William H. Robinson,
Henry E. Neuman de Vegvar,
Susan S. Prohaska,
Joon W. Rhee,
and Jane R. Parnes

Department of Medicine, Division of Immunology and Rheumatology and the Digestive Disease Center, Stanford University School of Medicine, Stanford, USA

Human CD6 possesses a large, alternatively spliced cytoplasmic domain

Human CD6 is a monomeric 105/130-kDa T cell surface glycoprotein that is involved in T cell activation. The apparent discrepancy between the size of the cytoplasmic domain in human (44 amino acids) and mouse (243 amino acids) CD6, led us to use reverse transcriptase-polymerase chain reaction of human peripheral blood lymphocyte mRNA to isolate cDNA clones that include the carboxyl-terminal coding region of human CD6. The nucleotide sequence of the longest human cDNA clone, CD6-PB1, predicts a protein of 668 amino acids with a 244-amino acid cytoplasmic domain similar in size to and possessing 71.5% amino acid sequence identity with the cytoplasmic domain of mouse CD6. This previously unrecognized 244-amino acid cytoplasmic domain does not have significant homology to any other known protein (except mouse CD6), but does possess two proline-rich motifs containing the SH3 domain-binding consensus sequence, a serine-threonine-rich motif repeated three times, three protein kinase C phosphorylation-site motifs, and 10 casein kinase-2 phosphorylation-site motifs. These sequences are likely to play a role in the ability of CD6-specific monoclonal antibodies to stimulate T cell proliferation. Full-length CD6 cDNA containing this cytoplasmic domain sequence encodes a monomeric 105/130-kDa protein that can be immunoprecipitated from the surface of transfected cells and comigrates upon SDS-PAGE with wild-type CD6 immunoprecipitated from PBL. We also isolated two alternatively spliced forms of human CD6 cDNA lacking sequences encoding membrane-proximal regions of the cytoplasmic domain which maintain the same reading frame as CD6-PB1. The short cytoplasmic domain of the previously reported human CD6-15 cDNA clone results from a deletion of a 20-bp segment through use of an alternative 3′ splice site, resulting in a frame shift and premature termination of translation relative to the clones we have isolated. These data demonstrate that human CD6 possesses a large cytoplasmic domain containing sequence motifs that are likely to be involved in signal transduction upon stimulation of T cells through CD6 ligation.

1 Introduction

Human CD6 is a T cell surface glycoprotein expressed at high levels by peripheral blood T cells and medullary thymocytes [1–3]. Most human CD6-specific mAb can enhance T cell proliferation induced by suboptimal doses of CD3-specific mAb, but they can differentially enhance proliferation when combined with other stimuli (*e.g.* PMA, macrophages) [3–7]. the cytoplasmic domain of human CD6 is constitutively phosphorylated on serine and threonine residues, and T cell activation by cross-linking CD3 with specific mAb results in phosphorylation of tyrosine and hyperphosphorylation of serine residues in the CD6 cytoplasmic domain [5, 8–10]. Human CD6 was recently shown to bind to the 100-kDa activated leukocyte cell-adhesion molecule (ALCAM) expressed on the surface of thymic epithelial cells [11].

Human CD6 exists on the cell surface as a heavily glycosylated monomeric 105 or 130-kDa type 1 integral membrane glycoprotein with intrachain disulfide bonds [5, 8, 9, 12, 13]. The difference in size between the two observed forms was thought to be related to protein degradation [5] or glycosylation differences [5, 8]; however a recent study indicates that this difference can be accounted for by phosphorylation of the larger relative to the smaller form [9]. From the cell line HPB-ALL, Aruffo et al. [13] isolated the human CD6 cDNA clone CD6-15 that encodes a mature 444-amino acid polypeptide homologous to CD5 and to the family of proteins containing cysteine-rich domains defined by that of the type 1 macrophage scavenger receptor. The predicted mass for this mature polypeptide of 47.4-kDa is significantly smaller than the 88-kDa nascent, unmodified CD6 polypeptide observed by Swack et al. [8] in T cells and thymocytes. Furthermore, immunoprecipitation of CD6 from COS cells transiently transfected with the human CD6-15 cDNA in an expression vector produced a 90-kDa product, which was significantly smaller than the 105- and 130-kDa CD6 glycoproteins immunoprecipitated concomitantly by Aruffo et al. from HPB-ALL [13], as well as by other investigators from peripheral blood T cells [5, 8, 9, 12]. Aruffo et al. [13] attributed this difference in molecular mass to incomplete glycosylation of the transfected CD6-15 protein product, a phenomenon which they have previously observed with other proteins expressed in COS cells [14].

