Amino Acids (2005) 28: 1–12
DOI 10.1007/s00726-004-0152-y

Amino Acids
Printed in Austria

Allosteric mechanisms in ACT domain containing enzymes involved in amino acid metabolism

Review Article

J. S. Liberles, M. Thórólfsson*, and **A. Martínez**

Department of Biomedicine, University of Bergen, Bergen, Norway

Received November 23, 2004
Accepted November 24, 2004
Published online January 18, 2005;

Summary. An important sequence motif identified by sequence analysis is shared by the ACT domain family, which has been found in a number of diverse proteins. Most of the proteins containing the ACT domain seem to be involved in amino acid and purine synthesis and are in many cases allosteric enzymes with complex regulation enforced by the binding of ligands. Here we explore the current understanding of the ACT domain function including its role as an allosteric module in a selected group of enzymes. We will further describe in more detail three of the proteins where some understanding is available on function and structure: i) the archetypical ACT domain protein *E. coli* 3PGDH, which catalyzes the first step in the biosynthesis of L-Ser, ii) the bifunctional chorismate mutase/prephenate dehydratase (P-protein) from *E. coli*, which catalyzes the first two steps in the biosynthesis of L-Phe, and iii) the mammalian aromatic amino acid hydroxylases, with special emphasis on phenylalanine hydroxylase, which catalyzes the first step in the catabolic degradation of L-phenylalanine (L-Phe). The ACT domain is commonly involved in the binding of a small regulatory molecule, such as the amino acids L-Ser and L-Phe in the case of 3PGDH and P-protein, respectively. On the other hand, for PAH, and probably for other enzymes, this domain appears to have been incorporated as a handy, flexible small module with the potential to provide allosteric regulation *via* transmission of finely tuned conformational changes, not necessarily initiated by regulatory ligand binding at the domain itself.

Keywords: ACT domain – Allosterism – Amino acid biosynthesis – Regulatory domain

Abbreviations: AAAH, aromatic amino acid hydroxylases; ACT, domain which names derives from aspartokinase, chorismate mutase, and TyrA (prephenate dehydrogenase); CM, chorismate mutase; PAH, phenylalanine hydroxylase; PDT, prephenate dehydratase; 3PGDH, 3-phosphoglycerate dehydrogenase; TH, tyrosine hydroxylase; TPH, tryptophan hydroxylase

* Current address: Protein Biophysics & Delivery, Novo Nordisk A/S, Novo Allé, 2880 Bagsværd, Denmark.

1 Introduction

The ACT domain is a sequence motif that was first identified as a regulatory module found in a number of diverse proteins (Aravind and Koonin, 1999). The name originates from three of the proteins in the domain family: aspartokinase, chorismate mutase, and TyrA (prephenate dehydrogenase) (Aravind and Koonin, 1999). Most of the proteins containing the ACT domain seem to be involved in amino acid and purine synthesis and are in many cases allosteric enzymes with complex regulation enforced by the binding of ligands (Aravind and Koonin, 1999; Chipman and Shaanan, 2001). By using the crystal structure of *E. coli* 3-phosphoglycerate dehydrogenase (3PGDH) as a framework, the ACT domain was suggested to be an evolutionarily mobile ligand binding regulatory module that has been fused to different enzymes at various times (Aravind and Koonin, 1999; Anantharaman et al., 2001). Thus, the ACT domain might be a provider of allosteric regulation in these otherwise very different enzymes (Aravind and Koonin, 1999). Allosteric cooperativity[1] is a very important and extended regulatory mechanism operating in metabolic and signal pathways. Direct control

[1] Allosteric effects generated from identical molecules, e.g. an enzyme substrate altering the consecutive binding of other substrates, are termed homotropic *vs* heterotropic, which are effects generated by effectors altering the binding of different molecules (Monod et al., 1965). The result of both allosteric effects is usually a cooperative response, which may be positive as well as negative, resulting in an increase or decrease, respectively, of affinity and/or activity.

of enzyme function via allosteric regulation is usually achieved through ligand-induced conformational changes of a given protein structure. There seems to be no common structural explanation for the cooperative response and similar proteins seem to have adapted individual mechanisms (Perutz, 1990), including the allosteric enzymes containing the ACT domain. In fact all dynamic proteins seem to have a potential for allosterism

