Minireview

Diversification and evolution of L-*myo*-inositol 1-phosphate synthase[1]

Arun Lahiri Majumder*, Anirban Chatterjee, Krishnarup Ghosh Dastidar, Manoj Majee

Plant Molecular and Cellular Genetics, Bose Institute (Centenary Building), P-1/12 CIT Scheme VII M, Kolkata 700054, India

Received 15 July 2003; revised 26 August 2003; accepted 26 August 2003

First published online 10 September 2003

Edited by Robert B. Russell

Abstract L-*myo*-Inositol 1-phosphate synthase (MIPS, EC 5.5.1.4), the key enzyme in the inositol and phosphoinositide biosynthetic pathway, is present throughout evolutionarily diverse organisms and is considered an ancient protein/gene. Analysis by multiple sequence alignment, phylogenetic tree generation and comparison of newly determined crystal structures provides new insight into the origin and evolutionary relationships among the various MIPS proteins/genes. The evolution of the MIPS protein/gene among the prokaryotes seems more diverse and complex than amongst the eukaryotes. However, conservation of a 'core catalytic structure' among the MIPS proteins implies an essential function of the enzyme in cellular metabolism throughout the biological kingdom.

Key words: *myo*-Inositol; L-*myo*-Inositol 1-phosphate synthase; NAD binding; Oxidoreductase; Core structure; Protein evolution

1. Introduction

The current impetus on structural and functional genomics along with concurrent developments in bioinformatics has added a new dimension to our understanding of the evolution of proteins/genes across divergent prokaryotic and eukaryotic taxa. Evolutionarily conserved proteins may now be traced via nucleotide sequences and by analysis of the functional catalytic domains encrypted in their structural organization.

The biosynthesis of inositol has been known as an evolutionarily conserved pathway and its importance in the biological kingdom has been recognized for a long time. The cyclitol and its metabolites are involved in growth regulation, membrane biogenesis, osmotolerance and many other biological functions apart from its role as a second messenger in signal transduction pathways. *myo*-Inositol, physiologically the most common stereoisomer among the eight possible geometric isomers of inositol, also enters into an array of biochemical reactions having diverse functions in cellular metabolism and abundantly distributed throughout the hierarchy of the biological system [1–5].

All *myo*-inositol-producing organisms studied to date produce the cyclitol via dephosphorylation of *myo*-inositol 1-phosphate generated from glucose 6-phosphate by an internal oxidoreduction and aldol cyclization reaction catalyzed by L-*myo*-inositol 1-phosphate synthase (MIPS; EC 5.5.1.4) (reviewed in [6]). The MIPS enzyme has been reported from a host of diverse sources, such as higher plants and animals, parasites, fungi, green algae, bacteria and archaea, and has been considered to be an ancient protein/gene [6–18]. Chloroplast-specific isoforms of MIPS from algae and higher plants have also been reported [11,12,19–21]. The distribution of MIPS across the domains of life is summarized in Fig. 1. This ubiquitous distribution and antiquity of the MIPS protein/gene makes it an ideal model for studying evolution throughout the biological kingdom.

2. Key features of MIPS

From biochemical and biophysical studies done on different MIPS proteins, the following have emerged as important key features of the MIPS enzyme:

- MIPS is the first and the rate-limiting enzyme in the synthesis of all inositol-containing compounds.
- MIPS converts glucose 6-phosphate to L-*myo*-inositol 1-phosphate in an NAD^+-dependent manner.
  The NAD^+ binding in eukaryotic MIPS involves a Rossman fold characterized by a GXGGXXG motif typical of an oxidoreductase [22].
- The enzymatic conversion involves three partial reactions in which two enzyme-bound intermediates are implicated.
- MIPS has been characterized as a type II (archaeal) or type III (yeast) aldolase based on its requirement for either divalent or monovalent cation.
- The recent crystal structure analysis for the MIPS protein from *Saccharomyces cerevisiae* exemplifies a case of 'induced fit' model for binding of the substrate with the catalytic domain of the enzyme.

3. Eukaryotic and prokaryotic MIPS gene sequences: a phylogenetic analysis

Identification and cloning of the structural gene for MIPS (termed INO1) and determination of its nucleotide sequences were first reported in *S. cerevisiae* [7,8,23,24]. To date more than 60 INO1 genes have been reported from evolutionarily diverse organisms, both prokaryotic and eukaryotic (Table 1). When different representative MIPS amino acid sequences (available at http://www.ncbi.nlm.nih.gov/entrez) are compared against each other with a multiple alignment tool

*Corresponding author. Fax: (91)-33-334 3886.
E-mail address: lahiri@bic.boseinst.ernet.in (A.L. Majumder).

[1] Dedicated to Dr. Frank Eisenberg, Jr., who introduced A.L.M. to this fascinating enzyme/protein.

