Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 8550–8554, November 1990
Biochemistry

Molecular structure of a major insulin/mitogen-activated 70-kDa S6 protein kinase

(growth factor action/protein kinase cascades/regulatory model)

PAPIYA BANERJEE, MIR F. AHMAD, J. RUSSELL GROVE, CARL KOZLOSKY, DANIEL J. PRICE, AND JOSEPH AVRUCH*

Diabetes Unit and Medical Services, Massachusetts General Hospital and Department of Medicine, Harvard Medical School, Boston, MA 02129

Communicated by Raymond L. Erikson, August 2, 1990 (received for review June 22, 1990)

ABSTRACT **The molecular structure of a rat hepatoma 70-kDa insulin/mitogen-stimulated S6 protein kinase, obtained by molecular cloning, is compared to that of a rat homolog of the 85-kDa *Xenopus* S6 protein kinase α; both kinases were cloned from H4 hepatoma cDNA libraries. The 70-kDa S6 kinase (calculated molecular mass of 59,186 Da) exhibits a single catalytic domain that is most closely related in amino acid sequence (56% identity) to the amino-terminal, kinase C-like domain of the rat p85 S6 kinase (calculated molecular mass of 82,695 Da); strong similarity extends through a further 67 residues carboxyl-terminal to the catalytic domain (40% identity), corresponding to a region also conserved among the kinase C family. Outside of this segment of ≈330 amino acids, the structures of the p70 and p85 S6 kinases diverge substantially. The p70 S6 kinase is known to be activated through serine/threonine phosphorylation by unidentified insulin/mitogen-activated protein kinases. A model for the regulation of p70 S6 protein kinase activity is proposed wherein the low activity of the unphosphorylated enzyme results from the binding of a basic, inhibitory pseudosubstrate site (located carboxyl-terminal to the extended catalytic domain) to an acidic substrate binding region (located amino-terminal to the catalytic domain); substrate binding is thereby prevented. S6 kinase activation requires displacement of this inhibitory segment, which is proposed to occur consequent to its multiple phosphorylation. The putative autoinhibitory segment contains several serine and threonine residues, each followed directly by a proline residue. This motif may prevent autophosphorylation but permit transphosphorylation; two of these serine residues reside in a maturation promoting factor (MPF)/cdc-2 consensus motif. Thus, hormonal regulation of S6 kinase may involve the action of MPF/cdc-2 or protein kinases with related substrate specificity.**

Phosphorylation of multiple serine residues on ribosomal protein S6 occurs as a ubiquitous component of the cellular response to many hormones and mitogens (1). This response is mediated by an increase in S6 protein kinase activity, which persists after cell disruption (2, 3). Considerable effort has focused on understanding the biochemical steps by which receptor occupancy leads to S6 kinase activation. Much evidence now indicates that the proximate step in S6 kinase activation, regardless of the initiating stimulus, is serine/threonine-specific phosphorylation of the S6 kinase polypeptide itself mediated by as yet unidentified protein kinases (4, 5).

Characterization of mitogen-stimulated S6 kinase at a molecular level was first achieved in *Xenopus* oocytes, wherein the dominant forms of S6 kinase detected after mitogen stimulation (namely, S6 kinases I and II) were purified as single polypeptide chains of 90–92 kDa (6). Amino acid sequence derived from S6 kinase II was employed to isolate a *Xenopus* cDNA encoding an 85-kDa polypeptide (S6 kinase α), whose amino acid sequence contained two kinase catalytic domains (7). An antiserum to the recombinant *Xenopus* S6 kinase polypeptide is reactive with the mitogen-stimulated S6 kinases (I and II) in *Xenopus* oocytes (4) and also immunoprecipitates mitogen-stimulated S6 kinase activity from avian and mammalian cells (8); an endogenous 85- to 90-kDa ^{32}P-labeled polypeptide, presumably the S6 kinase itself, is precipitated as well. Purified *Xenopus* S6 kinase II can be deactivated with protein phosphatase 2A and partially reactivated by phosphorylation with an insulin-stimulated p42 microtubule-associate protein 2 (MAP-2) protein kinase prepared from 3T3-L1 adipocytes (9).

Most efforts at purification of mitogen-stimulated S6 kinase activity from avian and mammalian sources have recovered 65- to 70-kDa polypeptides (10–12). We purified an hepatic 70-kDa S6 kinase, activated by cycloheximide treatment of the rat; this very low abundance enzyme cochromatographed with the major hepatic S6 kinase activated during liver regeneration and after insulin treatment of H4 hepatoma cells (13). The latter enzyme was also purified extensively from ^{32}P-labeled H4 hepatoma cells and found to correspond to a 70-kDa ^{32}P-labeled polypeptide, which is phosphorylated on serine residues in response to insulin, concomitant with activation (5). Like the *Xenopus* S6 kinase II, the hepatic p70 S6 kinase is also deactivated by protein phosphatase 2A; these two S6 kinases, however, are immunochemically distinct (D.J.P., J.A., E. Erikson, and J. Maller, unpublished results). Moreover, in experiments where the purified *Xenopus* S6 kinase II was phosphorylated and reactivated by an insulin-stimulated p42 MAP-2 (serine/threonine) protein kinase, the 70-kDa rat liver S6 kinase did not serve as a substrate for the p42 MAP-2 kinase.

