p59OASL, a 2'-5' oligoadenylate synthetase like protein: a novel human gene related to the 2'-5' oligoadenylate synthetase family

Rune Hartmann, Henrik Steen Olsen, Stefanie Widder, René Jørgensen and Just Justesen*

Department of Molecular and Structural Biology, University of Aarhus, C. F. Møllers allé Building 130, DK-8000 Aarhus C, Denmark and ¹Human Genome Sciences Inc., 9410 Key West Avenue, Rockville, MD 20850, USA

Received June 22, 1998; Revised and Accepted July 29, 1998

DDBJ/EMBL/GenBank accession nos AJ225089 and AJ225090

ABSTRACT

The 2'-5' oligoadenylate synthetases form a well conserved family of interferon induced proteins, presumably present throughout the mammalian class. Using the Expressed Sequence Tag databases, we have identified a novel member of this family. This protein, which we named p59 2'-5' oligoadenylate synthetase-like protein (p59OASL), shares a highly conserved N-terminal domain with the known forms of 2'-5' oligoadenylate synthetases, but differs completely in its C-terminal part. The C-terminus of p59OASL is formed of two domains of ubiquitin-like sequences. Here we present the characterisation of a full-length cDNA clone, the genomic sequence and the expression pattern of this gene. We have addressed the evolution of the 2'-5'oligoadenylate synthetase gene family, in the light of both this new member and new 2'-5' oligoadenylate synthetase sequence data from other species, which have recently appeared in the databases.

INTRODUCTION

The 2'-5' oligoadenylate (2–5A) system is a regulated RNA decay pathway, consisting of a number of 2'-5' oligoadenylate synthetases (2–5A synthetases) and a 2–5A activated ribonuclease, normally referred to as RNase L. The 2–5As were originally described as low molecular weight inhibitors of protein synthesis, which were produced in cell free extracts of interferon treated cells, especially following incubation with Poly I-Poly C (1,2). The ability of 2–5As to bind to and activate RNase L, resulting in a general RNA degradation, makes them potent inhibitors of protein synthesis (3,4).

The 2–5As are produced by a family of enzymes, the 2'-5' oligoadenylate synthetases (2-5A synthetase or OAS, EC 2.7.7.-). Hitherto, two human genes encoding highly homologous 2–5A synthetases have been described, referred to as the p42OAS gene and the p69OAS gene. Both genes are situated on chromosome 12 (5). The p42OAS gene encodes two splice variants, which differ only at their C-termini, having a molecular weight of 42 and 46 kDa (6,7). Also the p69OAS gene encodes two splice variants differing at their C-termini, having theoretical molecular weights

of 79 and 83 kDa, but migrating with apparent molecular weights of 69 and 71 kDa, respectively (8). Most likely, the p69OAS gene arose from a duplication of an ancestral 2–5A synthetase gene, since it harbours two repeated regions, each with ~40% homology to the p42OAS form (8). Furthermore, immunological evidence exists for a 2–5A synthetase with an apparent molecular weight of 100 kDa (9), which has also been found by purification procedures (10).

The biosynthesis of all the known 2–5A synthetases is stimulated by interferon (reviewed in 11). 2–5A synthetase is produced as a latent enzyme which is activated by certain classes of RNA, mainly double stranded RNA (dsRNA) (12). Several studies have dealt with stimulation of semipurified 2–5A synthetases by Poly I-Poly C and analogues thereof (reviewed in 13). Recently, we addressed the structural requirements of the RNA activator of pure 2–5A synthetase (p46 isoform) (12).

The involvement of the 2-5A system in the interferon induced antiviral effect has been inferred from several sets of experiments. For instance, over-expression of the p42OAS in Chinese hamster ovary (CHO) cells gives resistance to infection by picornavirus, but not to vesicular stomatitis virus (VSV) (14). In addition, transgenic plants expressing 2–5A synthetase as well as RNase L show a remarkable increase in their resistance to viral infections (15,16). The 2–5A system has also been suggested to play a role in the growth suppression engendered by interferons (17). Expression of a dominant negative RNAse L mutant in murine cells resulted in suppression of both the interferon mediated protection against picornavirus and the antiproliferative effect otherwise exerted by interferon (18). Recent evidence from RNase L deficient mice strongly suggests that apoptosis can be induced via the 2-5A system. The RNase L -/- mice had an enlarged thymus, caused by the lack of normal apoptosis of thymocytes and fibroblasts in the thymus (19).

In general, RNase L is expressed in all tissues and only weakly induced by interferon, (reviewed in 20), therefore the synthesis of 2–5As seems to be the controlling step in inducing RNA decay by the 2–5A pathway. The different isoforms of 2–5A synthetase apparently have similar enzymatic capacities, but the isoforms seem to have different tissue specific expression patterns and to respond differently to treatment with various cytokines including interferons (reviewed in 11).

The recent development of Expressed Sequence Tag (EST) databases and the increasing complexity of these databases

*To whom correspondence should be addressed. Tel: +45 89422682; Fax: +45 89422637; Email: jj@mbio.aau.dk

(reviewed in 21) gives an opportunity for the identification of new members of a given protein family. We therefore searched the EST database for sequences displaying homology to 2–5A synthetase. In the following, we thus present the cloning and characterisation of a novel member of the human 2–5A synthetase family and the corresponding gene.

MATERIALS AND METHODS

Cloning and sequencing

An EST clone (HG# 1173451) encoding a putative novel 2–5A syntethase was identified in an EST database by searching with the known human 2–5A family members P69OAS (M87284) and p42OAS (P00973) using the BLAST search algorithm (22). This clone was obtained from Human Genome Sciences and sequenced in its full length by the primer walking approach. Sequencing was performed on an Applied Biosystems 373A sequencer, using the ABI Prism Dye Terminator Cycle Sequencing kit (Perkin Elmer). Each strand was sequenced twice.