0014-5793/03/$22.00 © 2003 Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies.
doi:10.1016/S0014-5793(03)00974-8

Table 1
Diversity of MIPS across the spectrum of life, with key features of the gene/enzyme (names of organisms arranged alphabetically)

	Organism	Accession number		Gene and protein size		Remarks
		Protein	Nucleotide	Gene (kb)	Protein (subunit/ holoenzyme) (kDa)	
1.	*Actinidia arguta*	AAF97409	AY005128	~1.5	~58/–	(partial)
2.	*Aeropyrum pernix*	F72632	APE1517	~1.1	~40/–	(putative)
3.	*Anopheles gambii*	EAA00329	AAAB01008986	~1.5	~60/–	(putative)
4.	*Arabidopsis thaliana*	T50021	AY065415	~1.5	~60/~180	(native/exp)
5.	*Archaeoglobus fulgidus*	AAB89456	AE000979	~1.2	~45/–	(exp)
6.	*Aster tripolium*	BAC57963	AB090886.1			(partial)
7.	*Avena sativa*	BAB40956	AB059557	~1.5	~60/–	(putative)
8.	*Avicennia marina*	AAK21969	AY028259			(partial)
9.	*Bacillus cereus*	AAP09606	AE017006			(putative)
10.	*Bacteroides thetaiotaomicron*	AAO76633	AE016932.1			(partial)
11.	*Branchiostoma belcheri*	AAL02140	AY043320	~0.5		(partial)
12.	*Brassica napus*	AAB06756	U66307	~1.5	~60/~180	(putative)
13.	*Caenorhabditis elegans*	T18569	NM_064098	~1.6	~60/~180	(putative)
14.	*Candida albicans*	S45452	L22737	~1.6	~65/~240	(exp)
15.	*Chlamydomonas reinhardtii*	Chlamy database	20021010.7198.1	~1.5	~60/–	(putative)
16.	*Citrus paradisi*	CAA83565	Z32632	~1.5	~60/~180	(putative)
17.	*Corynebacterium glutamicum*	BAC00390	AP005283	~1.2	~45/–	(putative)
18.	*Drosophila melanogaster*	AAD02819	AF071104	~1.6	~60/–	(exp)
19.	*Entamoeba histolytica*	CAA72135	Y11270	~1.5	~60/~180	(exp)
20.	*Giardia lamblia*	EAA38884.1	AACB01000086			(putative)
21.	*Glycine max*	AAK49896	AF293970	~1.5	~60/~180	(native/exp)
22.	*Homo sapiens*	AAF26444	AF220530	~1.7	~70/–	(native)
23.	*Hordeum vulgare*	T04399/AAC17133	AF056325	~1.5	~60/~180	(native/exp)
24.	*Leishmania amazonensis*	AAB51376	U91965	~1.6	~65/–	(exp)
25.	*Leishmania major*	CAB94019	AL358652	~1.6	~65/–	(putative)
26.	*Leishmania mexicana*	CAC69873	AJ344544	~1.6	~65/–	(partial)
27.	*Lolium perenne*	AAN52772	AY154382	~1.5	~60/–	(putative)
28.	*Lycopersicon esculentum*	AAG14461	AF293460_1	~0.5		(partial)
29.	*Magnetospirillum magnetotacticum*	ZP_00048843	NZ_AAAP01001385			(partial)
30.	*Mesembryanthemum crystallinum*	AAB03687	U32511	~1.5	~60/~180	(native/exp)
31.	*Methanosarcina acetivorans*	AAM03529	AE010664	~1.1	~40/–	(putative)
32.	*Methanosarcina mazei*	AAM31066	AE013370	~1.1	~40/–	(putative)
33.	*Methanothermobacter thermoautotrophicus*	NP_276233	NC_000916	~1.1	~40/–	(putative)
34.	*Mus musculus*	AAF90201	AF288525	~1.6	~65/~200	(exp)
35.	*Mycobacterium leprae*	AAC43244	U00015	~1.1	~40/~160	(putative)
36.	*Mycobacterium tuberculosis*	P71703, NP_334460	Z80775	~1.1	~40/~160	(crystal) [28]
37.	*Mycobacterium bovis*	NP_853716	BX248334			(putative)
38.	*Nicotiana paniculata*	BAA84084	AB032073	~1.5	~60/~180	(putative)
39.	*Nicotiana tabacum*	BAA95788	AB009881	~1.5	~60/~180	(putative)
40.	*Neurospora crassa*	CAD70896	BX294019			(putative)
41.	*Novosphingobium aromaticivorans*	ZP_00096038	NZ_AAAV01000170			(partial)
42.	*Oryza sativa*	BAA25729	AB012107	~1.5	~60/~180	(native/exp)
43.	*Phaseolus vulgaris*	T10964/AAA91164	U38920.1	~1.5	~60/~180	(exp)
44.	*Pichia pastoris*	AAC33791	AF078915	~1.6	~65/~240	(exp)
45.	*Pinus taeda*	Pine Genomic Seq	Contig 7989	~1.5	~60	(putative)
46.	*Plasmodium falciparum*	CAD51482	AL929352	~1.8	~70/–	(putative)
47.	*Plasmodium yoelli*	EAA15800	AABL01001197	~1.8	~70/–	(putative)
48.	*Porteresia coarctata*	AAP74579	AF412340	~1.5	~60/~180	(native/exp)
49.	*Pyrobaculum aerophilum*	AAL63705	AE009838	~1.1	~40/~160	(putative)
50.	*Pyrococcus abyssi*	NP_126250	AJ248284	~1.1	~40/–	(putative)
51.	*Pyrococcus furiosus*	AAL81740	AE010261	~1.1	~40/–	(putative)
52.	*Pyrococcus horikoshii*	B75175	PAB1989	~1.1	~40/–	(putative)
53.	*Saccharomyces cerevisiae*	A30902	L23520	~1.6	~65/~230	(native/exp/cryst) [27]
54.	*Sesamum indicum*	AAG01148	AF284065	~1.5	~60/~180	(native/exp)
55.	*Solanum tuberosum*	AAK26439	AF357837_1	~1.5	~60/~180	(partial)
56.	*Spirodela polyrrhiza*	P42803	Z11693	~1.5	~60/~180	(native/exp)
57.	*Streptomyces coelicolor*	CAB38887	AL939115	~1.1	~40	(putative)
58.	*Suaeda maritima*	AAL28131	AF433879	~1.5	~60/~180	(putative)
59.	*Sulfolobus solfataricus*	AAK41169	AE006710.1			(partial)
60.	*Thermotoga maritima*	CAC21207	AJ401010.1	~1.1	~40/–	(putative)
61.	*Thermotoga neapolitana*	CAC21211	AJ401014	~1.1	~40/–	(partial)
62.	*Triticum aestivum*	AAD26332	AF120146	~1.5	~60/~180	(putative)
63.	*Xenopus laevis*	AAH44073.1	BC044073			(putative)
64.	*Xerophyta viscosa*		AY323824	~1.5	~60/~180	(exp)
65.	*Zea mays*	AAC15756	AF056326	~1.5	~60/~180	(putative)

