JOURNAL OF BACTERIOLOGY, Apr. 1996, p. 2161–2171 Vol. 178, No. 8
0021-9193/96/$04.00+0

MINIREVIEW

Evolutionary Recruitment of Biochemically Specialized Subdivisions of Family I within the Protein Superfamily of Aminotransferases†

ROY A. JENSEN* AND WEI GU

Department of Microbiology and Cell Science, University of Florida, Gainesville, Florida 32611

GENE RECRUITMENT AND THE GENERATION OF GENE FAMILIES

It is generally accepted that ancient organisms must have possessed small genomes producing fewer gene products than contemporary organisms. Since the evolution of high specificity is surely a demanding process, primitive enzymes are likely to have been broad-specificity catalysts, utilizing a family of related substrates and producing a family of related products (67). Events of gene duplication could then have facilitated the differential narrowing of substrate specificities to produce the expanded repertoire of specialized enzymes seen in contemporary organisms. The gene recruitment hypothesis (27) states that the latter scenario is the basis for the emergence of gene families whose members catalyze a variety of related enzymatic reactions. The ever-enlarging database of deduced amino acid sequences is revealing the outlines of many large and still-growing gene families whose common ancestor may often have been a fundamental multipurpose unit in the primitive systems referred to above.

Although numerous expected homology relationships have been confirmed, sometimes reasonable predictions of homology relationships have not been correct. For example, since the beta subunit of tryptophan synthase (TrpB) in *Escherichia coli* is a pyridoxal 5′-phosphate (PLP)-dependent enzyme that can be replaced functionally by a constitutively expressed tryptophanase (also PLP dependent) and because both catalyze β-elimination reactions, homology was predicted (27). However, TrpB and tryptophanase appear to be analogous rather than homologous (1, 47). It is additionally striking that wholly unanticipated homologs of TrpB have surfaced (47), namely, threonine synthase, threonine dehydratase, and D-serine dehydratase (all PLP-dependent enzymes).

Homolog relationships that are perhaps even more surprising have also emerged. For example, pyruvate oxidase (a peripheral-membrane redox flavoprotein) and the large catalytic subunit of acetohydroxyacid synthase isoenzymes (cytoplasmic proteins in the common pathway of isoleucine-valine biosynthesis) are homologs (4). In such instances of unforeseen homology, it has inevitably been possible to rationalize a common ancestor in terms of related mechanistic features of catalysis. Indeed, the elucidation of such relationships has been quite illuminating at both the biochemical and evolutionary levels.

In this minireview the model of gene recruitment is used to analyze one section of aminotransferase homologs whose membership includes a large gene family of orthologous and paralogous proteins.

RELATIONSHIP OF AMINOTRANSFERASES WITH OTHER PLP-DEPENDENT PROTEINS

Genes encoding aminotransferases illustrate a large and rapidly growing gene superfamily. These aminotransferases separate into four families which do not exhibit obvious homology when amino acid sequences are analyzed by standard alignment programs such as the PILEUP program in the University of Wisconsin Genetics Computer Group sequence analysis software package (11). However, a more sophisticated analysis technique for evaluation of distantly related proteins indicates that all four families belong to one divergent superfamily (39). Even more striking, the homology relationships of this superfamily have been expanded (1) to include a large number of other PLP-dependent enzymes which share the property that covalent alterations of substrate are almost always targeted to the α carbon attached to the amino substituent forming an imine linkage with the coenzyme. It is thus apparent that this huge, recently recognized division of PLP-dependent proteins possesses a common regiospecificity but encompasses a multiplicity of catalytic capabilities.

In addition to the aminotransferases, this alpha division of PLP-dependent enzymes includes glycine hydroxymethyltransferase, glycine *C*-acetyltransferase, 5-aminolevulinate synthase, 8-amino-7-oxononanoate synthase, a class of amino acid decarboxylases, tryptophanase, and tyrosine phenol-lyase (1). One family of amino acid decarboxylases is related to aminotransferases but not to any of three other families of PLP-dependent amino acid decarboxylases (54). The relatively facile mechanistic requirements of decarboxylation reactions presumably have occasioned an increased probability of independent evolutionary origins of the same reaction (54).

Figure 1 illustrates the immense homology network to which the aminotransferase superfamily belongs. It is, on the one hand, only one of numerous superfamilies within the alpha division of PLP-dependent enzymes; on the other hand, the aminotransferase superfamily has a breadth of membership that subdivides into families which in turn are being rapidly resolved into subfamily units. Aminotransferase family IV consists of serine aminotransferase and phosphoserine aminotransferase. Aminotransferase family III consists of D-alanine aminotransferase and branched-chain aminotransferase. Aminotransferase family II proteins catalyze reactions in which the distal amino moiety of basic amino acid substrates such as L-ornithine undergoes transamination.

This minireview is focused upon the single family boxed in Fig. 1 to show how a protein phylogeny, for which an enlarging

* Corresponding author. Phone: (904) 392-9677. Fax: (904) 392-5922.

† Florida Agricultural Experiment Station journal series no. R-05010.

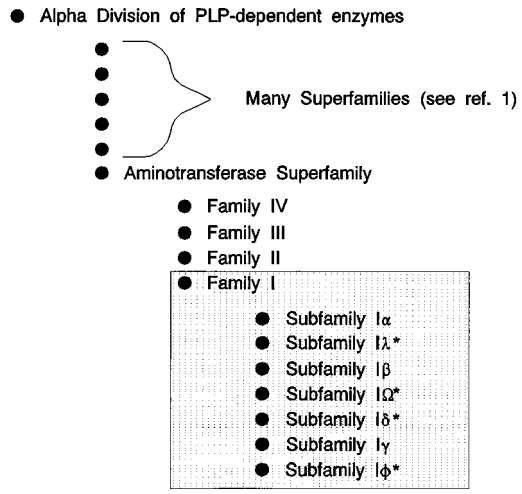


FIG. 1. Hierarchical relationships of aminotransferase homologs. The most recent additions are indicated by asterisks.

hierarchical breadth is currently being recognized, is also expanding internally to yield fine-tuned subdivisions.

FAMILY I AMINOTRANSFERASES

A dendrogram constructed by multiple alignment of deduced amino acid sequences (Fig. 2) depicts the evolutionary relationships of 47 proteins that belong to aminotransferase family I. A number of published sequences were not included (e.g., a large number of vertebrate and higher plant sequences) in order to avoid overrepresentation of closely related sequences in the data analysis. These proteins were previously segregated into three subfamilies (39) which we have designated Iα, Iβ, and Iγ (19). In this minireview we have designated four additional subfamilies: Iλ, IΩ, Iδ, and Iφ. As new sequences enter the database, the most outlying ones have the potential to be the initial members of new subgroups. For example, earlier, when many of the sequences shown were not yet available, *E. coli* HisC (Eco HisC) (subfamily Iβ) and rat Tat(c) (subfamily Iγ) were merely outlying lineages of the then-known sequences within the present subfamily Iα.

The most quantitatively active aminotransferases in nature involve the three keto acids generated during glycolytic and respiratory cycles and their amino counterparts, but at least 60 aminotransferases are thus far known (5). Many of these enzymes catalyze specialized reactions and are low-abundance proteins, e.g., the histidine biosynthesis enzymes in subfamily Iβ. Crystallographic studies of mammalian mitochondrial and cytosolic species of aspartate aminotransferase have provided an advanced understanding of the relationship of primary amino acid sequence to the spatial and covalent structures (5). Later-published X-ray crystallographic studies of *E. coli* AspC (subfamily Iα) revealed extensive similarity of three-dimensional structure in comparison with the *E. coli* and vertebrate aminotransferases (29, 57). It will be most interesting to obtain the spatial structures of proteins representing the remaining subfamilies in view of the varied substrate specificities and the varied patterns of conserved amino acid residues found in the different subfamilies.

The vertebrate enzyme is a homodimer. Each subunit consists of an extended amino-terminal segment (residues 1 to 14, numbered according to porcine cytosolic aspartate aminotransferase), a small domain (residues 15 to 47 and residues 326 to 412), and a large domain (residues 48 to 325). The binding of substrate facilitates a substantial conformational change, and the small domain located in one subunit migrates toward the bound PLP on the other subunit, closing the entrance to the active site. The amino-terminal segment is tethered to the large domain of the neighboring subunit and is apparently necessary for proper mobility of the small domain. The established mechanism of action for the vertebrate enzymes in subfamily Iα is closely tied to spatial interactions between domains and between subunits. Since amino acid residues shown to function in intersubunit and interdomain interactions are highly conserved within subfamily Iα, most members of this subfamily are undoubtedly homodimers that share a very similar overall mechanism. Given the common protein scaffold shared by alpha division proteins (1), the fine-tuned differences and similarities of interdomain interactions and of intersubunit interactions and the eventual elucidation of their relationship to mechanism variation at the subfamily level should be most intriguing.

FUNCTIONAL DIVERGENCE

Paralogous and orthologous homologs. Protein homologs are termed paralogous if they originated via gene duplication. When homologs were produced by speciation, their relationship is termed orthologous. If a set of homologs under comparison contains only orthologs (the simplest case), the dendrogram obtained by multiple alignment of amino acid sequences should roughly parallel the phylogenetic dendrogram of the organisms producing them. The homologs displayed in Fig. 2 contain both orthologs and paralogs. *Haemophilus influenzae*, an organism whose entire genome has been sequenced (12), possesses a paralogous set of four family I homologs: *H. influenzae* AspC (Hin AspC) (Iα), Hin HisC (Iβ), Hin HisH (Iβ), and Hin Alt (Iδ). *H. influenzae* is a relatively close phylogenetic neighbor of *E. coli*, and one therefore expects the close orthologous relationships in fact seen in Fig. 2 when comparing the Eco AspC-Hin AspC and Eco HisC-Hin HisC pairs. Consider, however, that if only the Eco AspC and Hin Alt homologs were known, Hin Alt would appear to be out of register with its expected phylogenetic position. One might erroneously conclude that Hin Alt arose by horizontal transfer from another organism (i.e., a xenologous relationship). This illustrates that in cases where a set of paralogous proteins became established in an ancestral cell long before divergence caused by speciation, a comparison of mismatched orthologous proteins can lead to evolutionary misinterpretations.

