Cloning, expression and characterization of YSA1H, a human adenosine 5′-diphosphosugar pyrophosphatase possessing a MutT motif

Lakhdar GASMI, Jared L. CARTWRIGHT and Alexander G. McLENNAN[1]

School of Biological Sciences, Life Sciences Building, University of Liverpool, P.O. Box 147, Liverpool L69 7ZB, U.K.

The human homologue of the *Saccharomyces cerevisiae* YSA1 protein, YSA1H, has been expressed as a thioredoxin fusion protein in *Escherichia coli*. It is an ADP-sugar pyrophosphatase with similar activities towards ADP-ribose and ADP-mannose. Its activities with ADP-glucose and diadenosine diphosphate were 56% and 20% of that with ADP-ribose respectively, whereas its activity towards other nucleoside 5′-diphosphosugars was typically 2–10%. cADP-ribose was not a substrate. The products of ADP-ribose hydrolysis were AMP and ribose 5-phosphate. K_m and k_{cat} values with ADP-ribose were 60 μM and 5.5 s^{-1} respectively. The optimal activity was at alkaline pH (7.4–9.0) with 2.5–5 mM Mg^{2+} or 100–250 μM Mn^{2+} ions; fluoride was inhibitory, with an IC_{50} of 20 μM. The YSA1H gene, which maps to 10p13–p14, is widely expressed in all human tissues examined, giving a 1.4 kb transcript. The 41.6 kDa fusion protein behaved as an 85 kDa dimer on gel filtration. After cleavage with enterokinase, the 24.4 kDa native protein fragment ran on SDS/PAGE with an apparent molecular mass of 33 kDa. Immunoblot analysis with a polyclonal antibody raised against the recombinant YSA1H revealed the presence of a protein of apparent molecular mass 33 kDa in various human cells, including erythrocytes. The sequence of YSA1H contains a MutT sequence signature motif. A major proposed function of the MutT motif proteins is to eliminate toxic nucleotide metabolites from the cell. Hence the function of YSA1H might be to remove free ADP-ribose arising from NAD^+ and protein-bound poly- and mono-(ADP-ribose) turnover to prevent the occurrence of non-enzymic protein glycation.

Key words: ADP-ribose, nucleotide sugar, nudix hydrolase, protein glycation.

INTRODUCTION

The MutT sequence motif proteins are a recently described family of mostly small, soluble nucleotide pyrophosphatases named after the 16 kDa *Escherichia coli mutT* gene product, a hydrolase with a strong preference for the mutagenic nucleotide 8-oxo-dGTP [1]. They all possess the sequence signature motif GX_5EX_5[UA]XRE[UT]XEEXGU (single-letter codes, where U is a hydrophobic amino acid, and the square brackets indicate that the position could be occupied by any of the amino acids within the brackets) [2]. The *E. coli* genome sequence reveals 11 further proteins with perfect or closely related MutT motifs, several of which have now been cloned and characterized [3]. The substrate specificities of these other proteins varies from quite specific to relatively broad; known substrates include dATP, NADH, nucleoside 5′-diphosphosugars (NDP-sugars) and dinucleoside polyphosphates. Because all these substrates have the general structure of a nucleoside diphosphate linked to another moiety, X, the name 'nudix' hydrolase has been proposed for this family. The function of members of this protein family might be to cleanse the cell of potentially deleterious endogenous nucleotide metabolites and to modulate the accumulation of metabolic intermediates by diverting them into alternative pathways in response to biochemical need [3].

MutT motif proteins are also found in eukaryotes, including humans [4]. Known human proteins include the MutT orthologue MTH1 [5], diadenosine 5′,5‴-P^1,P^4-tetraphosphatase (Ap_4A hydrolase) [6] and diphosphoinositol polyphosphate phosphohydrolase (DIPP) [7]. DIPP is homologous with the recently described Ap_6A hydrolases from budding and fission yeasts and, like them, is remarkable in its ability to hydrolyse the non-nucleotide diphosphoinositol polyphosphates in addition to the structurally unrelated diadenosine polyphosphates [8]. A survey of the GenBank expressed sequence tag (EST) database suggests the existence of at least six further human MutT motif proteins, including a protein encoded in an anti-sense transcript of the basic fibroblast growth factor gene [9], a protein closely related to DIPP (DIPP2) [7], and a homologue of the *Saccharomyces cerevisiae* YSA1 protein. YSA1 has been reported to have ADP-ribose pyrophosphatase activity (EC 3.6.1.13), although no details of this activity have been published [3,10].

ADP-ribose is the most significant ADP-sugar in mammalian cells. It is generated by the turnover of NAD^+ and protein-bound poly- and mono-(ADP-ribose) and, if allowed to accumulate, could be hazardous to the cell by virtue of its ability to modify protein histidine, lysine and cysteine residues by non-enzymic glycation [11–13] and to bind to ATP-activated K^+ channels [14]. The elimination of free ADP-ribose therefore seems to fall within the category of 'housecleaning' functions proposed for the MutT motif family. An ADP-ribose pyrophosphatase has recently been isolated from human erythrocytes [12,15], whereas three cytoplasmic and one mitochondrial ADP-ribose pyrophosphatases in rat liver have been described [16]. A possibly distinct ADP-sugar pyrophosphatase (EC 3.6.1.21) has also been isolated from

Abbreviations used: Ap_4A, diadenosine 5′,5‴-P^1,P^4-tetraphosphate (other diadenosine polyphosphates are abbreviated similarly); DIPP, diphosphoinositol polyphosphate phosphohydrolase; EST, expressed sequence tag; IPTG, isopropyl β-D-thiogalactoside; LB, Luria–Bertani; NDP-sugar, nucleoside 5′-diphosphosugar; NTA, nitrilotriacetate; ORF, open reading frame; Trx, thioredoxin; YSA1H, YSA1 homologue.

[1] To whom correspondence should be addressed (e-mail agmclen@liv.ac.uk).

The nucleotide sequence data reported will appear in DDBJ, EMBL and GenBank Nucleotide Sequence Databases under the accession number AF155832.

mammalian liver [17]. However, the amino acid sequences of none of these enzymes have been determined. Here we report the cloning, expression and characterization of YSA1H, the human homologue of the yeast YSA1 protein, and show that it is an ADP-sugar pyrophosphatase with its highest activity towards ADP-ribose, ADP-mannose and ADP-glucose.