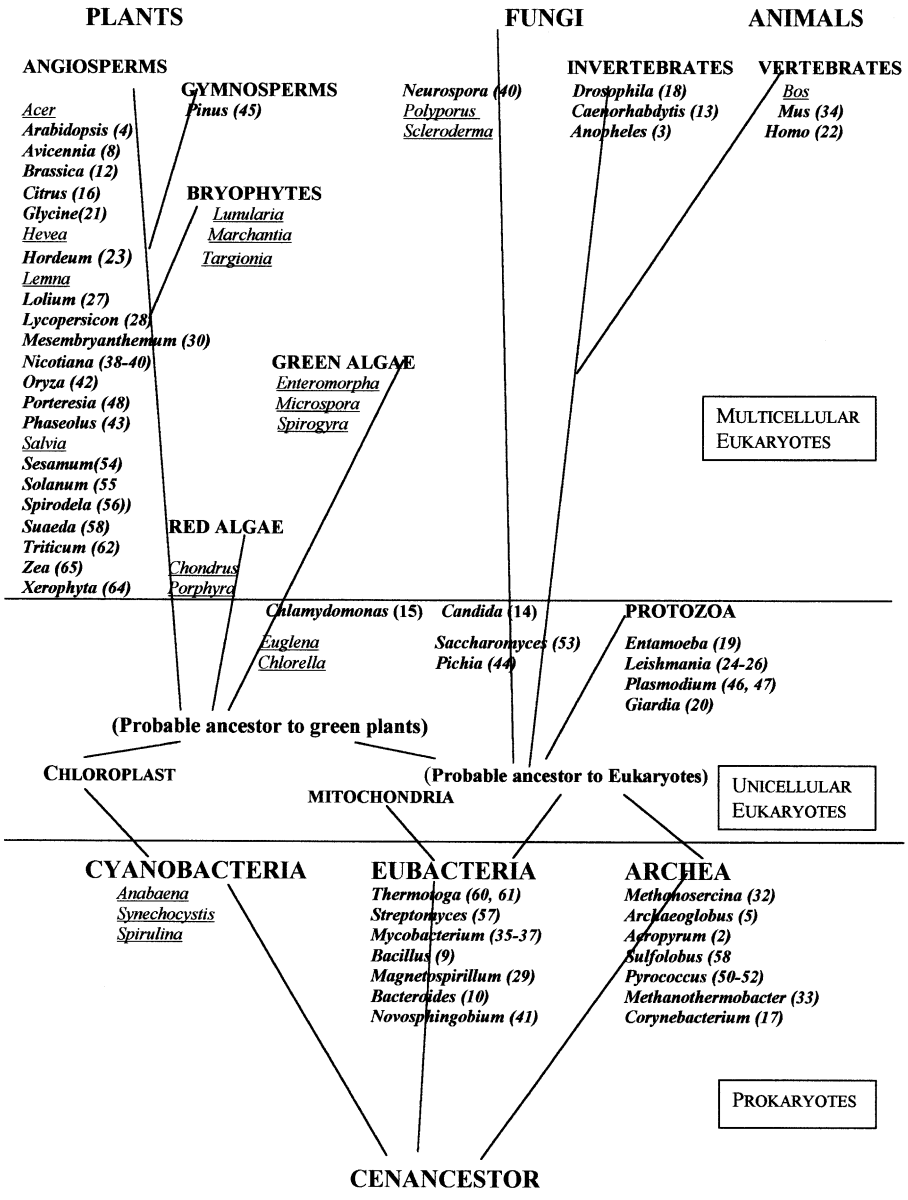

Fig. 1. Distribution of MIPS across the spectrum of life. The underlined names indicate organisms with a report of MIPS protein only, while names in bold indicate organisms for which a gene encoding MIPS has been identified (complete or partial). The numbers in parentheses correspond to the serial number of the organism in Table 1. The lines indicate the possible gene lineages. (Modified from [6].)

such as MULTALIN (http://prodes.toulouse.inra.fr/multalin/multalin.html) [25], a clear difference between the prokaryotic and the eukaryotic sequences is evident (alignment not shown). The eukaryotic sequences are clustered together with a very high degree of sequence similarity. The higher plants form one close subgroup, while the higher animals, the protozoa and the fungi form the other three subgroups in the eukaryotic cluster. As discussed earlier [6,13], stretches of amino acid residues such as GWGGNNG, LWTANTERY, NGSPQNTFVPGL and SYNHLGNNDG are conserved in MIPS proteins of all eukaryotes. Among the plants, larger stretches of the amino acid residues are conserved throughout the length of the protein showing a much higher degree of preservation of sequence identity irrespective of

←
Cryst = X-ray crystallographic data available of protein products; exp = protein overexpressed and purified from heterologous sources and characterized for MIPS activity; native = protein purified from native sources and characterized for MIPS activity; partial = where full gene sequences are not available, these are also putative annotations; putative = theoretical annotation of full-length cDNA/contig sequences, no experimental proof as to whether the protein product does have MIPS activity.
The *Chlamydomonas* MIPS sequence is available from the *Chlamydomonas* genome sequence database maintained in http://www.biology.duke.edu/chlamy_genome.
The Pine MIPS sequence is available from the Loblolly Pine genome sequence database maintained in http://pinetree.ccgb.umn.edu.