Functional divergence achieved in a set of paralogous proteins is not limited to alterations of substrate specificity. Aspartate aminotransferase exemplifies the existence in animals and plants of paralogs which exhibit similar substrate specificities but which are spatially compartmented in different organelles or different specialized cell types. In concert with other enzymes (e.g., as with the malate-aspartate shuttle), such separately compartmented isoenzymes are essential for transfer of reducing equivalents between organelles, for transfer of fixed carbon between cell types in C_4 plants, and for assimilation of reduced nitrogen into aspartate and asparagine in legume root modules.

In *E. coli*, AspC and TyrB are coexisting paralogs in which *tyrB* achieved sensitivity to repression control by L-tyrosine. The retention of a degree of overlapping substrate specificity exemplifies the basis for "backup" catalytic capability. TyrB can fully compensate for a state of *aspC* deficiency when repression of *tyrB* is eliminated physiologically or by mutation (15, 28). Thus, because of their common ancestry, paralogous

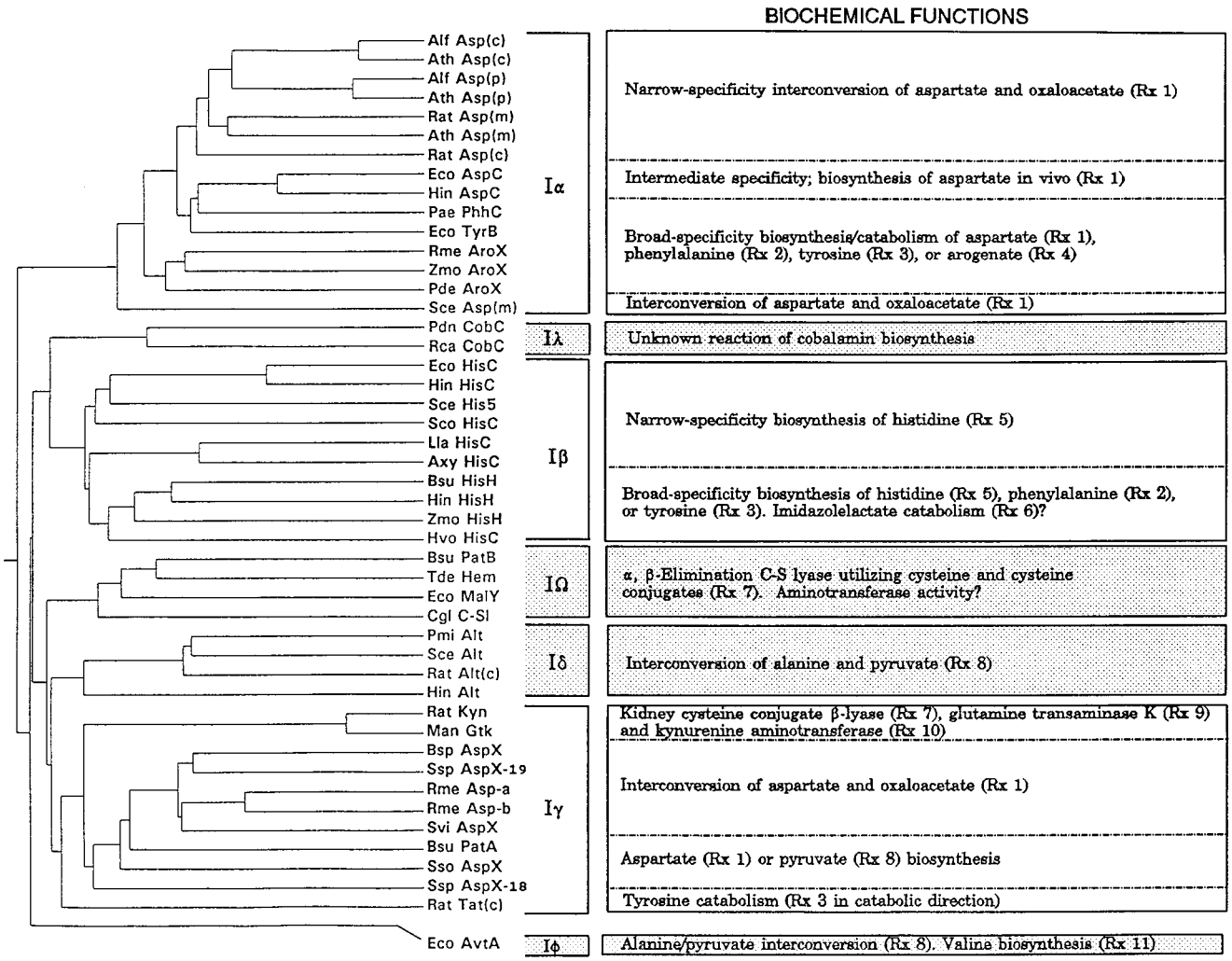


FIG. 2. Dendrogram showing amino acid homology relationships of family I aminotransferases. References are given below in parentheses after each designated aminotransferase (AT). For uniformity, AT proteins whose published designations have not been widely used are named AroX or AspX in cases where the in vivo function appears to be aromatic AT or aspartate AT, respectively. Lineages belonging to new subfamilies that are currently being recognized are assigned provisional subfamily designations of Iλ, IΩ, Iδ, and Iφ. Alf Asp(p) and Alf Asp(c) are selected representatives of lineages which have many other members (see reference 55); Rat Asp(c) and Rat Asp(m) also represent lineages which have many other vertebrate members (see reference 4). Functions of enzymes or clusters of enzymes are indicated on the right. Enzyme reactions (Rx) specified are listed in Table 1. Abbreviations: Alf Asp(c), alfalfa cytosolic aspartate AT (65); Ath Asp(c), *Arabidopsis thaliana* cytosolic aspartate AT (55); Alf Asp(p), alfalfa plastidial aspartate AT (14); Ath Asp(p), *A. thaliana* plastidial aspartate AT (55); Rat Asp(m), rat mitochondrial aspartate AT (24); Ath Asp(m), *A. thaliana* mitochondrial aspartate AT (55); Rat Asp(c), rat cytosolic aspartate AT (20); Eco AspC, *E. coli* aspartate AT (13); Hin AspC, *H. influenzae* aspartate AT (12); Pae PhhC, *Pseudomonas aeruginosa* aromatic AT (68; unpublished data); Eco TyrB, *E. coli* aromatic AT (13, 34), Rme AroX, *Rhizobium meliloti* aromatic AT (66); Zmo AroX, *Zymomonas mobilis* aromatic AT (16a); Pde AroX, *Paracoccus denitrificans* phenylalanine AT (62); Sce Asp(m), *Saccharomyces cerevisiae* mitochondrial aspartate AT (41); Pdn CobC, *Pseudomonas denitrificans* CobC (8); Rca CobC, *Rhodobacter capsulatus* CobC (51); Bsu PatB, *B. subtilis* unknown AT (64); Tde Hem, *Treponema denticola* hemolysin (6); Eco MalY, *E. coli* regulator of maltose metabolism (52); Cgl C-Sl, *Corynebacterium glutamicum* C-S lyase (53); Pmi Alt, *Panicum miliaceum* alanine AT (59); Sce Alt, *S. cerevisiae* alanine AT (EMBL accession no. Z48758); Rat Alt(c), rat cytosolic alanine AT (25); Hin Alt, *H. influenzae* alanine AT; Rat Kyn, rat kynurenine AT (42, 50); Man Gtk, human glutamine transaminase K (49); Bsp AspX, *Bacillus* species aspartate AT (60); Ssp AspX-19 and Ssp AspX-18, *Synechocystis* sp. aspartate AT proteins (30); Rme Asp-a (66) and Rme Asp-b (2), *R. meliloti* aspartate AT; Svi AspX, *Streptomyces virginiae* aspartate AT (31); Bsu PatA, *B. subtilis* aspartate AT (48); Sso AspX, *Sulfolobus solfataricus* aspartate AT (9); Rat Tat(c), rat cytosolic tyrosine AT (17); and Eco AvtA, *E. coli* alanine/valine AT (58). Human cytosolic alanine aminotransferase is nearly identical to rat Alt(c) (26). HisH, HisC, and His5 are synonymous designations for imidazole acetol phosphate AT from *Z. mobilis* (Zmo) (19), *B. subtilis* (Bsu) (22), *Halobacterium volcanii* (Hvo) (7), *E. coli* (Eco) (18), *S. cerevisiae* (Sce) (44), *Streptomyces coelicolor* (Sco) (35), *Acetobacter xylinum* (Axy) (63), *Lactococcus lactis* (Lla) (10), and *H. influenzae* (Hin) (12). In this group, a partial mycobacterial sequence is also available (23).

proteins provide cells with enhanced evolutionary potential through suppressor mutations.