Since this paper was submitted the Human Genome Nomenclature Committee (HGNC) has approved gene symbols for the human MutT motif proteins. Accordingly theYSA1H gene is designated NUDT5. Full details can be found on http://www.gene.ucl.ac.uk/users/hester/npym.html or by contacting the committee (e-mail nome@galton.ucl.ac.uk.).

EXPERIMENTAL

Materials

Mononucleotides, dinucleoside polyphosphates, NDP-sugars and nucleoside 5′-diphosphoalcohols were from Sigma. [^{32}P]UTP (800 Ci/mmol) was from ICN; NAD^+, NADH, $NADP^+$, NADPH, calf intestinal alkaline phosphatase, yeast inorganic pyrophosphatase, *Eco*RI and *Nco*I were from Boehringer Mannheim; pET 32b(+) was from Novagen; *Pfu* DNA polymerase was from Stratagene; I. M. A. G.E. clone 310860 was obtained from the UK Human Genome Mapping Project Resource Centre.

Cloning of YSA1H cDNA

I.M.A.G.E. clone 310860 contains an insert corresponding to a YSA1H cDNA in the polylinker-modified plasmid pT7T3D. The sequence of the insert was confirmed (see Figure 1) and the insert was then amplified by PCR. The 29-mer oligonucleotide forward and reverse primers 5′-d(TCGTTTGACCATGGAGAGCC AAGAACCAA)-3′ and 5′-d(TAAGTTGAATTCAAAGTCTT-AGTGAAGAA)-3′ were synthesized to provide an *Nco*I restriction site at the start of the amplified insert and an *Eco*RI site at the end. After amplification with *Pfu* DNA polymerase, the DNA was recovered by extraction with phenol/chloroform/3-methylbutan-1-ol (25:24:1, by vol.) and then digested with *Nco*I and *Eco*RI. The digest was gel-purified and the restriction fragment was ligated between the *Nco*I and *Eco*RI sites of the pET32b(+) thioredoxin (Trx) fusion vector (Novagen). The resulting construct, pET-YSA1H, generated a fusion of YSA1H downstream of the cleavable 109-residue Trx fusion protein and His-tag and S-tag sequences. This plasmid was used to transform *E. coli* XL1-Blue cells for propagation.

Protein expression and purification

E. coli strain BL21(DE3) was transformed with pET-YSA1H. A single colony was picked from a Luria–Bertani (LB) agar plate containing 60 μg/ml ampicillin and inoculated into 10 ml LB medium containing 60 μg/ml ampicillin. After overnight growth at 37 °C, the cells were transferred to 1 litre of LB medium containing 60 μg/ml ampicillin and grown at 30 °C to a D_{600} of 0.8. Isopropyl β-D-thiogalactoside (IPTG) was added to 1 mM and the cells were incubated for 4 h. The induced cells were harvested, washed and resuspended in 50 ml of breakage buffer [50 mM Tris/HCl (pH 8.0)/0.1 M NaCl]. The cell suspension was sonicated and the resulting lysate was cleared by centrifugation at 15000 ***g*** and 4 °C for 10 min. The supernatant was recovered and applied to a 15 mm × 50 mm column of Ni^{2+}/nitrilotriacetate (NTA)-agarose (Sigma) equilibrated with 20 mM Tris/acetate (pH 7.0)/0.3 M NaCl/10 mM 2-mercaptoethanol at a flow rate of 0.5 ml/min. After elution of the unbound proteins, a linear gradient of 0–50 mM histidine in equilibration buffer was applied at a flow rate of 1 ml/min; fractions of 1 ml were collected and analysed by SDS/PAGE. Those containing pure Trx–YSA1H fusion protein were collected and concentrated by ultrafiltration.

Enzyme assays and product identification

Substrates were screened by measuring the P_i released in a coupled assay by co-incubation of substrate with YSA1H protein and either alkaline phosphatase or inorganic pyrophosphatase [16,18]. The standard assay (200 μl) for phosphodiester substrates was incubation for 10 min at 37 °C with 50 mM Tris/HCl (pH 7.5)/5 mM $MgCl_2$/1 mM dithiothreitol/0.3 mM substrate/0.25 μg of Trx–YSA1H fusion protein and 0.5 μg (1 unit) of alkaline phosphatase. Phosphomonoester substrates were assayed as above, except that 0.5 μg (100 m-units) of inorganic pyrophosphatase was used instead of alkaline phosphatase. The P_i released in each case was measured colorimetrically. P_i released from $NADP^+$ and NADPH in control assays without YSA1H was subtracted. Kinetic parameters for ADP-ribose hydrolysis were determined by measuring the P_i released in the alkaline-phosphatase-coupled assay as described. Non-specific alkaline phosphodiesterase (nucleotide pyrophosphatase) activity was assayed with thymidine 5′-monophospho-*p*-nitrophenyl ester as substrate [19].

Reaction products generated from the different substrates were identified by high-performance anion-exchange chromatography. Reaction mixtures containing 50 mM Tris/HCl, pH 7.5, 5 mM $MgCl_2$, 1 mM dithiothreitol, 0.3 mM substrate and 3 μg of Trx–YSA1H fusion protein were incubated at 37 °C for 30 min in a volume of 200 μl. Samples (100 μl) of the reaction mix were applied to a 1 ml Resource-Q column (Pharmacia) equilibrated with 45 mM ammonium acetate (pH adjusted to 4.6 with H_3PO_4) and eluted with a linear gradient from 0 % to 100 % (v/v) 0.5 M NaH_2PO_4 (adjusted to pH 2.7 with acetic acid) for 10 min at a flow rate of 1 ml/min [15]. Elution was monitored at 260 nm and peaks were identified with the aid of standards.

Northern and dot-blot analyses

A ^{32}P-labelled anti-sense RNA to the YSA1H cDNA was generated from I.M.A.G.E. clone 310860 by using an *in vitro* transcription kit (Ambion), [^{32}P]UTP and T3 RNA polymerase in accordance with the manufacturer's instructions. A Multiple Choice™ Northern blot (OriGene) containing poly(A)$^+$-selected RNA from various human tissues (2 μg per lane) was probed overnight with the RNA probe (10^6 c.p.m./ml) at 64 °C. The blot was washed twice for 15 min each with SSC (0.15 M NaCl/0.015 M sodium citrate) containing 0.5 % SDS at 64 °C. The bound ^{32}P-labelled probe was detected with a PhosphorImager (Bio-Rad). The blot was then stripped for re-use by using the Strip-EZ solutions provided with the transcription kit. A Human RNA Master Blot™ dot-blot (Clontech) containing normalized loadings of poly(A)$^+$ RNA from 50 different human tissues was processed as above, except that the ExpressHyb™ hybridization buffer provided with the blot was used.