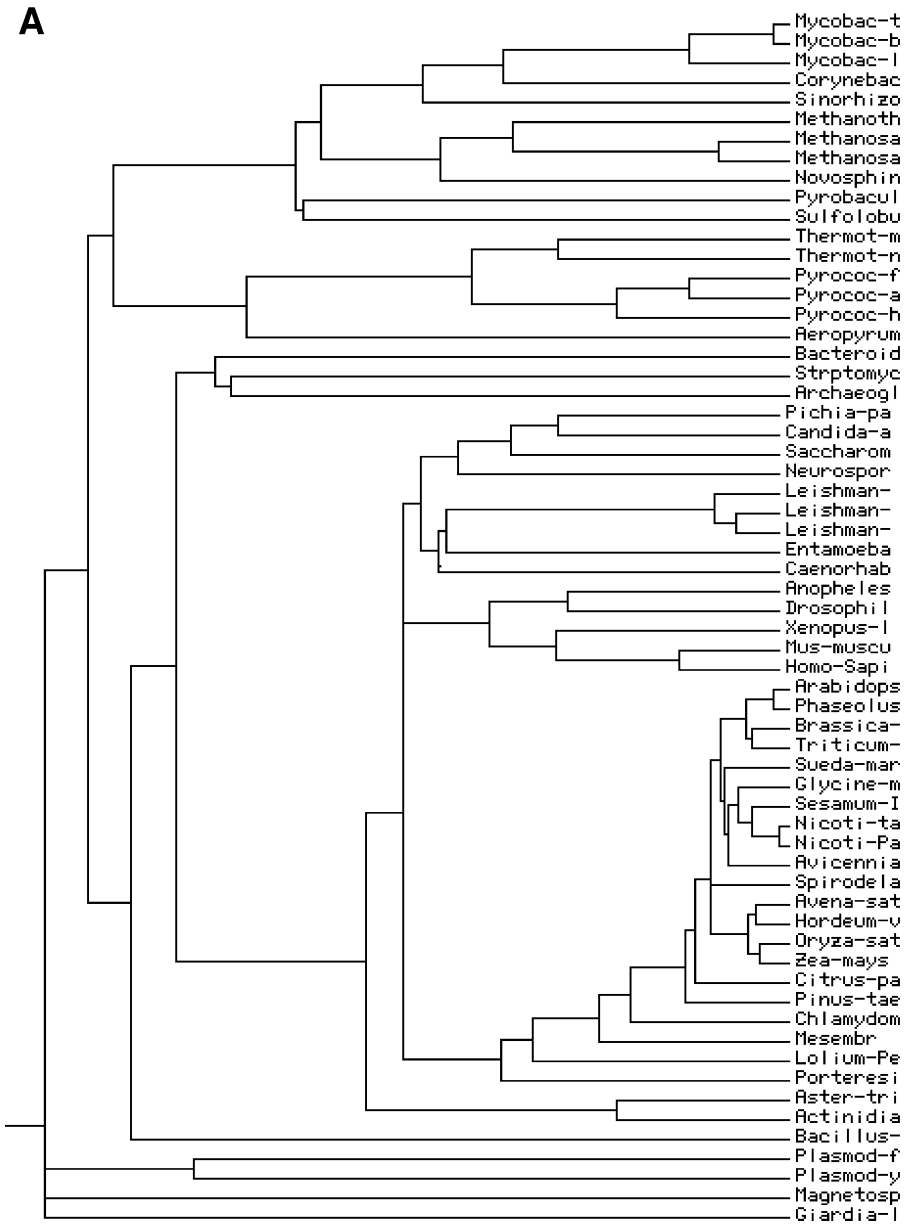

Fig. 2. A: Phylogenetic tree of all known MIPS amino acid sequences known to date. For names and corresponding accession numbers see Table 1. The distances were measured using the Phylip-type tree options of the 'TREETOP' phylogenetic tree generation software, available at http://www.genebee.msu.su/services/phtree_reduced.html. The tree was drawn using the Phylodendron interface, from http://iubio.bio.indiana.edu/treeapp/treeprint-form.html. A hundred rounds of bootstrapping were done to ensure the validity of the tree. B: Multiple sequence alignment of the probable active site regions from different MIPS sequences showing the remarkable conservation of the amino acid residues considered to be important for MIPS activity. The numbers at the top indicate the corresponding numbers and the asterisks correspond to the proposed or identified active site amino acid residues for the *Mycobacterium tuberculosis* MIPS sequence and those at the bottom indicate the same corresponding to the *S. cerevisiae* MIPS sequence.

whether the plant is a monocot, dicot, green alga or gymnosperm. This certainly indicates a monophyletic origin of the higher plant MIPS. In the case of fungi, such as *S. cerevisiae*, *Pichia pastoris* and *Candida albicans*, there is an extra amino acid stretch at the N-terminal end, unique to this group. This sequence is highly conserved amongst the fungi and is probably due to a later addition in the early fungal ancestral sequence after its divergence from the main eukaryotic stock. In the case of animals, the MIPS sequence has an extra C-terminal sequence, the significance of which is not clear as yet.

