RESEARCH COMMUNICATION

Human diadenosine 5′,5‴-P^1,P^4-tetraphosphate pyrophosphohydrolase is a member of the MutT family of nucleotide pyrophosphatases

N. Matthew H. THORNE, Steven HANKIN, Mark C. WILKINSON, Carmen NUÑEZ, Roger BARRACLOUGH and Alexander G. McLENNAN*

Department of Biochemistry, University of Liverpool, P.O. Box 147, Liverpool L69 3BX, U.K.

The cDNA and derived amino acid sequence of human diadenosine 5′,5‴-P^1,P^4-tetraphosphate pyrophosphohydrolase have been determined with the aid of the GenBank Expressed Sequence Tag database. This enzyme possesses a modification of the MutT sequence motif found in certain nucleotide pyrophosphatases. It is unrelated to the enzymes of diadenosine tetraphosphate catabolism found in prokaryotes and fungi.

INTRODUCTION

Diadenosine 5′,5‴-P^1,P^4-tetraphosphate (Ap_4A) is synthesized in all cells from ATP by aminoacyl-tRNA synthetases and is found at basal concentrations ranging from 0.05 to 5 μM [1,2]. Since the concentrations of Ap_4A and related nucleotides, such as diadenosine 5′,5‴-P^1,P^3-triphosphate (Ap_3A), increase after exposure of cells to various metabolic stresses, they have been implicated in the regulation of the cellular response to such stresses [3,4]. A possible role in DNA replication has also been proposed [5]. Alternatively, Ap_4A may be an inevitable by-product of the activation step of protein synthesis which must be degraded before it accumulates to levels that are either wasteful of ATP or are potentially toxic, through interaction with important adenine nucleotide-binding proteins such as protein kinases. Extracellularly, it is evident that Ap_4A and other diadenosine polyphosphates are important signalling molecules in the cardiovascular and nervous systems [6,7].

Intracellular accumulation is prevented (or regulation achieved) by specific catabolic enzymes [8]. All prokaryotes studied so far possess an Ap_4A (*symmetrical*) pyrophosphohydrolase ($Ap_4A \rightarrow ADP + ADP$; EC 3.6.1.41) of approx. 33 kDa. The *apaH* genes of *Escherichia coli* and *Klebsiella aerogenes* have been cloned and sequenced [9,10]. Among the lower eukaryotes, the acellular slime mould *Physarum polycephalum* also has a symmetrically-cleaving hydrolase [11]. However, the fission yeast *Schizosaccharomyces pombe* has an Ap_4A (*asymmetrical*) pyrophosphohydrolase ($Ap_4A \rightarrow AMP + ATP$; EC 3.6.1.17), which is a dimer of a 22 kDa polypeptide [12]. All other lower eukaryotes studied so far have a quite distinct enzyme, namely a reversible Ap_4A phosphorylase ($Ap_4A + P_i \leftrightarrow ATP + ADP$; EC 2.7.7.53) [8].

Saccharomyces cerevisiae has two phosphorylases, namely the 36–38 kDa *apa1* and *apa2* gene products, both of which have been cloned and found to share 60% sequence identity [13,14]. The *apa2* gene from *Kluyveromyces lactis* has also been sequenced [15]. Curiously, certain green algae, e.g. *Scenedesmus obliquus*, have an *asymmetrical* Ap_4A hydrolase in addition to Ap_4A phosphorylase [16].

In contrast to this variable pattern among lower eukaryotes, higher eukaryotes (vertebrates, invertebrates and plants) all appear to possess a single, 16–21 kDa Ap_4A (*asymmetrical*) pyrophosphohydrolase [8,17–19]. Here we report for the first time the cDNA and derived amino acid sequences of such an enzyme: the human Ap_4A hydrolase. Unlike any of the other enzymes it possesses a modification of the signature motif characteristic of the recently described 'MutT family' of nucleotide pyrophosphatases [20].