Immunoblotting

The Trx and His tags were cleaved from the fusion protein with thrombin and the resulting YSA1H protein with a residual N-terminal S-tag was used to raise a rabbit polyclonal antiserum by standard procedures. Cell lysates were separated by SDS/PAGE [13 % (w/v) gel]. The gels were then equilibrated immediately in transfer buffer [10 mM Caps/NaOH (pH 11.0)/10 % (v/v)

methanol] for at least 10 min before electrophoretic transfer of the separated proteins to a nitrocellulose membrane at 150 mA and 4 °C for 2 h. The membrane was blocked overnight at 4 °C with 3 % (w/v) fat-free powdered milk and 0.2 % (v/v) Tween 20 in PBS and then probed with a 1:1000 dilution of whole anti-YSA1H antiserum followed by a 1:5000 dilution of horseradish peroxidase-conjugated goat anti-rabbit IgG (Bio-Rad). After being washed, the membrane was developed with the peroxidase substrate diaminobenzidine tetrahydrochloride (Sigma).

Other methods

Protein concentrations were estimated by the Coomassie Blue binding method [20]. N-terminal sequencing and electrospray mass spectrometric analysis were performed as described previously [6,18].

RESULTS

Cloning, expression and purification of YSA1H

A BLAST 2.0 search of the GenBank EST database with the *S. cerevisiae* YSA1 sequence yielded a large number of overlapping homologous human clones that, when assembled, revealed a 973-base sequence with a 657-base open reading frame (ORF) potentially encoding a 219-residue, 24.3 kDa protein, including the N-terminal methionine residue (Figure 1). This ORF had 63 % amino acid sequence identity with the 26.1 kDa YSA1 protein over a 69-residue region encompassing the MutT motif (see Figure 6). We have named this putative protein human YSA1 homologue, or YSA1H. That the predicted sequence represents the full-length protein is indicated by the presence of multiple stop codons both upstream and downstream of the ORF and by the highly favourable sequence context of the proposed initiator codon, AAAATGG (Figure 1) [21]. Human sequence-tagged sites corresponding to these ESTs have been mapped to chromosome 10p13–p14 between the markers D10S189 and D10S191 by the International Radiation Hybrid Mapping Consortium [22]. The UniGene entry number is Hs.11817.

```
1                                                   31
CGC CCA TCC TTT TAG CAC CGC GGG AGG CGC CGG TGT TTC GAG CCG TGG ACC GGA TCG GCT
                ***
61                                                  91
GAC ACT GCT GCC TCC AGC TAG TTA TTT CGT CCT CTT CCG TTC TTC ACC CCT ACA CCT TGG
                        ***
121                                                 151                 163/1
AGG TGA ACT TCT CAC CTG AGG GCT GTA AAG ACT CGT TTG AAA ATG GAG AGC CAA GAA CCA
    ***                                                     M   E   S   Q   E   P
181/7                                               211/17
ACG GAA TCT TCT CAG AAT GGC AAA CAG TAT ATC ATT TCA GAG GAG TTA ATT TCA GAA GGA
T   E   S   S   Q   N   G   K   Q   Y   I   I   S   E   E   L   I   S   E   G
241/27                                              271/37
AAA TGG GTC AAG CTT GAA AAA ACA ACG TAC ATG GAT CCT ACT GGT AAA ACT AGA ACT TGG
K   W   V   K   L   E   K   T   T   Y   M   D   P   T   G   K   T   R   T   W
301/47                                              331/57
GAA TCA GTG AAA CGT ACA ACC AGG AAA GAG CAG ACT GCG GAT GGT GTC GCG GTC ATC CCC
E   S   V   K   R   T   T   R   K   E   Q   T   A   D   G   V   A   V   I   P
361/67                                              391/77
GTG CTG CAG AGA ACA CTT CAC TAT GAG TGT ATC GTT CTG GTG AAA CAG TTC CGA CCA CCA
V   L   Q   R   T   L   H   Y   E   C   I   V   L   V   K   Q   F   R   P   P
421/87                                              451/97
ATG GGG GGC TAC TGC ATA GAG TTC CCT GCA GGT CTC ATA GAT GAT GGT GAA ACC CCA GAA
M   G   G   Y   C   I   E   F   P   A   G   L   I   D   D   G   E   T   P   E
481/107                                             511/117
GCA GCT GCT CTC CGG GAG CTT GAA GAA GAA ACT GGC TAC AAA GGG GAC ATT GCC GAA TGT
A   A   A   L   R   E   L   E   E   E   T   G   Y   K   G   D   I   A   E   C
541/127                                             571/137
TCT CCA GCG GTC TGT ATG GAC CCA GGC TTG TCA AAC TGT ACT ATA CAC ATC GTG ACA GTC
S   P   A   V   C   M   D   P   G   L   S   N   C   T   I   H   I   V   T   V
601/147                                             631/157
ACC ATT AAC GGA GAT GAT GCC GAA AAC GCA AGG CCG AAG CCA AAG CCA GGG GAT GGA GAG
T   I   N   G   D   D   A   E   N   A   R   P   K   P   K   P   G   D   G   E
661/167                                             691/177
TTT GTG GAA GTC ATT TCT TTA CCC AAG AAT GAC CTG CTG CAG AGA CTT GAT GCT CTG GTA
F   V   E   V   I   S   L   P   K   N   D   L   L   Q   R   L   D   A   L   V
721/187                                             751/197
GCT GAA GAA CAT CTC ACA GTG GAC GCC AGG GTC TAT TCC TAC GCT CTA GCG CTG AAA CAT
A   E   E   H   L   T   V   D   A   R   V   Y   S   Y   A   L   A   L   K   H
781/207                                             811/217
GCA AAT GCA AAG CCA TTT GAA GTG CCC TTC TTG AAA TTT TAA GCC CAA ATA TGA CAC TGG
A   N   A   K   P   F   E   V   P   F   L   K   F   ***             ***
841/281                                             871/291
CCA TTT TTG TAA ACG AGA CCA CCA GGC CTT CTT CAC TAA GAC TTT GTA TTC AAC TTA GTT
            ***                                     ***
901                                                 931
TAA TGT AGA TTT GCC ATT AGC TTT TTC GTA AAA TAA AAG CAC AGA ACA GAA AAA AAA AAA
***                                                 ***
961
AAA AAA AAA AAA A
```

Figure 1 Nucleotide sequence of the YSA1H cDNA and predicted amino acid sequence of the YSA1H protein

In-frame stop codons upstream and downstream of the YSA1H ORF are indicated with three asterisks. The presumed initiating ATG in its strong initiating context AAAATGG is underlined with dashes. The MutT motif is underlined; the invariant amino acids are shown in bold. The likely polyadenylation signal AAATAAAA is doubly underlined.