We and others have recently isolated cDNA encoding mouse CD6 (W.H.R. and J.R.P., manuscript submitted; [15]). The predicted mouse CD6 protein sequence from these clones is similar to that of the previously reported

[I 14595]

Correspondence: Jane R. Parnes, Division of Immunology and Rheumatology, Medical School Lab. Surge Building, Room P306, Stanford University Medical Center, Stanford, CA 94305-5487, USA (Fax: 415-723-7509)

Key words: Human CD6 / T cells / Cytoplasmic domain / CD6

0014-2980/95/1010-2765$10.00 + .25/0

human CD6-15 [13] in its NH_2-terminal signal sequence, extracellular domain, and transmembrane domain. However, the predicted cytoplasmic domain of our mouse CD6-4 cDNA clone possesses 243 amino acids compared to the 44-amino acid cytoplasmic domain reported by Aruffo et al. [13] for human CD6-15. We now report the isolation of human PBL cDNA clones representing three alternatively spliced forms of the 3′ region of CD6 mRNA. These clones contain two insertions (20 and 105 bp) relative to the previously reported human CD6-15 sequence. The more proximal insertion (20 bp) resulted from use of an alternative 3′ splice site and produces a frame-shift that yields (for the longest clone) a predicted protein of 668 amino acids, with a 244-amino acid cytoplasmic domain similar in size and sequence to that of mouse CD6. This cytoplasmic tail contains several motifs that may be involved in CD6-mediated signaling.

2 Materials and methods

2.1 PCR amplification and isolation of CD6 3′ cDNA sequences

Total RNA was isolated from human PBL using RNAzol (Cinna/Biotec, Friendswood, TX). Human CD6 cDNA was generated and then enzymatically amplified using the PCR as described elsewhere [16] (primers: 5′-CCTCCATCGTTCTGGGA-3′ (representing nucleotide positions 1331–1347 of CD6-15 just 5′ of the Eco R1 restriction enzyme site at position 1346) and 5′-ATGAGATCTCCTGGCGGACTTGGAGTGTCTCTG-G- 3 (representing nucleotide positions 2668–2692 of CD6-15 and including a Cla I restriction enzyme cleavage site). The amplified CD6 cDNA fragments were cloned into the Bluescript SK^+ plasmid (Stratagene, La Jolla, CA) at the Eco R1 and Cla I restriction enzyme cleavage sites and sequenced fully in both directions using the dideoxy chain termination method [17].

2.2 cDNA constructs and transfection

The hybrid-CD6-PB1 cDNA molecule was generated by combining the 3′ CD6-PB1 cDNA Eco R1-Cla I 1469-bp fragment isolated using PCR with the 5′ Xho I-Eco R1 CD6 cDNA fragment including bp 1 through 1348 of the CD6-15 cDNA isolated by Aruffo et al. [13] by subcloning them into the Xho I and Cla I restriction enzyme sites of the pBJ-neo expression vector [18] in a three-way ligation. Transfection of L cells with the pBJ-neo-human CD6-PB1 hybrid cDNA expression construct was performed using the Lipofectin Reagent (Gibco BRL). L cell-CD6 cDNA transfectants were selected by growth in the presence of 1.0 mg/ml G418 (Gibco BRL) and screened for CD6 expression by flow cytometry.

2.3 Immunoprecipitation

Human PBL and transfected L cells were surface-labeled with ^{125}I using the lactoperoxidase method as described [19]. L cell and PBL lysates (lysis buffer: 1% Nonidet P-40, 10 mM Tris-HCl pH 7.5, 1 mM EDTA, 0.15 M NaCl, 1 mM phenylmethylsulfonyl fluoride, 1 mM iodoacetamide, 25 μg/ml soybean trypsin inhibitor, 25 μg/ml pepstatin, 5 μg/ml leupeptin, 50 μM sodium fluoride and 0.04 U/ml aprotinin) were precleared twice with normal mouse serum and then with the isotype-matched control mAb B6-27 (a mouse IgG1) specific for (4-hydroxy-3-nitrophenyl)acetyl (NP) [20]. Immunoprecipitation of human CD6 was performed by adding CD6-specific mAb (PharMingen, San Diego, CA) to the lysates followed by protein G-Sepharose. The precipitates were washed and electrophoresed on 8% polyacrylamide gels containing SDS as described [19].