Rapid amplification of cDNA ends (RACE)

The RACE was performed as described in (23). Total RNA from HeLa cells was used as the RNA source. The primers for the second nested PCR contained a *BamHI/XhoI* restriction site, respectively, which allowed cloning of the resulting PCR product in the *BamHI/XhoI* sites of the pBluescript SK vector (Stratagene). The clones were sequenced using the general sequencing primers M13-21 and M13 reverse (DNA technology, Aarhus, Denmark). Primer sequences were as follows:

Gene Specific Primer 1 (GSP1) TCCATATCAGCCTCAGAAC; GSP 2 (including a *Bam*HI site) CGCGGATCCGGAAGCTGT-GGAAACAGCTC; Linker primer 1 CAATCAAGAATCCCTG-CTCAGCGTAA; and Linker primer 2 (including a *Xho*I site) CGCCCTCGAGCCTCAACACCTACCCTATC.

Alignment

The alignment was established using the following sequences: Human P69OAS (M87284), Human p59OASL (AJ225089), Human p42OAS (Swissprot P00973), Mouse OAS (L3) (M33863), Rat OAS (Z18877), Porcine OAS (AJ225090) and Chicken OAS (AB002585) (accession numbers in parentheses). Alignment was performed with the *clustalw* programme (*clustalw*, an interactive service at the http://www2.ebi.ac.uk/clustalw) using default parameters (24). The alignment was resistant to changes of parameters. The output was embellished using the BOXSHADE software (http://ulrec3.unil.ch/software/BOX_form.html).

Construction of phylogenetic trees

The phylogenetic tree was constructed using the software supplied with the Phylip package (version 3.57) (25). The tree shown was created using maximum parsimony analysis on multiple aligned sequences (program PROTPARS, default parameters). The bootstrap re-sampling method was used (100 replicates) to assess the confidence of each node in the maximum parsimoni tree (programs SEQBOOT and CONSENSE). Phylogenetic trees constructed using the neighbour-joining algorithm on the same set of data gave a congruent phylogeny [programs PROTDIST (setting DAYHOFF matrix) and NEIGHBOR (default parameters),

or using Maximum Likelihood analysis (program PROTML, default settings)].

Northern blotting

Northern blots were performed using MTN filters from Clontech. The filters were hybridised in Hybrisol (Oncor) overnight at 42°C, at a probe concentration of ~1.5 × 10⁶ c.p.m./ml. After hybridisation the blots were washed at high stringency at 65°C and exposed to a film. The probe was prepared by isolating a *Bam*HI/*Xho*I fragment of the p59 coding sequence (nucleotides 374–919) or a 1.3 kb *Eco*RI fragment from the p42 cDNA (nucleotides 1–1275), from an agarose gel (Qiagen Gel Extraction kit). The isolated fragment was radioactively labelled using Stratagene Prime-it II kit and [α -³²P]dCTP.

RNase protection assay

A *XhoI/KpnI* fragment of the p59 cDNA (280 bp) was cloned in to the appropriate sites in the pBluescript SK(+) vector. After linearisation with *XhoI*, T₇ transcription was performed according to standard protocols using $[\alpha^{-32}P]$ UTP as incorporated label. The specific activity of the probe was ~1 × 10²⁰ c.p.m./mol. The γ -actin probe was prepared as described by Gunning *et al.* (26).

RNase protection was performed using the Ambion RPA II kit (Cat. #1410), following the manufacturer's instruction. p59 probe $(3 \times 10^{-16} \text{ mol})$ and $1 \times 10^{-15} \text{ mol}$ of the γ -actin probe was hybridised to 10 µg of total RNA from amniotic fluid (AMA) cells, treated as indicated.

Cell culturing

AMA cells (gift from J. E. Celis, Aarhus, Denmark) were grown in glutamax medium containing 10% newborn calf serum and penicillin/streptomycin, induced with 500 U/ml interferon- α or 100 U/ml interferon- γ for the indicated time. RNA samples were prepared by the acid guanidinium thiocyanate/phenol/chloroform extraction method (27).

RESULTS AND DISCUSSION

EST identification and sequencing

We searched the EST database for clones showing similarity to already known forms of 2–5A synthetase as described by Adams *et al.* (28). An EST clone appeared that revealed homology to the known forms of 2–5A synthetase and this clone was sequenced in its full length. The clone displayed a long open reading frame (ORF) and the corresponding amino acid sequence showed strong homology to known 2–5A synthetase proteins (Figs 1 and 2). The ORF had no obvious start codon, so we assumed that the clone was truncated at its 5' end.

To obtain the missing 5' part of the cDNA clone we performed the new RACE procedure on total RNA from HeLa cells (Fig. 3). The resulting PCR fragment was extracted from the gel and cloned in the *BamHI/XhoI* sites of the pBluescript SK vector. Ten individual clones were sequenced of which six contained an additional 19 bp, compared with the original EST sequence. The other three clones were truncated in a manner similar to the original EST clone and one contained no insert. Since no longer bands were observed in the PCR, even after extending the elongation time to 4 min, we assume that the band observed represents the authentic 5' end. The additional 19 bp that were found by 'new RACE' contained two in frame start codons. As the two AUG codons are only separated by 6 nt it is possible that either one could be used as an initiation codon (29). In this case, however, the second AUG codon is placed in a context making it highly unlikely to be used as the principal start codon (30) (the normal adenine in position –3 and the normal guanine in position +4 are both substituted by a cytosine). The complete cDNA clone contains 1819 bp and has one ORF coding for a 514 amino acid protein with a predicted molecular mass of 59 226 Da, hence called p59OASL (59 kDa 2–5 oligoadenylate synthetase-like protein) (accession number AJ225090). The size of the cDNA agrees well with the observed size of the messenger (1.8 kb) in northern blots (Fig. 4).