These results and others suggest that at least two classes of structurally distinct, independently regulated, mitogen-activated S6 protein kinases are present in avian and mammalian cells, corresponding to the 85-kDa and 70-kDa enzymes. In this report, we describe the molecular structure of the p70 S6 kinase obtained by means of molecular cloning and compare its structure to a cloned rat p85 S6 kinase†. A model for the regulation of the p70 S6 kinase activity is proposed.

MATERIALS AND METHODS

Protein Sequence. Two preparations of 70-kDa S6 kinase, purified from the livers of cycloheximide-treated rats through

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

Abbreviations: PCR, polymerase chain reaction; MPF, maturation promoting factor; nt, nucleotide(s); MAP-2, microtubule-associated protein 2.

*To whom reprint requests should be addressed at: Diabetes Unit, Massachusetts General Hospital East, Building 149, 13th Street, Charlestown, MA 02129.

†The sequence reported in this paper has been deposited in the GenBank data base (accession no. M37777).

Mono Q chromatography (about 30–40 μg each) (12), were autophosphorylated with [γ-^{32}P]ATP and subjected to SDS/PAGE. The ^{32}P-labeled 70-kDa polypeptide was electro-eluted, reduced, *S*-carboxymethylated, and subjected to complete tryptic digestion or treatment with excess CNBr. The peptides were applied to a Brownlee C_8 (2.1 mm × 3.0 cm) microbore column equilibrated with NH_4HCO_3 (≈20 mM) at pH 6.0 and eluted with a gradient of acetonitrile. Major peaks detected at OD_{214} were acidified, reduced in volume, reinjected on a second C_8 column equilibrated with 0.1% trifluoroacetic acid, and eluted with acetonitrile. Automated Edman degradation was performed on an Applied Biosystems 470A gas-phase sequencer. Partial amino acid sequence was obtained on five tryptic peptides and five CNBr peptides (Fig. 1).

cDNA Library Construction. Total RNA was isolated from rat hepatoma H4 cells by a modified guanidinium thiocyanate method, and poly(A)$^+$ RNA was selected. Double-stranded cDNA was synthesized by using the cDNA synthesis kit from Bethesda Research Laboratories; the primers for first-strand cDNA synthesis included oligo(dT), random hexanucleotides (14) or, in certain instances, specific sequences of DNA. After blunting of the cDNA ends with T4 DNA polymerase and ligation of appropriate adaptors or *Eco*RI linkers, cDNA libraries were constructed either in the λgt11 (Promega) or in the λZap II (Stratagene) vectors.

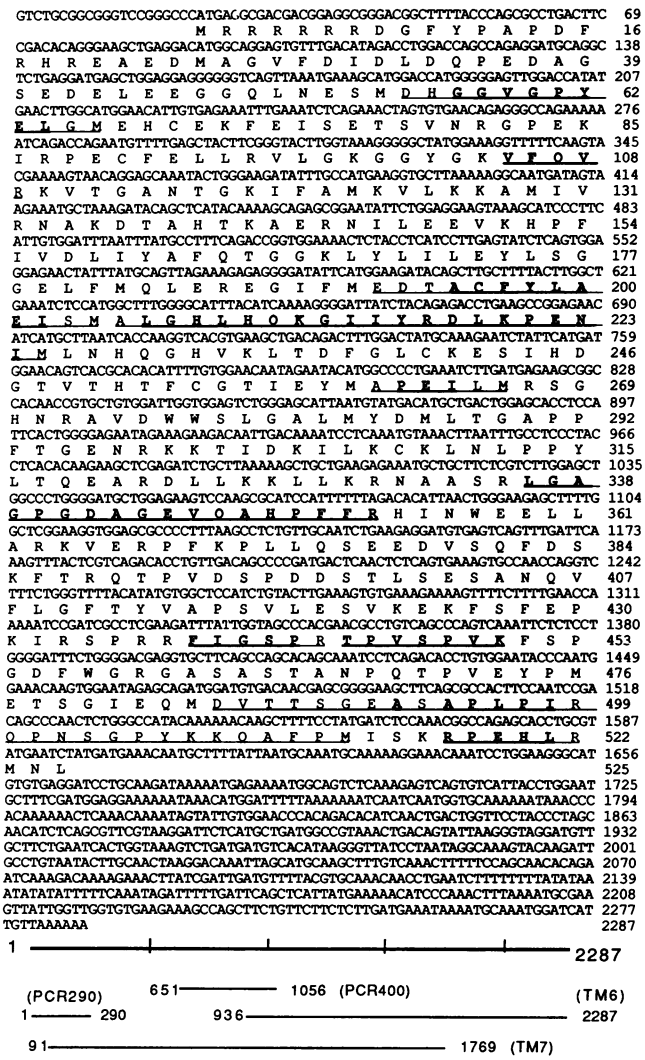


FIG. 1. DNA sequence and deduced amino acid sequence of the p70 S6 kinase. Each tryptic and CNBr peptide actually sequenced is underlined. Amino acids were identified at the residues shown in boldface type and were identical to the deduced amino acids at 81 of 83 residues; the two differences (Ser for Cys-196 and Thr for Arg-517) appear to reflect errors in amino acid sequence identification. The overlapping cDNA clones from which the sequence was derived are drawn to scale. The numbers indicate the termini of the clones.