Endosymbiotic origins of organellar aminotransferases? According to the endosymbiotic hypothesis of organelle origin, plastids and mitochondria arose via independent endosymbiotic events. If so, one might expect mitochondrial aminotransferases to cluster with that from an organism such as *Paracoccus denitrificans* (Pde AroX), whereas plastid-localized aminotransferases might perhaps cluster with those of cyanobacteria (e.g., *Synechocystis* sp. [Ssp AspX-18 or Ssp AspX-19]). However, Fig. 2 shows that all of the animal and plant aminotransferases in subfamily Iα cluster together, regardless of whether the spatial location is chloroplast, mitochondrion, or cytosol [indicated by (c), (m), or (c), respectively, immediately after the enzyme designation]. This clustering indicates that all of the plant and animal aminotransferase genes in

TABLE 1. Established aminotransferase I family reactions

Designation in Fig. 2	Enzyme reaction[a]
Rx 1	Oxaloacetate + GLU ⇌ L-aspartate + 2-KG
Rx 2	Phenylpyruvate + GLU ⇌ L-phenylalanine + 2-KG
Rx 3	4-Hydroxyphenylpyruvate + GLU ⇌ L-tyrosine + 2-KG
Rx 4	Prephenate + GLU ⇌ L-arogenate + 2-KG
Rx 5	Imidazoleacetol-P + GLU → histidinol-P + 2-KG
Rx 6	Imidazolepyruvate + GLU → L-histidine + 2-KG
Rx 7	Cysteine-R[b] → ammonia + pyruvate + HSR
Rx 8	L-Alanine + 2-KG ⇌ pyruvate + GLU
Rx 9	L-Glutamine + pyruvate (or phenylpyruvate) ⇌ 2-ketoglutaramate + L-alanine (or L-phenylalanine)
Rx 10	L-Kynurenine + 2-KG ⇌ 2-aminobenzoyl-pyruvate + GLU
Rx 11	L-Alanine + 2-keto-isovalerate ⇌ pyruvate + L-valine

[a] GLU, glucose; 2-KG, 2-ketoglutarate; P, phosphate.
[b] R is any conjugate with the sulfur moiety. If R is hydrogen, then cysteine-R is cysteine and HSR is H_2S.

subfamily Iα radiated from a common ancestor at a time following divergence from prokaryote genes. These observations are not easily reconciled with the endosymbiotic hypothesis of organelle evolution. This dilemma is further exacerbated in that the mitochondrial species of aspartate aminotransferase from *Saccharomyces cerevisiae* [Sce Asp(m)] represents the most outlying lineage of subfamily Iα and clusters with neither the prokaryote proteins nor the animal and plant proteins.

Biochemical specialization at the subfamily level. Figure 2 shows that a diversity of functional roles is carried out by family I aminotransferases. The substrate specificities and catalytic properties of many of the proteins considered here have not been adequately studied in many cases. In spite of this limitation, distinct functional roles can be associated with different subfamily groupings. Subfamily Iα mediates the interconversion of L-aspartate and oxaloacetate. The plant, animal, and (probably) yeast enzymes exhibit narrow specificity in catalyzing the aspartate aminotransferase reaction. Prokaryote enzymes fall into two groups. The aspartate aminotransferase reaction is physiologically relevant for *E. coli* AspC (and probably *H. influenzae* AspC), but it accepts aromatic amino acids as poor alternative substrates. The remaining prokaryote enzymes probably function equally well as aspartate aminotransferase or aromatic aminotransferase, judging from the rigorous work done with *E. coli* TyrB (21).

Three subfamilies seem to be specialized for a single function. Subfamily Iλ is presumably dedicated to an unknown aminotransferase reaction required for cobalamin biosynthesis (37). Subfamily Iδ is specialized for mediating the interconversion of L-alanine and pyruvate. Essentially the same function may exist for subfamily Iϕ, currently represented by a single member.

Subfamily Iβ consists wholly of imidazoleacetol phosphate aminotransferases, proteins that catalyze a step of L-histidine biosynthesis. The latter proteins are highly specific enzymes in the sense that no other aminotransferases are known to be capable of utilizing imidazoleacetol phosphate, and deficient *hisC* (also called *hisH* and *his5*) mutants are inevitably auxotrophic. However, *Zymomonas mobilis* HisH and *Bacillus subtilis* HisH have been shown to be capable additionally of utilizing aromatic amino acid substrates in vitro (19, 43) and in vivo (43). HisH might also perform the aminotransferase step of imidazolelactate catabolism in some organisms (19).

Subfamily IΩ is represented by four members, only one of which has been characterized at the enzymatic level. *Corynebacterium glutamicum* C-S lyase catalyzes an α-β elimination reaction with cysteine or cysteine conjugates (Table 1). It is interesting that the recently described *Treponema denticola* hemolysin (6) requires cysteine for hemolysin activity and generates H_2S. It is currently unknown whether any of these proteins catalyzes an aminotransferase reaction.

Subfamily Iγ, as presently constituted, contains a diverse membership which splits into the four groupings shown in Fig. 2. Three catalytic activities of crucial importance in mammals (kidney cysteine conjugate β-lyase, kidney glutamine transaminase, and kynurenine aminotransferase)—though studied separately for years—have proven to be synonymous (42, 49, 50). These enzymes had previously been known to have broad specificities that include aromatic amino acids. In striking contrast, another cluster contains members that are highly specific aspartate aminotransferases, as best documented by *Bacillus* species AspX (Bsp AspX) (61). Another group contains members that accept pyruvate in addition to aspartate, as demonstrated for *Sulfolobus solfataricus* AspX (Sso AspX) (3). Finally, rat Tat(c) is specialized for the essential function of tyrosine catabolism in mammals. Recently, a homolog (39% identity) of rat Tat(c) was found (3a) in *Trypanosoma cruzi* (not shown in Fig. 2). The *T. cruzi* enzyme recognizes all three aromatic amino acids quite well and utilizes oxaloacetate, 2-ketoglutarate, and pyruvate (the favored substrate) as alternative amino acceptors (40). This broad-specificity enzyme would seem to be a strong candidate for a contemporary "dinosaur" similar to the common ancestor of the superfamily. In view of the distinctly different specificities of the kynurenine aminotransferase (Kyn) and glutamine transaminase K (Gtk) grouping and rat Tat from each other and from the remaining independent subfamily Iγ members, the Kyn/Gtk and rat Tat lineages may very well merit subfamily status in the near future.

EVOLUTIONARY SCENARIO FOR FAMILY I AMINOTRANSFERASES

The recruitment hypothesis for evolutionary acquisition of a family of specialized, narrow-specificity enzymes from a broad-specificity forerunner suggests the following scenario. The common family I aminotransferase ancestor was probably a broad-specificity enzyme that could function as an aspartate aminotransferase, an aromatic aminotransferase, and an alanine aminotransferase. This all-purpose enzyme may very likely have possessed relatively poor additional catalytic capabilities that could have been sufficient for low-level output pathways (e.g., cobalamin, histidine) or that could have been fortuitous and initially nonfunctional (e.g., cysteine conjugate β-lyase).

An early gene duplication initiated the Iα branch, which narrowed its specificity to the aspartate aminotransferase and aromatic aminotransferase reactions. Further specificity changes led to specialization in vivo as either aspartate aminotransferases or aromatic aminotransferase. The contemporary members of subfamily Iα are all aspartate aminotransferases that accept aromatic amino acids to a negligible extent (e.g., vertebrate enzymes), to a moderate extent (e.g., *E. coli* AspC), or very well (e.g., *E. coli* TyrB). Since all known prokaryote subfamily Iα members are gram-negative bacteria, the first ancestor of subfamily Iα may have originated in gram-negative bacteria some time after divergence from other prokaryotes.

The remaining ancestral gene underwent successive gene duplications that allowed the evolution of narrowed specificities to achieve the following biochemical specializations: L-

TABLE 2. Functions of anchor residues in family I aminotransferases

Hierarchical level of invariance	Residue[a]	Functional role[b]	Exceptions
Alpha division of PLP-dependent proteins	K-258	Schiff base with PLP	None
	D-222 or E-222	Salt bridge via H bond with the pyridine N-1 of PLP[c]	None
Aminotransferase superfamily	G-197	Accommodates the turn at the interdomain interface	*L. lactis* HisC (Iβ)
	R-386	Salt bridge via H bonding with α-anionic moiety of substrate	None
Family I aminotransferases	*Y-70	H bonding with phosphate OP2 of PLP	Subfamily Iλ
	G-110	Unknown role	None
	N-194	H bonding with hydroxyl O-3′ of PLP	*L. lactis* HisC and Axy (Iβ)
	P-195	Has *cis* conformation within span of interdomain interface	None
	Y-225	H bonding with hydroxyl O-3′ of PLP	Subfamilies Iλ and IΩ
	R-266	H bonding with phosphates OP2 and OP4 of PLP	Subfamilies IΩ and Iγ
	G-268	Unknown role	Subfamily IΩ

[a] Numbered according to conventional usage based upon the numbering of the prototype pig cytosolic aspartate aminotransferase sequence (46). Residue Y-70 contacts PLP from the opposite subunit (denoted with an asterisk by convention) with respect to all other residues that are PLP ligands.
[b] As elucidated by spatial structure determination (32).
[c] Note that residue 222 is always D within family I aminotransferases.

alanine and pyruvate interconversion (subfamily Iφ and subfamily Iδ), L-histidine biosynthesis (subfamily Iβ), cobalamin biosynthesis (subfamily Iλ), and L-aspartate and oxaloacetate interconversion (subfamily Iγ). In Subfamily Iβ, the usual aminotransferase recognition of the negatively charged alpha carboxyl moiety of the substrate was altered to recognize a negatively charged phosphate moiety instead. Within subfamily Iγ, the rat Kyn and human Gtk (Man Gtk) orthologs have retained a strikingly great breadth of substrate recognition and may most nearly resemble the common family I ancestor.