A 3' fragment of the p59 gene has previously been described as a thyroid receptor interactor (TRIP14) due to a yeast two-hybrid screening (31), where a part of p59 (amino acids 260–413) is reported (Q15646). We have not investigated the possible link of p59OASL to the thyroid hormone system.

Gene structure

Using the cDNA sequence as input we searched the High Throughput Genomic Sequences databases using the BLAST algorithm. We found a clear match to a bac-clone 92N15 currently under sequencing at the Sanger center (Cambridge, UK). This clone is part of the MODY3 (Maturity onset of diabetes 3) region, which maps to chromosome 12q24.2 (32). The highly homologous genes p42OAS and p69OAS genes (5) also map to chromosome 12 where they link to the marker d12s1718 (Unigene Hs.82396 and Hs.24815, NCBI) with the genetic map position 129 cM and Radiation Hybrid map position 576 cR. The bac-clone 92N15 contains the marker d12s2088 (R. Cox, personal communication), so the p59OASL gene maps in the MODY3 region at map positions 138 cM and 580 cR. The two markers d12s1718 and d12s2088 are both contained in the WC12.8 contig (singly linked) (Whitehead Institute, MA, USA).

Based upon the sequence data of bac-clone 92N15 from the Sanger centre, we determined the structure of the p59OASL gene to be composed of 6 exons. Exons 1–5 encode the 349 amino acid N-terminal domain that has high homology to other 2–5A synthetases, exon 6 encodes a 165 amino acid C-terminal domain that has no homology to any of the known 2–5A synthetases (Fig. 1A) but homology to two consecutive ubiquitins (see below). Exon 6 also harbours a putative poly A signal.

The gene structure of the p42OAS gene has previously been determined (6). This gene is composed of 7 exons with the translational start site situated in exon 3. By comparing the intron exon boundaries of the p59OASL gene to the p42OAS gene, it became evident that exons 3–7 of the p42OAS gene have a structure similar to that of exons 1–5 of the p59OASL gene (Fig. 1B). This suggests that the two genes have arisen from a duplication of an ancestral gene; exon 6 of p59OASL has thus been fused to the duplicated gene by exon shuffling. The two untranslated exons of the p42OAS gene have either been lost in the p59OASL gene or added to the p42OAS gene after it diverged from p59OASL.

Both the gene structure of p59OASL and the primary structure of the p59OASL protein suggest a two domain structure of the p59OASL protein, consisting of an N-terminal 2–5A synthetase domain of ~350 amino acids encoded by exons 1–5 and a C-terminal domain on 165 amino acids encoded by exon 6.

A: Gene structure of the p59 gene



Figure 1. (A) Organisation of the p59OASL gene. Upper numbers refer to the first and last nucleotide in an exon, according to the cDNA for p59OASL. Numbers in the lower position indicate the size of a given intron. (B) The amino acid sequence of the p42OAS and p59OASL proteins were aligned using the *clustalw* program. The first and last amino acid in each exon is highlighted. The single highlighted leucine indicates the possible alternative splice site in the p42 gene, which is used to give the p46 variant.

The promoter region 3000 bases upstream of the initiation site was analysed for the presence of transcription factor binding sites. Among the numerous putative sites several putative ISRE sites around -150, -290, -330 and -960 nt can be linked to classical interferon induction pathways including STAT1 and IRF1 (33). However, we have also found one putative IRF1 and two IRF2 sites in the middle of the ~ 10 kb intron 1, which might explain the moderate interferon induction detected in the RNase protection assays (see later). However, the activity of these sites has to be determined.

Primary structure of the p59OASL protein

We aligned the amino acid sequence of seven different 2–5A synthetases, available in the Swissprot or GenBank databases (Fig. 2). As previously mentioned, the human p69OAS is composed of a repeat of two domains each showing homology to the p42OAS. Therefore, the p69OAS gene is represented by two sequences in the alignment p69N and p69C, corresponding to the N-terminal and C-terminal domains, respectively. Since the C-termini of various 2–5A synthetase isoforms are poorly conserved, and apparently not important for enzymatic function (34), we have only applied sequences N-terminal to the CFK motif (amino acids 331–333, p42OAS numbers), recently shown to be essential for enzymatic activity (35).

Recent mutational studies showed that a LXXXP motif at the N-terminus is important for enzymatic function. If either the