cDNA Screening Strategies. The sequence of tryptic peptide 2 and CNBr peptides 2 and 3 (numbered from the deduced amino terminus) aligned easily in the protein kinase catalytic domain homology segments designated XI and VI by Hanks *et al.* (15) (see Fig. 1). Synthetic oligonucleotides [17 base pairs (bp)] designed from the sequence of these peptides were employed as polymerase chain reaction (PCR) primers using deoxyinosine substitution to reduce degeneracy (P.B. and J.A., unpublished results). Specifically primed or randomly primed (hexanucleotide) (14) first-strand cDNA served as template; two rounds of amplification with nested primers yielded a dominant 406-bp DNA fragment (PCR400, Fig. 1). Subcloning and sequencing indicated that the correct amino acid sequences internal to the primers were present, and the motifs expected for a protein kinase catalytic domain were evident. This 406-bp insert was then employed to screen an oligo(dT)-primed cDNA library; a single positive clone, designated TM6, was identified. A second oligo(dT)-primed cDNA library was screened independently with the *Eco*RI insert from TM6. Ten positive plaques were purified; the longest (TM7) was 1.6 kilobases. DNA sequences extending 5′ to TM7 were isolated by PCR, employing as template either H4 hepatoma total RNA or a randomly primed H4 hepatoma cDNA library in λgt11. Using the cDNA template, the first PCR employed the λgt11 forward (λ_F) and reverse (λ_R) 24-mer primers and an oligonucleotide (ISET) based on nucleotides (nt) 229–242 (Fig. 1) in the antisense orientation. After 40 cycles of PCR, a dominant product of ≈320 bp was observed with λ_F but not with λ_R. A second round of PCR employed λ_F and an internal oligonucleotide (PGVG) corresponding to nt 181–204 (Fig. 1) in the antisense orientation. A major band of ≈290 bp (PCR290, Fig. 1) was subcloned in pBS (+/−); colonies were selected by hybridization to the *Eco*RI insert from TM7. DNA sequence analysis of several positive clones demonstrated the presence of the λ_F primer, adjacent λ DNA sequences, an *Eco*RI cloning site, and 90 bp of new sequences, followed by sequences corresponding to the 5′ end of TM7 extending to the site of PGVG priming. The same 90 bp of additional S6 kinase sequence were obtained by anchored PCR from H4 total RNA (data not shown).

The cDNA corresponding to the rat liver 85-kDa S6 kinase homolog of the *Xenopus* S6 kinase α was isolated using the *Xenopus* S6 kinase α cDNA (a gift from S. Jones and R. Erikson, Harvard University) as probe for screening H4 hepatoma λgt11 cDNA libraries at low stringency; details of the cDNA cloning and structure will be presented separately.

DNA sequencing was performed by the dideoxynucleotide chain-termination method as described by Chen and Seeburg (16). The sequence data were obtained from a combination of specific oligonucleotide priming and undirectional-nested deletions (Erase-a-Base system; Promega) with Sp6 or T7 primers. All sequences presented were determined at least twice in each strand. The University of Wisconsin Genetics Computer Group Programs, version 6.1, were used for sequence assembly analysis and manipulation (17).

RESULTS AND DISCUSSION

Molecular Structure of the 70-kDa S6 Protein Kinase. The nucleotide sequence of the H4 hepatoma cDNAs encoding the 70-kDa S6 kinase is shown in Fig. 1. The sequence encompasses 2287 bp; the first ATG starts at nt 22 and is followed by an open reading frame encoding 525 amino acids (calculated molecular mass of 59,186 Da), terminating with a tandem pair of TGA triplets (starting at nt 1597). A poly-

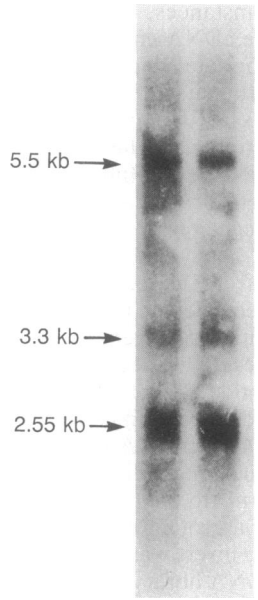


FIG. 2. Northern blot analysis of poly(A)$^+$ mRNAs from H4 hepatoma cells. The RNAs were denatured, electrophoresed, transferred to nitrocellulose filters, and hybridized with the insert from PCR400, which was ^{32}P-labeled by random priming (14).