EXPERIMENTAL

Materials

Sequencing-grade endoproteinase Glu-C and trypsin were from Sigma. cDNA clone 108448 was obtained from the IMAGE Consortium of the Human Genome Center at the Lawrence Livermore National Laboratory, Livermore, CA, U.S.A. Two-dimensional-SDS/PAGE standards were from Bio-Rad. Sequenase sequencing kit was from Amersham International.

Sequencing of peptides

Ap_4A hydrolase was purified from human placenta by chromatography on DEAE-Sephacel, Ultrogel AcA44 and Procion Red-P3BN–Sepharose as previously described [17]. Activity was monitored using a luminometric assay [18]. The final fraction, containing about 40–80 μg of protein, was freed from minor contaminants by reversed-phase chromatography on a Bio-Rad HiPore C_4 column (250 mm × 4.6 mm) at 1 ml/min using the following solvent system: 0–40% B at 3%/min, 40–75% B at

Abbreviations used: Ap_4A, diadenosine 5′,5‴-P^1,P^4-tetraphosphate; Ap_3A, diadenosine 5′,5‴-P^1,P^3-triphosphate; 8-oxo-dGTP, 8-oxo-7,8-dihydro-2′-deoxyguanosine 5′-triphosphate; TFA, trifluoroacetic acid; ORF, open reading frame; EST, expressed sequence tag; TFA, trifluoroacetic acid; bFGF, basic fibroblast growth factor.

* To whom correspondence should be addressed.

The nucleotide sequence of the cDNA and upstream ORF of human Ap_4A hydrolase (Figure 2) has been deposited with the EMBL/GenBank/DDBJ Nucleotide Sequence Databases under the accession number U30313.

1%/min and 85% B for 10 min, where A is 0.15% trifluoroacetic acid (TFA) in 10% (v/v) acetonitrile and B is 0.1% TFA in acetonitrile. Ap_4A hydrolase, which retained activity, was eluted at 55% (v/v) acetonitrile and was concentrated by centrifugal evaporation.

Endoproteinase Glu-C peptides were generated from 2 μg of homogeneous Ap_4A hydrolase by SDS/PAGE of the hydrolase, isolation and dehydration of the protein-containing band and rehydration of the gel slice in the presence of Glu-C and 0.02% Tween-20 [21]. After overnight incubation at room temperature, peptides were eluted from the gel slice with three aliquots of 0.1% TFA/60% acetonitrile. Combined aliquots were concentrated to 20 μl by centrifugal evaporation and injected on to a narrow-bore 10 mm × 2.1 mm C_{18} reversed-phase cartridge column (Brownlee) and eluted at 0.3 ml/min as follows: 0–40% B for 60 min then 40–75% B for 30 min, where A is 0.15% TFA and B is 0.1% TFA/80% (v/v) acetonitrile. Tryptic peptides were generated by liquid-phase digestion [22] and separated as described above. All peptides were concentrated, dried on to Polybrene–GF/C discs then sequenced using an Applied Biosystems 471A protein sequencer.

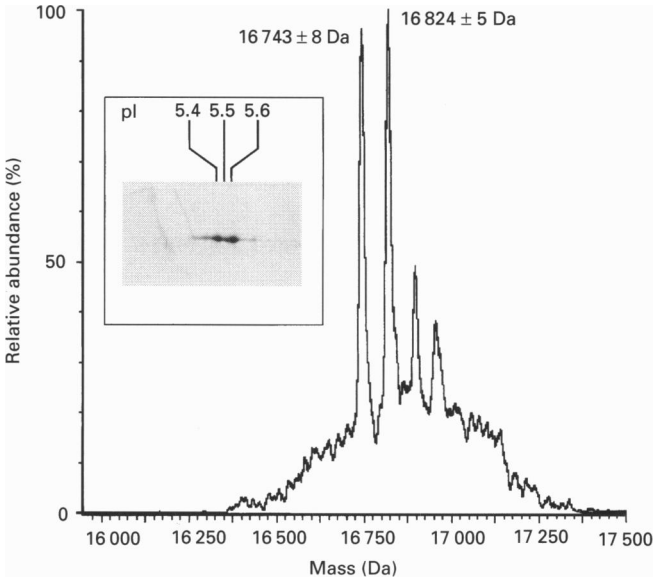

Figure 1 Electrospray MS and (inset) two-dimensional gel electrophoresis of human placental Ap_4A hydrolase

