PROTEINS: Structure, Function, and Genetics 45:318–324 (2001)

Conserved Core Structure and Active Site Residues in Alkaline Phosphatase Superfamily Enzymes

Michael Y. Galperin[1,†] and Mark J. Jedrzejas[2]

[1]*National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, Maryland 20894*

[2]*Children's Hospital Oakland Research Institute, 5700 Martin Luther King, Jr. Way, Oakland, California 94609*

ABSTRACT **Cofactor-independent phosphoglycerate mutase (iPGM) has been previously identified as a member of the alkaline phosphatase (AlkP) superfamily of enzymes, based on the conservation of the predicted metal-binding residues. Structural alignment of iPGM with AlkP and cerebroside sulfatase confirmed that all these enzymes have a common core structure and revealed similarly located conserved Ser (in iPGM and AlkP) or Cys (in sulfatases) residues in their active sites. In AlkP, this Ser residue is phosphorylated during catalysis, whereas in sulfatases the active site Cys residues are modified to formylglycine and sulfatated. Similarly located Thr residue forms a phosphoenzyme intermediate in one more enzyme of the AlkP superfamily, alkaline phosphodiesterase/nucleotide pyrophosphatase PC-1 (autotaxin). Using structure-based sequence alignment, we identified homologous Ser, Thr, or Cys residues in other enzymes of the AlkP superfamily, such as phosphopentomutase, phosphoglycerol transferase, phosphonoacetate hydrolase, and GPI-anchoring enzymes (glycosylphosphatidylinositol phosphoethanolamine transferases) MCD4, GPI7, and GPI13. We predict that catalytical cycles of all the enzymes of AlkP superfamily include phosphoenzyme (or sulfoenzyme) intermediates. Proteins 2001;45:318–324.** © 2001 Wiley-Liss, Inc.*

Key words: alkaline phosphodiesterase; nucleotide pyrophosphatase PC-1; autotaxin; glycosylphosphatidylinositol phosphoethanolaminetransferase; MCD4; Pig-O; phosphonoacetate hydrolase; phosphonate monoesterase

INTRODUCTION

Alkaline phosphatase (AlkP) is one of better studied and widely used enzymes, whose structure and catalytic mechanism are relatively well established. Its active site contains a binuclear cluster of Zn^{2+} ions, which are involved in phosphate binding, and a conserved Ser residue, which is phosphorylated in the course of catalysis.[1] A very similar structural fold and a similar organization of the active sites have been found in two human sulfatases, human *N*-acetylgalactosamine-4-sulfatase and cerebroside sulfatase.[2,3] In sulfatases, however, the corresponding active site residue is Cys, which is posttranslationally converted into formylglycine, and there appears to be a single metal atom, possibly Ca^{2+} or Mg^{2+}.[2,3] Based on the conservation of metal-binding residues, it was suggested that cofactor-independent phosphoglycerate mutase (iPGM) could be related to AlkP.[4] In 1998, Galperin et al. found same conserved metal-binding residues in several more enzymes, including phosphodeoxyribomutase, phosphoglycerol transferase and nucleotide pyrophosphatase and suggested that they all belong to the same superfamily of metalloenzymes, referred to as AlkP superfamily.[5] A recently updated alignment of the metal-binding portions of these enzymes is available, for example, in Pfam database[6] as domain PF01676. However, in spite of the well-known experimental data indicating formation of a phosphoenzyme intermediate in the catalytic cycle of iPGM,[7,8] the phosphorylated residue in iPGM has not been identified from sequence analysis, due to the relatively low sequence conservation in that area.[5,9]

After the crystal structure of iPGM had been determined,[10,11] it became possible to use structural comparisons to verify the predicted features of the AlkP superfamily and pinpoint the active site residues in its members. Surprisingly, a comparison of the iPGM structure (PDB code: 1EJJ) to the library of known protein structures using DALI[12] did not recognize any similarity of iPGM to AlkP, although the active sites of these enzymes could be almost perfectly superimposed.[10,11] Here, we reinvestigated the possible relation of iPGM and AlkP, identified their common structural core, and used the structure-based sequence alignment of iPGM, AlkP, and sulfatases to align other members of the AlkP superfamily and predict their active site residues.

MATERIALS AND METHODS

Sequence database searches against the nonredundant protein database maintained at the NCBI (Bethesda, MD) were performed using the PSI-BLAST[13] and HMMer2[14] (http://hmmer.wustl.edu) programs with various members

Abbreviations: AlkP, alkaline phosphatase; iPGM, cofactor-independent phosphoglycerate mutase; dPGM, cofactor-dependent phosphoglycerate mutase.

Grant sponsor: DARPA; Grant number: N66001-01-C-8013.