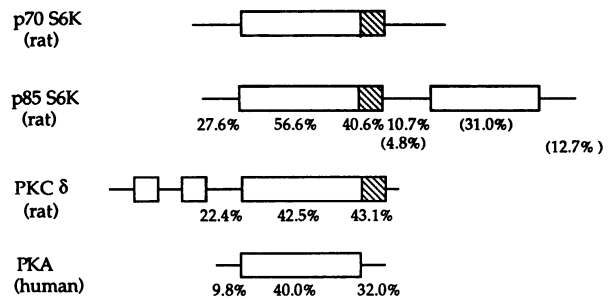
adenylylation signal (AAATAA, starting at nt 2259) is followed by a poly(A) tail starting 27 bp downstream. The protein encoded in the open reading frame contains amino acid sequences identical to all 10 of the tryptic and CNBr peptides analyzed (see Fig. 1). Thus, the cDNA encodes the protein corresponding to the 70-kDa S6 kinase purified from the livers of cycloheximide-treated rats. We have shown previously that this rat liver kinase is very closely related and probably identical to the insulin-stimulated 70-kDa S6 kinase from H4 hepatoma cells, based on similarities in size, tryptic maps after autophosphorylation, regulation, and immunochemical crossreactivity (5). H4 hepatoma poly(A)$^+$ RNA exhibited mRNA species of 2.55, 3.3, and 5.5 kb (Fig. 2).

The 5′-most AUG (starting at nt 22) may not be the sole or preferred translational start site.‡ The DNA sequence preceding this AUG lacks a purine at −3, whereas a purine at −3 is present preceding the next AUG (starting at nt 91). Such an arrangement may allow initiation to occur at the latter AUG (19). Our usual isolates of rat liver S6 kinase do contain two (occasionally more) major polypeptide bands between 65 and 70 kDa whose apparent molecular masses can be reduced to 60–62 kDa by treatment with phosphatase 2A. Given the susceptibility of the enzyme to proteolytic attack, we have been reluctant to conclude that this size heterogeneity reflects an intrinsic feature of the enzyme polypeptide, as opposed to the result of natural or artifactual posttranslational modifications; coupled with the present data, however, the possibility of multiple start sites merits further consideration. Attempts to identify the amino terminus of the rat liver 70-kDa S6 kinase directly by Edman degradation were unsuccessful, presumably due to blockage by an unidentified modifying group. The most amino-terminal peptide sequence obtained corresponds to the CNBr fragment following the third methionine in the coding region. Thus, the present data do not permit unambiguous assignment of the translational start site.

‡The sequence shown in Fig. 1 has been compared to that of a cDNA encoding a closely related 70-kDa S6 kinase of rat origin (18). Aside from sequencing errors now corrected, the two sequences are identical 3′ to nt 91 through the coding region but are completely divergent 5′ to nt 91.

Organization of the 70-kDa S6 Kinase and Homology with Other Protein Kinases: The Catalytic Domain. The sequence of the 70-kDa S6 kinase reveals a single, centrally located protein kinase catalytic domain of 267 residues (from Glu-89 to Ile-355) flanked by an amino-terminal segment of 88-amino acid residues (65 if the second AUG is employed) and a carboxyl-terminal segment of 170 residues (Figs. 1 and 3).