Pure enzyme (83 pmol) in 1% formic acid/50% acetonitrile was introduced at 10 μl/min by loop injection into a Waters 600MS HPLC system coupled to a VG Quattro mass spectrometer. The mass spectrometer was operated in the positive-ion electrospray mode. The operating parameters were: capillary voltage, 3.4 kV; high-voltage lens, 0.1 kV; focus, 30 V; skimmer, 35 V; and source temperature, 70 °C. An MCA scan was performed over the mass range 500–2000 Da with a scan time of 5 s. The raw data was processed to give molecular-mass information using VG MassLynx software. Pure enzyme (60 pmol) was also subjected to two-dimensional SDS/PAGE [48] (inset). The gel was calibrated with two-dimensional-SDS/PAGE standards.

DNA manipulation and sequencing

Clone 108448 was isolated at random from a normalized cDNA library [23] constructed from the liver and spleen of a 20-week post-conception male foetus. First-strand cDNA was primed with a *Pac*I–oligo(dT) primer (5′-AACTGGAAGAATTAATT-AAAGATCTTTTTTTTTTTTTTTTTT-3′). Double-stranded cDNA was ligated to *Eco*RI adaptors, digested with *Pac*I, cloned into the *Eco*RI and *Pac*I sites of a pT7T3 phagemid vector (Pharmacia, modified to include a *Pac*I site in the polylinker) with the 5′ end of the cDNA at the *Eco*RI site, and amplified in *E. coli* host strain DH10B. Clone 108448 was re-cloned in strain DH5a and DNA prepared by standard methods. Automated *Taq* cycle sequencing of both strands of the plasmid was performed twice on an Applied Biosystems 373 DNA sequencer using 1 μg of plasmid DNA and 20 ng each of three different sets of primers: (i) T3 (5′-AATTAACCCTCACTAAAGGG-3′) and M13 (5′-GTAAAACGACGGCCAGT-3′) vector primers; (ii) 5′-ACCATGGCCTTGAGAGCA-3′ and 5′-CCAGTCAGCTC-AGGCCTC-3′, which overlap the putative start and stop codons of the coding sequences identified in expressed sequence tags (ESTs) T77765 and T77766 respectively; and (iii) 5′-TGCAGG-CATCAGATGGCA-3′ and 5′–TTGAACTGAGCCAACTGG-3′, which are located towards the 3′ and 5′ ends respectively of the coding sequences identified in ESTs T77765 and T77766. The sequence of the coding region was confirmed three times by manual sequencing of both strands with primer sets (ii) and (iii) and a Sequenase kit.

RESULTS AND DISCUSSION

Homogeneous human placental Ap_4A hydrolase was found to comprise two major and two minor isoforms of approx. 18 kDa when analysed by two-dimensional gel electrophoresis, the two major species having pI values of 5.45 and 5.65 (Figure 1, inset). Since the human enzyme was previously sized at 17.5, 18 and 19.2 kDa [8,17], an accurate molecular mass was determined by electrospray MS to aid sequence identification. The spectrum revealed two major species with masses 16824±8 and 16743±5 Da (Figure 1). The higher values estimated by SDS/PAGE must be due to anomalous electrophoretic migration of the protein.