Among the prokaryotes the different MIPS sequences are quite divergent amongst themselves, and are far more distinct from any of the known eukaryotic sequences. The amino acid sequence alignment of the prokaryotic MIPS fails to show any striking similarity across different prokaryotic taxa as exemplified by a comparison between eubacterial and archaebacterial MIPS sequences. The *Archaeoglobus* MIPS [16] has more sequence similarity to the eukaryotic MIPS than the other known prokaryotic ones. The distribution of the MIPS sequences in the prokaryotic cluster shows that some archaeal

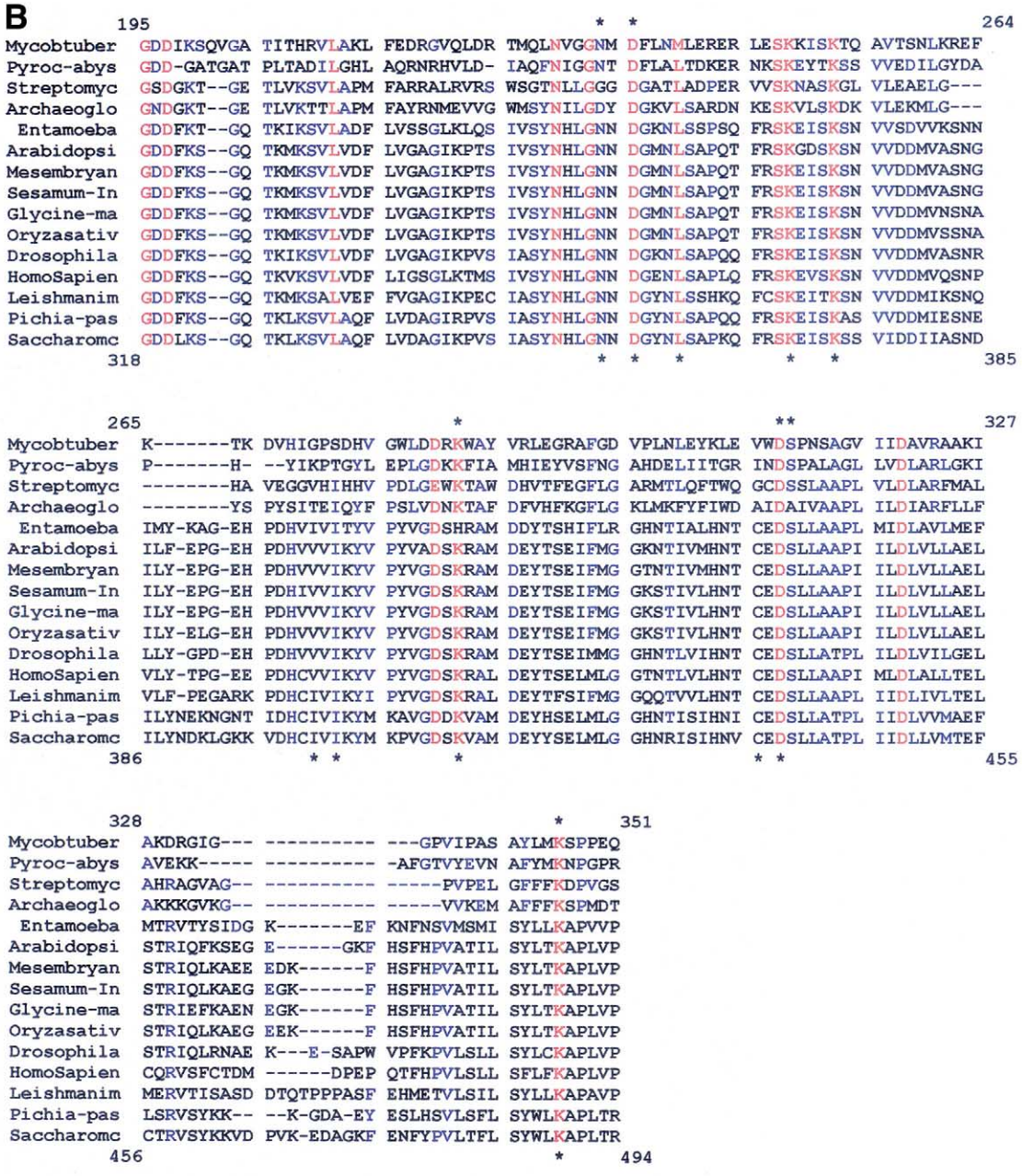

Fig. 2 (*Continued*).

(e.g. *Aeropyrum*, *Pyrococcus*) sequences share closer homology with eubacterial (e.g. *Thermotoga maritima* and *T. horikoshii*) sequences than with the other archaeal sequences such as *Archaeoglobus* or *Methanosarcina* and *Methanothermobacter*. The eubacterial MIPS sequences of *Mycobacterium* and *Streptomyces* share closer homology with each other than with other eubacteria such as *Thermotoga*. Based on such analysis of the MIPS sequences, a phylogenetic tree has been drawn by the topological algorithm of the Tree-Top interface (http://www.genebee.msu.su/services/phtree_full.html) (Fig. 2A) [26]. It is revealed that the prokaryotic stock diverged early and has since undergone profound changes in the amino acid sequence. In contrast, the eukaryotic stock of MIPS sequences are far more closely related and probably remained stable in its monophyletic origin.