The catalytic domain exhibits a consensus ATP-binding site (GKGGYG) followed by a lysine residue 20 residues thereafter. Other characteristic motifs include YRDLKPEN, DFG, and APE, corresponding to domains VI, VII, and VIII of Hanks *et al.* (15). The catalytic domain of the 70-kDa S6 kinase exhibits 57% identity to the more amino-terminal "kinase C-like" catalytic domain of the p85 S6 kinase,




```
p70 S6K    MRRRRRRD GFYPAPDFRH REAEDMAGVF DIDLDQPEDA GSEDELEEGG   48
p85 S6K          MPLAQL KEPWPLMELV PLDPENGQAS G......EEA   30
PKC     ...QGLKCED CGMNVHHKCR EKVANLCGIN QKLLAEALNQ VTQKASRKPE  303

p70 S6K  QLNESMDHGG VGPYELGMEH CEKFEISET. SVNRGPEKIR PECFELLRVL   97
p85 S6K  GLQPSKDEGI LK..... ....EISITH HVKAGSEKAD PSHFELLKVL   68
PKC      TPETVGIYQG FEKKTAVSGN DIPDNNGTYG KIWEGSNRCR LENFTFQKVL  353
PKA      MGNAAAAKKG SEQESVKEFL AKAKEDFLKK WESPAQNTAH LDQFERIKTL   50

p70 S6K  GKGGYGKVFQ VRKVTGANTG KIFAMKVLKK AMIVRNAKDT AHTKAERNIL  147
p85 S6K  GQGSFGKVFL VRKVTRPDNG HLYAMKVLKK ATL..KVRDR VRTKMERDIL  116
PKC      GKGSFGKVLL ...AELKGKE RYFAIKYLKK DVVLIDDDVE CTMVEKRVLA  400
PKA      GTGSFGRVML V.KH..KETG NHYAMKILDK QKVVK.LKQI EHTLNEKRIL   96

p70 S6K  EEVKHPFIVD LIYAFQTGGK LYLILEYLSG GELFMQLERE GIFMEDTACF  197
p85 S6K  ADVNHPFVVK LHYAFQTEGK LYLILDFLRG GDLFTRLSKE VMFTEEDVKF  166
PKC      LAWENPFLTH LICTFQTKDH LFFVMEFLNG GDLMFHIQDK GRFELYRATF  450
PKA      QAVNFPFLVK LEFSFKDNSN LYMVMEYVPG GEMFSHLRRI GRFSEPHARF  146

p70 S6K  YLAEISMALG HLHQKGIIYR DLKPENIMLN HQGHVKLTDF GLCKESIHDG  247
p85 S6K  YLAELALGLD HLHSLGIIYR DLKPENILLD EEGHIKLTDF GLSKEAIDHE  216
PKC      YAAEIICGLQ FLHGKGIIYR DLKLDNVMLD KDGHIKIADF GMCKENIFGE  500
PKA      YAAQIVLTFE YLHSLDLIYR DLKPENLLID QQGYIQVTDF GFAK...RVK  193

p70 S6K  TVTHTFCGTI EYMAPEILMR SGHNRAVDWW SLGALMYDML TGAPPFTGEN  297
p85 S6K  KKAYSFCGTV EYMAPEVVNR QGHTHSADWW SYGVLMFEML TGSLPFQGKD  266
PKC      NRASTFCGTP DYIAPEILQG LKYSFSVDWW SFGVLLYEML IGQSPFHGDD  550
PKA      GRTWTLCGTP EYLAPEIILS KGYNKAVDWW ALGVLIYEMA AGYPPFFADQ  243

p70 S6K  RKKTIDKILK CKLNLPPYLT QEARDLLKKL LKRNAASRLG AGPGDAGEVQ  347
p85 S6K  RKETMTLILK AKLGMPQFLS TEAQSLLRAL FKRNPANRLG SGPDGAEEIK  316
PKC      EDELFESIRV DTPHYPRWIT KESKDIMEKL FERDPAKRLG V....TGNIR  596
PKA      PIQIYEKIVS GKVRFPSHFS SDLKDLLRNL LQVDLTKRFG NLKNGVNDIK  293

p70 S6K  AHPFFRHINW EELLARKVER PFKPLLQSEE DVSQFDSKFT RQTPVDSPDD  397
p85 S6K  RHIFYSTIDW NKLYRREIKP PFKPAVAQPD DTFYFDTEFT SRTPRDSPGI  366
PKC      LHPFFKTINW NLLEKRKVEP PFKPKVKSPS DYSNFDPEFL NEKPQLSFSD  646
PKA      NHKWFATTDW IAIYQRKVEA PFIPKFKGPG DTSNFDDYEE EEIRVSINEK  343

p70 S6K  STLSESANQV .FLGFTYVAP SVLESVKEKF SFEPKIRSPR RFIGSPRTPV  446
p85 S6K  PP.SAGAHQL .FRGFSFVAT GLMEDDSKPR ATQAPLHSVV QQLHGKNLVF  414
PKC      KNLIDSMDQT AFKGFSFVNP K.YEQFLE                        673
PKA      CGKEFSEF                                              351

p70 S6K  SPVKFSPGDF WGRGASASTA NPQTPVEYPM ETSGIEQMDV TTSGEASAPL  496
p85 S6K  SDGYIVKETI GVGSYSVCKR CVHKATNMEY AVKVIDKSKR DPSEEIEILL  464

p70 S6K  PIRQPNSGPY KKQAFPMISK RPEHLRMNL                        525
p85 S6K  RYGQHPNIIT LKDVYD.DS. KHVYLVTELM RGGELLDK...           499
```

FIG. 3. Comparison of p70 S6 kinase with p85 S6 kinase, PKC δ, and the catalytic subunit of the cAMP-dependent protein kinase. (*Upper*) Schematic structure of p70 S6 kinase (K), comparing the percent identity in the amino-terminal segment (residues 1–88), catalytic domain (open bar, residues 89–355), catalytic domain extension (hatched bar, residues 356–420), and carboxyl-terminal tail (residues 421–525) to the corresponding segments in the rat p85 S6 kinase, rat kinase C δ, and human kinase A catalytic subunit. The numbers in parentheses represent the identities observed for the corresponding segments after alignment of the more carboxyl-terminal catalytic domain of p85 S6 kinase with the catalytic domain of the p70 S6 kinase. (*Lower*) The deduced amino acid sequences of the four protein kinases depicted in *Upper* are aligned. The brackets demarcate the domains described in *Upper*.

whether compared to the mouse p85 sequence reported by Alcorta *et al.* (20) or the H4 hepatoma p85 sequence that we isolated (Fig. 3). The mouse and rat p85 sequences are 97% identical to each other over their entire length, and the rat p85 sequence is 84% identical to the *Xenopus* p85 S6 kinase α. The catalytic domains next most similar to that of the p70 S6 kinase are those of the kinase C family (≈40–44% identity), followed closely by the cGMP- and cAMP-dependent protein kinases (≈40% identity with each).