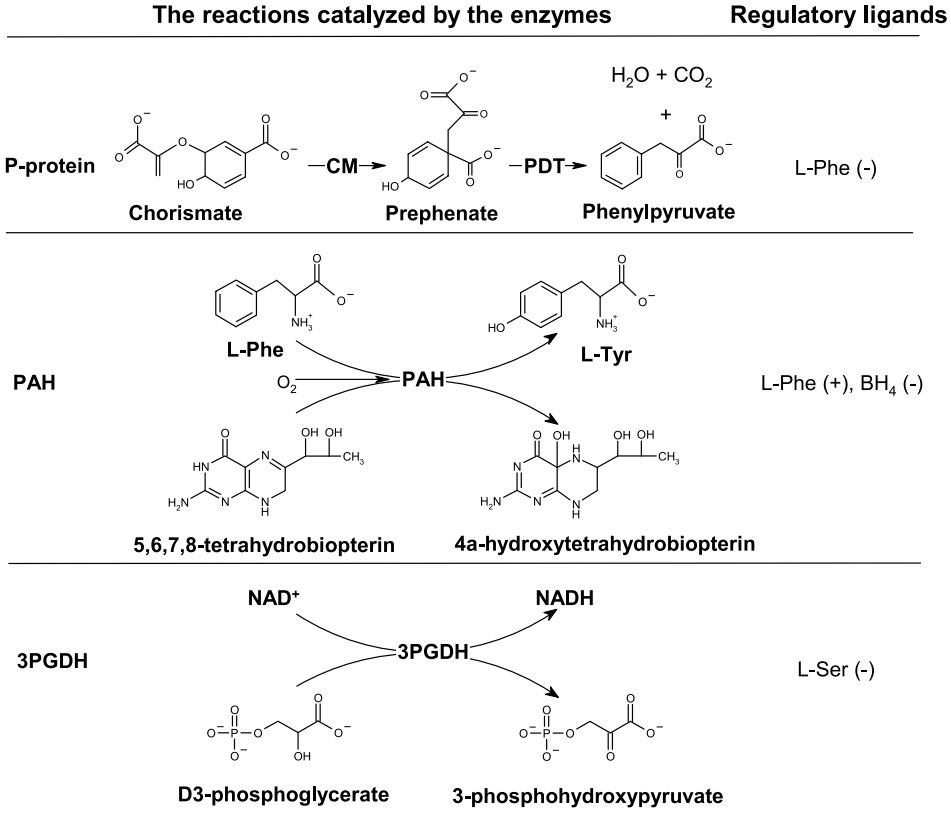


Fig. 1b.

Fig. 1. Domain organization and reactions of P-protein, PAH and 3PGDH. **a** Schematic representation of the domain structure, oligomeric organization, and sequence alignment of the ACT domains. *CM*, chorismate mutase; *PDT*, prephenate dehydratase; *RD*, regulatory domain; *A*, autoregulatory sequence; *CD*, catalytic domain; *OD*, oligomerization domain; *SD*, substrate binding domain; *ND*, nucleotide binding domain. The sequence alignment was constructed using a combination of T-Coffee (Notredame et al., 2000) for sequence alignment and combinatorial extension (*CE*) (Shindyalov and Bourne, 1998) for structural alignment. The secondary structure elements are shown in italic (*a*: alpha helix, *b*: beta sheet) and were obtained from the crystal structures of PAH (PDB code 1PHZ), marked above, and 3PGDH (PDB code 1PSD), marked below the alignment. Strictly conserved residues are marked with yellow, pair-wise conserved are marked with green for a conservation between the P-protein and PAH, blue for PAH and 3PGDH, and grey for the P-protein and 3PGDH. **b** Enzyme reactions and respective regulatory amino acid ligands

Fig. 2. Ribbon model of monomeric (**a**) and tetrameric (**b**) 3PDGH based on the crystal structure of the enzyme from *E. coli* (PDB code 1PSD). **a** Color code as in Fig. 1a. *ND*, nucleotide binding domain; *SD*, substrate binding domain; *RD*, regulatory domain. Bound NAD and serine are in stick representation. **b** The arrows point at the C-terminal regulatory (ACT) serine binding domains. Bound serine is in stick representation

Fig. 3. Ribbon model of monomeric (**a**) and tetrameric (**b**) PAH. Composite model using the structure of rat PAH (PDB code 2PHM) and human PAH (PDB code 2PAH). **a** Color code as in Fig. 1a. *A*, autoregulatory sequence; *RD*, regulatory domain; *CD*, catalytic domain; *OD*, oligomerization domain. The catalytic iron is shown in yellow. **b** The arrows point at the N-terminal regulatory (ACT) domains