†Correspondence to: Michael Galperin, National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894. E-mail: galperin@ncbi.nlm.nih.gov

Received 3 May 2001; Accepted 6 July 2001

of the AlkP superfamily[5] as queries. Structure-based sequence alignment of 1EJJ, 1ALK, and 1AUK was generated by VAST[15] (http://www.ncbi.nlm.nih.gov/Structure/VAST/vast.shtml) and DALI[12] (http://www.ebi.ac.uk/dali/) programs. It was then used as a seed alignment in additional PSI-BLAST and HMMer2 searches with default parameters. For the final alignment, various members of the AlkP superfamily were assembled into a special database, which was used to run PSI-BLAST searches to convergence, using each sequence in the database as query. Secondary structure assignments for iPGM, AlkP, and sulfatase were extracted from the PDB entries. Secondary structure assignments for other members of the AlkP superfamily were obtained using PhD.[16]

RESULTS AND DISCUSSION

Monophosphoglycerate mutase (EC 5.4.2.1), a key glycolytic enzyme, is found in nature in two forms that differ in their requirement for 2,3-bisphosphoglycerate.[17,18] The 2,3-bisphosphoglycerate-dependent form (dPGM) found in some bacteria, yeast, and animals (including humans) belongs to the same structural superfamily of enzymes as acid phosphatases and fructose-2,6-bisphosphatase.[12,19] The structure of the 2,3-bisphosphoglycerate-independent form of phosphoglycerate mutase (iPGM, PDB: 1EJJ) found in many bacteria, plants, and invertebrates[5,9,20] was resolved only recently.[10,11] This enzyme consists of two distinct structural domains connected by two long loops. The "phosphotransferase" domain is formed by residues 78 to 308, whereas the "phosphatase" domain contains ~70 N-terminal and ~200 C-terminal residues.[10] Previous sequence analysis of iPGM revealed its similarity to alkaline phosphatase (AlkP), including the conservation of metal-binding residues in the AlkP active site, but failed to find additional similarities such as the AlkP catalytic triad Asp-Ser-Ala.[4,5] Threading of the *Trypanosoma brucei* iPGM sequence against the AlkP structure allowed the alignment of some of the predicted structural elements.[9] It appeared therefore that sequence-based identification of the AlkP superfamily could be supported by a structural alignment of iPGM, AlkP, and sulfatases. After the crystal structure of iPGM has been determined and the active site Ser residue has been identified,[10,11] it became possible to use structural comparisons to verify the predicted features of the AlkP superfamily and pinpoint the active site residues in its other members.

Conserved Structural Core of the AlkP Superfamily Enzymes

When the newly solved iPGM structure (1EJJ) was submitted for the DALI[12] analysis, AlkP and sulfatases were not reported to be among the 170 closest structural neighbors of 1EJJ.[10,11] On the other hand, VAST[15] analysis identified AlkP and the sulfatases as close structural neighbors of one of the two structural domains of iPGM. Indeed, when only this "phosphatase" domain of iPGM[10] was submitted for DALI analysis, it was found to have a statistically significant similarity ($Z > 18.0$) to AlkP and sulfatase structures (Fig. 1). Structural alignment of the "phosphatase" domain of iPGM with AlkP and cerebroside sulfatase showed a common conserved core of 9 α-helices and 8 β-strands [Fig. 1(a,b)]. This alignment showed a remarkable consistency; most nonaligned residues were found in only one of the sequences, which allowed us to treat them as insertions into the conserved core structure of these enzymes [Fig. 1(a,c)]. In the case of iPGM, this insertion was 230 amino acids long and formed a well-separated domain, whereas in AlkP and sulfatase there were several insertions, all located at the surface of the conserved core [Fig. 1(a,c)]. It appears therefore that this two-domain organization of iPGM was the reason for the failure of DALI to recognize the structural similarity of iPGM and AlkP.

All three enzymes, AlkP, cerebroside sulfatase, and iPGM, have a characteristic α/β topology and consist of a central β-sheet flanked on both sides by α-helices [Fig. 1(c)]. The core β-sheet is built from 8 to 10 predominantly parallel β-strands; one or two β-strands in every enzyme are antiparallel to the rest [Fig. 1(b)]. Up to 10 α-helices flanking this β-sheet are nearly equally divided between its two sides (three to six helices on each side).

The conserved core of all three enzymes consists of a very well-aligned 8-stranded β-sheet with six strands in the flat region and two strands in the curled region [Fig. 1(a)]. In AlkP and sulfatase, there are two additional strands in the β-sheet that do not belong to the core structural part of these enzymes but comprise additional insertions as shown in the sequence alignment [Fig. 1(b)]. In Figure 1(a), four well-conserved α-helices, including two longer ones, are present behind the central β-sheet. In front of the central β-sheet, five overlapping α-helices are present, three of which can be perfectly superimposed. Two other α-helices in front [on the right side in Fig. 1(a)] have a lesser degree of overlap and can be represented only by a helix-like loop. In general, the degree of structural overlap of the core is strikingly high. The largest differences among all structures occur in the uppermost region [as seen in Fig. 1(a)] where elaborate loop structures are present.