Noncatalytic Domains of the 70-kDa S6 Protein Kinase. The first methionine is followed by six consecutive arginine residues. Comparable polybasic sequences are found in a variety of proteins that bind to polynucleotides, such as histones, protamines, viral proteins, etc; rat liver S6 itself contains several runs of four or five consecutive basic residues (21). This very basic amino-terminal segment of the p70 may interact with ribosomal RNA and contribute to the association of the kinase with 40S subunits. In contrast, the aspartic acid at residue 25 is followed by a highly acidic segment, wherein 10 out of 18 residues are aspartic acid or glutamic acid and basic residues are absent; a role for this segment in substrate binding will be considered below. The presence of these two oppositely charged segments in the amino-terminal region may underlie the ability of the p70 S6 kinase to bind avidly to both anion and cation-exchange matrices.

In regard to the region carboxyl-terminal to the catalytic domain, the homology between the p70, the p85 S6 kinase, and kinase C is maintained for ≈67 residues beyond the p70 S6 kinase catalytic domain through Glu-420; the corresponding segments in the p85 S6 kinase and the kinase C δ isoform are 40% and 43% identical to this region of the p70 S6 kinase. Thereafter, the sequences of these three kinases diverge. This point of divergence between p70 and p85 also represents the carboxyl terminus of the C_3 conserved domain of the kinase C family (22), which in that kinase family is then followed by a short segment of variable structure (e.g., four amino acid residues in kinase C δ), which is the protein kinase C carboxyl terminus. In contrast to kinase C, which terminates shortly after this conserved (C_3) segment, the p85 S6 kinase extends beyond the region homologous to kinase C to encode a second complete kinase catalytic domain, most closely related to the phosphorylase b kinase γ subunit. The p70 S6 kinase, in turn, extends from this point of sequence divergence with a unique 105-amino acid structure, which ends at the protein carboxyl terminus. This 105-amino acid segment is only 10.7% identical to the corresponding segment of the p85 S6 kinase. Comparison of this 105-amino acid carboxyl-terminal tail of the p70 S6 kinase with the National Biomedical Research Foundation protein data bank (release 21.0, June 1989) yielded no homologous sequences; nevertheless, inspection of this carboxyl-terminal tail revealed several interesting features. The 25-amino acid residues starting at Lys-423, just carboxyl-terminal to the kinase C homology domain, constitute a moderately basic (seven arginine/lysine and two glutamic acid) serine/threonine-rich (five residues) segment whose composition (except for the presence of four proline residues) is reminiscent of the carboxyl-terminal region of rat liver S6 itself (21), which is the region known to encompass all the S6 kinase-directed phosphorylation sites (23) (Fig. 4).

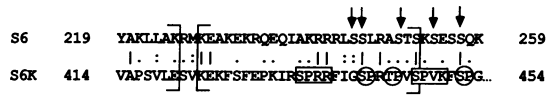


FIG. 4. Comparison of sequences in rat liver S6 to rat liver S6 kinase (K). The carboxyl-terminal 41 amino acid residues of rat liver S6 (21), which encompass all the reported (23) sites of S6 phosphorylation (indicated by arrows), are aligned with sequences from the carboxyl-terminal tail of rat liver p70 S6 kinase. Sites of potential phosphorylation of S6 kinase by proline-directed kinases are enclosed by circles or boxes; the boxes enclose maturation promoting factor (MPF)/cdc-2 consensus recognition motifs (24). The leftmost bracket indicates the end of the catalytic domain extension (see Fig. 3 and text). The p70 S6 kinase sequence enclosed in brackets is the putative pseudosubstrate site. Identical amino acids are indicated by vertical lines; dots are employed to indicate amino acids of increasing similarity, scored by the mutational difference matrix of Schwartz and Dayhoff (25) as provided in ref. 17. Identities score 1.5, one dot is ≥0.1, and two dots are ≥0.5.

The determinants of the substrate specificity of the S6 kinase have not yet been identified. The most amino-terminal of the five or more serine phosphorylation sites and probably the site of initial S6 phosphorylation is contained in the palindromic sequence RRLSSLR (23). We have found that synthetic peptides whose amino terminus begins with the palindromic sequence RRLSSLR are poorly phosphorylated by the 70-kDa S6 kinase, even if they are extended to contain the entire carboxyl-terminal S6 sequence. In contrast, peptides whose amino terminus extends back 14 residues, corresponding to sequence of the carboxyl-terminal CNBr fragment of rat liver S6, are high-affinity substrates ($K_m \approx 10\ \mu M$) for the 70-kDa S6 kinase (12). We therefore infer that these 14 residues, together with the palindromic sequence that follows, contain the elements in S6 critical for its high-affinity interaction with the p70 S6 kinase. The carboxyl-terminal segment of S6 can be aligned with the basic region of the p70 S6 kinase between Lys-423 and Ser-447, with an identity of 28% over these 25 residues (Fig. 4); on this basis, we speculate that this segment of p70 S6 kinase represents a pseudosubstrate site.