3 Results and discussion

3.1 Isolation of human CD6 3′ cDNA sequences

Initially we unsuccessfully attempted to re-clone the entire human CD6 cDNA using two different sets of oligonucleotide primers representing sequences in the 5′ and 3′ untranslated regions of the CD6-15 cDNA. The human CD6-15 cDNA contains large GC-rich regions in the 5′ coding sequence which we believe significantly increases the melting temperature of its corresponding double-stranded DNA, and thereby hinders PCR amplification of this region of the cDNA. We then successfully used oligonucleotide primers 3′ of this GC-rich region to amplify by PCR and clone the cDNA sequence encoding the 3′ coding region of human CD6. At least three distinct bands were identified upon agarose gel electrophoresis of the PCR products. The largest major band as well as the smaller minor bands (pooled together) were isolated and cloned (Fig. 1). The 1469-bp CD6-PB1 clone is our longest 3′ human CD6 cDNA clone. As compared to CD6-15, it contains a 20-bp insertion following nucleotide position 1507 of CD6-15, resulting in a frame-shift bypassing the predicted termination codon of CD6-15 (Fig. 1). CD6-PB1 also possesses a 105-bp insertion relative to CD6-15 starting at position 1957; this insertion encodes 35 amino acids that are conserved with respect to mouse CD6 (Figs. 1 and 2). Combining the sequence of the 5′ region of CD6-15 with the sequence of CD6-PB1 results in a predicted open reading frame coding for a 668-amino acid polypeptide chain (Figs. 1 and 2). The predicted hybrid-CD6-PB1 protein diverges from the predicted protein encoded by CD6-15 in the cytoplasmic domain after amino acid position 462, and contains an additional 200 amino acids in its cytoplasmic domain. The position and sequence surrounding the predicted termination codons for mouse CD6-4 and human CD6-PB1 are conserved (Fig. 2). We now predict a mature human CD6 protein of 644 amino acids, with a 244-amino acid cytoplasmic domain and a predicted molecular weight for the mature polypeptide of 69.4 kDa.

We submitted the amino acid sequence of the cytoplasmic domain of full-length human CD6 containing the additional 200 amino acids to the National Center for Biotechnology Information (Bethesda, Maryland) for comparison with all the protein sequences available in the SWISS-PROT and GenBank databases using the BlastP program [21]. We found no significant homology to any known proteins other than mouse CD6. We compared the cytoplasmic domain encoded by CD6-PB1 with the Prosite 12 database using the Quest program (IntelliGenetics, Mountain View, CA) to search for amino acid sequence motifs.

```
        Eco RI                                  TM ←|→CY                |=============
hCD6-PB1   IleLeuLeuLeuGlySerLeuIlePheIleAlaPheIleLeuLeuArgIleLysGlyLysTyrAlaLeuProValMet   435
hCD6-PB1 GAATTCTCCTCCTTGGCTCCCTCATCTTCATAGCCTTCATCCTCTTGAGAATTAAAGGAAAATATGCCCTCCCCGTAATG  1425

         ==========================================================================
hCD6-PB1   ValAsnHisGlnHisLeuProThrThrIleProAlaGlySerAsnSerTyrGlnProValProIleThrIleProLys   461
hCD6-PB1   GTGAACCACCAGCACCTACCCACCACCATCCCGGCAGGGAGCAATAGCTATCAACCGGTCCCCATCACCATCCCCAAA  1503

         =====----------------------------------------------*--^-----^-----^----------
hCD6-PB1   GluValPheMetLeuProIleGlnValGlnAlaProProProGluAspSerAspSerGlySerAspSerAspTyrGlu   487
hCD6-PB1   GAAGTTTTCATGCTGCCCATCCAGGTCCAGGCCCCGCCCCCTGAGGACTCAGACTCTGGCTCGGACTCAGACTATGAG  1581
hCD6-15    GluG                        lyProGlyProAlaProTER

         ----------------------------------------------|  +                ^
hCD6-PB1   HisTyrAspPheSerAlaGlnProProValAlaLeuThrThrPheTyrAsnSerGlnArgHisArgValThrAspGlu   513
hCD6-PB1   CACTATGACTTCAGCGCCCAGCCTCCTGTGGCCCTGACCACCTTCTACAATTCCCAGCGGCATCGGGTCACAGATGAG  1659

hCD6-PB1   GluValGlnGlnSerArgPheGlnMetProProLeuGluGluGlyLeuGluGluLeuHisAlaSerHisIleProThr   538
hCD6-PB1   GAGGTCCAGCAAAGCAGGTTCCAGATGCCACCCTTGGAGGAAGGACTTGAAGAGTTGCATGCCTCCCACATCCCAACT  1737

                                                          ^       *
hCD6-PB1   AlaAsnProGlyHisCysIleThrAspProProSerLeuGlyProGlnTyrHisProArgSerAsnSerGluSerSer   565
hCD6-PB1   GCCAACCCTGGACACTGCATTACAGACCCGCCATCCCTGGGCCCTCAGTATCACCCGAGGAGCAACAGTGAGTCGAGC  1815

           ^  ^                +                                  +
hCD6-PB1   ThrSerSerGlyGluAspTyrCysAsnSerProLysSerLysLeuProProTrpAsnProGlnValPheSerSerGlu   591
hCD6-PB1   ACCTCTTCAGGGGAGGATTACTGCAATAGTCCCAAAAGCAAGCTGCCTCCATGGAACCCCCAGGTGTTTTCTTCAGAG  1893

              ^
hCD6-PB1   ArgSerSerPheLeuGluGlnProProAsnLeuGluLeuAlaGlyThrGlnProAlaPheSerAlaGlyProProAla   617
hCD6-PB1   AGGAGTTCCTTCCTGGAGCAGCCCCCAAACTTGGAGCTGGCCGGCACCCAGCCAGCCTTTTCAGCAGGGCCCCCGGCT  1971

           *                ^                  P
hCD6-PB1   AspAspSerSerSerThrSerSerGlyGluTrpTyrGlnAsnPheGlnProProProGlnProProSerGluGluGln   643
hCD6-PB1   GATGACAGCTCCAGCACCTCATCCGGGGAGTGGTACCAGAACTTCCAGCCACCACCCCAGCCCCCTTCGGAGGAGCAG  2049

              P                          ^          Y
hCD6-PB1   PheGlyCysProGlySerProSerProGlnProAspSerThrAspAsnAspAspTyrAspAspIleSerAlaAlaTER   668
hCD6-PB1   TTTGGCTGTCCAGGGTCCCCCAGCCCTCAGCCTGACTCCACCGACAACGATGACTACGATGACATCAGCGCAGCCTAG  2127
```