Fig. 4. Schematic illustration of the model for the activation of PAH by L-Phe. The modeled full-length dimer is shown, with one subunit in wire frame (gray) and the other in schematic representation (color code for domains as in Figs. 1a and 3a). The loop Arg68-Asp75 and the hinge region Arg111-Thr117 are shown in green and important residues are shown as ball and stick representations. Steps (*i*), (*ii*) and (*iii*) leading to the activation of PAH are described in the text

Fig. 1a

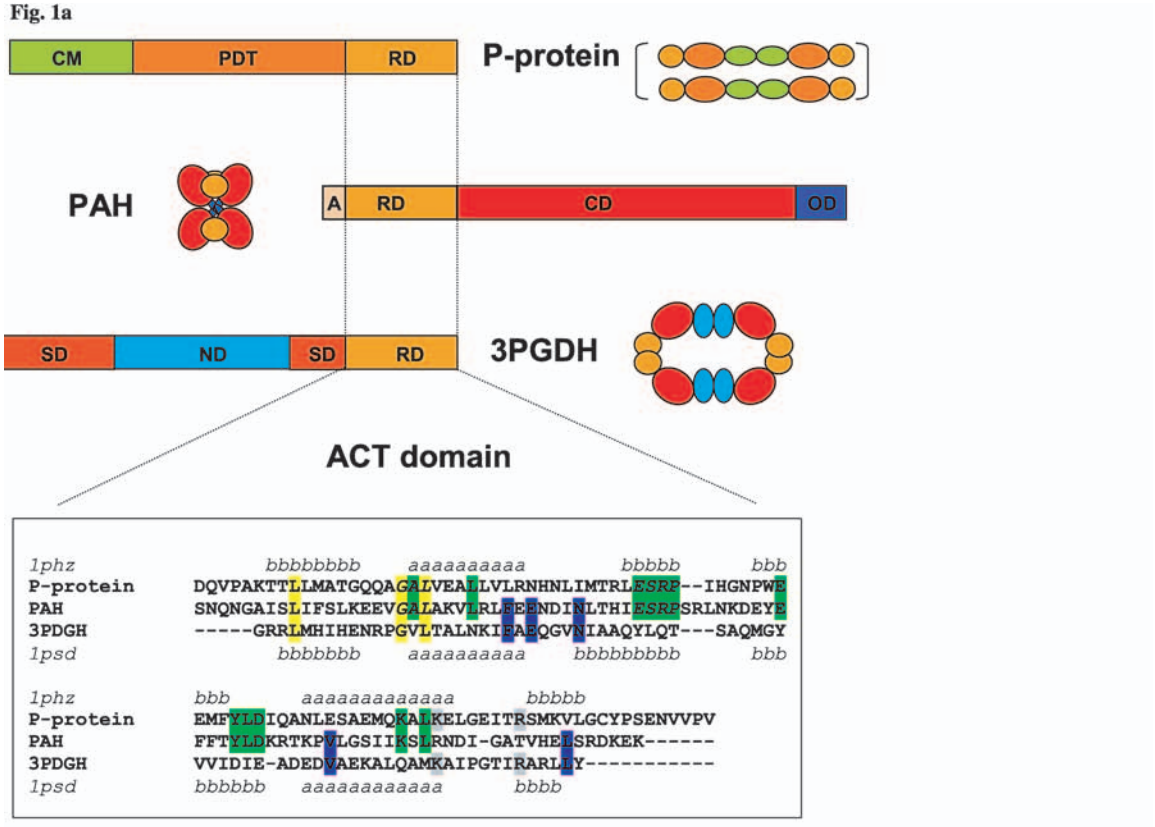


Fig. 2a

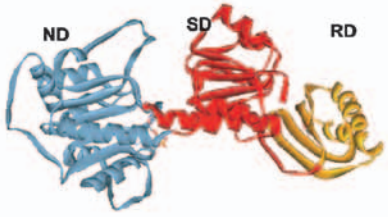


Fig. 2b

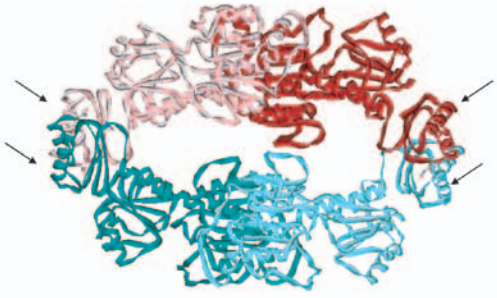


Fig. 3a

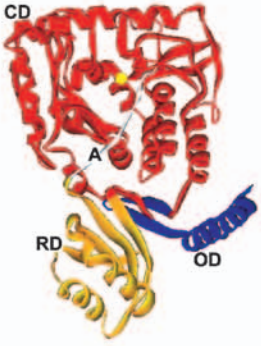


Fig. 3b

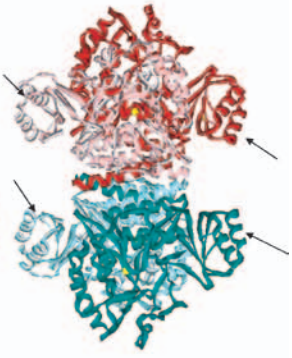
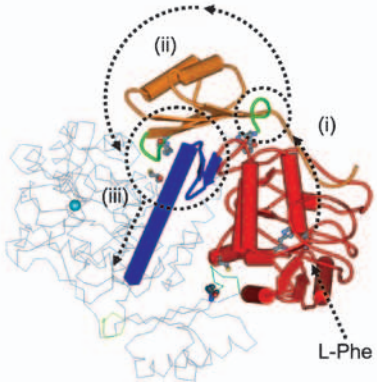


Fig. 4



(Gunasekaran et al., 2004). Here we review the current understanding on the regulatory function of the ACT domain and we analyze its role as an allosteric module in a selected group of enzymes. Most of the proteins that contain an ACT domain have not been studied in detail and structural information remains scarce. We have chosen to consider in more detail three of the proteins where some understanding is available on function and structure: i) the archetypical ACT domain protein 3PGDH, which catalyzes the first step in the biosynthesis of L-Ser (Grant et al., 1996), ii) the bifunctional chorismate mutase/prephenate dehydratase (P-protein) from *E. coli*, which catalyzes the first two steps in the biosynthesis of L-Phe (Zhang et al., 1998), and iii) the aromatic amino acid hydroxylases (AAAH), with special emphasis on phenylalanine hydroxylase (PAH), which catalyzes the first step in the catabolic degradation of L-phenylalanine (L-Phe) (Flatmark and Stevens, 1999; Fitzpatrick, 2000). Figure 1a shows a sequence alignment of the ACT domain of 3PGDH and P-protein from *E. coli* with PAH from rat. The catalytic and regulatory