Sequence Alignment of the AlkP Superfamily Enzymes

Using the structure-based sequence alignment of iPGM, AlkP, and sulfatase, generated by VAST[15] and DALI[12] [Fig. 1(c)], as a seed alignment for iterative database searches using PSI-BLAST[13] and HMMer[14] retrieved all the previously identified members of the AlkP superfamily[5] (Fig. 2). These searches resulted in a substantial improvement in resolution for the previously unaligned region that included Ser-102 of the AlkP (phosphorylated in the AlkP catalytic cycle) and the corresponding Cys residues, posttranslationally modified into formylglycine in the active sites of sulfatases[2,3] In spite of the well-known experimental data indicating formation of a phosphoenzyme intermediate in the catalytic cycle of iPGM,[7,8] previous attempts to pinpoint the phosphorylated residue in iPGM through sequence analysis have been hampered by relatively low sequence conservation in this area.[5] An

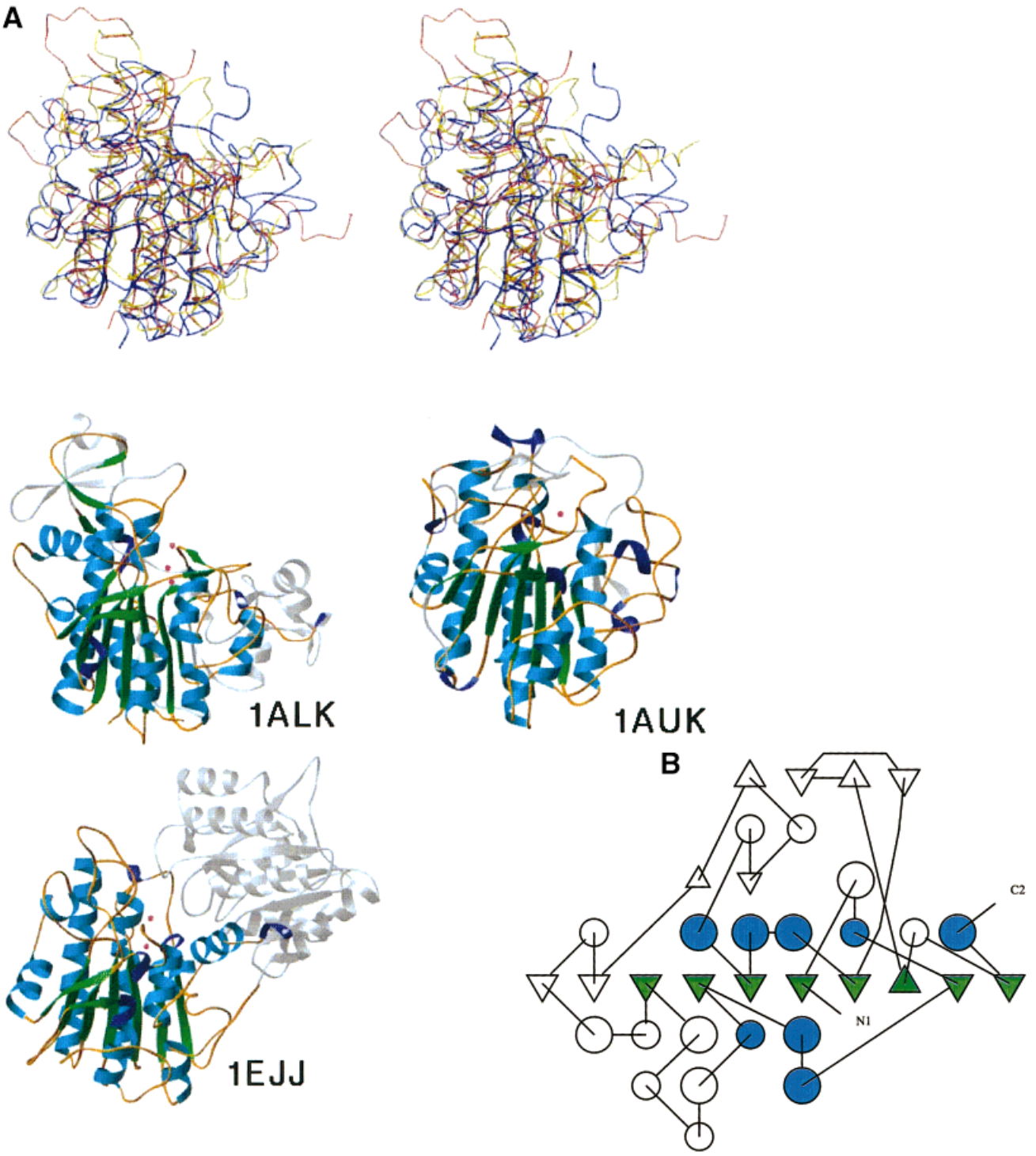
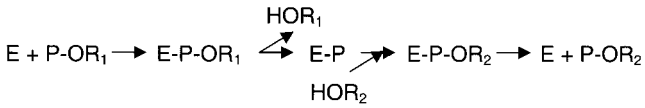