Several biochemical specializations have occurred independently in different subfamilies, e.g., alanine and pyruvate interconversion and aspartate and oxaloacetate interconversion. Specialization for biosynthesis or catabolism of aromatic amino acids has been particularly interesting, having occurred independently in three subfamily lineages: Iα, Iβ, and Iγ. Within subfamily Iβ, one cluster of enzymes include *B. subtilis* HisH (Bsu HisH) and *Z. mobilis* HisH (Zmo HisH). These proteins double as aromatic aminotransferases, probably explaining why the cognate genes are appropriately excluded as members of the histidine-controlled histidine operon. Note that *H. influenzae* has both a histidine operon gene specifying Hin HisC and a nonoperonic gene specifying Hin HisH. It would be interesting to compare the substrate specificities of these two proteins, one specialized for L-histidine biosynthesis and the other presumably for aromatic amino acid biosynthesis. In comparison to *H. influenzae*, the closely related *E. coli* possesses an operon-specified *E. coli* HisC protein for L-histidine biosynthesis but lacks an *H. influenzae* HisH counterpart. Instead aromatic aminotransferase (TyrB) was recruited from the Iα subfamily via gene duplication of *aspC*, followed by broadened substrate specificity and acquisition of repression control by *tyrR*.

SIGNATURE AMINO ACID MOTIFS AT TWO DIFFERENT HIERARCHICAL LEVELS

Family and subfamily patterns. Family I aminotransferases (and indeed the entire alpha division of PLP-dependent enzymes) share a common protein scaffold configuration which is directly related to the positioning of the PLP component (1).

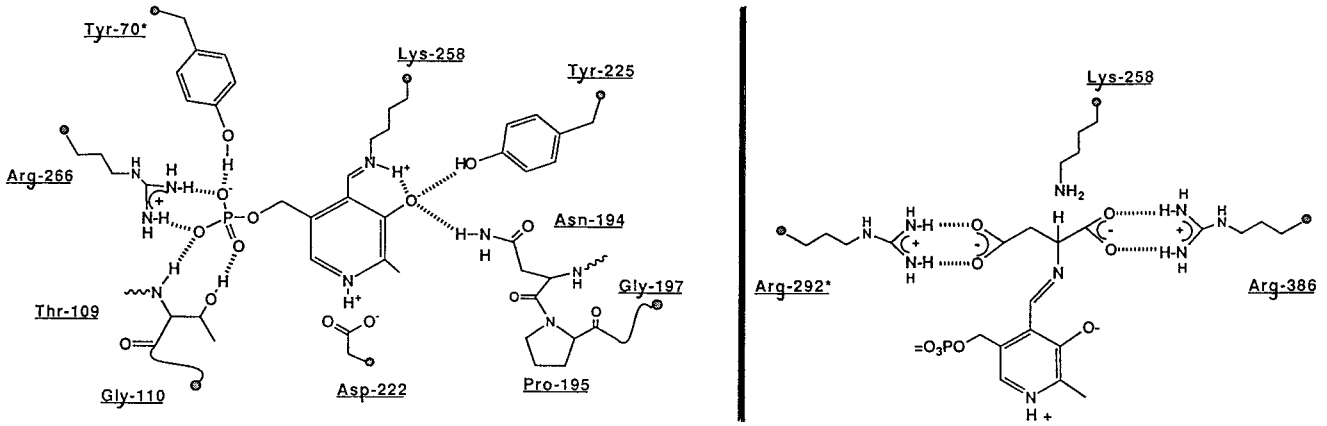
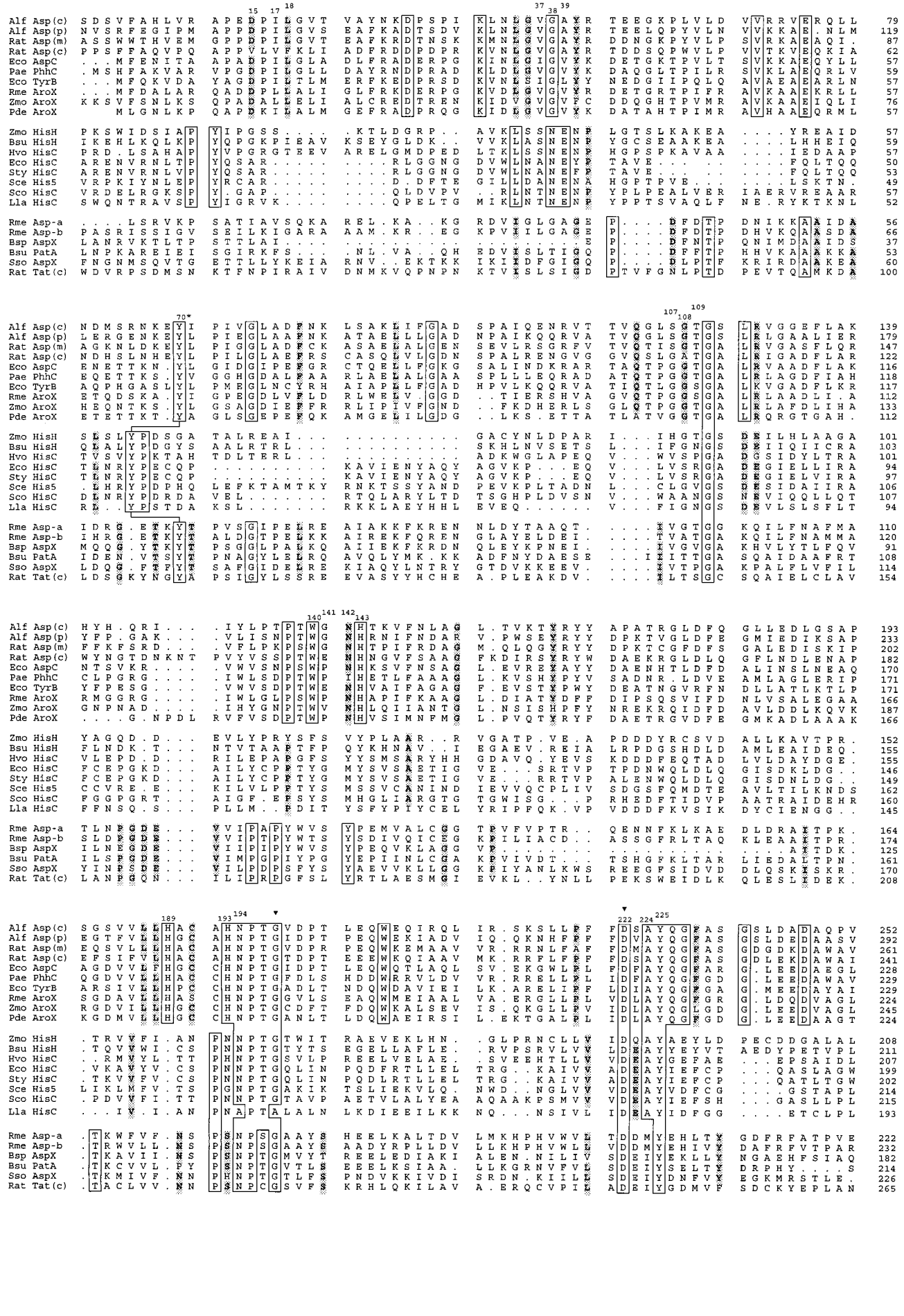


FIG. 3. Schematic diagrams of the active center of a family I aminotransferase showing the contacts of anchor residues with PLP (left) or with the L-aspartate substrate (right). Hydrogen-bonding and salt bridge interactions between the PLP cofactor and the side chain and main chain atoms of all of the anchor residues mentioned in Table 2 (except for Gly-268) are shown. The function of Gly-268 is unknown. The contacts of L-aspartate, bound as a Schiff base with PLP, with Arg-292* and Arg-386 (an anchor residue) are also shown. Asterisks denote residues located on the adjacent subunit.