properties of these enzymes are discussed in a structural perspective, seeking to get some insights on the putative role of the ACT domain as a truly regulatory module involved in allosteric regulation at the structural level, either by specific ligand binding and/or by transmission of conformational changes. We have also explored a potential evolutionary relationship between P-protein and PAH.

2 D-3-phosphoglycerate dehydrogenase (3PGDH) from *E. coli*

3PGDH (EC 1.1.1.95) is an enzyme that belongs to the D-isomer specific 2-hydroxyacid dehydrogenase family and catalyzes the oxidation of D-3-phosphoglycerate to 3-phosphohydroxypyruvate, which is the first step in the biosynthesis of L-Ser, using NAD^+ as the oxidizing agent (Grant et al., 1996; Achouri et al., 1997) (Fig. 1b). In bacteria, 3PGDH is feedback controlled by the end product L-Ser in an allosteric manner (Grant et al., 1996) and the C-terminal regulatory domains of the enzymes from lower plants, yeast and bacteria belong to the ACT domain family (see Pfam database, http://www.sanger.ac.uk/Software/Pfam/). On the other hand, mammalian 3PGDHs do not appear to be regulated by L-Ser and their C-terminal domains do not correspond to an ACT domain (Pind et al., 2002). It is common for small molecule binding domains, such as the ACT domain, to be linage specific, and analogs, not homologs, commonly perform the same function (Anantharaman et al., 2001). Moreover, there is very little information available on the enzyme from eukaryotic origin and the catalytic and the regulatory properties remain unknown. However, the enzyme from *E. coli* has been thoroughly examined in the last decade and Grant and coworkers have presented a structural explanation for the function and mechanism of the enzyme (Schuller et al., 1995; Al-Rabiee et al., 1996; Grant et al., 1996; Grant et al., 1999). Hereafter, we will focus on the properties of the *E. coli* enzyme.

