PROTEIN SEQUENCE MOTIFS

Exopolyphosphate phosphatase and guanosine pentaphosphate phosphatase belong to the sugar kinase/actin/hsp70 superfamily

Inorganic polyphosphates are linear polymers of orthophosphate which probably serve as cellular reservoirs of phosphate and energy in both prokaryotes and eukaryotes[1,2]. They are synthesized in *Escherichia coli* by a membrane-associated polyphosphate kinase[3] and are degraded by an exopolyphosphatase[4].

Guanosine pentaphosphate (pppGpp) is a cytoplasmic signaling molecule which, together with ppGpp, controls the 'stringent response', an adaptive process that allows bacteria to respond to amino acid starvation, resulting in the coordinated regulation of numerous cellular activities[5]. pppGpp is synthesized by synthetases and degraded by both a 3'-pyrophosphohydrolase and a 5'-phosphohydrolase[5].

Recently, the sequence of the *E. coli* polyphosphate phosphatase, Ppx, was published[4]. We here demonstrate that this protein, previously believed to be unique, is homologous to the *E. coli* guanosine pentaphosphate phosphohydrolase, (guanosine 5'-triphosphate, 3'-diphosphate γ-nucleotidase), GppA[6,7] (SWISS-PROT identifier GPPA_ECOLI) and that both of these enzymes belong to a large superfamily that includes sugar kinases, actin, heat shock protein hsp70, and prokaryotic cell cycle proteins.

As shown in Fig. 1, these two enzymes exhibit striking sequence similarity: 39% identity over an overlapping region of 492 residues. The amino-terminal regions exhibit greater similarity than the remainder of the molecules. The comparison score obtained using the RDF2 program with 200 random shuffles[8] was 97 standard deviations, a score far in excess of that required to establish that these two phosphatases share a common evolutionary origin. These two proteins are of about the same length (513 residues for the exopolyphosphatase and 494 for pppGpp phosphohydrolase).

Hydropathy profiles[9] (not shown) revealed that both enzymes possess one hydrophobic region, 18 amino acid residues long, which in both proteins is found 270 residues from the amino terminus. This region undoubtedly corresponds to an interior α-helix connecting two globular domains as shown in the three-dimensional structure of glycerol kinase[10]. Codon Adaptation Index (CAI) determination[11] revealed that expression of both genes, is similar [*ppx* (CAI=0.30) and *gppA* (CAI=0.32)] and well within the range calculated for genes expressed at low levels[11].

No other proteins in the current databanks, including numerous other phosphatases, were demonstrably homologous with the pppGpp and polyphosphate phosphatases. However, hexokinase type II of rat[12] showed weak sequence similarity with Ppx (6 SD; 26% identity in an overlapping region of 66 amino acid residues) and with GppA (5 SD; 21% identity in an overlapping region of 71 amino acid residues). Most of