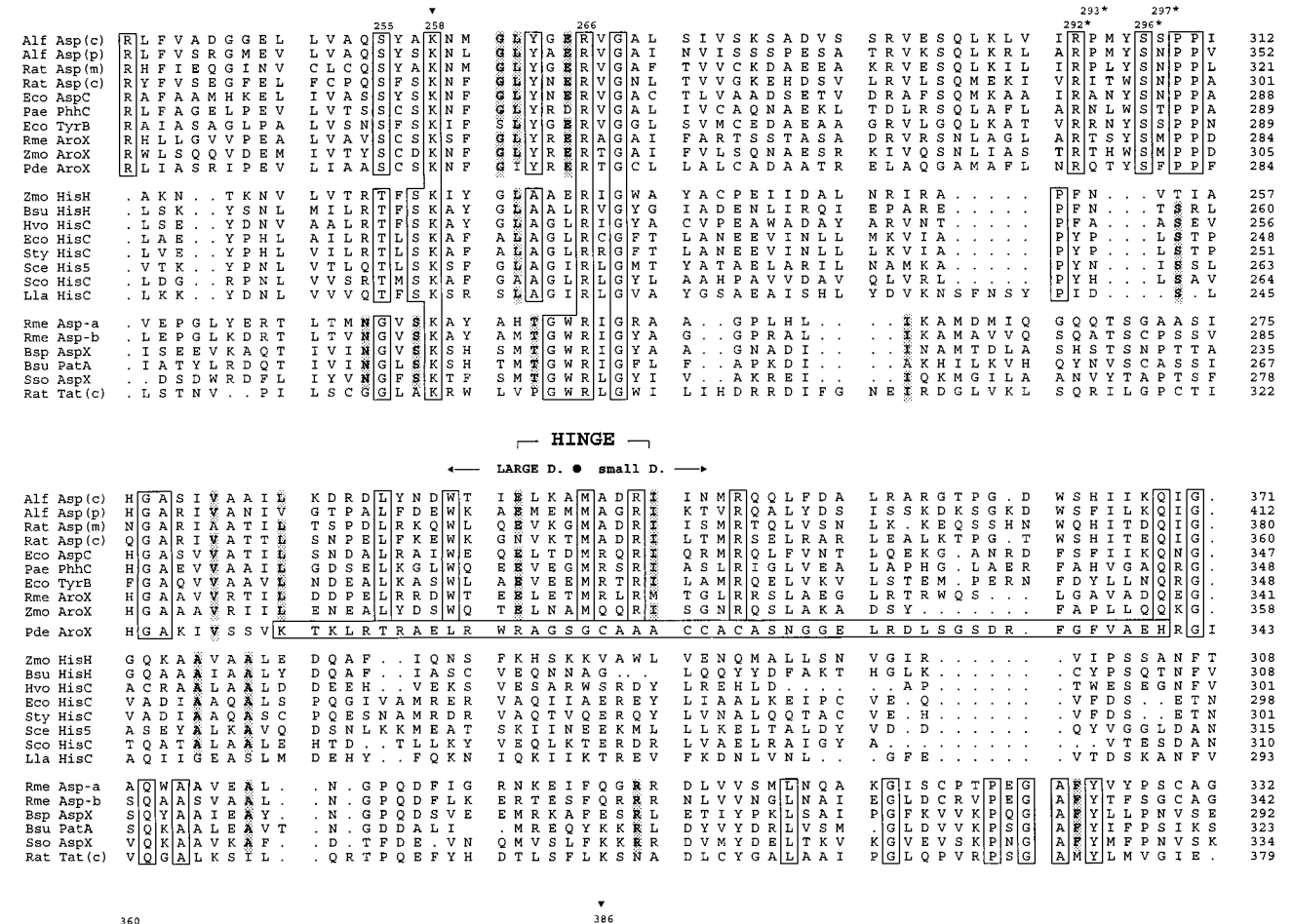


FIG. 4. Multiple alignment of family I aminotransferases. Subfamilies Iα, Iβ, and Iγ (see dendrogram of Fig. 2) are shown from top to bottom in each block. The four emerging subfamilies, represented by relatively few sequences in the database at this time, are not shown. Residues numbered above each block are those identified as catalytic-site residues for porcine cytosolic aspartate aminotransferase. Asterisks (e.g., 70*) indicate residues contributed by the adjacent homodimer subunit. Residues conserved within each subfamily are boxed, and residues conserved in all but one member of a subfamily are shaded. Completely conserved but out-of-alignment residues thought to be equivalent throughout are joined by lines. Residues conserved throughout all four aminotransferase families are indicated with small solid triangles, one of which is the active-site lysine (K-258). The hinge region at the point of juxtaposition (●) of the large domain (LARGE D.) and small domain (small D.) is shown. The long horizontal box enclosing 47 amino acid residues of *P. denitrificans* AroX (Pde AroX) highlights a sequence which varies markedly from its homologs and appears to be a sequencing error (see text). The published amino acid sequence for *Lactococcus lactis* HisC (Lla HisC) between residues 5 and 18 (NSGSKSLYWQVK) was considered a frameshift error, because it overlapped a gap present in other HisC sequences and because it lacked the highly conserved residues, PY. When translated in another reading frame, the sequence TRAVSPYLGRVK obtained aligned much better with its closest homologs as shown and included the conserved PY residues. For abbreviations of aminotransferases, see the legend to Fig. 2.

Eleven "anchor" residues are conserved throughout family I, and the functions of these residues are given in Table 2. Of these residues, four are common to the entire aminotransferase superfamily (one exception), and two of these four are invariant throughout the alpha division of PLP-dependent proteins. In family I, most of the invariant residues are ligands of PLP. Schematic diagrams are given (Fig. 3) that illustrate the interaction of anchor residues with PLP in the absence of

substrate, as well as with L-aspartate bound as a Schiff base with PLP in the active site.

Figure 4 provides a multiple alignment of amino acid sequences of selected members of those subfamilies (Iα, Iβ, and Iγ) which enjoy the most extensive documentation in the database thus far. Although residues 70 and 386 are not perfectly aligned in Fig. 4, they are considered to be functionally equivalent (38).

Each subfamily exhibits a different pattern of conservation around the 11 anchor residues, thus providing signature amino acid motifs that discriminate the seven subfamilies (Fig. 5). The signature motifs given in Fig. 5 reflect relatively subtle variations of atomic interactions, since they center around common anchor residues and exhibit considerable sequence overlap. Figure 4 reveals that, in addition, each subfamily possesses patterns of residue conservation that are highly individualistic. Each subfamily pattern is likely to reflect important specific differences in spatial contacts. For example, the subfamily Iα hinge region [$W(x_6)M(x_2)R(x_4)R$] shown in Fig. 4 is highly conserved and has a thoroughly established role in the mechanism utilized by Iα-type enzymes (see "substrate specificity" below). There is no trace of these conserved residues in any of the remaining subfamilies. Atomic and spatial interactions have thus far been elucidated only for subfamily Iα, where most highly conserved residues have indeed been identified as important mechanism-of-action entities. The unique sets of highly conserved residues in each of the remaining subfamilies will surely correspond to important differences in specificity and mechanism destined to be revealed by X-ray crystallographic analysis. Sequences representing subfamilies Iλ, IΩ, Iδ, and Iϕ were not included in Fig. 4, because more sequences are needed to identify subfamily patterns of residue conservation.

THE PARADIGM SUBFAMILY WITHIN THE AMINOTRANSFERASE SUPERFAMILY

Catalytic mechanism. The extensive and elegant work which has been done with vertebrate aminotransferases and (to a lesser extent) with *E. coli* aminotransferases within subfamily Iα has provided a superb basis for comparative work—both at a fine-tuned level within the subfamily and at a level of expansion to homolog groups beyond subfamily Iα. A brief summary of the type of information currently available for subfamily Iα members follows; analogous information can be anticipated for remaining subfamilies of family I in the near future.

Fifty-one residues are invariant within subfamily Iα, and many more are highly conserved. X-ray crystallographic analyses of the vertebrate aspartate aminotransferases have established very substantial information at the atomic level (5), and the spatial configuration of *E. coli* AspC has been shown to be very similar (29). Both substrate and coenzyme contact each subunit of the homodimer. The α-carboxyl group of substrate interacts with R-386 (small domain), while the distal-carboxyl group interacts with R-292 (large domain) of the neighboring subunit. The residues acting as PLP ligands (Table 2) contact a common subunit, except for Y-70 which contacts the adjacent subunit. Y-70 (catalytic function) and R-292 (substrate-binding function) from adjacent subunits contribute to a common active site.

There are two classes of critical interactions between amino acid residues: domain interactions and subunit interactions. One set of major surface contacts for domain interactions (using *E. coli* AspC residues as an example) are between G-38 to D-42 (small domain) and Y-263, D-325, and M-326 to I-330 (large domain). Another set of surface contacts is between N-357 to F-362 and R-386 (small domain) and C-192 to P-196 (large domain). Second, two types of intersubunit contacts have been demonstrated. One is between large-domain residues Y-263 to R-266 and V-50 to Q-58 of one subunit and *Y-295 to H-301 of the other subunit (denoted by asterisk). The other is between the N-terminal residues M-5 to A-11 of one subunit and residues *F-118 to K-126, *G-183, and *A-283 of the other subunit.

Of the 29 active-site residues identified for vertebrate aspartate aminotransferases and noted as such in Fig. 4, only G-38, *Y-70, W-140, H-143, H-189, H-193, N-194, D-222, A-224, Y-225, S-255, K-258, R-266, *R-292, *S-296, F-360, and R-386 are invariant throughout subfamily Iα. A portion of the published sequence for *P. denitrificans* AroX (Pde AroX) which appears to be in error is shown in Fig. 4 by enclosure within the long horizontal box. This region fails to match any of the residues that are invariant elsewhere within the subfamily, an unexpected observation, since this region contains the hinge residues which define rotation of the small domain upon substrate binding (see "substrate specificity" below). Codon usage of the questioned region was compared with the remainder of the sequence. Six rare codons were used eight times in the questioned region, but not at all in the remainder of the sequence. Codon preference analysis, using a database containing 41 *P. denitrificans* genes, confirmed that the sequencing of this region must be erroneous.

Inferences about other subfamilies. Aside from the highly conserved PLP-interacting residues and the R-386 residue which contacts the α-anionic substituent of the substrate, few regions are conserved when the subfamilies are compared with one another. The interdomain interface (residues 192 to 197 of the large domain), is a highly conserved region throughout the family. However, note (Fig. 4) that the small-domain region with which it interacts in subfamily Iα is not evident in either subfamily Iβ or subfamily Iγ. Thus, functional divergence has led to a variation in the particular spatial contacts which specify domain interactions and subunit interactions. Subfamily Iλ is unique among subfamilies in lacking the otherwise completely conserved Y-70. Since Y-70 contacts the adjacent subunit within the homodimer unit, this might mean that Iλ proteins are not homodimers. This would be a striking difference, because all family I aminotransferases that have been characterized for physical properties have thus far proven to be homodimers.

Substrate specificity. Among the well-characterized proteins within subfamily Iα, it has been possible to consider the impact of the spatial orientations of active-site residues upon substrate specificity. Rat Asp(c), rat Asp(m), *E. coli* AspC, and *E. coli* TyrB represent an ordered continuum of proteins with respect to an increased ability for utilization of aromatic amino acids. Rat Asp(c) is essentially specific for aspartate; Rat Asp(m) utilizes aromatic amino acids very poorly; *E. coli* AspC prefers aspartate by 3 orders of magnitude (21) but can nevertheless utilize aromatic amino acids in vivo under some conditions; and *E. coli* TyrB utilizes aromatic amino acids and aspartate equally well in vitro (21).