2.1 Structural organization

E. coli 3PGDH is organized into three functional domains: i) an N-terminal nucleotide binding domain (residues 108–294), ii) a middle substrate binding domain (residues 7–107, 295–336), and iii) a C-terminal regulatory domain which corresponds to the ACT domain (residues 337–410), and includes the L-Ser binding site (Figs. 1a and 2).

As seen in the crystal structure of the L-Ser inhibited *E. coli* 3PDGH (Fig. 2) the enzyme is arranged into a tetrameric toroidal structure (Schuller et al., 1995). The dimer is formed through the nucleotide binding domains of two subunits while two dimers interact by the regulatory domains forming the tetramer with four binding sites for L-Ser (Schuller et al., 1995) (Fig. 2). The four active sites in the tetramer lie in a cleft between the substrate binding domain and the nucleotide binding domain in each subunit (Schuller et al., 1995). According to Grant and coworkers the binding of L-Ser stabilizes the regulatory domain interface in the tetramer and imparts rigidity to the enzyme structure leading to the inhibition of the catalytic activity (Grant et al., 2003). The recent crystal structure of the mutant W139G 3PGDH (Bell et al., 2004), which is a low activity form of the enzyme responsive to L-Ser but not in a cooperative manner, has provided additional information on the structure of the non-inhibited enzyme and on the inhibiting conformational change induced by L-Ser (see below).

2.2 Regulation of 3PGDH involving the ACT domain module

3PGDH displays normal Michaelis-Menten kinetics with respect to substrate (D3-phosphoglycerate), product (3-phosphohydroxypyruvate), and cofactor (NAD^+), while L-Ser is a heterotropic allosteric negative effector of the enzyme activity. The binding of L-Ser shows characteristics of both positive and negative cooperativity. Thus, in the absence of NADH the two first L-Ser molecules are bound with positive cooperativity and the two last bind

with negative cooperativity (Grant et al., 1996). When NADH is present the negative cooperativity is reduced significantly (Grant et al., 2002). The positively cooperative inhibition of enzymatic activity is completely dependent on the positive cooperative binding of the two first L-Ser molecules (Grant et al., 2004). By using asymmetric hybrid tetramers of 3PGDH it has been shown that those two L-Ser binding sites are placed at diagonally opposite interfaces at the regulatory domains (Fig. 2). Although binding to all four sites takes place, the negative cooperativity of L-Ser binding to the remaining two binding sites was not of functional importance (Grant et al., 2004).

At the subunit level there are important regions for the transmission of the inhibitory effect between the regulatory and substrate binding domains. One of these is an α-helix connecting the nucleotide binding domain to the regulatory domain, and which has important contacts with residues in the substrate binding domain (Fig. 2). The helix is flanked on each side by Gly–Gly motifs that create hinge regions crucial for allosteric control, i.e. Gly294–Gly295 on the nucleotide binding domain and Gly336–Gly337 on the regulatory domain (Grant et al., 2001a; Bell et al., 2004). In addition, there are important contacts between three arginine residues in the regulatory domain and side chain carbonyls in the substrate binding domain, and mutations of these residues reduced the cooperativity of inhibition and increased the sensitivity to L-Ser concentration (Grant et al., 2001b). The residue Trp139 is located close to the subunit interface and the active site cleft and the structure of the W139G-3PGDH mutant has provided important insights into the allosteric mechanism of the enzyme (Bell et al., 2004). Thus, the structure has revealed another important hinge, the Pro105-Thr109, which connects the substrate binding and nucleotide binding domains and has manifested the unusual flexibility of this allosteric enzyme. The binding of L-Ser induces the rotation of the nucleotide binding domain around the hinges and the disruption of several hydrophobic subunit contacts in the tetramer, while new subunit-to-subunit contacts are created. The allosteric regulation of 3PGDH appears to be mediated mainly by changes occurring at the regulatory (ACT) domain interface and the active site cleft, but the structure of the ACT domain itself does not seem to be significantly changed upon L-Phe binding (Bell et al., 2004).

3 The bifunctional P-protein of *E. coli*

The P-protein of *E. coli* is a bifunctional chorismate mutase-prephenate dehydratase that catalyzes the conversion of chorismate to prephenate and then the decarboxylation and dehydration to form phenylpyruvate (Zhang et al., 1998). These are the first two steps in the biosynthesis of L-Phe and L-Tyr via the shikimate pathway in microorganisms and plants (Herrmann and Weaver, 1999).

