation. This specific phosphorylation is linear at lower concentrations of the recombinant P-32/HA-binding protein with inhibition at higher concentrations of recombinant P-32/HA-binding protein. Significantly, the P-32 phosphorylation was found to be enhanced in presence of HA. HA, though predominantly present in extracellular space is also present in cytosol and nucleus. Thus, CKII phosphorylation seemed to be regulated by not only HA-binding protein but also nuclear/cytosolic level of HA in cells. Specific phosphorylation of
DISCUSSION

P-32/HA-binding protein by CKII may further explain its association with splicing factor SF2 in HeLa cells. CKII is ubiquitously present in cytosol and nucleus of eukaryotic cells (Hathaway and Traugh, 1982; Edelman et al., 1987; Krebs et al., 1988) and can also behave as RNA binding protein kinase (Knondror and Stepanov, 1984). The C group hnRNP protein, implicated in splicing is known to get phosphorylated \textit{in vivo} by a CKII type activity (Holcomb and Friedman, 1984; Friedman et al., 1985). Besides C group hnRNP, other pre-mRNA binding proteins, mainly UA2F, the Mr52,000 protein of trimeric U4/U6/U5 and SF2 are phosphoproteins. Mayrand et al (1993) proposed that a cascade of critical phosphorylation and dephosphorylation directs their sequential binding and release for participation in the precatalytic state of splicing reaction. In the context of complex association of SF2 and P-32/HA-binding protein, the presence of multiple CKII sites in the protein and also the specific CKII phosphorylation of recombinant P-32/HA-binding protein, we assume that phosphorylation of this protein may be regulating the splicing ability of SF2 and HA being present in the nucleus may be playing a definite role in this regulation. Krainer et al (1991) also did not rule out the possibility of a role of P-32 in splicing and rather have shown the \textit{in vitro} interaction between P-32 and splicing factor SF2 (A.R. Krainer, personal communication). In this respect, our immunocytochemistry data showing the localization of the HA-binding protein/P-32 in nucleus and nuclear membrane in HeLa cells is of particular importance which indirectly supports the possibility of specific interaction of SF2 and P-32. At this stage, we are unable to clarify any functional role of this SF2 and P-32 interaction in pre-mRNA splicing. However, in this context, the recent report of the identification of galectin-3 (galactose/lactose specific lectin) as a factor in pre-mRNA splicing should be discussed. Dagher et al (1995) documented that HeLa cells pre-mRNA spliceosomal complex contains a high affinity saccharide (galactose/lactose) binding lectin called galectin-3 present in both cytoplasmic and nuclear fractions of cells. They considered galectin-3 as a splicing factor because firstly, the saccharide ligands that bind to galectin-3, inhibit splicing reaction; secondly, nuclear extract depleted of galectin-3 fail to form spliceosomal
complex with pre-mRNA and does not carry out the *in vitro* splicing reaction and finally the addition of recombinant galectin-3 to galectin-3 depleted nuclear extract restored the splicing reaction. The role of HA available in the nucleus as a modulator of activities of P-32 in pre-mRNA splicing will be an interesting study of the future.