Α

В

p59 Chicken p42 Pig Mouse Rat p69C P69N consensus	LXXXP 1 MALMOETYSTPASRLESETAQ.LQPHRETKEEVLDAVRTTEFFLETENERGEKRGLDQDVRULKVVK 1MGUESVSSRQLGGTAAHLQPSTEESTAVKQTVKDCFFLETERGEGSVLKTVK 1MEURNTPARSLDKFIEGTLUPDTCSRMQDHADCGFLKERCFRGSSYPVCVSKVVK 1MEURNTPARSLDKFIEGTLUPDTCSRTQVKEADCGFLKERCFGGAAHPVRVSKVVK 1MEURNTPARSLDKFIEGTLUPDTCSRTQVKEADCGFLKERCFGGAAHPVRVSKVVK 1MEURNTPARSLPATLDKFIEGTLUPDTCSRTQVKEADCGFLKERCFGGAAHPVRVSKVVK 3HPQEIRSTPWKLDKFIEGTLUPDTCSRTQVKAANVCFFLKERCFGGAAHPVRVSKVVK 3LPAPUFTTPHLLDKFIKBTQPNKCGLEQDSAVNRTFLKERCFRQSTAKIQVI 1MGNGSSCLSSVPAQKEGWFTQCKKYVK
p59 Chicken p42 Pig Mouse Rat p69C P69N consensus	P-LOOD 7 VGSFENSTVLRSTREVELVAFLSCFHSTCHAAK KUVIRLIWITWQSOLLLDLGBOLRM 56 GGSAGKGTALRINSDAD WIFINCF SECOLDARKGIP QEINOLEACO BALSVKFE GAP 60 GGS GKGTTLRGRSDADLVVFLSPL TFODOLINRGEP QEINOLEACO BALSVKFE GAP 60 GGS GKGTTLRGRSDADLVVFLNL SFEDOLIRRGEP QEINOLEACO BALSVKFE GAP 60 GGS GKGTTLGRSDADLVVFLNL SFEDOLIRRGEP VEINOLEACO BALSVKFE GAP 60 GGS GKGTTLGRSDADLVVFLNNL SFEDOLIRRGEP VEINOLEACO BALSVKFE GAP 60 GGS GKGTTLGRSDADLVVFLNNL SFEDOLIRRGEP VEINOLEACO BALSVKFE GAP 60 GGS GKGTTLGRSDADLVVFLNNL SFEDOLIRRGEP VEINOLEACO BALSVKFE GAP 60 GGS KGTLGRSDADLVVFLNNL SFEDOLIRRGEP VEINOLEACO BALSVKFE GAP 60 GGS KGTLGRSDADLVVFLNNL SFEDOLIRRGEP VEINOLEACO BALSVKFE GAP 60 GGS KGTLGRSDADLVVFLNNL SFEDOLIRRGEP VEINOLEACO BALSVKFE GAP 60 GGS KGTL GRSDADLVVFLNNL SFEDOLIRRGEP VEINOLEACO BALSVKFE GAP 60 GGS KGTL GRSDADLVVFLNNL SFEDOLIRRGEP VEINOLEACO BALSVKFE GAS 62 GGS GKGTL GRSDADLVVFLNNL SFEDOLIRRGEP VEINOLEACO BALSVKFE GAS 62 GGS GKGTL GRSDADLVVFLNNL SFEDOLIRRGEP VEINOLEACO BALSVKFE GAS 62 GGS GKGTL GRSDADLVVFLNNL SFEDOLIRRGEP VEINOLING FENOLEACO BALSVKFE GAS 63 GSG KGTL GRSDADLVVFLNNL SFEDOLIRRGEP VEINOLING FEDOLIRRGEP VEINOLING 64 GGS KGTL GRSDADLVVFLNNSKS TSOKNERHN KEI FOLKAFWEB EE E-VSF PP 60 GGSYG KTVLRGNSDTLVFFSDLKOFODGKRSQRFI DKTGDKFKFCLFTMLKNNFE GKS 71 ** **
p59 Chicken p42 Pig Mouse Rat p69C P69N consensus	29 EQRV DALVET O RGTAEP TVT NPARALGPSLPNSQ PPER YVSLIKASGGPG-NBCP 22 SRDRPRSLSLTLS SASSES VD LPAMDALGQ IQDARPDPA YVKLPATRCGPG-NBCP 25 RWGNPRALSFVLS LQLGEGV PDVLPAPDALGQETGSYKPNPQTVVKLIEECTDLQK-EG-EFST 24 RRENPRALSFVLS PQLQQEV PDVLPAPDALGQWTPGYKPNPEIYVOLIKECKSRGK-EG-EFST 26 WWPNRSLSFKLSAPHLQQEV PDVLPAPDVLGHNTSSKPDPRTVAILIEECTSLGK-DG-EFST 26 WWPNRALSFKLSAPHLQQEV PDVLPAPDVLGHNTSSKPDPRTVAILIEECTSLGK-DG-EFST 26 WWPNRALSFKLSAPHLQQEV PDVLPAPDVLGHNSLYSNPDRTVAILIEECTSLGK-DG-EFST 26 WWPNRALSFKLSAPHLQEV PDVLPAPTVLGHNSLYSNPDRTVAILIEECTSLGK-DG-EFST 26 WWPNRALSFKLSAPHLQEV PDVLPAPTVLGHNSLYSNPDRTVAILIEECTSLGK-DG-EFST 26 WWPNRALSFKLSAPHLQEV PDVLPAPTVLGHNSLYSNPDRTVAILIECTSLGK-DG-EFST 27 KWKAEPVLSFSLKKVLNESGSPDVLPAPAALGUSSGSTESPEWAGLIDLYKSSDL-PGGEFST 24 LDGSTIQVFTKNQRISTVLAFNAL-SLND-NBSEWTYRELKRSLDKTNASPG-EFFV 41