A
1ALK
1AUK
1EJJ
B
C2
N1
C
12
62
403
407
445
462
1EJJ 2 SKKPVALYILDGFALR----DETYGNAVAQAN---------KPNFDRYWNLYPHTTLRACGEAVGLP-EGQMGNSEVGHLNIGAGRIVYQSLTRINIAIREGEFD-217- 92
1ALK 41 PAKNIILLIGDGMGDS----------EITAARNYAEGAGGFFKGIDAL---PLTGQYTHYALNKKTGKPDYVTDSAASATAWSTGVKTYNGALGVDIH---------- 125
1AUK 19 RPPNIVLIFADDLGYGDLGCYGHP-----SST---------TPNLDQLAAGGLRFTDFYVP----------VSLCTPSRAALLTGRLPVRMGMYPGVLVPSSRGG----- 99
1EJJ 310 PTNLDNTIGEVLSQHGLRQLRIATE-KYPHVTFFMSGGR----EEEFPG-EDRILIN----SPKVPTYDLKPE----MSAYEVTDALLKEIEADK----YDAIILNYANPDMVGHSGK---
1ALK 125 -EKDHPTILEMAKAAGLATGNVSIAELQDATPA-ALVAHV-31-LLNAR--ADVTLGG-52-FADG---NMPVR-27----PTLAQMTDKAIELLSKNEKGFFLQVEGASIDKQDHAAN---
1AUK 100 LPLEEVTVAEVLAARGYLTGMAG----KWHL-GVGPEGAF----LPPHQGFHRFLGIP----YSHDQGPCQNLT-36-GLEARYNAFAHDLMADAGRQDRFFLYYASHHTHYPQFSGQ-7-
1EJJ 411 LEPTIKAVEAVDECLGKVVDALLKG----GIAITADHGNADEVLTPDGK---PQTAH-----TTNPVPVIVTKKG-IK---LRDGGILGDLAPTMLDLLGLPQPKEMTGKSLIVK 511
1ALK 335 FCGQIGETVDLDEAVQRALEFAKKEG---NTLVIVTADHAHASQIVAPDTK-26-SQEH-----TGSQLRIAAYGPHAAN---VVGLTDQTDLFYTMKAALGLK---------- 449
1AUK 244 RGPFGDSLMELDAAVGTLMTAIGDLGLLEETLVIFTADNGPETMRMSRGGCSGLLRCGKGTTYEGGVREPALAFWPGHIAPGVTHELASSLDLLPTLAALAGAPLPNVTLDGFDLSP 360

Figure 1.

analysis of the iPGM structure allowed the identification of Ser-62 as the phosphorylation site in iPGM,[10] equivalent to the active site Ser and Cys residues in AlkP and sulfatase, respectively. Recently, phosphorylation of this residue has been directly demonstrated in the iPGM from *T. brucei* (J.-F. Collet, V. Stroobant, and E. van Schaftingen, submitted for publication). The PSI-BLAST and HMMmer searches allowed us to align the corresponding regions in other enzymes of the AlkP superfamily, such as alkaline phosphodiesterase/nucleotide pyrophosphatase (autotaxin) PC-1,[21,22] streptomycin-6-phosphatase,[23] Ca^{2+}-dependent ATPase,[24] phosphonoacetate hydrolase,[25] phosphonate monoesterase,[26] and GPI-anchoring enzymes (glycosylphosphatidylinositol phosphoethanolaminetransferases) MCD4, GPI7, and GPI13[27–30] (Fig. 2). In each of these enzymes, in addition to previously identified conserved metal-binding residues,[5] similarly positioned Ser, Thr, or Cys residues were detected, homologous to the active-site Ser-102 of AlkP and Ser-62 of iPGM (Fig. 2). For two more members of the AlkP superfamily, phosphopentomutase[31] and phosphoglycerol transferase,[32] homologous Ser or Thr residues could not be unambiguously identified from PSI-BLAST results, but could be tentatively predicted from secondary structure assignments.