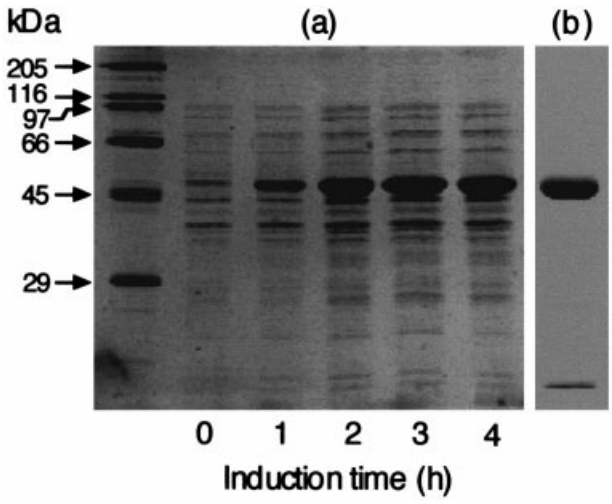

Figure 2 Overexpression and purification of Trx–YSA1H fusion protein

(**a**) *E. coli* BL-21(DE3) cells transformed with pET-YSA1H were induced with 1 mM IPTG for up to 4 h. Aliquots were taken at hourly intervals, boiled in sample buffer, analysed by SDS/PAGE [13 % (w/v) gel] and stained with Coomassie Blue. (**b**) YSA1H protein after purification on Ni^{2+}/NTA-agarose as described in the Experimental section. The protein standards were myosin (205 kDa), β-galactosidase (116 kDa), phosphorylase *b* (97 kDa), BSA (66 kDa), ovalbumin (45 kDa) and carbonic anhydrase (29 kDa).

A YSA1H cDNA was amplified by PCR from I.M.A.G.E. clone 310860, which was predicted to contain the full-length cDNA. For expression, the PCR fragment was inserted into the pET32b(+) vector and the recombinant plasmid (pET-YSA1H) was then used to transform *E. coli* BL-21(DE3) cells to generate a His-tagged thioredoxin fusion protein of predicted mass 41.6 kDa. After induction with IPTG, cell samples were analysed by SDS/PAGE, which showed that the soluble fraction contained a major IPTG-inducible protein band migrating with an apparent molecular mass of 47 kDa, which was therefore presumed to be the desired product (Figure 2a). The expressed Trx–YSA1H fusion protein was then purified in a single step to apparent homogeneity on a Ni^{2+}/NTA-agarose column (Figure 2b).

Substrate specificity

All MutT motif proteins studied so far have nucleotide pyrophosphatase activity that hydrolyses compounds containing an NDP-linked to another moiety. Therefore a range of nucleotides was assayed to determine the substrate specificity of Trx–YSA1H. In the presence of 5 mM Mg^{2+} ions, Trx–YSA1H was most active against ADP-sugars and some compounds of general structure ADP-sugar-X, e.g. Ap_2A and NADPH (Table 1). Lower activity was obtained with other NDP-sugars and CDP-alcohols, whereas no activity was obtained with (deoxy)nucleoside 5′-triphosphates, nucleoside 5′-diphosphates or nucleoside 5′-monophosphates or diadenosine polyphosphates, even when the enzyme concentration was increased

Table 1 Substrate specificities of YSA1H, human erythrocyte ADP-ribose pyrophosphatase, rat liver ADP-ribase II and calf liver ADP-sugar hydrolase

Results are expressed as the percentage activity obtained with ADP-ribose under the assay conditions used. Values in columns three to six are published (references in square brackets) and are based on different assay conditions from those used for YSA1H. Values for YSA1H are derived from assays containing 5 mM Mg^{2+} ions and are the averages of triplicate determinations. Abbreviation: n.t., not tested.

	Activity (% of value with ADP-ribose)				
		Human ADP-ribose pyrophosphatase			
Substrate	YSA1H	Erythrocyte [15]	Erythrocyte [12]	Rat liver ADP-ribase II [16]	Calf liver ADP-sugar hydrolase [17]
ADP-ribose	100	100	100	100	100
ADP-mannose	103	70	n.t.	n.t.	163
ADP-glucose	56	21	0	25	170
cADP-ribose	0	n.t.	0	n.t.	n.t.
Ap_2A	20	n.t.	n.t.	31	n.t.
NADPH	12	0	16	n.t.	n.t.
NADH	7	0	21	30	n.t.
NAD^+	7	0	42	< 2	5
Deamino-NAD^+	7	n.t.	n.t.	n.t.	n.t.
$NADP^+$	5	0	2	n.t.	n.t.
IDP-ribose	9	59	n.t.	31	n.t.
GDP-glucose	7	0	n.t.	n.t.	0
GDP-mannose	5	40	n.t.	< 2	n.t.
GDP-α-fucose	2	n.t.	n.t.	n.t.	n.t.
UDP-glucose	2	5	n.t.	< 2	0
UDP-galactose	7	0	n.t.	n.t.	n.t.
UDP-*N*-acetylglucosamine	< 1	n.t.	n.t.	n.t.	n.t.
CDP-glucose	3	5	n.t.	< 2	0
CDP-glycerol	3	n.t.	n.t.	2	n.t.
CDP-choline	< 1	n.t.	n.t.	< 2	n.t.
CDP-ethanolamine	< 1	n.t.	n.t.	< 2	n.t.
CoA	3	n.t.	n.t.	n.t.	n.t.
FAD	3	n.t.	n.t.	32	5
Ap_nA ($n = 3$–6)	0	0	n.t.	n.t.	n.t.
(d)NTP, (d)NDP, (d)NMP	0	0	n.t.	0	n.t.