p59 Chicken p42 Pig Mouse Rat p69C P69N consensus	ATP binding domain YALELLT YAWE 90 SF ELQRNFYKHRPTKLKSLIRLVKHWYQQYVKARSPRANLPPLYALELLT YAWEMG EEDENEM 83 SFTELQR FYKNRPAKLKDL RLVKHWYQQYVKARSPRANLPPLYALELLT YAWEMG EEDENEM 84 SFTELQR FYKNRPAKLKDL RLVKHWYQQYVKARSPRANLPPLYALELLT YAWEMG EEDENEM 85 CFTELQR FYKNRPAKLKDL RLVKHWYQN KKKIGKLPPQ YALELLT YAWENG GR-EFFS 86 CFTELQR FYKNRPAKLKSLIRLVKHWYQN KKKIGKLPPQ YALELLTYYAWENG GRK-THFS 90 CFTELQRNF KORPTKLKSLIRLVKHWYQLCKEK GKPLPPQ YALELLTYYAWENG GRC-YFN 90 CFTELQRNF KORPTKLKSLIRLVKHWYQLCKEK GKPLPPQ YALELLTYYAWENGGNGC-YFN 90 CFTELQRNF KORPTKLKSLIRLVKHWYQLCKEK GKPLPPQ YALELLTYYAWENGCNGC-YFN 90 CFTELQRNF KORPTKLKSLIRLVKHWYDLCKEK GKPLPPQ YALELLTYYAWENGCNGC-YFN 90 CFTELQRNF KORPTKLKSLIRLVKHWYDLCKEK GKPLPPQ YALELLTYYAWENGONGC-YFN 910 CFTELQRNF KORPTKLKSLIRLVKHWYDLCKEK GKPLPPQ YALELLTYYAWENGONG-YFN 920 CFTELQRNF KORPTKLKSLIRLVKHWYDLCKEK GKPLPPQ YALELLTYYAWENGONG-YFN 930 CFTELQRNF KORPTKLKSLIRLVKHWYDLCKEK GKPLPPQ YALELLTYYAWENGONG-YFN 940 CFTELQRYF KORPTKLKSLIRLVKHWYDLCKEK GKPLPPG YALELLTYYAWENGONG-YFN 941 ************************************
p59 Chicken p42 Pig Mouse Rat p69C P69N consensus	RPVILDPADPT AMAEGRETVLEL LEY V CIYWTKYYTLHNAI EDC // OL K KERP ILDPADPTLNV EGYR MAEGRETVLEL GURD CIYWEKYYSLEDEWYGAL EOL ORRVILDPADPTGIL COS-KNN 251 TAOGRETVLELVINYOOLCIYWEKYYDEKNPI EXIL OLTKPREVILDPADPTGN GGDPKRM 253 TAOGRETVLELVINYOHLRIYWTKYYDENPEVG CULOOL KPREVILDPADPTGNV GGDPKRM 253 TAOGRETVLELVINYOHLRIYWTKYYDEOHPOVG CULOOL KPREVILDPADPTGNV GGDPKRM 253 TAOGRETVLELVINYOHLRIYWTYYDEOHPOVG UL OL KAREVILDPADPTGNV GGDPKGM 253 TAOGRETVLELVINYOOLRIYWTKYYDEOHPOVG UL OL KAREVILDPADPTGNV GGDPKGM 254 TAOGRETULEUVCYOOLGII KVNYNEEDETVRK LIGIL OLKREVILDPADPTGNV GGDRGM 255 TAOGRETVLELVCOYOOLGII KVNYNEEDET RNILL OLOKREVILDPADPTGNV GGDRCM 254 IAEGVRTVLEL KCOCKLCIYMMVNYNEEDET RNILL OLOSARVILDPVDTNNVS-GDKICM 255 TAOGRETULELVCOYOOLGII KVNYNEEDET RNILL OLOSARVILDPVDTNNVS-GDKICM 256 TAOGRETVLEL KCOCKLCIYMMVNYNEEDET RNILL OLOSARVILDPVDTNNVS-GDKICM 257 TAOGRETULEL KCOCKLCIYMMVNYNEEDET RNILL OLOSARVILDPVDTNNVS-GDKICM 258 TAOGRETVLEL KCOCKLCIYMMVNYNEEDET RNILL OLOSARVILDPVDTNNVS-GDKICM 259 TAOGRETVLEL KCOCKLCIYMMVNYNEEDET RNILL OLOSARVILDPVDTNNVS-GDKICM 259 TAOGRETVLEL KCOCKLCIYMMVNYNEEDET RNILL OLOSARVILDPVDTNNVS-GDKICM
p59 Chicken p42 Pig Mouse Rat p69C P69N consensus	320 DIWAORRSQGLKQDCCY 336 312 DLWAKEAAESRWSLPCVSDVCS 333 317 RQLAQEARA-WLNYPCPK 333 317 QRLAQEARAV-WLGYPCCK 333 319 RRLAEEADV-WLWYPCFI 335 319 RRLASEARL-WLQYPCFM 335 554 HLLDKEAKV-RLSSPCFK
UBIQ_HUMAN HUMIFN15K P59_EXON6 UBIL_HUM consensus	1MOLFVKTUTGKTITLEVEPSOTIENVKAKICDKEGTEPDOORLIFAGKOLEDGR 1MGWDLTVKMLAGNEFOVSISSSMSVSEIKAOITCKIGVHAFOORLAVHPSGVALODRV 1 RARD HLTVEQRGYPDENLINNEYEPIRKVKEKIRRTKGYSS-LORLSEOVPGSEECLLS 1MOLTVKALOGRECSLOVPEDELVSTIKOLVSEK NVPVROORLIFKGKALADGK 1*. **
UBIQ_HUMAN HUMIFN15K P59_EXON6 UBIL_HUM consensus	55 ILSDYNLOKESTIHLVIRLRGGXXXKOIFVKTLTGKTITLEVEDSDTIENWKAKIQ 59 PLASOGIGHGSTVLLVVDKCDEPSILVENNKGESSTYEVELUQTVAHLKQQVS 60 SRCSLAKYGIFSHTHIYLLETIPSELOVFVKNFDCGSVAYAINENSFILG <u>LKQO</u> IE 55 RLSDYSLGFNSKUNLVVKPLEKVLLEGGAGRLADSPPPQVMQLISKVLARHFSAADASR 61
UBIQ_HUMAN HUMIFN15K P59_EXON6 UBIL_HUM consensus	111 DKEGIPPDOORLIFAGROIEDGRT SDYNIOKESTLHUVIRLRGG 113 GLEGYDDLFWLTFEGRPLEDO PLGEYGERDISTYFWNIRLRGGETEFEGRS 116 DOOGLPKKOCCLEFOCOVLODWEGGIYGFODSDTLIFSKK-GELFPAS 115 VLEQLORDYERSISR-LTIDDIERLASRFHPPDVTETYEKGFSK