Common Catalytic Mechanisms of the AlkP Superfamily Enzymes

Remarkably, in every experimentally studied member of the AlkP superfamily, these conserved Ser, Thr, or Cys residues have been found to be phosphorylated or sulfatated. In bovine nucleotide phosphodiesterase, for example, phosphorylation of the active site Thr residue has been experimentally demonstrated, and this residue was identified from a peptide map.[21,33] Recently, metal-binding residues of mouse nucleotide phosphodiesterase have been shown to be needed for catalytic activity.[34] Ca^{2+}-dependent ATPase was also shown to be phosphorylated, although the phosphorylated residue has been proposed to be an Asp residue.[24,35] On the basis of the sequence and structural conservation, we propose that catalytic cycles of all the members of this superfamily include phosphorylation (or sulfatation, or phosphonation) of these conserved Ser/Thr/Cys residues. This would mean that all enzymes of the AlkP superfamily have the same reaction scheme as was originally proposed for AlkP[1,36]:



where the phosphate (sulphate or phosphonate) acceptor can be either water (R_2 = H, in phosphatases, sulfatases, phosphonate hydrolases) or a second substrate (in phosphoglycerol and phosphoethanolamine transferases) or just a different hydroxyl group of the same original substrate R_1 (in iPGM and phosphopentomutase).

It is important to note that phosphonopyruvate decarboxylase is not a member of the AlkP superfamily, even though MJ0010, MJ1612, and several related proteins (Fig. 2) are annotated in GenBank as phosphonopyruvate decarboxylases. The original annotation of those proteins was based on their similarity to an uncharacterized protein BcpC from *Streptomyces hygroscopicus* (GenBank accession number D37809.1, gi|520857), originally thought to encode a phosphonopyruvate decarboxylase. After the true TPP-dependent phosphonopyruvate decarboxylase (GenBank accession number D37809.2, gi|5545271) was sequenced,[37,38] the annotation of the original BcpC protein as phosphonopyruvate decarboxylase was withdrawn and replaced by OrfZZ (gi|7416070). Nevertheless, mis-annotations of MJ0010, MJ1612, and related proteins stay in the database and continue to spawn new erroneous annotations. We have previously suggested that these enzyme comprise ancient phosphomutases and are responsible for the phosphoglycerate mutase activity in archaea.[5] In any case, all the characterized enzymes of the AlkP superfamily comprise either phosphatases (phosphonatases, sulfatases) or phosphomutases.

Practical Aspects of the AlkP Superfamily Studies

The availability of the structure-based alignment of the AlkP superfamily enzymes should stimulate further studies of the detailed catalytic mechanisms of these enzymes, including GPI-anchoring enzymes and peptidoglycan phosphoglycerol transferase, which are notoriously difficult to study *in vitro* because of the complexity of their substrates. The AlkP family also includes important enzymes that are involved in biosynthesis of antibiotics, such as streptomycin,[23] bialaphos,[37] and mitomycin.[39] To simplify future studies, positions of predicted active-site residues in representative enzymes of AlkP superfamily are marked in Figure 2 (residues of the *Escherichia coli* AlkP PPB_ECOLI are numbered as in the mature enzyme).

One promising avenue for future studies would be a search for potential inhibitors of iPGM. This enzyme has never been found in any vertebrates,[20] leaving cofactor-dependent form of PGM (dPGM) as the only form of this enzyme in humans.[17,18] Indeed, dPGM mutations are known to cause hereditary myopathies showing the importance of this enzyme in human matabolism.[40] In contrast,

Fig. 1. Conserved structural core of alkaline phosphatase superfamily enzymes. (**A**) Structural alignment of *Escherichia coli* alkaline phosphatase (PDB: 1ALK), human cerebroside sulfatase (PDB: 1AUK), and cofactor-independent phosphoglycerate mutase from *Bacillus stearothermophilus* (iPGM, PDB: 1EJJ). Top: superimposition of C_α traces of 1ALK (red), 1AUK (blue), and 1EJJ (yellow). Bottom: conserved core elements of 1ALK, 1AUK, and 1EJJ are colored as follows: α-helices = light blue, 3_{10} helices = dark blue, β-strands = green, and loops = brown. Structural elements that are unique for each enzyme are shown in light gray. The red dots indicate metal atoms of the active sites. (**B**) TOPS cartoon of 1ALK topology.[50] The α-helices and β-strands that are conserved in other enzymes of the AlkP superfamily are colored as shown in **A**. Secondary structure elements that are specific for AlkP are left blank. (**C**) Structure-based sequence alignment of iPGM, AlkP, and cerebroside sulfatase. Numbers indicate the positions of the first and last amino acid residues and the sizes of gaps between aligned segments. Blue and green shading indicates conserved α-helices and β-strands, respectively, as shown in **A**. Active site Ser and Cys residues are shown as red letters on yellow background. White letters on red background indicate amino acid residues that are involved in metal binding; positions of these residues in iPGM are indicated above the line. Residues that are aligned in all three enzymes are shown in bold; identical residues are indicated by asterisks; similar residues are indicated by dots.