In order to determine the amino acid sequence of the enzyme, tryptic and Glu-C peptides were generated and sequenced. Tryptic peptides obtained were (i) LSHEHQAYR, (ii) ELNYVAR, (iii) TVIYXLAEVK, (iv) WLGLEEAXXLAQFK and (v) VDNNAIEFLLLQASD while Glu-C peptides were (vi) GHQFLXSIE, (vii) LNYVARNKPKTVI, (viii) FLLLQASDGIHH, (ix) XQLAQFKE and (x) AGQLTIIE. Overlaps suggested the existence of the following peptides: (xi) VDNNAIEFLLLQASDGIHH, (xii) WLGLEEAXQLAQFKE and (xiii) ELNYVARNKPKTVIYXLAEVK. A search of the SWISS-PROT and Brookhaven protein databases revealed no matches to any of these peptides. However, a search of the GenBank expressed sequence tag database (dbEST), using the NCBI BLAST server and program tblastn, revealed a perfect match to peptide (xi) in four single-pass partial cDNA sequences (GenBank accession numbers are given in parentheses): (a) human lung (T35530); (b) human eye (T30642); (c) human testis (T19141); and (d) human foetal liver + spleen (T77765). A match to peptide (xii) with X = C was also found in a human foetal liver + spleen EST (T77766). T77765 and T77766 represent the 5′ and 3′ ends respectively of the IMAGE clone 108448 (GDB ID: G00-464-065) from the Washington University–Merck EST project [24]. The origin and structure of this clone are described in the Experimental section.

Clone 108448 was obtained and its complete sequence determined (Figure 2). Two open reading frames (ORFs) are present. The first short ORF commencing at nucleotide 64 could potentially encode a 19-amino-acid peptide. Two further stop codons precede the second ORF, which commences with the ATG at nucleotide 175 and encodes a 147-amino-acid polypeptide of 16829 Da and predicted pI 5.26, containing all of the

```
1                                                   31
CCT CCT ACC TCC TTC TGC TTC GGT GCG TTT GCT TCT GAG GTT CTC CAG TGT CAC AAC AAA

61                                                  91
CAC ATG CCA GCC CTG TTT TAC AGG GAG CCC TGG AGG AGT TGG GAT AGA GGC CAC ATT GAC
    Met pro ala leu phe tyr arg glu pro trp arg ser trp asp arg gly his ile asp
121                                                 151                 175/1
TGA GGG TAG TTG CCA GGG TCC TGC AGT TAT ACA CAA AGT CCT TAG GAT AAG ACC ATG GCC
***     ***                                                 ***             Met ala
181/3                                               211/13
TTG AGA GCA TGT GGC TTG ATC ATC TTC CGA AGA TGC CTC ATT CCC AAA GTG GAC AAC AAT
leu arg ala cys gly leu ile ile phe arg arg cys leu ile pro lys val asp asn asn
241/23                                              271/33
GCA ATT GAG TTT TTA CTG CTG CAG GCA TCA GAT GGC ATT CAT CAC TGG ACT CCT CCC AAA
ala ile glu phe leu leu leu gln ala ser asp gly ile his his trp thr pro pro lys
301/43                                              331/53
GGC CAT GTG GAA CCA GGA GAG GAT GAC TTG GAA ACA GCC CTG AGG GAG ACC CAA GAG GAA
gly his val glu pro gly glu asp asp leu glu thr ala leu arg glu thr gln glu glu
361/63                                              391/73
GCA GGC ATA GAA GCA GGC CAG CTG ACC ATT ATT GAG GGG TTC AAA AGG GAA CTC AAT TAT
ala gly ile glu ala gly gln leu thr ile ile glu gly phe lys arg glu leu asn tyr
421/83                                              451/93
GTG GCC AGG AAC AAG CCT AAA ACA GTC ATT TAC TGG CTG GCG GAG GTG AAG GAC TAT GAC
val ala arg asn lys pro lys thr val ile tyr trp leu ala glu val lys asp tyr asp
481/103                                             511/113
GTG GAG ATC CGC CTC TCC CAT GAG CAC CAA GCC TAC CGC TGG CTG GGG CTG GAG GAG GCC
val glu ile arg leu ser his glu his gln ala tyr arg trp leu gly leu glu glu ala
541/123                                             571/133
TGC CAG TTG GCT CAG TTC AAG GAG ATG AAG GCA GCG CTC CAA GAA GGA CAC CAG TTT CTT
cys gln leu ala gln phe lys glu met lys ala ala leu gln glu gly his gln phe leu
601/143         613/147                             631
TGC TCC ATA GAG GCC TGA GCT GAC TGG AGC AGA GTC ATT TGC TTC AGC AGG ATC CTT GTG
cys ser ile glu ala ***
661                                                 691
GGC CTT CTA AGA TGA AGC CAC CCT CAG GTC CAG GGA AGG TTG TGC TGG TAT TTG GCT CAT
                ***
721                                                 751
GAC AGC CAA GAG CAG ATT TGT GAA ATC GGC TCA ACT CCC AGG TGA GAG CAA GCA AAA ATC
                                                        ***
781                                                 811
TTG GCT GGG TGG AAA GGA AGG CAA AAG AGT AAA AAT TAA AAA GGC CAG GCC CAA GTA AGT

841                                                 871
GTA CCT TGT ACT TTA TAA ATA AAC CTC AAG CAG CTC AAA AAA AAA AAA AAA A
```