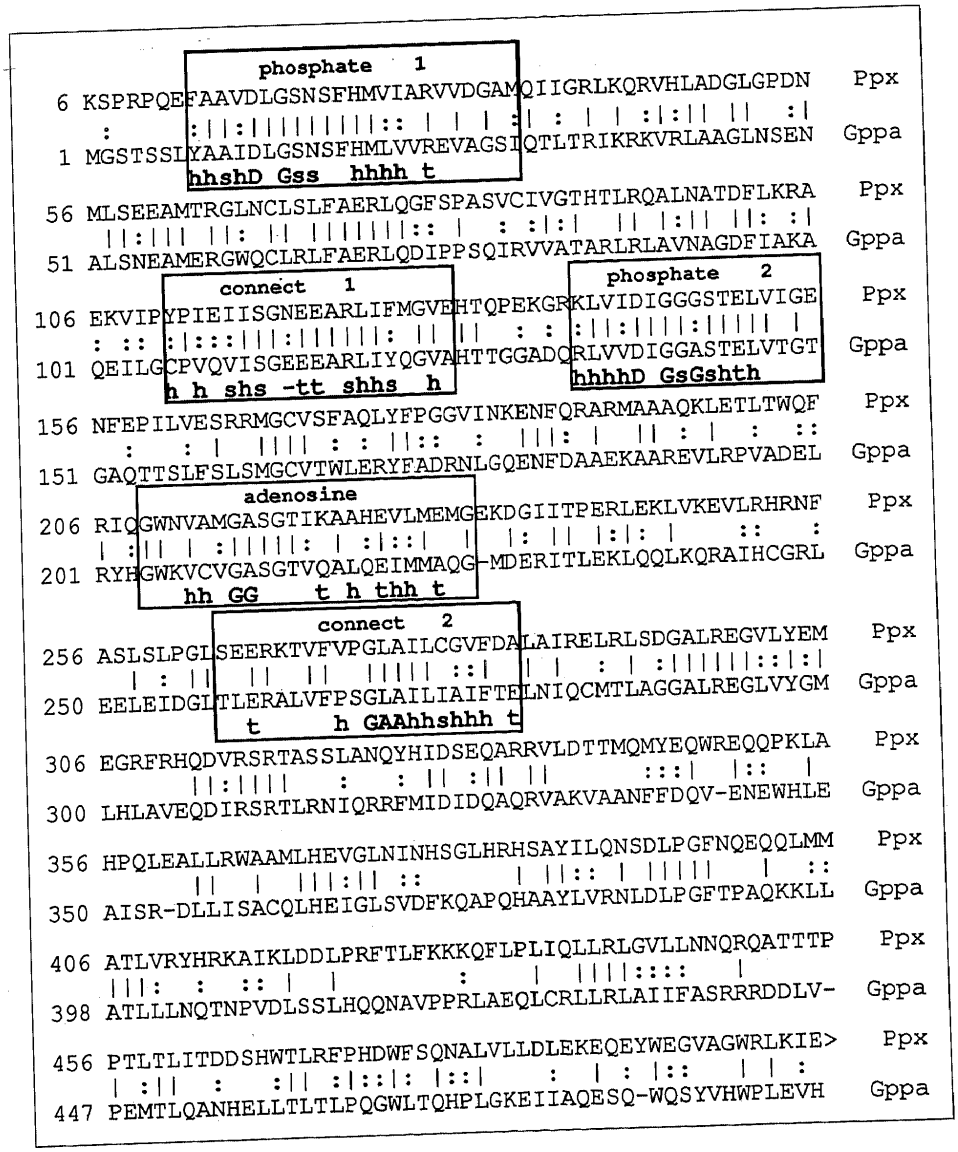

Figure 1

Aligned amino acid sequences of the *E. coli* exopolyphosphatase (Ppx) and guanosine pentaphosphate phosphohydrolase (GppA). The FASTA program[7], using the dipeptide mode (ktup=2) was used to align the proteins and to assess similarity. Numbers to the left of the two sequences indicate the first amino acid in the row. Identical (|) and similar (:) residues are indicated. The five boxes denote the conserved regions characteristic of the ATPase fold of the sugar kinase/actin/hsp70 superfamily[13]. Phosphate 1, phosphate 2 and the adenosine motifs were derived from the more closely related actin/hsp70 sub-branch of this family, whereas connect 1 and connect 2 correspond to the sugar kinase sequence patterns[14]. The consensus motif within each of the five boxes is denoted with the following amino acid groups: h, hydrophobic; s, small; t, tiny or polar; –, negatively charged. For further details, see Ref. 13.

the conserved regions in the alignment correspond to the five conserved boxes characteristic to the ATPase fold of sugar kinases, actin and members of the hsp70 family of heat shock proteins (see Fig. 1)[13,14]. The presence of these five conserved boxes in Ppx and in GppA supports the notion that in terms of the three-dimensional structure and apparent ATP binding properties, the two phosphatases belong to the recently described sugar kinase/actin/hsp70 superfamily[13,14].

The results show that the two *E. coli* phosphatases discussed here are closely related enzymes which have similar catalytic properties but entirely different physiological roles in the bacterial cell. They are the first identified phosphatases in the large and functionally diverse superfamily of proteins that includes sugar kinases, actin and hsp70.

References

1 Kulaev, I. S. (1979) *The Biochemistry of Inorganic Polyphosphates*, John Wiley & Sons
2 Wood, H. G. and Clark, J. E. (1988) *Annu. Rev. Biochem.* 57, 235–260
3 Akiyama, M., Crooke, E. and Kornberg, A. (1992) *J. Biol. Chem.* 267, 22556–22561
4 Akiyama, M., Crooke, E. and Kornberg, A. (1993) *J. Biol. Chem.* 268, 633–639
5 Cashel, M. and Rudd, K. E. (1987) in Escherichia coli *and* Salmonella typhimurium. *Cellular and Molecular Biology* (Ingraham, J. L., Low, K. B., Magasanik, B., Schaechter, M. and Umbarger, H. E., eds), American Society for Microbiology
6 Kalman, M., Murphy, H. and Cashel, M. (1991) *The New Biologist* 3, 886–895
7 Daniels, D. L., Plunkett, G., III, Burland, V. and Blattner, F. R. (1992) *Science* 257, 771–778
8 Pearson, W. R. and Lipman, D. J. (1988) *Proc. Natl Acad. Sci. USA* 85, 2444–2448
9 Kyte, J. and Doolittle, R. F. (1982) *J. Mol. Biol.* 157, 105–132
10 Hurley, J. H. *et al.* (1993) *Science* 259, 673–677
11 Sharp, P. M. and Li, W-H. (1987) *Nucleic Acids Res.* 15, 1281–1295
12 Thelen, A. P. and Wilson, J. E. (1991) *Arch. Biochem. Biophys.* 286, 645–651
13 Bork, P., Sander, C. and Valencia, A. (1992) *Proc. Natl Acad. Sci. USA* 89, 7290–7294
14 Bork, P., Sander, C. and Valencia, A. (1993) *Protein Science* 2, 31–40

JONATHAN REIZER, AIALA REIZER AND MILTON H. SAIER, JR

Department of Biology, University of California at San Diego, La Jolla, CA 92093-0116, USA.