Fig. 2. Sequence alignment of alkaline phosphatase superfamily enzymes. The proteins are listed under their SWISS-PROT names, where available, or as abbreviated gene and species symbols (left column); names of experimentally characterized proteins are shown in bold. Numbers indicate positions of the first and last amino acid residues and the sizes of the gaps between aligned segments. The numbers at the end of each line indicate the total length of the given protein (in parentheses) and its unique identifier (gi) in the NCBI protein database. The coloring of predicted active site residues is as in Figure 1(c). Yellow shading indicates conserved uncharged amino acid residues. Green background indicates conserved Pro and small (G, A, or S) residues. In the structure line, H indicates α-helix and E indicates β-strand.

iPGM is the only form of PGM in such human pathogens as *Mycoplasma genitalium*, *M. pneumoniae*, *Trypanosoma brucei*, the ulcer-causing bacterium *Helicobacter pylori* and the food-borne intestinal pathogen *Campylobacter jejuni*, which makes their iPGM an attractive drug target.[5,9,41] A comparison of iPGM, AlkP and sulfatase structures (Fig. 1) shows that these enzymes are sufficiently different to make a search for an iPGM-specific inhibitor feasible.

We have argued earlier that archaeal members of the AlkP superfamily (e.g., MJ0010 and MJ1612) might resemble ancestral phosphomutases.[5,18] Alignment of the diverse members of this superfamily (Fig. 2) shows that phosphonate hydrolases and alkaline phosphodiesterases do not contain any major insertions and thus appear most similar to the conserved core of AlkP superfamily. This observation suggests that the original function of this core enzyme was to catalyze the hydrolysis of polyphosphates and phosphonates that could have been abundant on this planet in its early years.

On a more general note, the AlkP superfamily represents yet another example when the members of a superfamily of structurally related enzymes with little sequence similarity turn out to have remarkably similar catalytic mechanisms. Other examples of such superfamilies, unifying diverse enzyme that have common chemical mechanisms, include Ntn hydrolase,[42] enolase,[43] amidohydrolase,[44] chloroperoxidase/phosphatase,[45] ATP-grasp,[46] HAD[47] superfamilies, and many others (see Refs. 48 and 49 for reviews). This has important implications for functional analysis of newly sequenced uncharacterized proteins. When sequence analysis identifies new protein sequence as a member of one of such superfamilies, it often becomes possible to predict the three-dimensional structure and even likely catalytic mechanism of that protein. As a result, superfamily classification of new proteins provides a quick way to assess their potential chemical and even biological functions, helping to make sense out of the rapidly growing number of predicted uncharacterized proteins.

ACKNOWLEDGEMENTS

The authors thank J-F. Collet for communicating unpublished results on the phosphorylation of the Ser residue in *T. brucei* iPGM and E. Koonin for helpful comments. This work was supported in part by DARPA contract N66001-01-C-8013.