Figure 2 cDNA and deduced amino acid sequence of human Ap_4A hydrolase and upstream ORF

Nucleotides are numbered from the 5′ end of the cDNA, and amino acids are numbered from the deduced N-termini of the ORFs. Stop codons are indicated by asterisks (***). Peptides identified by direct amino acid sequencing are underlined. DNA sequencing primers are dot-underlined. The putative polyadenylation signal is double-underlined. The amino acids of the MutT domain are in **bold**.

peptide sequences derived previously. Given the measured sizes and pI values of the major species in the enzyme preparation, this ORF clearly represents the human Ap_4A hydrolase. The smaller of the two species probably lacks the C-terminal alanine residue (78 Da). The initial ATG codon of this ORF is in the highly favourable initiation context ACC*ATG*G, while the ATG of the upstream ORF is in the unfavourable CAC*ATG*C context [25]. Short ORFs are often found upstream of eukaryotic genes and, given the poor initiation context of its ATG, this ORF is unlikely to be translated, although its presence may reduce the translational efficiency of the downstream ORF [25]. Near the 3′ end of the cDNA there is a typical polyadenylation signal (AATAAA).

A search of the BLOCKS database of ungapped conserved sequence segments [26] revealed that the derived amino acid sequence contains a previously unreported modification of the recently described MutT sequence motif GX_5EX_5[UA]XRE&XEE, where X is 'any amino acid', U is a