PEER BORK AND CHRIS SANDER

European Molecular Biology Laboratory, D-6900 Heidelberg, FRG.

LETTERS

Vesicular transporters join the major facilitator superfamily (MFS)

Marger and Saier have recently described the evolutionary relationships between membrane transport proteins from prokaryotes and eukaryotes[1]. The authors discuss the phylogenetic tree of over 50 proteins, all being part of a superfamily which they name MFS (major facilitator superfamily).

Recently, two small new families of mammalian transport proteins were identified. Both families are expected to play a role in vesicular transport. SV2, whose protein is a major component of the synaptic vesicle membrane, is represented by two related genes, SV2A[2,3] and SV2B[4]. No direct function has yet been assigned to SV2 proteins. The second family includes vesicular amine transporters which protect cells against MPP^+ toxicity[5]. Members of this family have been isolated from chromaffin granules[5] (CGAT) and from rat brain[5] (SVAT). Both groups were reported to be related to transporters in bacteria and lower eukaryotes, but their evolutionary relationship to the superfamily of facilitators has not yet been established.

An analysis based on pairwise alignments followed by multi-sequence alignments (similar to that presented by Marger and Saier) identified these proteins as genuine members of the MFS family. Furthermore, the SV2 family

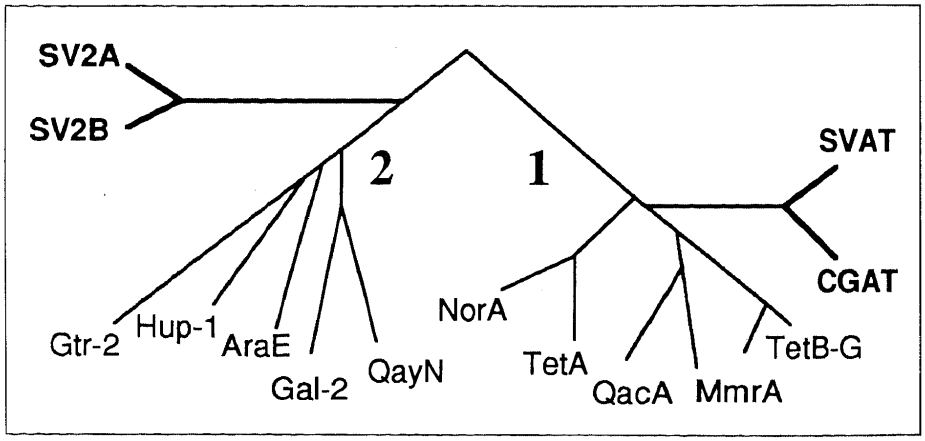

Figure 1
Phylogenetic tree of representative members of only two of the five clusters of MFS. Representatives of each cluster are marked as in Marger and Saier[1]. The two families of vesicular transporters (SV2A, SV2B, CGAT and SVAT) are in bold. Only the junction connecting cluster 1 and cluster 2 is presented. The branches of vesicular transporters are facing outside for clarity of presentation only. Branches lengths are drawn to scale.

(78.5% identity over 690 amino acids[4]) represents an early branching of MFS cluster 2 which includes, for example, bacterial sugar–H^+ symporters and sugar uniporters from lower eukaryotes and mammalians. The statistical analysis[6] using PAUP 3.0 (bootstrap approach, 100 samplings), indicates that the SV2s arose after the divergence of the five main clusters of the MFS superfamily but before the divergence among members of cluster 2 (Fig. 1). The vesicular transporters (SVAT and CGAT, 75% identity over 520 amino acids[5]), are members of cluster 1 which includes drug-resistant proteins. In this cluster, these amine transporter genes diverged just after and close to the branching of the quinolone-resistance efflux of *Staphylococcus aureus* (NorA) and tetracycline-resistance efflux of *E. coli* (TetA, Fig. 1). This evolutionary connection was supported by a statistical analysis (PAUP 3.0).

We have established the membership of these four genes representing two small subfamilies in MFS. It is likely that other genes related to vesicular transporters from lower eukaryotes and from prokaryotes are still to be found.