4. Eukaryotic and prokaryotic MIPS crystal structures: a clue towards conservation of an essential 'core catalytic domain' throughout evolution

Do MIPS proteins across evolutionarily divergent taxa conserve an essential 'core catalytic domain' in spite of the divergence among them? Recent analysis of the crystal structure of two of the known MIPS enzymes, one eukaryotic (*S. cerevisiae*) and the other prokaryotic (*Mycobacterium tuberculosis*) [27–29] along with the reported MIPS gene sequences (Table 1), provides an inkling towards the presence of a 'core structure' in all MIPS proteins conserved throughout evolution.

The *Saccharomyces* MIPS crystal structure shows a remarkable homotetrameric association, in which the protein shows a 222 symmetry with a non-crystallographic two-fold axis relat-

ing two monomers in an asymmetric unit and a crystallographic two-fold axis relating the two molecules at one end. In both cases the dimerization interface of the protomers and the tetramerization interfaces of the dimers bury a large surface area. Integrity of the total structure is maintained by a set of hydrophobic interactions including both aromatic (e.g. W) as well as non-aromatic (e.g. I, L) residues. These hydrophobic interactions are important in subunit interaction across the interfaces and also in maintaining the solution structure as evident from the non-dissociation of the subunits in aqueous environment. The MIPS structure from *Saccharomyces* has three well-defined domains, a central domain including the N- and C-terminal ends which is involved in subunit interactions, an NAD binding domain containing a modified Rossman fold and a catalytic domain which contains the active site amino acids and residues that occur at the tetramerization interface [27].

In the case of the MIPS structure from *Mycobacterium*, two well-defined domains (D1 and D2) linked through two hinged regions have been predicted. Of these, D1a contains a modified Rossman fold, and D1b is a C-terminal extension whereas D2 is the tetramerization domain. It is possible that the D1 domain of the prokaryotic MIPS has evolved to more complexity in the eukaryotes and developed into the two separate well-defined domains (NAD binding domain and the central domain) [27]. The D2 domain has in parallel undergone further complexity in the eukaryotic MIPS to give rise to the catalytic domain containing the tetramerization interface (as in *Saccharomyces* MIPS). For NAD binding, the putative Rossman fold signature (GXGXXG) [22] is quite conserved in most of the eukaryotic MIPS sequences while it is not so distinct in the case of the prokaryotic MIPS sequences. However, in both *Saccharomyces* and *Mycobacterium* MIPS crystal structure, similar conformation of the NAD ring with respect to the phosphodiester and *N*-glycosylic bonds is seen in the total redox cycle and the C4 (pro-S) and Si face is accessible for the hydride shift. In both cases a long looped structure followed by the GXGXXG signature establishes strong interactions with the pyridine ring of the NAD, which probably prevents the dissociation of the NAD molecule during the redox cycle into the medium. Hence, it is probable that the NAD binding site had evolved early and has been maintained across evolutionary timescales.

The most striking feature about the two MIPS structures is the absence of electron density for a stretch of amino acids in the crystal. In the *Mycobacterium* MIPS, it is the stretch of 26 amino acids in the D2 domain (241–267), whereas in case of the *Saccharomyces* MIPS it is the 58 amino acid stretch (351–409) of the tetramerization interface bearing the catalytic domain. However, when the *Saccharomyces* MIPS protein was crystallized in the presence of 2-deoxyglucitol-6 phosphate, a substrate analogue of the enzyme, nucleation around these 58 amino acids resulted in an observed electron density in their diffraction pattern. The results exemplify the 'induced fit' model hypothesized earlier by Koshland [30]. It is possible that the aforesaid unorganized region comprising 26 amino acids in the *Mycobacterium* MIPS would have been found to be similarly nucleated around the substrate analogue. It may be assumed that such a mechanism of 'induced fit' might have evolved quite early in evolution and remained largely unchanged from early prokaryotes to modern eukaryotes. In a more recent work [31], structures of NAD^+ and NADH-bound yeast MIPS were studied. The NADH-bound form was found to be ordered in the presence of a phosphate and a glycerol in the active site concomitant with a repositioning of the nicotinamide ring and a motion of a loop region to accommodate the bound phosphate. Interestingly, possible presence of a metal ion in the yeast MIPS was also suggested. Although the mechanistic role of this metal ion is undefined as yet this raises questions about the universality of the proposed reaction mechanism of the MIPS enzyme so far conjectured.

From the multiple sequence alignment of the 'active site' region by MULTALIN [25], it is revealed that the aforesaid stretch of 58 amino acids in *Saccharomyces* MIPS or that of the 26 amino acids of *Mycobacterium* MIPS is found to be remarkably conserved (~73%) in all MIPS from bacteria to human (Fig. 2B). An alignment of such selected amino acid sequences of MIPS genes from eubacteria, archaea, parasites, fungi, plant or animal sources reveal that the stretch harbors the critical amino acid residues (marked with asterisks in Fig. 2B) presumed to be essential for substrate binding to the MIPS enzyme. Moreover, an amino acid stretch SYNHLGNNDG, one of the four stretches of amino acids of eukaryotic MIPS identified as 'highly conserved' (i.e. GWGGNNG, LWTANTERY, NGSPQNTFVPGL and SYNHLGNNDG), also resides in this region. Such considerations point towards the possibility that a 'core structure' for catalytic activity is conserved throughout evolution among all the MIPS genes/proteins. It is only logical to assume that the protein, because of its essential function(s) in the biological system, retains its catalytic domain intact throughout evolution of organisms despite allowing diversification elsewhere in its structure. This contention can be verified by structure elucidation of MIPS from various representative phyla in future.