Proposed Mechanism for the Regulation of the 70-kDa S6 Protein Kinase. This pseudosubstrate site is envisioned as a domain that binds to and occludes the substrate-binding site and maintains the kinase in its basal-inhibited state (26). A candidate for the complementary substrate-binding site is the highly acidic segment located just amino-terminal to the catalytic domain, from Asp-29 through Glu-46. This segment could provide a suitable binding site for the highly basic ribosomal S6. Thus, in the basal-inhibited state of the p70 S6 kinase, the basically charged pseudosubstrate site, situated in the carboxyl-terminal tail of the enzyme, would bind to this acidic site, folding over and obstructing the catalytic domain, including the ATP-binding site (Fig. 5). Consistent with the

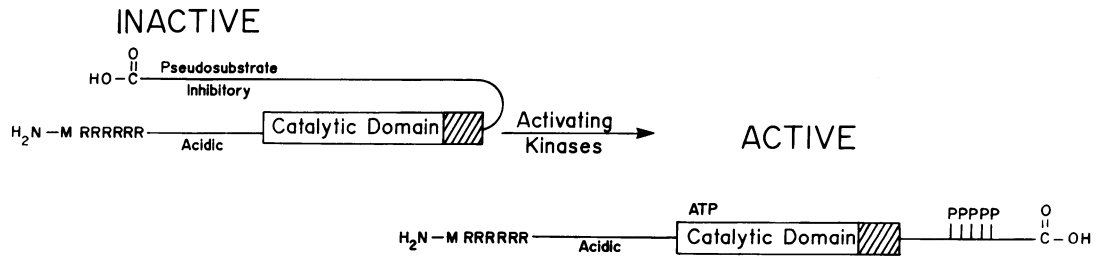


FIG. 5. A model for the regulation of p70 S6 kinase activity. See text for details.

occlusion of the ATP-binding site in the basal state is the observation that deactivation/dephosphorylation of the activated p70 S6 kinase with phosphatase 2A is accompanied by a parallel loss of autophosphorylation as well as substrate phosphorylation (5).

The p70 S6 kinase is known to be activated through phosphorylation of serine/threonine (predominantly serine) residues on the enzyme polypeptide catalyzed by previously unidentified serine/threonine protein kinases. The model proposed immediately suggests a molecular mechanism for disinhibition/activation of the p70 S6 kinase. Serine/threonine phosphorylation of p70 S6 kinase within or near the carboxyl-terminal basic pseudosubstrate segment would weaken the electrostatic binding of this inhibitory segment to the amino-terminal acidic domain and permit access of ATP and S6 to their respective binding sites. Inspection of the amino acid sequence context surrounding the serine/threonine residues located within and nearby the region of this putative pseudosubstrate site reveals the presence of a pair of virtually tandem MPF/cdc-2 consensus phosphorylation sites (24) as RSPRR and VSPVK (Fig. 4). Moreover, immediately contiguous to these sites within the putative pseudosubstrate segment are two more serine and one threonine residues, each also followed immediately by a proline residue. Such a motif, which is found at three further sites in the carboxyl-terminal 170 amino acids of the p70 S6 kinase (eight serine-proline/threonine-proline doublets altogether) may also be compatible with phosphorylation by MPF/cdc-2 (27) or kinases with related substrate specificity (28, 29).

This model of the regulation of the p70 S6 kinase differs from the usual inhibitory pseudosubstrate concept in that autoactivation by means of autophosphorylation does not occur in the basal state despite the presence of serine/threonine residues in the inhibitory segment. In fact, autophosphorylation is only observed with the activated enzyme and occurs primarily on tryptic peptides that are largely distinct from those phosphorylated in the insulin-stimulated cell (5). The inability of the basal enzyme to catalyze an activation through autophosphorylation is probably due to the presence of proline residues immediately carboxyl-terminal to the serine/threonine in the inhibitory segment, which may serve as a negative determinant of S6 kinase substrate specificity. This arrangement permits activation to be catalyzed by protein kinases that *require* a proline carboxyl-terminal to the serine/threonine, such as MPF/cdc-2. We have observed that MPF catalyzes the phosphorylation of the p70 S6 kinase on tryptic peptides that comigrate with a subset of the ^{32}P-labeled tryptic peptides found in ^{32}P-labeled p70 S6 kinase purified from insulin-stimulated, ^{32}P-labeled H4 hepatoma cells; a modest (25–100%) activation of S6 kinase activity is observed concomitantly (D.J.P., J.A., E. Erikson, and J. Maller, unpublished results). This finding suggests that MPF/cdc-2 itself or protein kinases with overlapping substrate specificity (alone or in concert with other protein kinases) may catalyze the phosphorylation and activation of p70 S6 kinase under insulin stimulation.

The amino acid sequence of the rat liver p85 S6 kinase does not exhibit the $\frac{S}{T}$PX$\frac{R}{K}$ motif apparently preferred by MPF/cdc-2. Nevertheless, the p85 S6 kinase does contain one serine-proline and four threonine-proline doublets, which are conserved from *Xenopus* to rat, and are located between the first and second catalytic domains (at Thr-359 and Ser-363) and toward the carboxyl terminus of the p85 polypeptide (Thr-612, Thr-621, and Thr-711). The role of these residues vis-a-vis the phosphorylation and activation of the p85 *Xenopus* S6 kinase is not known; however, it is of interest that activation of the p85 *Xenopus* S6 kinase II by p42 MAP-2 kinase involves phosphorylation of predominantly threonine residues on the p85 S6 kinase (9).