The vertebrate aspartate aminotransferases have been shown to undergo a conformational change with the binding of substrate, resulting in movement of the small domain closer to the substrate. The small domain of an enzyme such as rat Asp(c) rotates around residue T-325. The hinge residues (N-322 to I-330) which mediate rotation of the small domain for rat Asp(c) are labeled in Fig. 4. In this region, the key residues for small-domain rotation are the V-323–K-324 pair and A-327–D-328. The conformational change is related to the high degree of specificity for aspartate seen for vertebrate

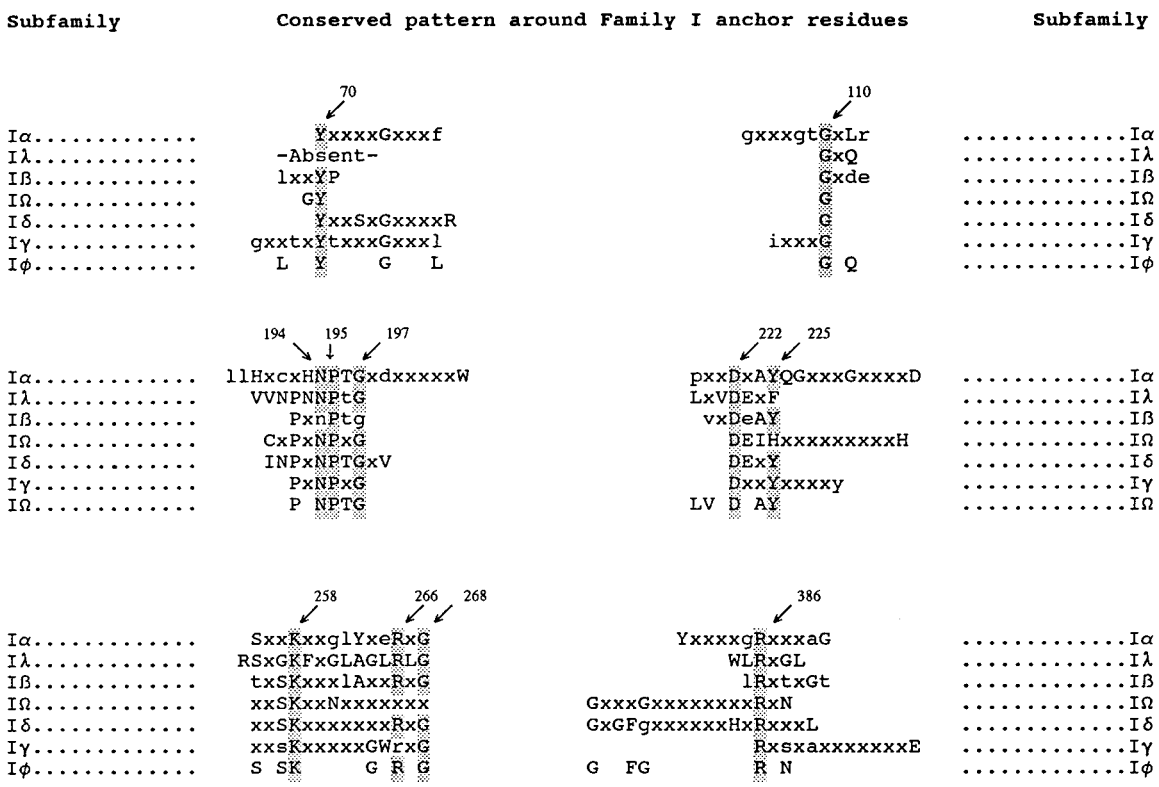


FIG. 5. Signature amino acid motifs that discriminate subfamilies within aminotransferase Family I. Since subfamily Iφ currently has only one member, only the anchor residue and any flanking residue that matches a residue conserved in another subfamily are shown. For all subfamilies, x denotes any amino acid residue, residues specified in lowercase lettering indicate near-total conservation, and residues specified in capital lettering indicate total conservation. In subfamily Iα, residue 192 is A in all known plant and animal sequences and C in the yeast and microbial sequences.

aspartate aminotransferases. Although *E. coli* AspC functions primarily as aspartate aminotransferase in vivo (15), it exhibits broad specificity and accepts aromatic substrates. In *E. coli* AspC, the bulky side chains prevent R-327–Q-328 from closely approaching the main chain atoms of the L-323–T-324 pair. *E. coli* AspC has been shown to have an active-site pocket displaying a hydrophobic character in contrast to that of vertebrate aspartate aminotransferases. The hydrophobic pocket of *E. coli* AspC consists of Y-70, Y-263 (large domain) and I-37, G-38, and V-39 (small domain). There is a strong contact via van der Waals interaction between residues Y-263 and V-39. In contrast, there is only a limited interaction between residues Y-263 and A-39 of vertebrate aspartate aminotransferases. *E. coli* TyrB, which accepts aromatic substrates much better, possesses a L-39 residue which should interact strongly with Y-263. Since the small domain moves relatively little upon binding substrate in the broad-specificity members of subfamily Iα, the small-domain residues forming the hydrophobic pocket must be near the K-258-linked PLP. These considerations are consistent with the four top members of the subfamily Iα cluster in Fig. 2 being highly specific for aspartate, and with the remaining members of subfamily Iα having broad specificity.

The differing substrate specificities of *E. coli* AspC and *E. coli* TyrB cannot be attributed to variation of a single or even several residues. Köhler et al. (33) showed that conversion of residues 39, 141, and 293 of *E. coli* AspC to those present in *E. coli* TyrB had little effect upon substrate specificity, suggesting that multiple side chain differences and minor background rearrangements must be involved. However, there has been a recent significant achievement showing that the substrate specificity of *E. coli* AspC can be broadened to that of *E. coli* TyrB by site-directed mutagenesis (45). This specificity alteration required six residue changes in *E. coli* AspC: V39L (V at position 39 changed to L), K41Y, T47I, N69L, T109S, and N297S. X-ray structure determination of the multiple mutant (36) supports the model proposed by Seville et al. (56) whereby the aromatic side chain of the substrate is able to displace the side chain of R-292. A new nonpolar binding site was shown to have been created by introducing the six mutations which enlarged the active-site pocket. However, it is surprising that if one compares these six residues of *Pseudomonas aeruginosa* PhhC (which utilizes aromatic amino acids like *E. coli* TyrB) with those of *E. coli* AspC and *E. coli* TyrB, they resemble *E. coli* AspC in five of six cases. Hence, within the framework of subfamily Iα structure, it must be that multiple combinations of residue substitutions allow broadened substrate specificity.

Even within the intensively studied subfamily Iα, the exact functions of some highly conserved residues are not clear. Gloss et al. (16) have focused upon C-191 which contacts the invariant residue Y-225. An in vivo conversion of C-191 to A-191 had been shown not to affect the kinetic properties of *E. coli* AspC significantly. However, in vivo conversion of C-191 to A-191 would need to go through an S-191 intermediate, and the 191 position was shown to yield an enzyme with unfavorable kinetic parameters when occupied by serine. Thus, it was proposed that since position 191 occupies a highly conserved three-dimensional environment, there is no neutral mutation pathway to a functionally equivalent A-191. Figure 4 shows that the single exception is the *Rhizobium meliloti* AroX en-

zyme which possesses S-191. Assuming that this is not a sequencing error, Gloss (15a) suggests the interesting possibility that a physical effect caused by the C191S change may be uniquely beneficial to *Rhizobium meliloti*. For example, one physical effect is an increase in the pH optimum, and perhaps *R. meliloti* AroX needs a higher internal aldimine pK_a. Another physical effect of C191S for *E. coli* AspC is to destabilize E-PLP relative to E-PMP, thus favoring keto acid production. Perhaps S-191 in *R. meliloti* AroX is a factor that promotes function in the catabolic direction.

CONCLUDING PERSPECTIVE

During divergent evolution of structural genes, enzyme homologs have radiated from their common ancestor as a complex function of elapsed time, differential functional specialization, and the constraints imposed by the particular enzymatic mechanism. The aminotransferase mechanism is sufficiently complex that it has perhaps only evolved once, giving rise to the single aminotransferase superfamily thus far known. This protein superfamily exemplifies a very large grouping with many different substrate-specialized members. The ready amplification of given proteins via cloning methodology has dramatically accelerated the ability to obtain crystal structures of even low-abundance enzymes. The prospects for the complete elucidation of spatial structures of aminotransferases at the atomic level in concert with variable capabilities of substrate recognition throughout a series of homologs offer a rich source of insight into how the evolutionary process can generate catalytic diversity.

Within family I, a histidine biosynthesis step has emerged as a capability which is uniquely confined to subfamily Iβ. The divergence allowed by speciation in the various orthologous subfamily members will contribute to an understanding of what features are critical for maintaining the ability to utilize imidazoleacetol phosphate. On the other hand, aminotransferases which are highly specific for aspartate have emerged in both subfamily Iα and subfamily Iγ. An excellent understanding of the molecular basis for this specificity is known in subfamily Iα because of X-ray crystallography data. While such data are not yet available for subfamily Iγ, it is clear from the multiple alignments of amino acid sequences that different atomic interactions are responsible for substrate specificity in the two subfamilies (albeit within a similar overall protein scaffold that reflects the close homology relationships).

Aromatic amino acids can be utilized in representatives of at least three subfamilies. A comparison of the atomic interactions which allow the utilization of aromatic amino acids in each of the three subfamilies will be possible when X-ray crystal structures of appropriate subfamily Iβ and Iγ members become available. Interesting functional redundancies may even extend across family boundaries, as exemplified by the ability of IlvE in aminotransferase family III to synthesize phenylalanine in vivo (15, 28).

This minireview has focused upon aminotransferase family I, where the most definitive work (in subfamily Iα) has been done, and emphasizes the recognition of more-refined and emerging subfamily divisions. When it is appreciated that this is merely one of many families within a superfamily, which in turn has membership in a massive division of PLP-dependent proteins, one can hardly fail to be overwhelmingly optimistic about the evolutionary insights and the feats of protein engineering that must be at hand.