After observing the specific CKII phosphorylation *in vitro* we studied the *in vivo* phosphorylation of the protein under HA stimulation in transformed macrophages. As most of the putative phosphorylatable sites are of serine/threonine kinase phosphorylation sites, we employed some kinase/phosphatase regulators of serine/threonine phosphorylation. As expected, the 34 kDa HA-binding protein undergoes enhanced phosphorylation by HA, serine/threonine kinase stimulator (PMA) and serine/threonine phosphatase inhibitor (calyculin A) whereas this is unaltered with staurosporine (PKC inhibitor). All these results point out the involvement of serine/threonine kinase(s) in HA-binding protein phosphorylation. But it will be difficult to say at this stage whether the PKC, CKII, ERK or cdc2 kinase each alone or in conjunction with each other takes part in this *in vivo* phosphorylation. Though CKII was earlier used to be considered as messenger independent molecule, evidences are pouring in demonstrating the activation of CKII by mitogenic stimulation like EGF, insulin and insulin like growth factor (Carroll and Marshak, 1989). CKII activity was shown to be increased by EGF in A-431 carcinoma cells correlated with an increase in the phosphorylation of the CKII (Sommercorn et al., 1987). Pepperkok et al (1991) have also demonstrated that oligonucleotides complementary to CKII mRNAs (of both α and β sub units) significantly inhibit the early phase of mitogenic stimulations of cells by EGF and serum. The *in vivo* phosphorylation 68 kDa secretory protein followed the same pattern of regulation as that of 34 kDa HA-binding protein. The anti-phosphoamino acid immunoblotting of anti-HA-binding protein antibodies immunoprecipitates revealed that 34 kDa HA-binding protein is phosphorylated at threonine residue whereas 68 kDa protein is phosphorylated at serine residue. This observation identifies the \(^{76}\text{T}, ^{231}\text{T}, ^{261}\text{T}\) (all by CKII) and \(^{165}\text{T}\) by (ERK/cdc2 kinase) as the probable sites of *in vivo* phosphorylation of 34 kDa HA-
DISCUSSION

binding protein. Similarly, $^{205}\text{S}$ (by PKC and CKII both) and $^{213}\text{S}$ (by CKII) are the probable sites of 68 kDa protein, provided the 68 kDa has the same protein backbone as that of 34 kDa HA-binding protein.

Multisite phosphorylation is a prevalent form of protein modification whose full implications are just beginning to be understood. Multiple phosphorylations can correlate with the generation of a variety of protein forms, in which one or more properties are altered. Different phosphorylation could thus be linked to distinct protein functions or graded effects on a single function. Alternatively, multiple phosphatase might be necessary for one critical conformational change. In this respect the net outcome of these multisite phosphorylations is a subject of future study. As all the kinases (cdc2, ERK, CKII, PKC) mentioned are cytosolic or nuclear enzymes, the \textit{in vivo} phosphorylation of 34 kDa HA-binding protein is expected to occur in cytoplasm or in nucleus. However, the phosphorylation form of cell surface 34 kDa HA-binding protein (Gupta et al., 1991) has also been studied in our laboratory. The enhanced phosphorylation of cell surface 34 kDa HA-binding protein has been observed in the isolated plasma membrane of EL4, a lymphoma cell line. It seems that cell surface 34 kDa HA-binding protein as well as secretory 68 kDa protein are phosphorylated in cytosol and then transported to the cell surface. The hydrophilicity study of the pro-protein revealed a stretch hydrophobic residues at the N-terminus of the pro-protein which may form the signal peptide for the translocation of the protein from cytosol to cell surface. This is in consistent with the observation of Ghebrehiwet et al. (1994) who identified the completely identical 282 amino acids SF2P-32 (Honore et al., 1993) as gC1q receptor (discussed later), that the residues 1-7 or 1-13 at the N-terminus of the pro-protein may form the signal peptide since these two-segments give scores of 4.5 or 5.5, respectively in the SIGCLEAVE command of GCG programme. They also proved the gC1q receptor as a cell surface protein of 209 amino acids. If the first 1-7 or 1-13 residues really act as a signal peptide, then most probably the whole 282 amino acids long pro-protein undergo proteolytic cleavage on the cell surface to remove the first 73 amino acids. Alternatively,
DISCUSSION

the translocation of the protein may be phosphorylation dependent as like as the
intracellular phosphorylation and extracellular translocation or secretion of IL2, which
does not possess a defined signal peptide (Benschel et al., 1988).

Further analysis of the role of this protein in signal transduction revealed that this
protein transduces HA-induced signal in PLC-γ phosphorylation in macrophages which
after activation probably leads to breakdown of phosphatidyl inositol 4,5-bisphosphate
(PIP₂) resulting in enhanced inositol 1,4,5-triphosphate (IP₃) formation. Both, PLC-γ
phosphorylation and IP₃ formation are inhibited when the cells are preincubated with
anti-HA-binding protein antibodies and then, stimulated with HA, which clearly indicated
the specific role of this protein in HA-induced signal transduction. This observation
establishes that HA does promote second messenger formation through 34 kDa
HA-binding protein and resembles in its effects other ECM molecules like collagen and
laminin, the cell adhesive proteins which promotes PLC-γ activation and IP₃ formation
(Watson et al., 1985; Liotta et al., 1988).