REFERENCES

1. Kim EE, Wyckoff HW. Reaction mechanism of alkaline phosphatase based on crystal structures. Two-metal ion catalysis. J Mol Biol 1991;218:449–464.
2. Bond CS, Clements PR, Ashby SJ, Collyer CA, Harrop SJ, Hopwood JJ, Guss JM. Structure of a human lysosomal sulfatase. Structure 1997;5:277–289.
3. Lukatela G, Krauss N, Theis K, Selmer T, Gieselmann V, von Figura K, Saenger W. Crystal structure of human arylsulfatase A: the aldehyde function and the metal ion at the active site suggest a novel mechanism for sulfate ester hydrolysis. Biochemistry 1998; 37:3654–3664.
4. Grana X, de Lecea L, el-Maghrabi MR, Urena JM, Caellas C, Carreras J, Puigdomenech P, Pilkis SJ, Climent F. Cloning and sequencing of a cDNA encoding 2,3-bisphosphoglycerate-independent phosphoglycerate mutase from maize. Possible relationship to the alkaline phosphatase family. J Biol Chem 1992;267:12797–12803.
5. Galperin MY, Bairoch A, Koonin EV. A superfamily of metalloenzymes unifies phosphopentomutase and cofactor-independent phosphoglycerate mutase with alkaline phosphatases and sulfatases. Prot Sci 1998;7:1829–1835.
6. Bateman A, Birney E, Durbin R, Eddy SR, Howe KL, Sonnhammer EL. The Pfam protein families database. Nucleic Acids Res 2000;28:263–266.
7. Breathnach R, Knowles JR. Phosphoglycerate mutase from wheat germ: studies with ^{18}O-labeled substrate, investigations of the phosphatase and phosphoryl transfer activities, and evidence for a phosphoryl-enzyme intermediate. Biochemistry 1977;16:3054–3060.
8. Blattler WA, Knowles JR. Phosphoglycerate mutases: stereochemical course of the phosphoryl group transfers catalyzed by the cofactor-dependent enzyme from rabbit muscle and the cofactor-independent enzyme from wheat germ. Biochemistry 1980;19:738–743.
9. Chevalier N, Rigden DJ, Van Roy J, Opperdoes FR, Michels PA. *Trypanosoma brucei* contains a 2,3-bisphosphoglycerate independent phosphoglycerate mutase. Eur J Biochem 2000;267:1464-–1472.
10. Jedrzejas MJ, Chander M, Setlow P, Krishnasamy G. Structure and mechanism of action of a novel phosphoglycerate mutase from *Bacillus stearothermophilus*. EMBO J 2000;19:1419–1431.
11. Jedrzejas MJ, Chander M, Setlow P, Krishnasamy G. Mechanism of catalysis of the cofactor-independent phosphoglycerate mutase from *Bacillus stearothermophilus*. Crystal structure of the complex with 2-phosphoglycerate. J Biol Chem 2000;275:23146–23153.
12. Holm L, Sander C. Mapping the protein universe. Science 1996;273: 595-603.
13. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zheng Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST—A new generation of protein database search programs. Nucleic Acids Res 1997;25:3389–3402.
14. Eddy SR. Profile hidden Markov models. Bioinformatics 1998;14: 755-763.
15. Gibrat JF, Madej T, Bryant SH. Surprising similarities in structure comparison. Curr Opin Struct Biol 1996;6:377–385.
16. Rost B, Sander C. Prediction of protein secondary structure at better than 70% accuracy. J Mol Biol 1993;232:584–599.
17. Fothergill-Gilmore LA, Watson HC. The phosphoglycerate mutases. Adv Enzymol Relat Areas Mol Biol 1989;62:227–313.
18. Jedrzejas MJ. Structure, function, and evolution of phosphoglycerate mutases: comparison with fructose-2,6-bisphosphatase, acid phosphatase, and alkaline phosphatase. Prog Biophys Mol Biol 2000;73:263–287.
19. Murzin AG, Brenner SE, Hubbard T, Chothia C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol 1995;247:536–540.
20. Grana X, Perez de la Ossa P, Broceno C, Stocker M, Garriga J, Puigdomenech P, Climent F. 2,3-Bisphosphoglycerate-independent phosphoglycerate mutase is conserved among different phylogenic kingdoms. Comp Biochem Physiol B Biochem Mol Biol 1995;112:287-–293.
21. Clair T, Lee HY, Liotta LA, Stracke ML. Autotaxin is an exoenzyme possessing 5′-nucleotide phosphodiesterase/ATP pyrophosphatase and ATPase activities. J Biol Chem 1997;272:996-–1001.
22. Bollen M, Gijsbers R, Ceulemans H, Stalmans W, Stefan C. Nucleotide pyrophosphatases/phosphodiesterases on the move. Crit Rev Biochem Mol Biol 2000;35:393–432.
23. Mansouri K, Piepersberg W. Genetics of streptomycin production in Streptomyces griseus: nucleotide sequence of five genes, *strF-GHIK*, including a phosphatase gene. Mol Gen Genet 1991;228: 459-–469.
24. Desrosiers MG, Gately LJ, Gambel AM, Menick DR. Purification and characterization of the Ca^{2+}-ATPase of *Flavobacterium odoratum*. J Biol Chem 1996;271:3945–3951.
25. Kulakova AN, Kulakov LA, Quinn JP. Cloning of the phosphonoacetate hydrolase gene from *Pseudomonas fluorescens* 23F encoding a new type of carbon-phosphorus bond cleaving enzyme and its expression in *Escherichia coli* and *Pseudomonas putida*. Gene 1997;195:49–53.
26. Dotson SB, Smith CE, Ling CS, Barry GF, Kishore GM. Identification, characterization, and cloning of a phosphonate monoester hydrolase from *Burkholderia caryophilli* PG2982. J Biol Chem 1996;271:25754–25761.
27. Gaynor EC, Mondesert G, Grimme SJ, Reed SI, Orlean P, Emr SD. MCD4 encodes a conserved endoplasmic reticulum membrane protein essential for glycosylphosphatidylinositol anchor synthesis in yeast. Mol Biol Cell 1999;10:627–648.
28. Benachour A, Sipos G, Flury I, Reggiori F, Canivenc-Gansel E, Vionnet C, Conzelmann A, Benghezal M. Deletion of GPI7, a yeast gene required for addition of a side chain to the glycosylphosphatidylinositol (GPI) core structure, affects GPI protein transport, remodeling, and cell wall integrity. J Biol Chem 1999;274:15251–15261.
29. Hong Y, Maeda Y, Watanabe R, Ohishi K, Mishkind M, Riezman H, Kinoshita T. Pig-n, a mammalian homologue of yeast Mcd4p, is involved in transferring phosphoethanolamine to the first mannose of the glycosylphosphatidylinositol. J Biol Chem 1999;274: 35099–35106.
30. Taron CH, Wiedman JM, Grimme SJ, Orlean P. Glycosylphosphatidylinositol biosynthesis defects in Gpi11p- and Gpi13p-deficient yeast suggest a branched pathway and implicate gpi13p in phosphoethanolamine transfer to the third mannose. Mol Biol Cell 2000;11:1611–1630.
31. Hammer-Jespersen K. Nucleoside catabolism. In: Munch-Petersen A, editor. Metabolism of nucleotides, nucleosides and nucleobases in microorganisms. London: Academic Press; 1983. p 203–258.
32. Jackson BJ, Bohin JP, Kennedy EP. Biosynthesis of membrane-derived oligosaccharides: characterization of *mdoB* mutants defective in phosphoglycerol transferase I activity. J Bacteriol 1984;160: 976–981.
33. Culp JS, Blytt HJ, Hermodson M, Butler LG. Amino acid sequence of the active site peptide of bovine intestinal 5′-nucleotide phosphodiesterase and identification of the active site residue as threonine. J Biol Chem 1985;260:8320-–8324.
34. Gijsbers R, Ceulemans H, Stalmans W, Bollen M. Structural and catalytic similarities between nucleotide pyrophosphatases/phosphodiesterases and alkaline phosphatases. J Biol Chem 2001;276:1361–1368.
35. Peiffer WE, Desrosiers MG, Menick DR. Cloning and expression of the unique Ca^{2+}-ATPase from *Flavobacterium odoratum*. J Biol Chem 1996;271:5095-5100.
36. Holtz KM, Stec B, Kantrowitz ER. A model of the transition state in the alkaline phosphatase reaction. J Biol Chem 1999;274:8351–8354.
37. Nakashita H, Watanabe K, Hara O, Hidaka T, Seto H. Studies on the biosynthesis of bialaphos. Biochemical mechanism of C-P bond formation: discovery of phosphonopyruvate decarboxylase which catalyzes the formation of phosphonoacetaldehyde from phosphonopyruvate. J Antibiot (Tokyo) 1997;50:212–219.
38. Nakashita H, Kozuka K, Hidaka T, Hara O, Seto H. Identification