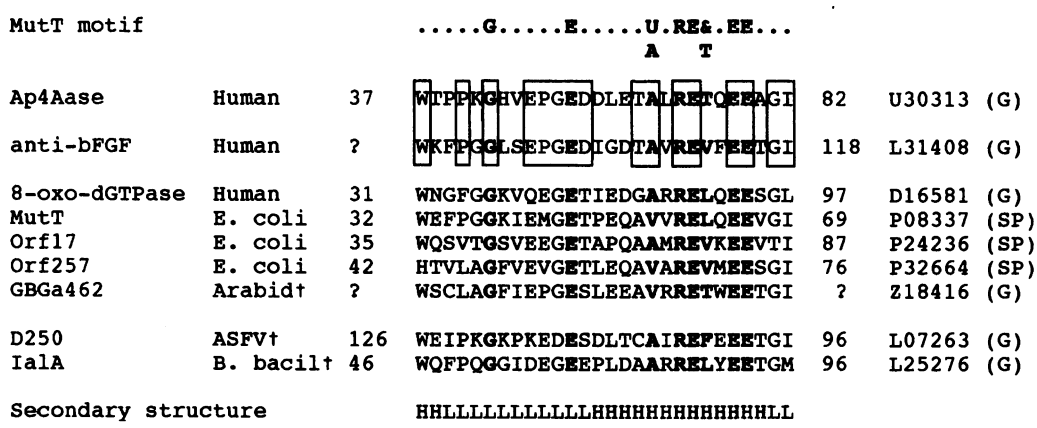

Figure 3 Sequence comparisons in the MutT family

The abbreviations used are: anti-bFGF, product of antisense bFGF mRNA; Arabid, *Arabidopsis thaliana*; ASFV, African-swine-fever virus; B. bacil., *Bartonella bacilliformis*; E. coli, *Escherichia coli*; 8-oxo-dGTPase, 8-oxo-dGTP pyrophosphohydrolase. Numbers to the right and left of sequences indicate numbers of amino acids on N- and C-terminal sides respectively. The notation '&' in the consensus MutT motif indicates a bulky hydrophobic residue. Accession numbers are GenBank (G) or SWISS-PROT (SP). Sequence identities between human Ap_4A hydrolase (Ap_4Aase) and anti-bFGF are boxed. Secondary structure [20]: H, α-helix, L, loop.

bulky aliphatic residue and & is a bulky hydrophobic residue [20]. The human Ap_4A hydrolase has T in place of & (Figure 3). The *E. coli* protein MutT is a nucleoside triphosphate pyrophosphohydrolase with a strong preference for the potentially mutagenic substrate 8-oxo-7,8-dihydro-2′-deoxyguanosine 5′-triphosphate (8-oxo-dGTP) [27]. Genes for corresponding proteins have been cloned and sequenced from other organisms, including humans, and possess the same motif [28,29]. Furthermore, a number of other proteins share this motif, including two pyrophosphohydrolases from *E. coli*, Orf17 (which prefers dATP) and Orf257 (which prefers NADH, but also degrades Ap_4A [28,30,31]), several related viral gene products, e.g. the putative D250 protein of African-swine-fever virus [20,32], a putative protein encoded by the antisense RNA of human and *Xenopus laevis* basic fibroblast growth factor (bFGF) [33,34] and the putative invasion protein IalA of *Bartonella bacilliformis* [35] (Figure 3). These sequences were aligned using the multiple sequence alignment program CLUSTAL W [36]. Between residues W38 and I65, the Ap_4A hydrolase shares 43% sequence identity with the *E. coli* MutT, Orf17, Orf257 and human 8-oxo-dGTPase proteins and 56% identity with the human bFGF antisense product.

The occurrence of T in place of & in the MutT domain is not unique to human Ap_4A hydrolase. The *Arabidopsis* EST GBGa462 also has T in place of & (Figure 3), suggesting that a hydrophobic residue in this position is not as essential as has been suggested from NOESY of MutT–nucleotide complexes [37]. It remains to be determined whether this EST also encodes an Ap_4A hydrolase.

NMR secondary-structure measurements for the conserved domain in MutT have shown that it comprises an α-helix flanked by two loops [38,39]. A predictive analysis [40,41] of the Ap_4A hydrolase sequence shows that the sequence ^{55}ALRETQEEAG^{64} is solvent-accessible and has a high probability (71%) of being α-helical, with the flanking sequences in loops. Hence, as in MutT, this sequence is probably an important part of the substrate-binding and/or catalytic site of the enzyme. Classification of Ap_4A hydrolase as a MutT domain protein shows that this domain is not in itself responsible for the unusual attack on the β-phosphorus of the polyphosphate chain observed with the *E. coli* MutT protein and 8-oxo-dGTP [42]. Previous studies have shown that water attacks the α-phosphorus atoms in Ap_4A [43,44].

No other motifs common to Ap_4A hydrolase and the other MutT domain-containing proteins as a group were found by BLOCKS or CLUSTALW. However, a comparison of the Ap_4A hydrolase and the *E. coli* MutT protein with the PROTOMAT block maker using both the MOTIF and the Gibbs algorithms to search for common sequences [45] revealed two other homology blocks: (i) ^{80}LNYVARNKPKTVIYWLAE^{97}, with 33% identity with (exact correspondence), and 72% similarity to (conservative substitution), the MutT sequence ^{71}LEYEFPDRHITLWFW-LVE^{88}; and (ii) ^{98}EHQAYRWLGL^{107}, with 40% identity with, and 50% similarity to, the MutT sequence ^{110}EGQPGEW-MSL^{119}. As matches are considered to be significant when detected by both algorithms used by PROTOMAT [45], this strengthens the relationship between these two enzymes.