Figure 1. Nucleotide and predicted carboxyl-terminal amino acid sequences of human CD6 cDNA. The predicted transmembrane (TM) and cytoplasmic (CY) domains of full-length CD6-PB1 are indicated. Human CD6 cDNA clones that lack sequences encoding membrane-proximal regions of the cytoplasmic domain were also isolated: CD6-PB2 lacks the 96 bp overscored by (= = =) and CD6-PB3 lacks the 219 nucleotides overscored by (= = =) and (---). The nucleotide sequences of CD6-PB1 and -PB2 contain a 20 bp insertion indicated in bold at position 1508, as compared to the previously published cDNA CD6-15 [13]. The nucleotide sequences of CD6-PB1, -PB2, and -PB3 also contain a 105 bp insertion relative to CD6-15 indicated in bold at position 1957. CD6-PB1, -PB2 and -PB3 have identical sequences 3′ of their predicted termination codons. Two proline-rich sequences which contain the SH3-domain binding consensus motif X-Pro-X-Pro-Y-X-Pro are indicated by (P). The three Asp-Ser-Ser-Ser/Thr-Y-Ser-X-X-Glu-X-Tyr motifs are indicated by (*). The Tyr-Asn-Asn-Ile motif is indicated by a (Y). The 10 motifs containing the casein kinase-2 phosphorylation site consensus sequence are indicated by (^). The three motifs containing the PKC phosphorylation site consensus sequence are indicated by (+). The accession number for the hybrid CD6-PB1 cDNA sequence is U34623 (U34624 for PB2 and U34625 for PB3) at the sequence data base at the national Center for Genome Resources.

As indicated in Fig. 1, we found 10 motifs containing the casein kinase-2 phosphorylation site consensus sequence Ser/Thr-X-X-Asp/Glu (where X represents a non-conserved residue, and serine or threonine is phosphorylated) [22] and three motifs containing the protein kinase C (PKC) phosphorylation site consensus sequence Ser/Thr-X-Arg/Lys (where X represents a non-conserved residue, and serine or threonine is phosphorylated) [23]. Both PKC and casein kinase-2 are known to phosphorylate multiple different transmembrane protein receptors [22, 23]. Notably, phorbol ester, serum treatment or both, of human PBL has been shown to increase the phosphorylation of CD6 on serine, and both of these agents are activators of PKC [8–10, 23, 24]. It is possible that PKC, casein kinase-2, or a different closely related kinase could phosphorylate CD6 at these sites.