20-fold. No activity was observed with thymidine-5′-monophospho-*p*-nitrophenyl ester as substrate; therefore YSA1H is devoid of alkaline phosphodiesterase 1 activity (EC 3.1.4.1). Comparative data are also presented in Table 1 for the human erythrocyte ADP-ribose pyrophosphatase, rat liver ADP-ribase II and calf liver ADP-sugar hydrolase. When assayed with 5 mM Mn^{2+} in place of Mg^{2+}, rat liver ADP-ribase II shows greatly increased activity with UDP-sugars and CDP-sugars and CDP-alcohols [16]. YSA1H activity towards CDP-glycerol was increased 4-fold when Mn^{2+} replaced Mg^{2+}, in comparison with the 50-fold increase observed with ADP-ribase II, whereas the activity of YSA1H towards ADP-sugars was decreased by 50–70% with Mn^{2+}. The UV-absorbing nucleotide reaction products generated from NDP-sugars were identified by HPLC and in each case were the corresponding NMP (Figure 3). The other product was assumed to be the sugar 5-phosphate.

Kinetic parameters and reaction requirements

Kinetic parameters were determined with ADP-ribose as substrate. K_m and V_{max} values were 60 μM and 8 μmol/min per mg respectively. The latter figure corresponds to a k_{cat} of 5.5 s^{-1}. With ADP-ribose as substrate, the enzyme displayed optimal activity at alkaline pH (7.4–9.0) with 2.5–5 mM Mg^{2+} or 100–250 μM Mn^{2+} ions (results not shown). In common with other MutT motif proteins, fluoride was inhibitory, with an IC_{50} of approx. 20 μM.

Tissue specificity of YSA1H mRNA expression

Northern blot and dot-blot analyses were performed to determine respectively the size and tissue distribution of YSA1H mRNA transcripts. A transcript of approx. 1.4 kb was detected in all human tissues present on the Northern blot (Figure 4a). A larger but less prominent transcript of approx. 5 kb was also detected, which might represent residual pre-mRNA as well as a smaller 1.1 kb species that was most prominent in placenta. Expression seemed stronger in lung, small intestine and stomach than in muscle, testis or placenta, but the blot was not normalized for differential rates of expression. A better comparison of tissue expression was obtained by dot-blot analysis with a blot that had been normalized with respect to eight different housekeeping genes. This showed that YSA1H was expressed in all 50 different human tissues present on the blot (Figure 4b). Although the amount of mRNA does not necessarily reflect expression at the protein level, such widespread expression is consistent with a role as a housecleaning enzyme required in all tissues. The level of mRNA expression was, however, higher in liver, kidney, pituitary, placenta and thymus than in other adult tissues and lowest in bone marrow, lymph nodes and certain areas of the brain.

Size and presence of YSA1H protein in human cell extracts

On gel-filtration analysis, the 41.6 kDa Trx–YSA1H fusion protein had a molecular mass of 85 kDa, suggesting that it

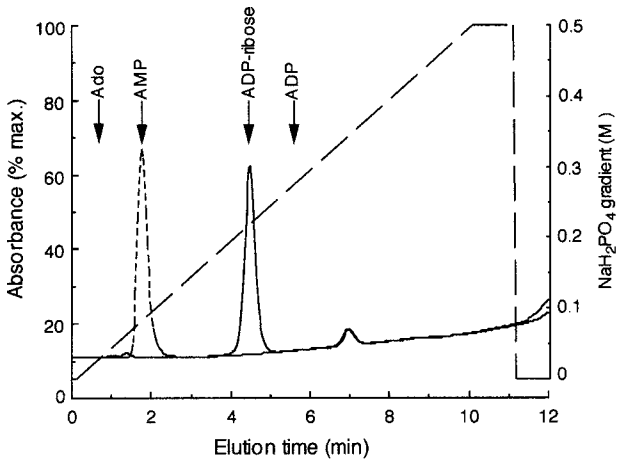

Figure 3 Identification of AMP reaction product from hydrolysis of ADP-ribose

Assays containing 0.3 mM ADP-ribose were incubated with or without 3 μg of Trx–YSA1H fusion protein for 30 min, then applied to a 1 ml Resource-Q anion-exchange column as described in the Experimental section. The positions of standards are indicated. The lines are as follows: solid line, without enzyme; short broken line, with enzyme; long broken line, gradient.

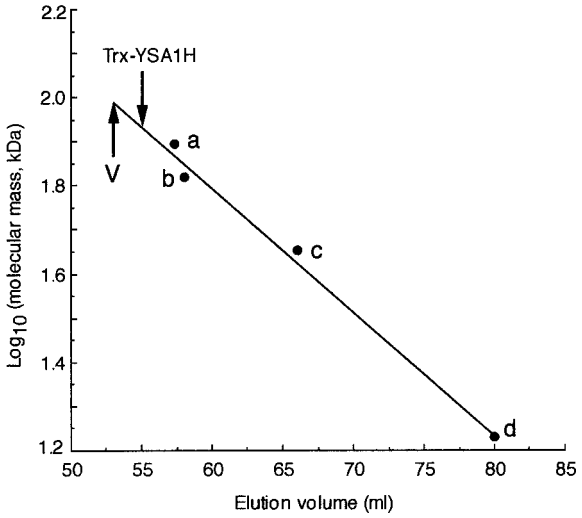

Figure 5 Gel-filtration analysis of purified recombinant Trx–YSA1H fusion protein

A sample of homogeneous Trx–YSA1H fusion protein was applied to a HiLoad 16/60 Superdex 75 column in 50 mM Tris/HCl (pH 7.5)/0.1 M NaCl and eluted at 1 ml/min in the same buffer. The column was calibrated with the following standards: a, conalbumin (77 kDa); b, BSA (66 kDa); c, ovalbumin (45 kDa); d, myoglobin (17 kDa). The void volume (*V*) and the elution position of Trx–YSA1H detected by its UV absorbance are indicated by arrows.

behaved as a dimer in solution (Figure 5). To show that the YSA1H protein was actually expressed in human cells, a rabbit polyclonal antibody was raised against recombinant YSA1H from which most of the vector fusion sequences had been removed by treatment with thrombin. Figure 6(a) shows an immunoblot analysis of the complete Trx–YSA1H fusion protein (lane 1) and after cleavage with thrombin, which yielded a product of theoretical molecular mass 27.5 kDa (lane 2), and enterokinase, which produced a near-native protein with a single alanine residue N-terminal to the initiating methionine residue (lane 3). Although the predicted molecular mass of the recombinant alanyl-YSA1H was 24.4 kDa, it ran on SDS/PAGE with