Similar sequence comparisons with the prokaryotic symmetrically cleaving Ap_4A hydrolases and fungal Ap_4A phosphorylases showed no significant similarities to the human Ap_4A hydrolase. Neither of these Ap_4A-degrading enzymes has a MutT domain. It appears that prokaryotes, lower eukaryotes and higher eukaryotes have arrived at separate solutions for the elimination and/or regulation of Ap_4A. Thus, it will clearly be of interest to determine the sequences of the specific Ap_3A hydrolases found in lower and higher eukaryotes [8,19,46,47]. The availability of sequence information on the human Ap_4A hydrolase will also permit its relationship to ectoenzymes involved in extracellular dinucleoside polyphosphate metabolism to be determined.

Finally, the present study highlights the value of the rapidly expanding EST database in aiding the identification of new protein sequences.

This work was supported by a grant from the Biotechnology and Biological Sciences Research Council (BBSRC). S.H. was a recipient of a BBSRC studentship. The technical assistance of E. Mayers, M. C. Prescott and A. Benfold is gratefully acknowledged.

REFERENCES

1 Brevet, A., Chen, J., Lévêque, F., Plateau, P. and Blanquet, S. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 8275–8279

2 Garrison, P. N. and Barnes, L. D. (1992) in Ap_4A and Other Dinucleoside Polyphosphates (McLennan, A. G., ed.), pp. 29–61, CRC Press, Boca Raton, FL
3 Baker, J. and Jacobson, M. K. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 2350–2352
4 Farr, S. B., Arnosti, D. N., Chamberlin, M. J. and Ames, B. N. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 5010–5014
5 Baxi, M. D., McLennan, A. G. and Vishwanatha, J. K. (1994) Biochemistry **33**, 14601–14607
6 Schlüter, H., Offers, E., Brüggeman, G. et al. (1994) Nature (London) **367**, 186–188
7 Pintor, J. and Miras-Portugal, M. T. (1993) Drug Dev. Res. **28**, 259–262
8 Guranowski, A. and Sillero, A. (1992) in Ap_4A and Other Dinucleoside Polyphosphates (McLennan, A. G., ed.), pp. 81–133, CRC Press, Boca Raton, FL
9 Mechulam. Y., Fromant, M., Mellot, P. et al. (1985) J. Bacteriol. **164**, 63-69
10 Azakami, H., Sugino, H. and Murooka, Y. (1992) J. Bacteriol. **174**, 2344–2351
11 Barnes, L. D. and Culver, C. A. (1982) Biochemistry **21**, 6123–6128
12 Robinson, A. K., de la Peña, C. E. and Barnes, L. D. (1993) Biochim. Biophys. Acta **1161**, 139–148
13 Kaushal, V., Avila, D. M., Hardies, S. C. and Barnes, L. D. (1990) Gene **95**, 79–84
14 Plateau, P., Fromant, M., Schmitter, J. M. and Blanquet, S. (1990) J. Bacteriol. **172**, 6892–6899
15 Mulder, W., Scholten, I. H., van Roon, H. and Grivell, L. A. (1994) Biochim. Biophys. Acta **1219**, 719-723
16 McLennan, A. G., Mayers, E., Hankin, S., Thorne, N. M. H., Prescott, M. and Powls, R. (1994) Biochem. J., **300**, 183–189
17 Lazewska, D., Starzynska, E. and Guranowski, A. (1993) Protein Express. Purif. **4**, 45–51
18 Prescott, M., Milne, A. D. and McLennan, A. G. (1989) Biochem. J. **259**, 831–838