ACKNOWLEDGMENTS

We appreciate the generosity of Lisa Gloss of Penn State University who not only provided Fig. 3 but also suggested some improvements to an early version of the manuscript.

REFERENCES

1. **Alexander, F. W., E. Sandmeier, P. K. Mehta, and P. Christen.** 1994. Evolutionary relationships among pyridoxal-5′-phosphate-dependent enzymes. Eur. J. Biochem. **219:**953–960.
2. **Alfano, J. R., and M. L. Kahn.** 1993. Isolation and characterization of a gene coding for a novel aspartate aminotransferase from *Rhizobium meliloti*. J. Bacteriol. **175:**4186–4196.
3. **Birolo, L., M. I. Arnone, M. V. Cubellis, G. Andreotti, G. Nitti, G. Marino, and G. Sannia.** 1991. The active site of *Sulfolobus solfataricus* aspartate aminotransferase. Biochim. Biophys. Acta **1080:**198–204.

3a. **Bontempi, E. J., J. Búa, L. Åslund, B. Porcel, E. L. Segura, J. Henriksson, A. Örn, U. Pettersson, and A. M. Ruiz.** 1993. Isolation and characterization of a gene from *Trypanosoma cruzi* encoding a 46-kilodalton protein with homology to human and rat tyrosine aminotransferase. Mol. Biochem. Parasitol. **59:**253–262.

4. **Chang, Y.-Y.** 1992. Common ancestry of *Escherichia coli* pyruvate oxidase and the acetohydroxy acid synthase of the branched-chain amino acid biosynthetic pathway, p. 81–104. *In* R. P. Mortlock (ed.), The evolution of metabolic function. CRC Press, Inc., Boca Raton, Fla.
5. **Christen, P., and D. E. Metzler (ed.).** 1985. Transaminases. John Wiley & Sons, New York.
6. **Chu, L., A. Burgum, D. Kolodrubetz, and S. C. Holt.** 1995. The 46-kilodalton hemolysin gene from *Treponema denticola* encodes a novel hemolysin homologous to aminotransferases. Infect. Immun. **63:**4448–4455.
7. **Conover, R. K., and W. F. Doolittle.** 1990. Characterization of a gene involved in histidine biosynthesis in *Halobacterium* (*Haloferax*) *volcanii*: isolation and rapid mapping by transformation in an auxotroph with cosmid DNA. J. Bacteriol. **172:**3244–3249.
8. **Crouzet, J., L. Cauchois, F. Blanche, L. Debussche, D. Thibaut, M.-C. Rouyez, S. Rigault, J.-F. Mayaux, and B. Cameron.** 1990. Nucleotide sequence of a *Pseudomonas denitrificans* 5.4-kilobase DNA fragment containing five *cob* genes and identification of structural genes encoding *S*-adenosyl-L-methionine:uroporphyrinogen III methyltransferase and cobyrinic acid *a*,*c*-diamide synthase. J. Bacteriol. **172:**5968–5979.
9. **Cubellis, M. V., C. Rozzo, G. Nitti, M. I. Arnone, G. Marino, and G. Sannia.** 1989. Cloning and sequencing of the gene coding for aspartate aminotransferase from the thermoacidophilic archaebacterium *Sulfolobus solfataricus*. Eur. J. Biochem. **186:**375–381.
10. **Delorme, C., S. D. Ehrlich, and P. Renault.** 1992. Histidine biosynthesis genes in *Lactobacillus lactis* subsp. *lactis*. J. Bacteriol. **174:**6571–6579.
11. **Devereux, J., P. Haeberli, and P. Marquess.** 1987. The program manual for the sequence analysis software package. Genetics Computer Group, Madison, Wis.
12. **Fleischmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J.-F. Tomb, B. A. Dougherty, J. M. Merrick, K. McKenney, G. Sutton, W. FitzHugh, C. Fields, J. D. Gocayne, J. Scott, R. Shirley, L.-I. Liu, A. Glodek, J. M. Kelly, J. F. Weidman, C. A. Phillips, T. Spriggs, E. Hedblom, M. D. Cotton, T. R. Utterback, M. C. Hanna, D. T. Nguyen, D. M. Saudek, R. C. Brandon, L. D. Fine, J. L. Fritchman, J. L. Fuhrmann, N. S. M. Geoghagen, C. L. Gnehm, L. A. McDonald, K. V. Small, C. M. Fraser, H. O. Smith, and J. C. Venter.** 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science **269:**496–512.
13. **Fotheringham, I. G., S. A. Dacey, P. P. Taylor, T. J. Smith, M. G. Hunter, M. E. Finlay, S. B. Primrose, D. M. Parker, and R. M. Edwards.** 1986. The cloning and sequence analysis of the *aspC* and *tyrB* genes from *Escherichia coli* K12. Biochem. J. **234:**593–604.
14. **Gantt, J., R. J. Larson, M. W. Farnham, S. M. Pathirana, S. M. Miller, and C. P. Vance.** 1992. Aspartate aminotransferase in effective and ineffective alfalfa nodules. Plant Physiol. **98:**868–878.
15. **Gelfand, D. H., and R. A. Steinberg.** 1977. *Escherichia coli* mutants deficient in the aspartate and aromatic aminotransferases. J. Bacteriol. **130:**429–440.

15a. **Gloss, L. M.** Personal communication.

16. **Gloss, L. M., D. E. Spencer, and J. F. Kirsch.**. Cysteine-191 in aspartate aminotransferases appears to be conserved due to a neutral mutation pathway to the functional equivalent, alanine-191. Proteins, in press.

16a. **Gomez, P., and L. O. Ingram.** Unpublished data.

17. **Grange, T., C. Guenet, J. B. Dietrich, S. Chasserot, M. Fromont, N. Befort, J. Jami, G. Beck, and R. Pictet.** 1985. Complete complementary DNA of rat tyrosine aminotransferase messenger RNA: deduction of the primary structure of the enzyme. J. Mol. Biol. **184:**347–350.
18. **Grisolia, V., M. S. Carlomagno, A. G. Nappo, and C. B. Bruni.** 1985. Cloning, structure, and expression of the *Escherichia coli* K-12 *hisC* gene. J. Bacteriol. **164:**1317–1323.
19. **Gu, W., G. Zhao, C. Eddy, and R. A. Jensen.** 1995. Imidazole acetol phosphate aminotransferase in *Zymomonas mobilis*: molecular genetic, biochem-