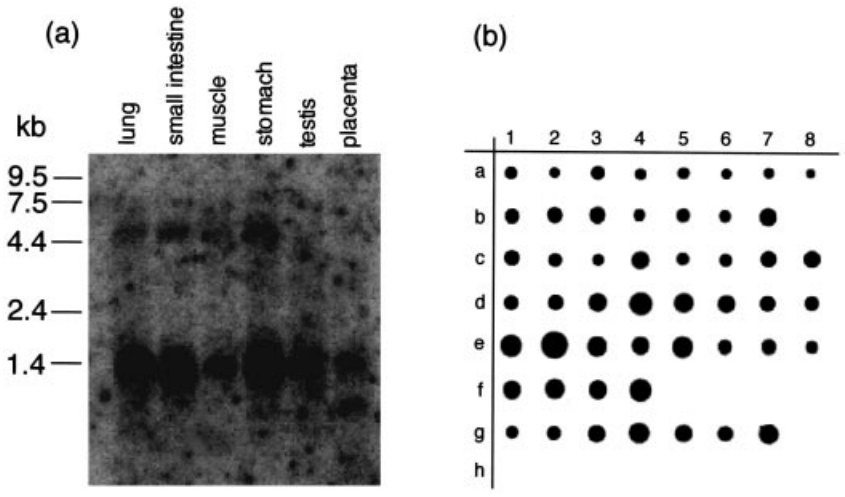

Figure 4 Expression of YSA1H mRNA in different human tissues

(**a**) Northern blot of poly(A)$^+$ RNA from five human tissues (2 μg per lane) probed with ^{32}P-labelled YSA1H anti-sense RNA transcribed *in vitro* from I. M. A. G.E. clone 310860 with T3 RNA polymerase. (**b**) RNA dot-blot containing normalized loadings of 89–514 ng of each poly(A)$^+$ RNA per dot from 50 different human tissues and 100 ng each of six control RNA and DNA species (500 ng for dot h8). The blot was probed as in (**a**). Dot identification: a1, whole brain; a2, amygdala; a3, caudate nucleus; a4, cerebellum; a5, cerebral cortex; a6, frontal lobe; a7, hippocampus; a8, medulla oblongata; b1, occipital lobe; b2, putamen; b3, substantia nigra; b4, temporal lobe; b5, thalamus; b6, nucleus accumbeus; b7, spinal cord; c1, heart; c2, aorta; c3, skeletal muscle; c4, colon; c5, bladder; c6, uterus; c7, prostate; c8, stomach; d1, testis; d2, ovary; d3, pancreas; d4, pituitary gland; d5, adrenal gland; d6, thyroid gland; d7, salivary gland; d8, mammary gland; e1, kidney; e2, liver; e3, small intestine; e4, spleen; e5, thymus; e6, peripheral leucocytes; e7, lymph node; e8, bone marrow; f1, appendix; f2, lung; f3, trachea; f4, placenta; g1, foetal brain; g2, foetal heart; g3, foetal kidney; g4, foetal liver; g5, foetal spleen; g6, foetal thymus; g7, foetal lung; h1, yeast total RNA; h2, yeast tRNA; h3, *E. coli* rRNA; h4, *E. coli* DNA; h5, poly(rA); h6, human C_0t 1 DNA; h7, human DNA; h8, human DNA.

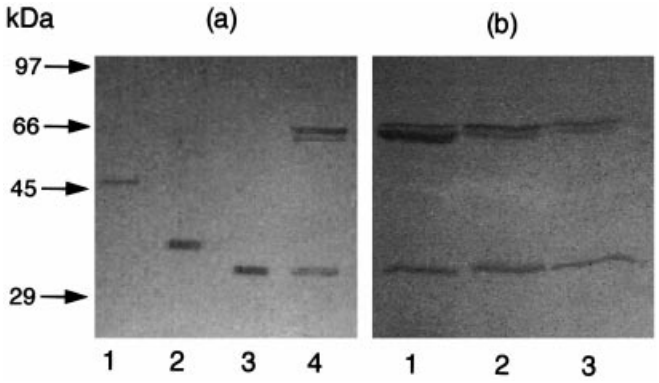

Figure 6 Immunoblot analysis of YSA1H protein expression in human cells

DAB-stained immunoblots were probed with anti-YSA1H as described in the Experimental section. (**a**) Lane 1, Trx–YSA1H fusion protein; lane 2, Trx–YSA1H cleaved with thrombin; lane 3, Trx–YSA1H cleaved with enterokinase; lane 4, KB cell extract. (**b**) Lane 1, U937 cell extract; lane 2, KB cell extract; lane 3, human erythrocyte extract.

an apparent molecular mass of 33 kDa. The N-terminal sequence of the enterokinase-cleaved recombinant protein was confirmed as AMESQEPT (including the vector alanine residue) by Edman degradation; electrospray mass spectrometric analysis of the same material confirmed its molecular mass as 24.4 kDa (results not shown). It therefore seems that recombinant YSA1H migrates anomalously on SDS/PAGE. This phenomenon of slower migration has been observed with other MutT motif proteins, including human Ap_4A hydrolase [6] and *S. cerevisiae* Ap_6A hydrolase [18] and might result from incomplete denaturation and SDS binding during sample preparation owing to the stability imparted by the mixed β-sheet core structure that might be a feature of this enzyme family [23].

The same species of apparent molecular mass 33 kDa was observed after examination of whole cell lysates prepared from human KB nasopharyngeal carcinoma cells (Figure 6a, lane 4, and Figure 6b, lane 2), human U937 promonocytic cells (Figure 6b, lane 1) and human erythrocytes (Figure 6b, lane 3). In addition, two extra bands of approx. 66 and 64 kDa were observed in the cell extracts. These might have been dimers resulting from incomplete sample denaturation; alternatively, they might have been unrelated proteins. We have previously identified two abundant proteins that cross-react strongly with an affinity-purified polyclonal antibody raised against another recombinant MutT motif protein, human Ap_4A hydrolase, as hsp60 (63 kDa) and glutamate dehydrogenase (61 kDa) (J. Kyte, J. L. Cartwright and A. G. McLennan, unpublished work). The purified human erythrocyte ADP-ribose pyrophosphatase has been reported to be a dimer of 34 kDa subunits [15]. If YSA1H is the same enzyme, our results would suggest that it is in fact a dimer of 24.3 kDa subunits.