5. MIPS diversification and evolution: some insights

The ubiquitous distribution of MIPS calls for some essential function(s) of the protein in the biological kingdom. Depletion of inositol has been known to induce 'inositol-less death' in a number of organisms and is well studied in the eukaryotic system [32]. Different inositol derivatives are also known to be essential metabolites for both prokaryotic and eukaryotic organisms [5,33]. Hence, it is not surprising that the MIPS gene/protein has been detected in virtually all life forms along the evolutionary lineage (Fig. 1). Although most of such studies have centered around the cytosolic MIPS, an organellar MIPS from algal and higher plant chloroplasts, regulated by light and salt, has been of interest to this laboratory [11,12] and others [20,21]. Recent studies have identified a yet unassigned open reading frame of *Synechocystis* as encoding a functional MIPS thus establishing the cyanobacterial origin of the chloroplast MIPS isoform (Chatterjee et al., communicated). It will be interesting to investigate if the eubacterial MIPS turns out to be a progenitor of a yet undetected mitochondrial MIPS as well!

It seems that for prokaryotes the cenancestral MIPS sequence has evolved in diverse paths from early stages of evolution (Fig. 2A) keeping a few core amino acids of the enzyme/protein conserved despite diverse organization of the gene(s) (Fig. 2B). It is reasonable to speculate that variability in the environment might have been a major contributing factor towards such widespread diversity among the prokaryotic MIPS sequences. An example of the environmental effects

that might have 'induced' the diverse design in the MIPS protein structure is the thermotolerant MIPS from the archaeon *Archaeoglobus fulgidus*. This enzyme is active at very high temperature ($>60°C$) and exhibits the highest catalytic activity among all MIPS proteins known to date [16]. Although no structural information is available as yet for the archaeal enzyme, the sequences certainly have contributed to the stable architecture of the enzyme/protein at high temperature. This design might have accrued from the need of this hyperthermophilic archaeon for synthesis of di-*myo*-inositol-1,1′-phosphate, an osmolyte found in this group of organisms inhabiting high salinity and temperature [34]. Moreover, the MIPS from *Archaeoglobus*, identified as a type II aldolase, is active in the presence of divalent cations such as Zn^{2+} or Mn^{2+} while all the eukaryotic MIPS proteins, known to be of type III aldolase, require NH_4^+ for their optimal activity. This is also justified in the sequences of *Archaeoglobus* and *Saccharomyces* proteins where D259 of the former is replaced by N354 in the latter. Similarly, in the case of the *Mycobacterium* MIPS, a Zn^{2+} cation is found to be an integral part rendering proper conformation of the presumed active site cleft [29]. In this regard, the recent report of the possible presence of a metal ion in yeast MIPS [31], however, is striking. What induced the change in the eukaryotic MIPS to a preference for NH_4^+ over a divalent cation for optimal catalytic activity or to a type III aldolase from a type II aldolase remains an intriguing question at this stage.

All eukaryotic MIPS seems to have evolved from one common stock, probably from the fusion of an archaebacterial and a eubacterial MIPS gene [35]. The similarity between the *Archaeoglobus* and the eukaryotic MIPS sequences might be a pointer to this major event. The ancestral eukaryotic MIPS gene should have had traces of both the eubacterial and archaebacterial MIPS genes in it. However, since then, the eukaryotic MIPS sequence has maintained striking conservation across the different eukaryotic groups as evident from the multiple sequence alignment (data not presented). Moreover, some of the important amino acid residues identified in the active site of the *S. cerevisiae* MIPS (namely Q325, L352, N354, D356, L360, K369, K373, I400, I402, K412, C436, D438 and K489) are all highly conserved in all the eukaryotic MIPS enzymes. These amino acids can be considered to be part of a 'eukaryotic core structure' which has remained largely the same during evolution, even though the rest of the protein sequence has changed over time. Even among the eukaryotes, environmental pressure might have worked for designing MIPS proteins with a unique regulatory character as evidenced by the recent report of a salt-tolerant MIPS from a halophytic wild rice, *Porteresia coarctata* (*PINO1*; GenBank accession number AF412340) reported from this laboratory. This gene is characterized by alterations and rearrangements in its nucleotide sequence compared to other eukaryotic INO1 genes and is believed to have assumed an oligomeric protein structure stable towards in vitro salt effects as exemplified by biochemical and biophysical experiments with bacterially expressed proteins as well as its functional expression in planta (Majee et al., communicated).

It is generally argued that for any protein evolving through different lineages, in addition to changes due to random, non-adaptive evolutionary forces, diversity might arise due to adaptive changes in its gene sequences to suit the environment of the organism. However, to preserve the original function of the protein, a core functional structure must remain conserved through all such changes [36,37]. An ancient protein as it is, MIPS may not be an exception.

Acknowledgements: A.L.M. is supported by grants from DBT, DST and CSIR, Government of India. A.C. and K.G.D. thank CSIR, Government of India for providing Senior Research Fellowships to them. Thanks are due to the Staff of the Bioinformatic Centre, Bose Institute for their cooperation and in providing the necessary infrastructure.