ical, and evolutionary analysis. J. Bacteriol. **177:**1576–1584.
20. **Hayashi, H., Y. Horio, T. Tanaka, M. Taketoshi, and H. Wada.** 1987. Rat cytosolic aspartate aminotransferase: molecular cloning of cDNA and expression in *E. coli*, p. 39–42. *In* T. Korpela and P. Christen (ed.), Biochemistry of vitamin B_6. Birkhäuser, Basel.
21. **Hayashi, H., K. Inoue, T. Nagata, S. Kuramitsu, and H. Kagamiyama.** 1993. *Escherichia coli* aromatic amino acid aminotransferase: characterization and comparison with aspartate aminotransferase. Biochemistry **32:**12229–12239.
22. **Henner, D. J., L. Band, G. Flaggs, and E. Chen.** 1986. The organization and nucleotide sequence of the *Bacillus subtilis hisH*, *tyrA* and *aroE* genes. Gene **49:**147–152.
23. **Hinshelwood, S., and N. G. Stoker.** 1992. Cloning of mycobacterial histidine synthesis genes by complementation of a *Mycobacterium smegmatis* auxotroph. Mol. Microbiol. **6:**2887–2895.
24. **Huynh, Q. K., R. Sakakibara, T. Watanabe, and H. Wada.** 1981. The complete amino acid sequence of mitochondrial glutamic oxaloacetic transaminase from rat liver. J. Biochem. (Tokyo) **90:**863–875.
25. **Ishiguro, M., M. Suzuki, K. Takio, T. Matsuzawa, and K. Titani.** 1991. Complete amino acid sequence of rat liver cytosolic alanine aminotransferase. Biochemistry **30:**6048–6053.
26. **Ishiguro, M., K. Takio, M. Suzuki, R. Oyama, T. Matsuzawa, and K. Titani.** 1991. Complete amino acid sequence of human liver cytosolic alanine aminotransferase (GPT) determined by a combination of conventional and mass spectral methods. Biochemistry **30:**10451–10457.
27. **Jensen, R. A.** 1976. Enzyme recruitment in evolution of new function. Annu. Rev. Microbiol. **30:**409–425.
28. **Jensen, R. A., and D. H. Calhoun.** 1981. Intracellular roles of microbial aminotransferases: overlap enzymes across different biochemical pathways. Crit. Rev. Microbiol. **8:**229–266.
29. **Kamitori, S., A. Okamoto, K. Hirotsu, T. Higuchi, S. Kuramitsu, H. Kagamiyama, Y. Matsura, and Y. Katsube.** 1990. Three-dimensional structures of aspartate aminotransferase from *Escherichia coli* and its mutant enzyme at 2.5 Å resolution. J. Biochem. (Tokyo) **108:**175–184.
30. **Kaneko, T., A. Tanaka, S. Sato, H. Kotani, T. Sazuka, N. Miyajima, M. Sugiura, and S. Tabata.** 1996. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC 6803. DNA Res. **2:**153–166.
31. **Katayama, M., Y. Sakai, S. Okamoto, F. Ihara, T. Nihira, and Y. Yamada.** Gene organization in the *ada-rplL* region of the *Streptomyces virginiae* chromosome. Gene, in press.
32. **Kirsch, J. F., G. Eichele, G. C. Ford, M. G. Vincent, and J. N. Jansonius.** 1984. Mechanism of action of aspartate aminotransferase proposed on the basis of its spatial structure. J. Mol. Biol. **174:**497–525.
33. **Köhler, E., M. Seville, J. Jäger, I. Fotheringham, M. Hunter, M. Edwards, J. N. Jansonius, and K. Kirschner.** 1994. Significant improvement to the catalytic properties of aspartate aminotransferase: role of hydrophobic and charged residues in the substrate binding pocket. Biochemistry **33:**90–97.
34. **Kuramitsu, S., K. Inoue, T. Ogawa, H. Ogawa, and H. Kagamiyama.** 1985. Aromatic amino acid aminotransferase of *Escherichia coli*: nucleotide sequence of the *tyrB* gene. Biochem. Biophys. Res. Commun. **133:**134–139.
35. **Limauro, D., A. Avitabile, C. Cappellano, A. M. Puglia, and C. B. Bruni.** 1990. Cloning and characterization of the histidine biosynthetic gene cluster of *Streptomyces coelicolor* A3(2). Gene **90:**31–41.
36. **Malashkevvich, V. N., J. J. Onuffer, J. F. Kirsch, and J. N. Jansonius.** 1995. Alternating, arginine-modulated substrate specificity in an engineered aminotransferase. Nature Struct. Biol. **2:**548–553.
37. **Mehta, P. K., and P. Christen.** 1993. Homology of pyridoxal-5′-phosphate-dependent aminotransferases with the *cobC* (cobalamin synthesis), *nifS* (nitrogen fixation), *pabC* (*p*-aminobenzoate synthesis) and *malY* (abolishing endogenous induction of the maltose system) gene products. Eur. J. Biochem. **211:**373–376.
38. **Mehta, P. K., T. I. Hale, and P. Christen.** 1989. Evolutionary relationships among aminotransferases. Eur. J. Biochem. **186:**249–253.
39. **Mehta, P. K., T. I. Hale, and P. Christen.** 1993. Aminotransferases: demonstration of homology and division into evolutionary subgroups. Eur. J. Biochem. **214:**549–561.
40. **Montemartini, M., J. A. Santomé, J. J. Cazzulo, and C. Nowicki.** 1993. Purification and partial structural and kinetic characterization of tyrosine aminotransferase from epimastigotes of *Trypanosoma cruzi*. Biochem. J. **292:**901–906.
41. **Morin, P. J., G. S. Subramanian, and T. D. Gilmore.** 1992. AAT1, a gene encoding a mitochondrial aspartate aminotransferase in *Saccharomyces cerevisiae*. Biochim. Biophys. Acta **1171:**211–214.
42. **Mosca, M., C. Liviana, J. Breton, C. Speciale, E. Okuno, R. Schwarcz, and L. Benatti.** 1994. Molecular cloning of rat kynurenine aminotransferase: identity with glutamine transaminase K. FEBS Lett. **353:**21–24.
43. **Nester, E. W., and A. L. Montoya.** 1976. An enzyme common to histidine and aromatic amino acid biosynthesis in *Bacillus subtilis*. J. Bacteriol. **126:**699–705.
44. **Nishiwaki, N., N. Hayashi, S. Irie, D.-H. Chung, S. Harashima, and Y. Oshima.** 1987. Structure of the yeast *his5* gene responsive to general control of amino acid biosynthesis. Mol. Gen. Genet. **208:**159–167.
45. **Onuffer, J., and J. F. Kirsch.** 1995. Redesign of the substrate specificity of *Escherichia coli* aspartate aminotransferase to that of *Escherichia coli* tyrosine aminotransferase by homology modeling and site-directed mutagenesis. Protein Sci. **4:**1750–1757.
46. **Ovchinnikov, Y. A., C. A. Egorov, N. A. Aldanova, M. Y. Feigina, V. M. Lipkin, N. G. Abdulaev, E. V. Grishin, A. P. Kiselev, N. N. Modyanov, A. E. Braunstein, O. L. Polyanovsky, and V. V. Nosikov.** 1973. The complete amino acid sequence of cytoplasmic aspartate aminotransferase from pig heart. FEBS Lett. **29:**31–34.
47. **Parsot, C.** 1987. A common origin for enzymes involved in the terminal steps of the threonine and tryptophan biosynthetic pathways. Proc. Natl. Acad. Sci. USA **84:**5207–5210.
48. **Perego, M., S. P. Cole, D. Burbulys, K. Trach, and J. A. Hoch.** 1989. Characterization of the gene for a protein kinase which phosphorylates the sporulation-regulatory proteins SpoOA and SpoOF of *Bacillus subtilis*. J. Bacteriol. **171:**6187–6196.
49. **Perry, S., H. Harries, C. Scholfield, T. Lock, L. King, G. Gibson, and P. Goldfarb.** 1995. Molecular cloning and expression of a cDNA for human kidney cysteine conjugate β-lyase. FEBS Lett. **360:**277–280.
50. **Perry, S. J., M. A. Schofield, M. MacFarlane, E. A. Lock, L. J. King, G. G. Gibson, and P. S. Goldfarb.** 1993. Isolation and expression of a cDNA coding for rat kidney cytosolic cysteine conjugate β-lyase. Mol. Pharmacol. **43:**660–665.
51. **Pollich, M., S. Jock, and G. Klug.** 1993. Identification of a gene required for the oxygen-regulated formation of the photosynthetic apparatus of *Rhodobacter capsulatus*. Mol. Microbiol. **10:**749–757.
52. **Reidl, J., and W. Boos.** 1991. The *malX malY* operon of *Escherichia coli* encodes a novel enzyme II of the phosphotransferase system recognizing glucose and maltose and an enzyme abolishing the endogenous induction of the maltose system. J. Bacteriol. **173:**4862–4876.
53. **Rossol, I., and A. Pühler.** 1992. The *Corynebacterium glutamicum aecD* gene encodes a C-S lyase with α,β-elimination activity that degrades aminoethylcysteine. J. Bacteriol. **174:**2968–2977.
54. **Sandmeier, E., T. I. Hale, and P. Christen.** 1994. Multiple evolutionary origin of pyridoxal-5′-phosphate-dependent amino acid decarboxylases. Eur. J. Biochem. **221:**997–1002.
55. **Schultz, C. J., and G. M. Coruzzi.** 1995. The aspartate aminotransferase gene family of *Arabidopsis* encodes isoenzymes localized to three distinct subcellular compartments. Plant J. **7:**61–75.
56. **Seville, M., M. G. Vincent, and K. Hahn.** 1988. Modeling the three-dimensional structures of bacterial aminotransferases. Biochemistry **27:**8344–8349.
57. **Smith, D. L., S. C. Almo, M. D. Toney, and D. Ringe.** 1989. 2.8-Å Resolution crystal structure of an active-site mutant of aspartate aminotransferase from *Escherichia coli*. Biochemistry **28:**8161–8167.
58. **Sofia, H. J., V. Burland, D. L. Daniels, G. Plunkett III, and F. R. Blattner.** 1994. Analysis of the *Escherichia coli* genome. V. DNA sequence of the region from 76.0 to 81.5 minutes. Nucleic Acids Res. **22:**2576–2586.
59. **Son, D., and T. Sugiyama.** 1992. Molecular cloning of an alanine aminotransferase from NAD-malic enzyme type C4 plant *Panicum miliaceum*. Plant Mol. Biol. **20:**705–713.
60. **Sung, M.-H., K. Tanizawa, H. Tanaka, S. Kuramitsu, H. Kagamiyama, K. Hirotsu, A. Okamoto, T. Higuchi, and K. Soda.** 1991. Thermostable aspartate aminotransferase from a thermophilic *Bacillus* species. J. Biol. Chem. **266:**2567–2572.
61. **Sung, M.-H., K. Tanizawa, H. Tanaka, S. Kuramitsu, H. Kagamiyama, and K. Soda.** 1990. Purification and characterization of thermostable aspartate aminotransferase from a thermophilic *Bacillus* species. J. Bacteriol. **172:**1345–1351.
62. **Takagi, T., T. Taniguchi, Y. Yamamoto, and T. Shibatani.** 1991. Molecular cloning of the L-phenylalanine transaminase gene from *Paracoccus denitrificans* in *Escherichia coli* K-12. Biotechnol. Appl. Biochem. **13:**112–119.
63. **Takemura, H., S. Horinouchi, and T. Beppu.** 1993. Suppression of an ethanol-sensitive mutation of *Acetobacter pasteurianus* by overexpression of the *his1* gene encoding histidinol phosphate aminotransferase. J. Ferment. Bioeng. **76:**224–228.
64. **Trach, K. A., and J. A. Hoch.** 1993. Multisensory activation of the phosphorelay initiating sporulation in *Bacillus subtilis*: identification and sequence of the protein kinase of the alternate pathway. Mol. Microbiol. **8:**69–79.
65. **Udvardi, M. K., and M. L. Kahn.** 1991. Isolation and analysis of a cDNA clone that encodes an alfalfa (*Medicago sativa*) aspartate aminotransferase. Mol. Gen. Genet. **231:**97–105.
66. **Watson, R. J., and V. K. Rastogi.** 1993. Cloning and nucleotide sequencing of *Rhizobium meliloti* aminotransferase genes: an aspartate aminotransferase required for symbiotic nitrogen fixation is atypical. J. Bacteriol. **175:**1919–1928.
67. **Yčas, M.** 1974. On earlier states of the biochemical systems. J. Theor. Biol. **44:**145–146.
68. **Zhao, G., T. Xia, J. Song, and R. A. Jensen.** 1994. *Pseudomonas aeruginosa* possesses homologues of mammalian phenylalanine hydroxylase and 4-a-carbinolamine dehydratase/DCoH as part of a three-component gene cluster. Proc. Natl. Acad. Sci. USA **91:**1366–1370.