Sequence comparisons

Compilation of mouse EST data revealed a 218-residue 24 kDa mouse homologue, mYSA1H, with overall 83% amino acid identity and 95% similarity to the human sequence. In addition to the *S. cerevisiae* YSA1 gene, closely related sequences are also present in rat and *Caenorhabditis elegans* (Figure 7). Significant matches to sequences from the prokaryotes *Treponema pallidum*, *Bacillus subtilis* (YQKG), *Synechocystis* sp. and to the *E. coli* YRFE gene (*orf186*) were also obtained (Figure 7). The *B. subtilis* YQKG gene has been reported to encode an ADP-ribose pyrophosphatase [10]; the *E. coli orf*186 gene product has a broad substrate specificity including ADP-ribose, NADH and Ap_3A [10]. The highly specific ADP-ribose pyrophosphatase from the archaeon *Methanococcus jannaschii* [13] and the *E. coli orf*209 gene product (YQIE, *rdsA*), which has also been reported to be an ADP-ribose pyrophosphatase [10], show only limited

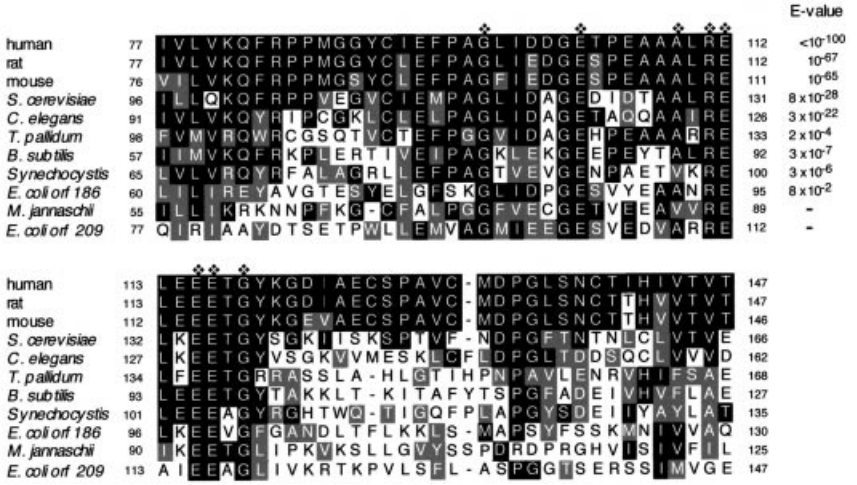

Figure 7 Multiple sequence alignment of YSA1H and related sequences

A 69-residue section of the YSA1H sequence encompassing the MutT motif was aligned by using the CLUSTAL W program with sequences from rat (AA858886), mouse (AI508660), *Saccharomyces cerevisiae* YSA1 (Q01976), *Caenorhabditis elegans* (AF067946), *Treponema pallidum* (O83713), *Bacillus subtilis* YQKG (P54570) and *Synechocystis* sp. (P72646), and the *Escherichia coli orf186* (P45799), *Methanococcus jannaschii* ADP-ribose pyrophosphatase (Q58549) and *E. coli orf209* (P36651) sequences; the numbers in parentheses are the accession numbers for these sequences. Amino acid identities with the human sequence are shaded black and similarities are shaded grey. 'Invariant' amino acids of the MutT motif are shown by ❖. The numbers to the right and left of the sequences are the amino acid positions in each complete sequence. Numbers in the right-hand column are the *E* values for these sequences from a BLAST 2.0 search of the GenBank EST and non-redundant databases, with the complete YSA1H as query sequence. The *E* value is a significance threshold representing the number of hits expected by chance when searching the database with the query sequence.

similarity to YSA1H overall but do show similarities in the region surrounding the MutT motif, in addition to the invariant amino acid residues of the motif itself, which might be involved in determining substrate specificity.

DISCUSSION

One mitochondrial and three cytosolic ADP-ribose pyrophosphatases have previously been partly purified from rat liver. The cytosolic ADP-ribase I and the mitochondrial enzyme are very similar in terms of low (micromolar) K_m for ADP-ribose and a stringent substrate specificity limited to ADP-ribose and IDP-ribose [16,24]. A similar activity has also been reported in embryonic cysts of the brine shrimp *Artemia franciscana* [25]. The cytosolic Mn^{2+}-dependent ADP-ribase has no activity at all with Mg^{2+} ions [16]. Therefore YSA1H does not seem to correspond to any of these enzymes. Its marked preference for ADP-sugar substrates indicates a closer similarity to the ADP-sugar pyrophosphatase activity detected in various mammalian liver extracts, including calf and rabbit, although the range of substrates analysed with this enzyme is rather limited [17]. There are also similarities to rat liver ADP-ribase II and the human erythrocyte ADP-ribose pyrophosphatase. These enzymes have a relatively high K_m for ADP-ribose (50–100 μM for rat, 170 μM for human erythrocyte and 60 μM for YSA1H) [15,16], whereas the native molecular mass of 47 kDa reported for rat ADP-ribase II [16] corresponds well to a dimer of 24.3 kDa subunits, the known subunit molecular mass of YSA1H. Apparent differences in substrate specificity include: (1) the high activity of the human erythrocyte enzyme with IDP-ribose and GDP-mannose [15]; (2) the lack of activity of the human erythrocyte enzyme with ADP-glucose [12] and (3) the high activity of the other enzymes with nicotinamide adenine dinucleotides. In the last case, significant contamination of commercial nucleotides with ADP-ribose was a likely cause of variation in results [16]. Because significant differences are apparent in the reported data for the same enzyme for human erythrocytes [12,15] and because rat ADP-ribase II exhibits a much broader specificity for NDP-sugars and nucleoside 5′-diphosphoalcohols in the presence of Mn^{2+} ions [16], differences in assay procedures and conditions might have a large role in the observed differences. Taken together, these results suggest that YSA1H, the human erythrocyte ADP-ribose pyrophosphatase, rat liver ADP-ribase II and mammalian liver ADP-sugar pyrophosphatase might be the same enzyme. However, this conclusion requires confirmation.