References

[1] Eisenberg Jr., F., Bolden, A.H. and Loewus, F.A. (1964) Biochem. Biophys. Res. Commun. 14, 419–424.
[2] Biswas, B.B., Ghosh, B. and Majumder, A.L. (1984) in: Subcellular Biochemistry, Vol. 10 (Roodyn, D.B., Ed.), pp. 237–280, Plenum Press, London.
[3] Loewus, F.A. (1990) in: Inositol Metabolism in Plants (Moore, D.J., Boss, W.F. and Loewus, F., Eds.), pp. 13–19, Wiley Liss, New York.
[4] Bohnert, H.J., Nelson, D.E. and Jensen, R.G. (1995) Plant Cell 7, 1099–1111.
[5] Loewus, F.A. and Murthy, P.P.N. (2000) Plant Sci. 150, 1–19.
[6] Majumder, A.L., Johnson, M.D. and Henry, S.A. (1997) Biochim. Biophys. Acta 1348, 245–256.
[7] Majumder, A.L., Duttagupta, S., Goldwasser, P., Donahue, T.F. and Henry, S.A. (1981) Mol. Gen. Genet. 184, 347–354.
[8] Donahue, T.F. and Henry, S.A. (1981) J. Biol. Chem. 256, 7077–7085.
[9] Tian, F., Migaud, M.E. and Frost, J.W. (1999) J. Am. Chem. Soc. 121, 5795–5796.
[10] Hegeman, C.E., Good, L.L. and Grabau, E.A. (2001) Plant Physiol. 125, 1941–1948.
[11] Adhikari, J., Majumder, A.L., Bhaduri, T.J., DasGupta, S. and Majumder, A.L. (1987) Plant Physiol. 85, 611–614.
[12] RayChaudhury, A., Hait, N.C., DasGupta, S., Bhaduri, T.J., Deb, R. and Majumder, A.L. (1997) Plant Physiol. 115, 727–736.
[13] Chun, J.-A., Jin, U.-H., Lee, J.-W., Yi, Y.-B., Hyung, N.-I., Kang, M.-H., Pyee, J.-H., Suh, C.-M., Kang, C.-W., Seo, H.-Y., Lee, S.-W. and Chung, C.-H. (2003) Planta 216, 874–880.
[14] Bachhawat, N. and Mande, S. (1999) J. Mol. Biol. 291, 531–536.
[15] Bachhawat, N. and Mande, S. (2000) Trends Genet. 16, 111–113.
[16] Chen, L., Zhou, C., Yang, H. and Roberts, M.F. (2000) Biochemistry 39, 12415–12423.
[17] Majumder, A.L., Hait, N.C., Deb, I., Majee, M., Chatterjee, A., Ghosh Dastidar, K., Bhattacharyya, S., Ghosh, S., Chatterjee, A., Maitra, S. and Pattanaik, S. (2003) in: Molecular Insight in Plant Biology (Nath, P., Mattoo, A.K., Ranade, S.A. and Weil, J.H., Eds.), pp. 67–76, Oxford and IBH, New Delhi.
[18] Lohia, A., Hait, N.C. and Majumder, A.L. (1999) Mol. Biochem. Parasitol. 98, 67–69.
[19] Hait, N.C., RayChaudhury, A., Das, A., Bhattacharyya, S. and Majumder, A.L. (2002) Plant Sci. 162, 559–568.
[20] Johnson, M.D. and Wang, X. (1996) J. Biol. Chem. 271, 17215–17218.
[21] Lackey, K.H., Pope, P.M. and Johnson, M.D. (2003) Plant Physiol. 132, 2240–2247.
[22] Kleiger, G. and Eisenberg, D. (2002) J. Mol. Biol. 323, 69–76.
[23] Klig, L.S. and Henry, S.A. (1984) Proc. Natl. Acad. Sci. USA 81, 3816–3820.
[24] Johnson, M.D. and Henry, S.A. (1989) J. Biol. Chem. 264, 1274–1283.
[25] Corpet, F. (1988) Nucleic Acids Res. 16, 10881–10890.
[26] Brodsky, L.I., Ivanov, V.V., Kalaidzidis, Ya.L., Leontovich, A.M., Nikolaev, V.K., Feranchuk, S.I. and Drachev, V.A. (1995) Biochemistry 60, 923–928.
[27] Stein, A.J. and Geiger, J.H. (2002) J. Biol. Chem. 277, 9484–9491.
[28] Kniewel, R., Buglino, J.A., Shen, V., Chadha, T., Beckwith, A. and Lima, C.D. (2002) J. Struct. Func. Genomics 2, 129–134.
[29] Norman, R.A., McAlister, M.S.B., Murray-Rust, J., Movahedzadeh, F., Stoker, N.G. and McDonald, N.Q. (2002) Structure 10, 393–402.

[30] Koshland, D.E. (1958) Proc. Natl. Acad. Sci. USA 44, 98–105.
[31] Jin, X. and Geiger, J.H. (2003) Acta Crystallogr. D59, 1154–1164.
[32] Henry, S.A., Atkinson, K.D., Kolat, A.I. and Culbertson, M.R. (1977) J. Bacteriol. 130, 472–484.
[33] Loewus, F.A. and Loewus, M.W. (1983) Annu. Rev. Plant Physiol. 34, 137–161.
[34] Chen, L., Spiliotis, E. and Roberts, M.F. (1998) J. Bacteriol. 180, 3785–3792.
[35] Margulis, L. (1996) Proc. Natl. Acad. Sci. USA 93, 1071–1076.
[36] Monizngo, A.F., Marcote, E.M. and Robertson, J.D. (1996) Nat. Struct. Biol. 3, 133–140.
[37] Chothia, C. and Gerstein, M. (1997) Nature 385, 579–581.