Figure 2. (A) Alignment of various 2–5A synthetases. The amino acid sequences of the N-terminal part of the indicated 2–5A synthetase proteins (see text), were aligned using the *clustalw* program. (B) Alignment of the C-terminal part (165 amino acids) of p59OASL coded by exon 6, with ISG15, Human Ubiquitin-like protein GDX (UBIL) and two consecutive ubiquitins separated by XXX. The aligned sequences were subsequently embellished using the BOXHADE program.



Figure 3. PCR products from the RACE procedure were separated on a 1% agarose gel. A strong band appears when the RNA was decapped prior to ligation, but not if decapping was omitted. Prolonging the extension time to 4 min did not lead to new larger bands. Lane 1, PCR using decapped RNA and an extension time of 1 min; lane 2, decapped RNA and a extension time of 4 min; lane 3, RNA not decapped prior to ligation, extension time 1 min; lane 4, same as lane 3 but the extension time was 4 min.



Figure 4. Northern blots. Filters containing 2 μ g of poly A⁺ RNA pr. lane from various tissues were probed with (**A**) a p59OASL specific probe or (**B**) a p42OAS/p46OAS specific probe. Lane 1, spleen; lane 2, thymus; lane 3, prostate; lane 4, testis; lane 5, uterus; lane 6, small intestine; lane 7, colon (mucosal lining); lane 8, peripheral blood leucocytes. The p42/p46 probe gives distinct bands corresponding to the p42 and p46 isoforms. The origin of the larger bands is unknown.

leucine or proline were mutated, the enzyme would lose its capability to form 2–5As but would still be able to bind both ATP and dsRNA (36). This motif is conserved except in the chicken 2–5A synthetase form.

Generally the amino acids from 190 to 320 are well-conserved (numbers according to the p59OASL sequence). Mutational and biochemical studies placed the high affinity ATP binding domain of p42OAS in the proximity of lysine 199 (p42 numbering) (37). A stretch of ~25 amino acids around lysine 199 is highly conserved, except for some relaxation in the p69N sequence. We also point out two other highly conserved motifs, the YALTLLT(v/i)YAWE motif and the RP(v/i)ILDPADPT motif. No functional data exist for those motifs.

Finally, there is a P-loop motif, presumably involved in phosphate binding, in the p42OAS family which is especially homologous to adenylate kinases (myokinases): VVXGXXSG-KGT. The last glycine (G) is unique in the P-loop of these two protein families compared to all other P-loops (38). This is a striking phenomenon considering that the substrates for these two very different enzymes actually are rather similar, ATP + ATP versus ATP + AMP. However, in p59OASL the crucial lysine (K) is replaced by an asparagine (N) which could be an indication of the essential differences in the activity of p59OASL compared to all the other 2–5A synthetases, including the most homologous one, chicken 2-5A synthetase. The same region of the p42OAS was proposed to share structural homologies to the rat DNA polymerase β phosphate binding loop (39). The aspartate (D) residues 75 and 77 are presumed to co-ordinate an active site magnesium ion. In p59OASL these aspartates are replaced by glutamates by single nucleotide substitutions. Glutamate remains capable of co-ordinating the crucial magnesium ion, but the change from aspartate to glutamate might lead to a difference in substrate specificity or in the reaction conditions required. Further mutational studies are necessary to shed light on the functional role of the P-loop region of p59OASL, as well as the other members of the 2-5A synthetase family.

The C-terminal domain, exon 6, of p59OASL shows homology to the ubiquitin like proteins ISG15 [M13755, (18)] (identities = 40/165) and the ubiquitin-like protein GDX [P11441, (40)] (identities = 31/165) (Fig. 2). All three proteins seem to have evolved from a duplication of a ubiquitin gene. The C-terminus of p59OASL exon 6 is 36% homologous to ubiquitin [P02248, (41)] (identities = 26/76) and the N-terminus of p59OASL exon 6 is 23% homologous to ubiquitin (identities = 22/76). There are examples of other proteins, ribosomal protein S30 (42) and large proline rich protein BAT3 (43), which contain intrinsic ubiquitin domains playing a role in the regulation of the activities and half-lives of these proteins. It is thus tempting to speculate that the C-terminal domain is involved in the regulation of either the 2-5A synthetase activity or the stability of the protein. One might even propose that a proteasome pathway is involved in a degradation process leading to smaller forms of p59OASL, and that a proteasomal activity is part of the post-translational control mechanism of the activity of the p59OASL protein.

Phylogenetic studies

Using the alignments presented above, we have constructed a hypothetical phylogenetic tree (Fig. 5), as described in Materials and Methods. The p42OAS forms from four different species (Rat, Mouse, Pig and Man) group in the same family (bootstrap



Figure 5. Phylogenetic tree for the different 2–5A synthetase forms. Bootstrap values are shown.

value 100%) whereas p59OASL groups together with a chicken form of 2–5A synthetase (bootstrap value 95%). The two sequences representing the p69OAS form either have a node between them (as on the tree shown) or they group to their own branch of the tree. Except for the slight variations in the placing of the p69OAS sequences, the use of different algorithms and change of parameters did not change the topology of the phylogenetic tree.

From the topology of the phylogenetic tree it is evident that the multiple isoforms seen in humans diverged before the radiation of the mammalian class, meaning that a similar number of isoforms are very likely to exist in other mammalian species.