Fluoride is a potent inhibitor of several nudix hydrolases involved in diadenosine polyphosphate metabolism [7,18,26]. Because it also inhibits YSA1H with an IC_{50} of 20 μM, such inhibition might be a more general feature of this protein family. Because the specific rat liver and *Artemia* ADP-ribase I enzymes are similarly inhibited by fluoride [25], it is possible that they too are nudix hydrolases related to the specific *M. jannaschii* enzyme. The very stringent substrate specificity of these last three enzymes suggests that activities of this type might form the first line of defence against protein glycation by free ADP-ribose. However, ADP-sugars other than ADP-ribose are not significant metabolites in mammalian cells. ADP-glucose is an important precursor for bacterial glycogen and plant starch synthesis [27], whereas ADP-mannose has no known physiological function, although the commercially available synthetic compound can replace ADP-heptoses in bacterial outer-membrane lipopolysaccharide synthesis *in vitro* [28]. ADP-ribose might therefore also be the most important substrate for human YSA1H *in vivo*. Of the other potential substrates, Ap_2A is the most recent of the diadenosine polyphosphates to be identified *in vivo*. It is present in specific granules in human myocardium and can attain an intragranular concentration of 9 mM [29]. When released extracellularly, it acts as a coronary vasodilator. Within the granules, however, it might not be accessible to YSA1H and so might not be a significant substrate *in vivo*.

In conclusion, this study describes the biochemical properties of a new human member, YSA1H, of the MutT motif family of nucleotide pyrophosphatases and suggests a possible function. However, unequivocal confirmation of this function must await the results of genetic knock-out studies in yeast, *C. elegans* and mouse.

We thank M. C. Wilkinson for N-terminal sequencing and M. C. Prescott for mass spectrometric analysis. This work was supported by project grant 053038 from the Wellcome Trust and grant no. F25BJ from the Leverhulme Trust.

REFERENCES

1 Maki, H. and Sekiguchi, M. (1992) Nature (London) **355**, 273–275
2 Koonin, E. V. (1993) Nucleic Acids Res. **21**, 4847
3 Bessman, M. J., Frick, D. N. and O'Handley, S. F. (1996) J. Biol. Chem. **271**, 25059–25062
4 McLennan, A. G. (1999) Int. J. Mol. Med. **4**, 79–89
5 Sakumi, K., Furuichi, M., Tzuzuki, T., Kakuma, T., Kawabata, S.-I., Maki, H. and Sekiguchi, M. (1993) J. Biol. Chem. **268**, 23524–23530
6 Thorne, N. M. H., Hankin, S., Wilkinson, M. C., Nunez, C., Barraclough, R. and McLennan, A. G. (1995) Biochem. J. **311**, 717–721
7 Safrany, S. T., Caffrey, J. J., Yang, X. N., Bembenek, M. E., Moyer, M. B., Burkhart, W. A. and Shears, S. B. (1998) EMBO J. **17**, 6599–6607
8 Safrany, S. T., Ingram, S. W., Cartwright, J. L., Falck, J. R., McLennan, A. G., Barnes, L. D. and Shears, S. B. (1999) J. Biol. Chem. **274**, 21735–21740
9 Li, A. W., Too, C. K. L. and Murphy, P. R. (1996) Biochem. Biophys. Res. Commun. **223**, 19–23
10 O'Handley, S. F., Frick, D. N., Dunn, C. A. and Bessman, M. J. (1998) J. Biol. Chem. **273**, 3192–3197
11 Moss, J. and Vaughan, M. (1990) ADP-ribosylating Toxins and G Proteins, American Society for Microbiology, Washington, DC
12 Zocchi, E., Guida, L., Franco, L., Silvestro, L., Guerinni, M., Benatti, U. and De Flora, A. (1993) Biochem. J. **295**, 121–130
13 Sheikh, S., O'Handley, S. F., Dunn, C. A. and Bessman, M. J. (1998) J. Biol. Chem. **273**, 20924–20928
14 Kwak, Y.-G., Park, S.-K., Kim, U.-H., Han, M.-K., Eun, J.-S., Cho, K.-P. and Chae, S.-W. (1996) Am. J. Physiol. **271**, C464–C468
15 Kim, J. S., Kim, W. Y., Rho, H. W., Park, J. W., Park, B. H., Han, M. K., Kim, U. H. and Kim, H. R. (1998) Int. J. Biochem. Cell Biol. **30**, 629–638
16 Canales, J., Pinto, R. M., Costas, M. J., Hernández, M. T., Miró, A., Bernet, D., Fernández, A. and Cameselle, J. C. (1995) Biochim. Biophys. Acta **1246**, 167–177
17 Rodriguez, P., Bass, S. T. and Hansen, R. G. (1968) Biochim. Biophys. Acta **167**, 199–201
18 Cartwright, J. L. and McLennan, A. G. (1999) J. Biol. Chem. **274**, 8604–8610
19 Gasmi, L., Cartwright, J. L. and McLennan, A. G. (1998) Biochim. Biophys. Acta **1405**, 121–127
20 Peterson, G. L. (1982) Methods Enzymol. **91**, 95–119
21 Kozak, M. (1996) Mammal. Genome **7**, 563–574
22 Deloukas, P., Schuler, G. D., Gyapay, G., Beasley, E. M., Soderlund, C., Rodriguez-Tomé, P., Hui, L., Matise, T. C., McKusick, K. B., Beckmann, J. S. et al. (1998) Science **282**, 744–746
23 Abeygunawardana, C., Weber, D. J., Gittis, A. G., Frick, D. N., Lin, J., Miller, A. F., Bessman, M. J. and Mildvan, A. S. (1995) Biochemistry **34**, 14997–15005
24 Bernet, D., Pinto, R. M., Costas, M. J., Canales, J. and Cameselle, J. C. (1994) Biochem. J. **299**, 679–682
25 Fernández, A., Ribeiro, J. M., Costas, M. J., Pinto, R. M., Canales, J. and Cameselle, J. C. (1996) Biochim. Biophys. Acta **1290**, 121–127
26 Guranowski, A. (1990) FEBS Lett. **262**, 205–208
27 Smith-White, B. J. and Preiss, J. (1992) J. Mol. Evol. **34**, 449–464
28 Kadrmas, J. L., Brozek, K. A. and Raetz, C. R. H. (1996) J. Biol. Chem. **271**, 32119–32125
29 Luo, J., Jankowski, J., Knobloch, M., van der Giet, M., Gardanis, K., Russ, T., Vahlensieck, U., Neumann, J., Schmitz, W. et al. (1999) FASEB J. **13**, 695–705

Received 8 July 1999/20 August 1999; accepted 15 September 1999