19 Jakubowski, H. and Guranowski, G. (1983) J. Biol. Chem. **258**, 9982–9989
20 Koonin, E. V. (1993) Nucleic Acids Res. **21**, 4847
21 Rosenfeld, J., Capdevielle, J., Guillemot, J. C. and Ferrara, P. (1992) Anal. Biochem. **203**, 173–179
22 Stone, K. L., LoPresti, M. B., Crawford, J. M., DeAngelis, R. and Williams, K. R. (1989) in A Practical Guide to Protein and Peptide Purification and Microsequencing (Matsudaira, P. T., ed.), pp. 31–47, Academic Press, San Diego
23 Soares, M. B., Bonaldo, M. D., Jelene, P., Su, L., Lawton, L. and Efstratiadis, A. (1994) Proc. Natl. Acad. Sci. U.S.A. **91**, 9228–9232
24 Lennon, G., Auffray, C., Polymeropoulos, M. and Soares, B. (1994) Genomics, in the press
25 Kozak, M. (1991) J. Cell Biol. **115**, 887–903
26 Henikoff, S. and Henikoff, J. G. (1994) Genomics **19**, 97–107
27 Maki, H. and Sekiguchi, M. (1992) Nature (London) **355**, 273–275
28 Bullion, L. C., Méjean, V., Claverys, J.-P. and Bessman, M. (1994) J. Biol. Chem. **269**, 12339–12344
29 Sakumi, K., Furuichi, M., Tsuzuki, T. et al. (1994) J. Biol. Chem. **268**, 23524–23530
30 Takahagi, M., Iwasaki, H., Nakata, A. and Shinagawa, H. (1991) J. Bacteriol. **173**, 5747–5753
31 Frick, D. N. and Bessman, M. J. (1995) J. Biol. Chem. **270**, 1529–1534
32 Dixon, L. K., Twigg, S. R. F., Baylis, S. A. et al. (1994) J. Gen. Virol. **75** 1655–1684
33 Murphy, P. R. and Knee, R. S. (1994) Mol. Endocrinol. **8**, 852–859
34 Volk, R., Koster, M., Poting, A., Hartmann, L. and Knochel, W. (1989) EMBO J. **8**, 2983–2988
35 Mitchell, S. J. and Minnick, M. F. (1995) Infect. Immun. **63**, 1552–1562
36 Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) Nucleic Acids Res. **22**, 4673-4680
37 Frick, D. N., Weber, D. J., Abeygunawardana, C., Gittis, A. G., Bessman, M. J. and Mildvan, A. S. (1995) Biochemistry **34**, 5577–5586
38 Abeygunawardana, C., Weber, D. J., Frick, D. N., Bessman, M. J. and Mildvan, A. S. (1993) Biochemistry **32**, 13071–13080
39 Weber, D. J., Abeygunawardana, C., Bessman, M. J. and Mildvan, A. S. (1993) Biochemistry **32**, 13081–13088
40 Rost, B. and Sander, C. (1994) Proteins **20**, 216–226
41 Rost, B. and Sander, C. J. (1993) J. Mol. Biol. **232**, 584–599
42 Weber, D. J., Bhatnagar, S. K., Bullions, L. C., Bessman, M. J. and Mildvan, A. S. (1992) J. Biol. Chem. **267**, 16939–16942
43 McLennan, A. G., Prescott, M. and Evershed, R. P. (1989) Biomed. Environ. Mass Spectrom. **18**, 450–452
44 Dixon, R. M. and Lowe, G. (1989) J. Biol. Chem. **264**, 2069–2074
45 Henikoff, S., Henikoff, J. G., Alford, W. J. and Pietrokovski, S. (1995) Gene **163**, GC17–26
46 Garcia-Agúndez, J. A., Cameselle, J. C., Costas, M. J., Sillero, M. A. G. and Sillero, A. (1991) Biochim. Biophys. Acta **1073**, 402–409
47 Brevet, A., Chen, J., Fromant, M., Blanquet, S. and Plateau, P. (1993) J. Bacteriol. **173,** 5275–5279
48 O'Farrell, P. H. (1975) J. Biol. Chem. **250**, 4007–4021

Received 17 July 1995/21 August 1995; accepted 29 August 1995