The expression pattern of p59OASL

To evaluate the expression of the p59OASL messenger RNA in different tissues we made northern blots using Multiple Tissue Northern from Clontech (Fig. 4). A wide variety of tissues were screened for p59OASL expression by this method. Low expression could be detected in most tissues, but high expression was detected in primary blood leukocytes (and other tissue related to the haematopoetic system), in colon, stomach and to some extent in testis, whereas the expression in small intestine was low (Table 1). An aspect common to tissues with high expression of p59OASL mRNA might be that they contain highly proliferating cells. The cells of the lymphoid system as well as the cells present in colon or stomach epithelium are rather short lived and are constantly replaced.

To our knowledge, the expression patterns of the p42OAS or p69OAS isoforms have not been addressed in detail, but

re-probing of the filter with a p42/p46 specific probe showed a more widespread expression of this isoform (Fig. 4) compared to p59OASL.

We investigated the induction of p59OASL by interferon- α and interferon- γ in AMA cells. The transcription of p59 is stimulated 2–5-fold by the addition of interferon- α and 2–3-fold by interferon- γ , as demonstrated by RNase protection assays (Fig. 6). Using RT–PCR we have seen a similar induction in HeLa cells (data not shown).

The data presented here suggest that the p59OASL expression is strictly controlled, and only expressed in certain tissues. Presumably this regulation correlates with the cellular function of p59OASL. We are currently investigating whether p59OASL is linked to the interferon system in a similar manner as the p42OAS and p69OAS isoforms, or whether p59OASL is part of other cytokine systems. In the view of the prevalence of p59OASL in highly proliferating cells, we believe that a role may exist for p59OASL as part of a general growth control pathway in those cells.

Table 1. Expression of p59OASL in various tissues

Tissue	Relative expression	Number of blots where
		tissue is represented
PBL	++++	3
Stomach	+++(+)	1
Spleen	+++	3
Lymph node	+++	1
Colon (muscosal lining)	+++	2
Fetal brain	+++	1
Thyroid	++(+)	1
Testis	++(+)	3
Thymus	++	3
Fetal liver	++	1
Pancreas	++	1
Adrenal medulla	++	1
Adrenal cortex	++	1
Fetal kidney	++	1
Uterus	++	1
Small intestine	+	3
Prostate	+	2
Ovary	+	1
Fetal lung	+	1

A number of northerns were made using MTN filters from Clonetech (see text for details). The relative expression was judged from the radiography after 5 days exposure. RNA from some of the tissues was present on several filters, as indicated.

The functions of p59OASL

Escherichia coli expressed histidine-tagged p59OASL does not have any 2–5A synthetase activity using conditions (44), where an equivalent histidine-tagged p42OAS was fully active. This is not surprising, taking into account the differences in the presumed active sites, notably the P-loop.

There are evidently numerous experiments that have to be performed to explore the functions of p59OASL, which could



Figure 6. Regulation of the p59OASL gene by interferon. The induction of the p59OASL gene by interferon- α or interferon- γ was measured by RNAse protection assays, using total RNA form induced AMA cells. Cells were induced as indicated. The protected fragments were separated on a 5% native polyacrylamide gel, and the interesting bands were quantified using a PhosphorImager. The p59OASL signal was normalised to the γ -actin signal prior to calculation of the induction. Each column represents a mean of four independent samples, the standard deviation is indicated.

constitute an important new regulatory feature in the 2–5A pathway and connect this to the ubiquitin systems including the proteasome control of cell cycle factors and transcription factors. In particular, p59OASL could be a factor in the 2–5A involvement in apoptosis (19,45).

ACKNOWLEDGEMENTS

We would like to thank the following people: Anette Bejder for excellent help with the RACE experiment and Lars Aagaard for giving invaluable help in using the Phylip package and for many fruitful discussions, Niels Ole Kjeldgaard and Tracey Flint for critical reading of the manuscript and many valuable comments, Andrew King at the Sanger Centre (Cambridge, UK) and R. D. Cox, Welcome Trust Centre for Human Genetics (Oxford, UK) for valuable help with the genomic sequence and finally Cecilia Rosada Kjeldsen for helpful discussion. This work was supported by the Danish Natural Science Research Council (BIOTEK 3 program), The Danish Cancer Society and the Karen Elise Jensen's Foundation.

REFERENCES

- Roberts,W.K., Hovanessian,A., Brown,R.E., Clemens,M.J. and Kerr,I.M. (1976) *Nature*, 264, 477–480.
- 2 Hovanessian, A.G., Brown, R.E. and Kerr, I.M. (1977) Nature, 268, 537-540.
- 3 Clemens, M.J. and Williams, B.R. (1978) Cell, 13, 565-572.
- Baglioni, C., Minks, M.A. and Maroney, P.A. (1978) *Nature*, 273, 684–687.
 Williams, B.R., Saunders, M.E. and Willard, H.F. (1986)
- Somatic Cell Mol. Genet., **12**, 403–408. 6 Benech, P., Mory, Y., Revel, M. and Chebath, J. (1985) *EMBO J.*, **4**, 2249–2256.
- Beneti, F., Mory, I., Revel, M. and Chevani, J. (1963) *EmBO J.*, 4, 2249–224
 Saunders, M.E., Gewert, D.R., Tugwell, M.E., McMahon, M. and
- Williams, B.R. (1985) *EMBO J.*, 4, 1761–1768.
 8 Marie, I. and Hovanessian, A.G. (1992) *J. Biol. Chem.*, 267, 9933–9939.
- Chebath, J., Benech, P., Hovanessian, A., Galabru, J. and Revel, M. (1987)
 J. Biol. Chem., 262, 3852–3857.
- 10 Dougherty, J.P., Samanta, H., Farrell, P.J. and Lengyel, P. (1980) J. Biol. Chem., 255, 3813–3816.