Visual inspection of the cytoplasmic domain of human CD6 revealed a conserved amino acid motif repeated three times with the consensus sequence Asp-Ser-Ser-Ser/Thr-Y-Ser-X-X-Glu-X-Tyr (where X represents a non-conserved residue, and Y represents a non-conserved residue or a gap in the alignment) (Figs. 1 and 2). Visual inspection also revealed the amino acid motif Tyr-Asp-Asp-Ile that represents a potential recognition specificity of the SH2 domain of c-fgr or a different Src kinase family member [25]. T cell activation results in phosphorylation of tyrosine and hyperphosphorylation of serine residues in the cytoplasmic domain of human CD6 [9,10], and these sequence motifs could represent targets for phosphorylation of the CD6 cytoplasmic domain on serine, tyrosine, or both residues. They could also represent binding sites for interaction with other proteins.

We identified two conserved proline-rich sequences which contain the X-Pro-X-Pro-Y-X-Pro motif (X represents a non-conserved residue, and Y represents proline or a non-conserved residue) characteristic of the class I SH3-binding motif consensus sequence [26] (Figs. 1 and 2). Notably, the proline-rich sequence Gln-Pro-Pro-Pro-Gln-Pro-Pro at position 633 of human CD6 has significant similarity to the chicken YAP65 SH3-binding motif Gln-Pro-Pro-Pro-Leu-Ala-Pro (Fig. 2) [26]. These proline-rich sequences may represent binding sites for SH3 domain-containing proteins involved in CD6 signaling.

Cloning of the smaller, minor CD6 PCR products from human PBL yielded two distinct forms of human CD6 cDNA lacking sequence encoding membrane-proximal regions of the cytoplasmic domain (Figs. 1 and 2). CD6-PB2 lacks 96 bp, encoding 32 amino acids, from nucleotide positions 1412 to 1507 of CD6-PB1. This 96-bp deletion in CD6-PB2 is immediately 5′ of the 20-bp deletion in CD6-15; those 20 bp are present in CD6-PB2. CD6-PB3 lacks 219 bp, encoding 73 amino acids, between nucleotide positions 1412 and 1630 of CD6-PB1. This 219-bp deletion in CD6-PB3 begins at the same 5′ site as the 96-bp deletion in CD6-PB2 and continues through and beyond the 20-bp deletion in CD6-15. The CD6-PB2 and -PB3 cDNA both maintain the same reading frame as CD6-PB1 after the gaps in their nucleotide alignment, and thus have identical cytoplasmic domain 3′ coding sequences and termination codons.

The CD6-PB2 and CD6-PB3 cDNA sequences we describe here most likely resulted from alternative splicing of the CD6 transcript. The deletions in both CD6-PB2 and CD6-PB3 relative to CD6-PB1 begin with nucleotide 1412,

```
                ==================================================*^-^-^-----
hCD6-PB1    RIKGKYALPVMVNHQHLPTTIPAGSNSYQPVPITIPKEVFMLPIQVQAPPPEDSDSGSDSDYEHY   489
mCD6-4      -A--Q----AS----Q-S-ANQ--I-N-H------A--AP--F--..PRV-A----S--------   485
hCD6-15     ------------------------------------GPGPAP*

            -------------- +      ^
hCD6-PB1    DFSAQPPVALTTFYNSQRHRVTDEEVQQSRFQMPPLEEGLEELHASHIPTANPGHCITDPPSLGP   554
mCD6-4      ---S-----------------E--A--N--------------V----A-D-RP-VA-V--R-S   550

                 *   ^                          +                      ~~~~~~*
hCD6-PB1    QYHPRSNSESSTSSGEDYCNSPKSKLPPWNPQVFSSERSSFLEQPPNLELAGTQPAFSAGPPADD   619
mCD6-4      ---V-N--D-----E-G---D-S--P----S-A-Y--K-PLT----------.S--VFS--S---   514

            -----^---------P~~~~~~~~~~~~P-                 Y
hCD6-PB1    SSSTSSGEWYQNFQPPPQPPSEEQFGCPGSPSPQPDSTDND..DYDDISAA*                668
mCD6-4      -----------------H-PA---E---P-G--T--I-D-EE-----G--*                 665
```

Figure 2. Comparison of the predicted amino acid sequences of full-length human and mouse CD6 cytoplasmic domains. Identical residues are denoted by (---) and gaps in the alignment are denoted by (···). The predicted 32 amino acids absent in CD6-PB2 are overscored by (= = =) and the 73 amino acids absent in CD6-PB3 are overscored by (= = =) and (---). The predicted 35 amino acids absent in the cytoplasmic domain predicted from the 3′ untranslated region of CD6-15 are indicated by (~ ~). Other motifs are indicated according to the legend to Fig. 1.