The widespread interest in the S6 kinases is due primarily to a desire to understand the mechanism of their regulation by mitogens and hormones. Both p70 and p85 S6 kinases exist in a virtually inactive form in unstimulated cells and are activated through serine/threonine-specific phosphorylation of the enzyme polypeptides. Thus, the relevant questions are the nature of the structure stabilizing the basal, inactive state, the mechanism by which phosphorylation activates the enzyme, and the identity of the serine/threonine protein kinases that mediate the mitogenic/hormonal activation/phosphorylation. The model proposed for the regulation of the p70 S6 protein kinase presents numerous testable features. The major significance of the model is that it offers strong clues to the identity of the mitogen-activated kinases that mediate S6 kinase phosphorylation and directs our attention to the class of serine-proline/threonine-proline-directed kinases as potential mediators of hormone and mitogen regulation of enzyme activity as well as DNA transcription and the cytoarchitecture.

We thank James Woodgett for stimulating discussions. Martha Chambers is acknowledged for excellent editorial and secretarial assistance. This work was supported in part by National Institutes of Health Grant DK17776.

1. Gordon, J., Nielsen, P. J., Manchester, K. L., Towbin, H., Jimenez de Asua, L. & Thomas, G. (1982) *Curr. Top. Cell. Regul.* **21**, 89–99.
2. Smith, C. J., Rubin, C. S. & Rosen, O. M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 2641–2645.
3. Novak-Hofer, I. & Thomas, G. (1984) *J. Biol. Chem.* **259**, 5995–6000.
4. Erikson, E. & Maller, J. L. (1989) *J. Biol. Chem.* **264**, 13711–13717.
5. Price, D. J., Gunsalus, J. R. & Avruch, J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 7944–7948.
6. Erikson, E. & Maller, J. L. (1986) *J. Biol. Chem.* **261**, 350–355.
7. Jones, S. W., Erikson, E., Blenis, J., Maller, J. L. & Erikson, R. L. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 3377–3381.
8. Sweet, L. J., Alcorta, D. A. & Erikson, R. L. (1990) *Mol. Cell. Biol.* **10**, 2787–2792.
9. Sturgill, T. W., Ray, L. B., Erikson, E. & Maller, J. L. (1988) *Nature (London)* **334**, 715–718.
10. Blenis, J., Kuo, C. J. & Erikson, R. L. (1987) *J. Biol. Chem.* **262**, 14373–14376.
11. Jeno, P., Ballou, L. M., Novak-Hofer, I. & Thomas, G. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 406–410.
12. Price, D. J., Nemenoff, R. A. & Avruch, J. (1989) *J. Biol. Chem.* **264**, 13825–13833.
13. Nemenoff, R. A., Price, D. J., Mendelsohn, M. J., Carter, E. A. & Avruch, J. (1988) *J. Biol. Chem.* **263**, 19455–19560.
14. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13.
15. Hanks, S. K., Quinn, M. A. & Hunter, T. (1988) *Science* **241**, 42–52.
16. Chen, E. Y. & Seeburg, P. H. (1985) *DNA* **4**, 165–170.
17. Devereux, H., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–407.
18. Kozma, S. C., Ferrari, S., Bassand, P., Siegmann, M., Totty, N. & Thomas, G. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 7365–7369.
19. Kozak, M. (1987) *Nucleic Acids Res.* **15**, 8125–8132.
20. Alcorta, D. A., Crews, C. M., Sweet, L. J., Bankston, L., Jones, S. W. & Erikson, R. L. (1989) *Mol. Cell. Biol.* **9**, 3850–3859.
21. Chan, Y. L. & Wool, I. G. (1988) *J. Biol. Chem.* **263**, 2891–2896.
22. Ono, Y., Fujii, T., Ogita, K., Kikkawa, U., Igarashi, I. K. & Nishizuka, Y. (1988) *J. Biol. Chem.* **263**, 6927–6932.
23. Krieg, J., Hofstennge, J. & Thomas, G. (1988) *J. Biol. Chem.* **263**, 11473–11477.
24. Maller, J. L. (1990) *Biochemistry* **29**, 3157–3166.
25. Schwartz, R. M. & Dayhoff, M. O. (1979) in *Atlas of Protein Sequence and Structure*, ed. Dayhoff, M. O. (Natl. Biomed. Res. Found., Washington, DC), pp. 353–358.
26. Soderling, T. R. (1990) *J. Biol. Chem.* **265**, 1823–1826.
27. Cisek, L. J. & Corden, J. L. (1989) *Nature (London)* **339**, 679–684.
28. Vulliet, P. R., Hall, F. L., Mitchell, J. P. & Hardie, D. G. (1989) *J. Biol. Chem.* **264**, 16292–16298.
29. Countaway, J. L., Northwood, I. C. & Davis, R. J. (1989) *J. Biol. Chem.* **264**, 10828–10835.