and expression of the gene encoding phosphonopyruvate decarboxylase of Streptomyces hygroscopicus. Biochim Biophys Acta 2000; 1490:159–162.
39. Mao Y, Varoglu M, Sherman DH. Molecular characterization and analysis of the biosynthetic gene cluster for the antitumor antibiotic mitomycin C from *Streptomyces lavendulae* NRRL 2564. Chem Biol 1999;6:251–263.
40. DiMauro S, Miranda AF, Khan S, Gitlin K, Friedman R. Human muscle phosphoglycerate mutase deficiency: newly discovered metabolic myopathy. Science 1981;212:1277–1279.
41. Galperin MY, Koonin EV. Searching for drug targets in microbial genomes. Curr Opin Biotechnol 1999;10:571–578.
42. Brannigan JA, Dodson G, Duggleby HJ, Moody PC, Smith JL, Tomchick DR, Murzin AG. A protein catalytic framework with an N-terminal nucleophile is capable of self-activation. Nature 1995; 378:416–419.
43. Babbitt PC, Hasson MS, Wedekind JE, Palmer DR, Barrett WC, Reed GH, Rayment I, Ringe D, Kenyon GL, Gerlt JA. The enolase superfamily: a general strategy for enzyme-catalyzed abstraction of the alpha-protons of carboxylic acids. Biochemistry 1996;35: 16489–16501.
44. Holm L, Sander C. An evolutionary treasure: unification of a broad set of amidohydrolases related to urease. Proteins 1997;28:72–82.
45. Neuwald AF. An unexpected structural relationship between integral membrane phosphatases and soluble haloperoxidases. Protein Sci 1997;6:1764–1767.
46. Galperin MY, Koonin EV. A diverse superfamily of enzymes with ATP-dependent carboxylate-amine/thiol ligase activity. Protein Sci 1997;6:2639-–2643.
47. Aravind L, Galperin MY, Koonin EV. The catalytic domain of the P-type ATPase has the haloacid dehalogenase fold. Trends Biochem Sci 1998;23:127–129.
48. Koonin EV, Tatusov RL, Galperin MY. Beyond the complete genomes: from sequences to structure and function. Curr Opin Struct Biol 1998;8:355–363.
49. Todd AE, Orengo CA, Thornton JM. Evolution of function in protein superfamilies, from a structural perspective. J Mol Biol 2001;307:1113–1143.
50. Westhead DR, Hatton DC, Thornton JM. An atlas of protein topology cartoons available on the World-Wide Web. Trends Biochem Sci 1998;23:35–36.