which corresponds in location to an exon/intron junction in the mouse CD6 gene [15], suggesting that this is the start of an exon whose sequence is spliced out in the mRNA corresponding to both of these clones. Similarly, the 3′ end of each of these deletions corresponds to an exon/intron junction in the mouse CD6 gene; the sequence missing in CD6-PB2 and -PB3 correspond to the splicing out of one and two exons, respectively, based on the mouse exon/intron structure [15]. The 96-bp deletion in CD6-PB2 relative to CD6-PB1 ends at position 1507, while the 20-bp deletion in CD6-15 as compared to CD6-PB1 starts at position 1508. 1507/1508 is also the site at which an intron is spliced out in the mouse CD6 transcript [15]. Notably, the 3′ end of the 20-bp deletion in CD6-15 cDNA contains a potential, though not perfect, 3′ splice site (CTGCCCATCCAG), suggesting that this deletion might result from inappropriate splicing out of these 20 nucleotides. The product of mRNA containing the 20-bp deletion is probably not expressed at significant levels *in vivo* because no smaller species of CD6 polypeptide corresponding in size to that of CD6-15 is observed in immunoprecipitates of wild-type CD6 from human PBL (Fig. 3, [9]). We do not know whether human CD6 mRNA with only the 3′ 105-bp deletion of clone CD6-15 is present *in vivo*. We isolated, but did not sequence, multiple cDNA clones that could correspond according to size either to such a cDNA (with or without the 20-bp deletion of the CD6-15 cDNA) or to the CD6-PB2 cDNA.

3.2 Comparison of the mouse and human CD6 cytoplasmic domains

We compared the predicted amino acid sequences of the full-length cytoplasmic domain of human CD6 with that of mouse CD6-4 (W.H.R. and J.R.P., manuscript submitted) (Fig. 2). These domains exhibit 71.5% (176/246) identity, compared with 70% identical amino acid residues in the full protein. One single and two double amino-acid gaps are present in the alignment of the cytoplasmic domains. There are 22 conserved serine and 9 conserved tyrosine residues in the cytoplasmic domains of mouse and human CD6. The three serine/threonine-rich motifs, the two proline-rich motifs, the Tyr-Asp-Asp-Ile motif, 6 of the 10 casein kinase-2 phosphorylation site motifs and two of the three PKC phosphorylation site motifs are conserved in mouse CD6 (Fig. 2).

3.3 Immunoprecipitation of human CD6 from PBL and L cell transfectants expressing the hybrid CD6-PB1 cDNA

To demonstrate that the hybrid-CD6-PB1 cDNA does in fact encode full-length human CD6, we performed immunoprecipitations from lysates of radioiodinated PBL as well as L cells expressing the hybrid CD6-PB1 cDNA. When analyzed under reducing conditions, the CD6-specific mAb precipitates polypeptide bands of 105 and 130 kDa from L cells expressing the hybrid CD6-PB1 cDNA molecule; these bands comigrate with wild-type CD6 immunoprecipitated from activated PBL (Fig. 3 A, lanes 3 and 5). When analyzed under non-reducing conditions, the CD6-specific mAb precipitates polypeptide bands of 100 and 125 kDa that also comigrate with wild-type CD6 immunoprecipitated from activated PBL (Fig. 3 B, lanes 3 and 5). As previously described, non-reduced CD6 migrates faster than reduced CD6 due to the presence of intrachain disulfide bonds [5, 8, 13]. Thus, hybrid CD6-PB1 encodes full-length CD6, and the additional 200 amino acids predicted by CD6-PB1 can account for the previous discrepancies in the molecular weights of CD6-15 as compared to wild-type human CD6. Cardenas et al. [9] demonstrated that treatment of resting PBL with FCS results in phosphorylation of a fraction of the resting 105-kDa CD6 polypeptide species to form the 130-kDa species. FCS has been shown to induce phosphorylation of