Recent work on the cellular signalling by ECM cell adhesive protein identified
a family of proteins known as cell adhesive receptors. Out of them integrin is the most
promising candidate to take part in biochemical signalling from ECM to cell interior.
Interaction of cell adhesion molecule with integrin initiates integrin clustering which in
turn leads to tyrosine auto phosphorylation of focal adhesion kinase
(pp125FAK). Autophosphorylation of pp125FAK results in stable binding of it to the SH2
domain of Src family kinases. It is also shown that PLC-γ is a potential target for
pp125FAK/ppSrc complex which after being activated produce IP₃ by PIP₂ breakdown.

The ECM interaction with integrin has been demonstrated in lymphocytes, where
B-cell and T-cell antigen receptor leads to phosphorylation of PLC-γ (Weiss et al., 1994)
resulting in PIP₂ breakdown and IP₃ formation. Morino et al (1995) proposed that
Src-family kinases might be strong candidate in this β₁ integrin signal transduction
pathway (as discussed before) leading to the MAP kinase activation. From our present
observation, the possibility that HA-binding protein may transduce the HA-induced
signalling event followed by integrin interaction is speculated due to following reasons; firstly, HA-binding protein was shown as cell adhesive protein (Gupta et al., 1991); secondly, HA induction is reported to initiate enhanced tyrosine phosphorylation (Ranganathan et al., 1995); thirdly, our previous report of the presence of two plasmamembrane proteins of 41 kDa and 37 kDa which may serve as connecting link between cell surface and cytosol (Gupta et al., 1993) and finally, the inhibition of HA induced PLC-γ phosphorylation and IP₃ formation by pretreatment of the cells with anti-HA-binding protein antibodies.

Thus, at the second phase of this work we have described the role of 34 kDa HA-binding protein in signal transduction showing specific HA induced specific phosphorylation and regulation of such phosphorylation by serine/threonine kinase/phosphatase regulators. All these observations suggest that HA binding protein may act as a serine/threonine kinase substrate. Moreover, our work also highlights the specific role of this protein in HA induced second messenger formation and PLC-γ phosphorylation which probably relates to mobilization of intracellular Ca+++, needed for cellular functions. Further work identifying this protein as specific substrate of CKII may explain its association with SF2 with a probable function in pre-mRNA splicing.

In support of our observations, few HA-binding proteins are already reported to be involved in cellular signalling. Serum, c-AMP and PMA, as for example, shown to initiate the signal transduction pathways, resulting in gene induction of HA-synthase. Klewes and Prehm (1989) have shown that HA-synthase could also be phosphorylated and activated by the tyrosine kinase pp60ˢᵉ from Rous sarcoma virus in chicken embryo fibroblasts and the result indicated that multiple kinases could directly phosphorylate the synthase changing its activity. CD38, a strong homologue of ADP-ribosyl cyclase protein having ecto NADase activity leads to the synthesis of cyclic ADP-ribose (cADPR), an endogenous second messenger regulating the mobilization of intracellular calcium (Lee et al., 1989). CD38 mediated signalling for activation and proliferation of human T cells, thymocyte and NK cells has also been reported (Funaro et al., 1990). Yoneda et al (1988
DISCUSSION