- 11 Chebath, J. and Revel, M. (1992) In Baron, B., Copperhaver, F., Dianzani, W.R., Hughes, T.K., Jr, Klimpel, G.R., Jr, Niesel, D.W., Stanton, G.J. and Tyring, S.K. (eds), *Interferon: Principles and Medical Applications*. The University of Texas Medical Branch at Galverston, Galverston, TX, pp. 225–236.
- 12 Hartmann, R., Noerby, P.L., Martensen, P.M., Joergensen, P., James, M.C., Jacobsen, C., Moestrup, S.K., Clemens, M.J. and Justesen, J. (1998) *J. Biol. Chem.*, **273**, 3236–3246.
- 13 Johnston, M.I. and Torrence, P.F. (1984) In Friedman, R.M. (ed.), *Interferon 3. Mechanisms of Production and Action*. Elsevier, Amsterdam, pp. 189–298.
- 14 Chebath, J., Benech, P., Revel, M. and Vigneron, M. (1987) Nature, 330, 587–588.
- 15 Mitra,A., Higgins,D.W., Langenberg,W.G., Nie,H., SenGupta,D.N. and Silverman,R.H. (1996) Proc. Natl Acad. Sci. USA, 93, 6780–6785.
- 16 Ogawa, T., Hori, T. and Ishida, I. (1996) Nature Biotechnol., 14, 1566–1569.
- 17 Rysiecki, G., Gewert, D.R. and Williams, B.R. (1989) J. Interferon. Res., 9, 649–657.
- 18 Narasimhan, J., Potter, J.L. and Haas, A.L. (1996) J. Biol. Chem., 271, 324–330.
- 19 Zhou, A., Paranjape, J., Brown, T.L., Nie, H., Naik, S., Dhong, B., Chang, A., Trapp, B., Fairchild, R., Colmenares, C. and Silverman, R.H. (1997) *EMBO J.*, 16, 6355–6363.
- 20 Silverman, R.H. (1997) Ribonucleases: Structure and Functions. Academic Press, Inc., NY, pp. 515–551.
- 21 Marra, M.A., Hillier, L. and Waterston, R.H. (1998) Trends Genet., 14, 4-7.
- 22 Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990)
- *J. Mol. Biol.*, **215**, 403–410. 23 Frohman,M.A. (1994) *PCR Meth. Appl.*, **4**, S40–58.
- 24 Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) Nucleic Acids Res., 22, 4673–4680.
- 25 Felsenstein, J. (1996) Methods Enzymol., 266, 418-427.
- 26 Gunning, P., Ponte, P., Blau, H. and Kedes, L. (1983) Mol. Cell Biol., 3, 1985–1995.
- 27 Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem., 162, 156-159.
- 28 Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O. *et al.* (1995) *Nature*, **377**, 3–174.
- 29 Kozak, M. (1995) Proc. Natl Acad. Sci. USA, 92, 2662–2666.
- 30 Kozak, M. (1997) EMBO J., 16, 2482–2492.
- 31 Lee, J.W., Choi, H.S., Gyuris, J., Brent, R. and Moore, D.D. (1995) Mol. Endocrinol., 9, 243–254.
- 32 Yamagata,K., Oda,N., Kaisaki,P.J., Menzel,S., Furuta,H., Vaxillaire,M., Southam,L., Cox,R.D., Lathrop,G.M., Boriraj,V.V. et al. (1996) Nature, 384, 455–458.
- 33 Heinemeyer, T., Wingender, E., Reuter, I., Hermjakob, H., Kel, A.E., Kel, O.V., Ignatieva, E.V., Ananko, E.A., Podkolodnaya, O.A., Kolpakov, F.A., Podkolodny, N.L. and Kolchanov, N.A. (1998) *Nucleic Acids Res.*, 26, 362–367.
- 34 Ghosh,S.K., Kusari,J., Bandyopadhyay,S.K., Samanta,H., Kumar,R. and Sen,G.C. (1991) J. Biol. Chem., 266, 15293–15299.
- 35 Ghosh,A., Sarkar,S.N., Guo,W., Bandyopadhyay,S. and Sen,G.C. (1997) J. Biol. Chem., 275, 33220–33226.
- 36 Ghosh,A., Desai,S.Y., Sarkar,S.N., Ramaraj,P., Ghosh,S.K., Bandyopadhyay,S. and Sen,G.C. (1997) J. Biol. Chem., 272, 15452–15458.
- 37 Kon, N. and Suhadolnik, R.J. (1996) J. Biol. Chem., 271, 19983-19990.
- 38 Saraste, M., Sibbald, P.R. and Wittinghofer, A. (1990) Trends Biochem. Sci., 15, 430–434.
- 39 Holm,L. and Sander,C. (1995) Trends Biochem. Sci., 20, 345-346.
- 40 Toniolo, D., Persico, M. and Alcalay, M. (1988) Proc. Natl Acad. Sci. USA, 85, 851–855.
- 41 Wiborg, O., Pedersen, M.S., Wind, A., Berglund, L.E., Marcker, K.A. and Vuust, J. (1985) *EMBO J.*, **4**, 755–759.
- 42 Kas,K., Michiels,L. and Merregaert,J. (1992) Biochem. Biophys. Res. Commun., 187, 927–933.
- 43 Banerji, J., Sands, J., Strominger, J.L. and Spies, T. (1990) Proc. Natl Acad. Sci. USA, 87, 2374–2378.
- 44 Justesen, J., Ferbus, D. and Thang, M.N. (1980) Nucleic Acids Res., 8, 3073–3085.
- 45 Kisselev,L.L, Justesen,J., Wolfson,A.D. and Frolova,L.Y. (1998) FEBS Lett., 427, 157–163.