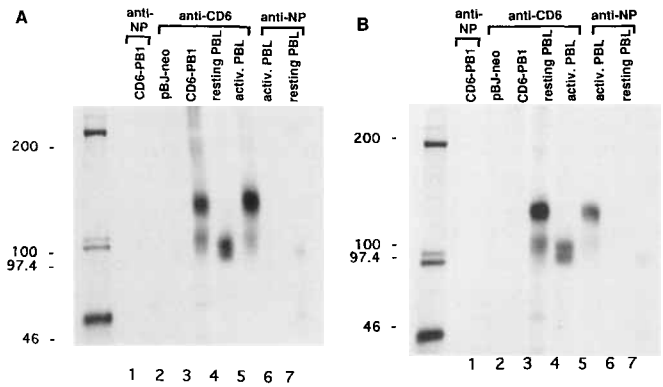

Figure 3. Immunoprecipitation of human CD6 from cells lysates of PBL and L cell transfectants expressing full-length hybrid CD6-PB1 cDNA. L cells transfected with the pBJ-neo expression vector alone (A and B, lane 2), L cells transfected with the hybrid CD6-PB1 cDNA in the pBJ-neo expression vector (A and B, lanes 1 and 3), resting peripheral blood lymphocytes (A and B, lanes 4 and 7), and PBL activated with 20 ng/ml PMA and 10% FCS for 2 h. (A and B, lanes 5 and 6) were surface-labeled with ^{125}I and lysed. The lysates were precleared twice with normal mouse serum before immunoprecipitation with NP-specific isotype-matched control mAb B6-27 (A and B, lanes 1, 6 and 7) and then with CD6-specific mAb (A and B, lanes 2–5). Samples were electrophoresed on 8% polyacrylamide gels containing SDS under reducing (A) or nonreducing (B) conditions. The migration positions of ^{14}C-labeled M_r standards ($\times 10^{-3}$) are shown in the left margin.

distinct proteins [27, 28] as well as to activate PKC in fibroblasts [24]. We have identified three potential PKC phosphorylation sites in the cytoplasmic domain of human CD6, and we cultured the L cell transfectants expressing the hybrid CD6-PB1 molecule in medium containing 10% FCS. It is likely that serum-induced PKC-mediated phosphorylation of CD6 is responsible for the presence of both the 105- and 130-kDa species in our CD6 immunoprecipitates from these transfectants.

Immunoprecipitation analysis of wild-type CD6 from resting human PBL revealed doublet bands of 99 and 103 kDa of equivalent intensity under reducing conditions (Fig. 3 A and B, lane 4). These bands are thought to represent less-phosphorylated forms of human CD6 (relative to the 130-kDa species) present on resting T cells [9]. This doublet could represent expression of two distinct CD6 polypeptide species, differences in post-translational modification of a single CD6 polypeptide chain, or the presence of a degradation product. If this doublet does in fact represent two distinct CD6 polypeptide species, then it is possible that the smaller could represent the polypeptide encoded by CD6-PB2, CD6-PB3, or a postulated mRNA specis containing the full cytoplasmic domain except for the 105-bp deletion in CD5-15.

Agarose gel electrophoresis of the human CD6 RT-PCR products from PBL demonstrated a major band corresponding in size to CD6-PB1 and minor bands that correspond in size to CD6-PB2, CD6-PB3, and potentially CD6-15 with or without the 20-bp deletion (data not shown). The 219 bp missing in CD6-PB3 and the 105 bp missing in the 3′ region of CD6-15 contain serine-threonine-rich motifs, casein kinase-2 phosphorylation site motifs, and a proline-rich motif. If these motifs represent substrates for phosphorylation or are involved in interaction with other proteins, then expression of mRNA corresponding to CD6-PB3, or the postulated mRNA lacking the 3′ 105 bp absent in CD6-15 through alternative splicing, might represent a mechanism by which T cells regulate the signaling function of CD6.

4 Concluding remarks

Human CD6 can provide signals to co-stimulate T cell proliferation and is phosphorylated on tyrosine and serine upon T cell activation [4–7, 9, 10]. We now demonstrate that human CD6 possesses a large cytoplasmic domain of 244 amino acids closely related to the 243-amino acid cytoplasmic domain of mouse CD6, but without significant homology to other known proteins. The newly identified cytoplasmic domain of human CD6 possesses conserved proline-rich motifs that are potential sites for interaction with SH3-domain containing proteins, a YDDI motif that is a potential site for interaction with SH2-domain containing proteins, as well as serine/threonine-rich motifs and both casein kinase-2 and PKC phosphorylation site motifs that are potential substrates for phosphorylation. To understand better the functional role of CD6 in T cells, further studies are necessary to elucidate the cytosolic molecules with which CD6 interacts and through which CD6 signals.

This work was supported by National Institutes of Health grants CA46507 and CA68675 (formerly AI30155) to J.R.P. W.H.R. and S.S.P. received support from training grant AI07290 from the National Institutes of Health. H.E.N.d.V. received support from Digestive Disease Center grant DK38707 from the National Institutes of Health. We thank Dr. B. Seed for clone CD6-15 [13].