a, b) proposed that serum 85 kDa protein after intercalating into plasma membrane mediates HA induced signal transduction leading to cell proliferation. T-cell activation via CD44 requires LFA-1 (ligand of ICAM-1) and thus, ICAM-1 is reported to be involved in the T-cell activation (Funaro et al., 1994). CD44 tail region possesses some similarity with the members of the G-protein superfamily and purified CD44 is reported to bind GTP and display some GTPase activity in \textit{in vitro} assays (Lokeshwar and Bourguignon, 1992). Binding of HA to CD44 is also reported to initiate intracellular signalling (Bourguignon et al., 1993). Hall et al (1994) suggested that HA stimulates cell locomotion via a rapid and transient protein tyrosine kinase signalling event mediated by RHAMM. The signalling event ultimately leads to IP$_3$ formation which mobilizes intracellular Ca++ likely to be required for cell locomotion (Turley, 1989). Under the light of above mentioned signalling events, the 34 kDa HA-binding protein being an endogenous substrate of protein kinases and being localized at different cellular compartment further highlights the role of HA-binding protein in cellular signalling.

We have already discussed the homology of HA-binding protein with the splicing factor associated protein. In addition, the analysis of the search data gave us the interesting information revealing that SF2P-32/34kDa HA-binding protein is homologous with two other proteins namely, gC1q receptor (Ghebrehiwet et al., 1994) and YL2 (Luo et al., 1994) and having significant homology with bacterial collagen adhesin protein (Patti et al., 1992) suggesting that HA-binding protein may have multifunctional activities.

The complement protein C1q, is reported to bind with gC1q receptor, a protein completely identical to P-32/HA-binding protein. This binding takes place by the first 24 amino acids of the mature gC1q-R and contributes to the diversity of C1q mediated response. As the HA-binding motif does not fall within the C1q-binding site, it is probable that binding of HA to gC1q-R regulates C1q binding. Prehm (1995) further revealed that C1q itself is a HA-binding protein and this binding is covalent in nature unlike other HA-binding protein-HA interactions which are ionic in nature. Thus, HA might be playing a definite role by regulating functions of both the C1q and C1q-R. In
chronic inflammation, continuous production of chemotactic factors is necessary to attract granulocytes in the inflamed area. These chemotactic factors are generally complement factors that are produced during inflammation. HA, being specifically enriched in the joints of rheumatoid arthritis patients and its covalent binding with Ig forms HA-Ig complex which itself become antigenic to elicit antibodies. Thus, complete cascade is initiated by the binding of C1q to antigen-Ig complex. The observations of Ghebreweit et al. (1994) and Prehm (1995) thus revealed that HA after binding to each of C1q, C1q-R and Ig may act as a potential modulator in immunological reactions. Thus, in future the role of C1q and HA regulation in gC1q-R mediated complex formation in chronic inflammatory processes could be studied in details.

Luo et al. (1994) identified a murine homologue (92%) of P-32/HA-binding protein called YL2 that interacts with basic domain of Rev, one of the earliest proteins expressed by HIV-1 responsible for post-transcriptional regulation of viral RNAs. Though, Rev functions are yet to be fully understood, the transportation of viral RNAs from nucleus to cytoplasm and its effect on RNA splicing have been studied. Luo et al. (1994) showed that both the activities of Rev (RNA splicing inhibition and facilitated transport of viral transcripts) are modulated by interaction of Rev and YL2. This work supports the assumptions of involvement of P-32 in splicing made by Krainer et al. (1991). In this respect, the effect of nuclear HA will further throw light on the modulations of YL2/HA-binding protein in Rev functions.

The collagen adhesin (also known as polypeptide having collagen binding activity) of Staphylococcus aureus also possesses significant homology with the 34 kDa HA-binding protein. This bacterial cell surface protein binds collagen with high affinity and specificity (Patti et al., 1992). The interaction between collagen and collagen adhesin is one of the potential mechanisms of bacterial adhesion to host cells as found in the pathogenesis of osteomyelitis and infectious arthritis. The 34 kDa HA-binding protein is significantly homologous to the repeat sequences of collagen adhesin which are probably resulted from gene duplication events. This observation is mainly important.
because the synovial fluid of arthritis patients contains large amount of HA (Dahl and Husby, 1985; Wells et al., 1992) and collagen binding protein (CRIII) has also been reported to bind HA (Carter and Wayner, 1988). The exact role played by HA and the implications of this homology is at present obscure which needs further investigations.