Received July 13, 1995; accepted July 28, 1995.

5 References

1 Reinherz, E. L., Meuer, S., Fitzgerald, K. A., Hussey, R. E., Levine, H. and Schlossman, S. F., *Cell* 1982. *30:* 735.
2 Kamoun, M., Kadin, M. E., Martin, P. J., Nettleton, J. and Hansen, J. A. *J. Immunol.* 1981. *127:* 987.
3 Morimoto, C., Rudd, C. E., Letvin, N. L., Hagan, M. and Schlossman, S. F., *J. Immunol.* 1988. *140:* 2165.
4 Walker, C., Bettens, F. and Pichler, W. J., *Eur. J. Immunol.* 1987. *17:* 873.
5 Swack, J. A., Gangemi, R. M., Rudd, C. E., Morimoto, C., Schlossman, S. F. and Romain, P. L., *Mol. Immunol.* 1989. *26:* 1037.
6 Gangemi, R. M., Swack, J. A., Gaviria, D. M. and Romain, P. L., *J. Immunol.* 1989. *143:* 2439.
7 Osorio, L. M., Garcia, C. A., Jondal, M. and Chow, S. C., *Cell. Immunol.* 1994. *154:* 123.
8 Swack, J. A., Mier, J. W., Romain, P. L., Hull, S. R. and Rudd, C. E., *J. Biol. Chem.* 1991. *266:* 7137.
9 Cardenas, L., Carrera, A. C., Yague, E., Pulido, R., Sánchez-Madrid, F. and le Landazuri, M. O., *J. Immunol.* 1990. *145:* 1450.
10 Wee, S. F., Schieven, G. L., Kirihara, J. M., Tsu, T. T., Ledbetter, J. A. and Aruffo, A., *J. Exp. Med.* 1993. *177:* 219.
11 Bowen, M. A., Patel, D. D., Li, X., Modrell, B., Malacko, A. R., Wang, W.-C., Marquardt, H., Neubauer, M., Pesando, J. M., Franke, U., Haynes, B. F. and Aruffo, A., *J. Exp. Med.* 1995. *181:* 2213.
12 Yssel, H., De Vries, J. E., Borst, J. and Spits, H., *Cell. Immunol.* 1987. *105:* 161.
13 Aruffo, A., Melnick, M. B., Linsley, P. S. and Seed B., *J. Exp. Med.* 1991. *174:* 949.
14 Seed, B. and Aruffo, A., *Proc. Natl. Acad. Sci. USA* 1987. *84:* 3365.
15 Whitney, G., Bowen, M., Neubauer, M. and Aruffo, A., *Mol. Immunol.* 1995. *32:* 89.
16 Robinson, W. H., Ying, H., Miceli, M. C. and Parnes, J. R., *J. Immunol.* 1992. *149:* 880.
17 Sanger, F., Nicklen, S. and Coulson, A. R., *Proc. Natl. Acad. Sci. USA* 1977. *74:* 5463.
18 Lin, A. Y., Davaux, B., Green, A., Sagerstrom, C., Elliott, J. F. and Davis, M. M., *Science* 1990. *249:* 677.
19 Zamoyska, R., Vollmer, A. C., Sizer, K. C., Liaw, C. W. and Parnes, J. R., *Cell* 1985. *43:* 153.
20 Boersch-Supan, M. E., Agarwal, A., White-Scharf, M. E. and Imanishi-Kari, T., *J. Exp. Med.* 1985. *161:* 1272.
21 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J., *J. Mol. Biol.* 1990. *215:* 403.
22 Pinna, L. A. *Biochim. Biophys. Acta* 1990. *1054:* 267.
23 Woodgett, J. R., Gould, K. L. and Hunter, T., *Eur. J. Biochem.* 1986. *161:* 177.
24 Rodriguez-Pena, A. and Rozengurt, E., *EMBO J.* 1985. *8:* 149.
25 Songyang, Z., Shoelson, S. E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W. G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R. J., Neel, B. G., Birge, R. B., Fajardo, J. E. Chou, M. M., Hanafusa, J., Schaffhausen, B. and Cantly, L. C., *Cell* 1993. *72:* 767.
26 Feng, S., Chen, J. K., Yu, H., Simon, J. A. and Schreiber, S. L., *Science* 1994. *266:* 1241.
27 Saksela, K., Makela, T. P., Evan, G. and Alitalo, K., *EMBO J.* 1989. *8:* 149.
28 Ohta, Y., Ohba, T., Fukunaga, K. and Miyamoto, E., *J. Biol. Chem.* 1988. *263:* 11540.