3.1 Structural organization

The *E. coli* P-protein has a three domain structure with i) an N-terminal domain with chorismate mutase (CM; EC 5.4.99.5) activity (residues 1–109), ii) a middle domain with prephenate dehydratase (PDT; EC 4.2.1.51) activity (residues 110–285), and iii) a regulatory C-terminal domain (residues 286–386), which belongs to the ACT domain family (Zhang et al., 1998) (Fig. 1a). This latter domain is essential to elicit the negative allosteric regulation by L-Phe binding (Zhang et al., 1998). L-Phe binds with positive cooperativity and the binding shifts the protein from dimeric to less active tetrameric and higher oligomeric forms (Baldwin et al., 1981; Pohnert et al., 1999), rather similar to the effect of L-Ser on 3PGDH (Schuller et al., 1995). The dimerization motif has been located in the CM domain while tetramerization appears to involve the regulatory domain (Zhang et al., 1998). Nevertheless, structural information is scarce since the only available crystal structure of the *E. coli* P-protein corresponds to a truncated form containing the CM domain (Lee et al., 1995), and the 3D structure of its ACT domain has not been determined yet.

3.2 Regulation of the P-protein involving the ACT domain module

The activities of both the CM and the PDT domains are negatively regulated by L-Phe, which is a feed-back inhibitor regulating the activities by heterotropic allosterism. L-Phe appears to affect the activities towards chorismate and prephenate differently, and only 10% inhibition for the CM domain activity was observed in comparison to over 90% inhibition of the PDT domain activity (Zhang et al., 1998). Also, calorimetric binding studies with deletion mutants have shown that the CM domain has little effect on the binding of L-Phe (Pohnert et al., 1999). In the same study it was also shown that a truncated form representing the regulatory domain spontaneously assembled into dimers in the absence of L-Phe and the binding of L-Phe to this fragment occurred with negative cooperativity contrary to the full-length P-protein (Pohnert et al., 1999).

The lack of 3D structural information for the full-length protein hampers the elucidation of the regulatory

mechanism and how the inhibitory effect is propagated between the domains. However, based on sequence similarity with the regulatory domains of PAH and 3PGDH and by mutational analyses, two short candidate regions, i.e. residues 309–312 (GALV) and 329–332 (ESRP), were predicted to be involved in L-Phe binding (Pohnert et al., 1999) (Fig. 1a and Appendix 1). Mutants in the GALV region show significantly decreased binding of L-Phe and diminished negative regulation of the prephenate dehydratase activity while mutations in the ESRP region only showed some reduction in binding of L-Phe and consequent feedback inhibition (Pohnert et al., 1999). Mutational studies also indicated that residue Cys374 is linked to the cooperative inhibition process and the propagation of conformational changes (Pohnert et al., 1999). The mutation of Trp338 abolished almost completely the L-Phe binding but had little effect on the feedback inhibition, which could pinpoint this residue as critical in the binding of L-Phe (Pohnert et al., 1999). The sequence similarity with PAH (Fig. 1a) and the location of a postulated allosteric L-Phe binding site in the regulatory domain of PAH will be discussed further (see below).

4 The aromatic amino acid hydroxylases (AAAH)

Phenylalanine hydroxylase (PAH, EC 1.14.16.1), tyrosine hydroxylase (TH, EC 1.14.16.2) and tryptophan hydroxylase (TPH, EC 1.14.16.4) constitute the family of non-heme iron-dependent AAAH. Recently, it has been recognized that the TPH activity in fact corresponds to two gene products, i.e. the peripheral TPH1 and the neuronal TPH2 (Walther et al., 2003). The enzymes share a number of physical, structural and catalytic properties (for review see (Kappock and Caradonna, 1996; Flatmark and Stevens, 1999; Martínez et al., 2001)).

4.1 Structural organization

The mammalian AAAH have a multi-domain structure with a common organization in an N-terminal regulatory domain, a middle catalytic domain and an oligomerization domain, with a dimerization motif and a helical tetramerization motif at the C-terminus (Figs. 1a and 3). The sequence segment that belongs to the ACT domain family corresponds to the N-terminal domain. For PAH and the TPHs the ACT domain identity was easily established (Aravind and Koonin, 1999; Chipman and Shaanan, 2001), while TH was not originally recognized as an ACT domain containing protein. It is clear, however, that TH is a homolog of PAH and the TPHs (Appendix 1), and its regulatory domain is therefore regarded as an ACT domain (from residues 78 to 162 in human TH1). The subunits are organized in the tetramer as a dimer of dimers and the interaction between dimers only takes place through the tetramerization motifs of the four subunits which are arranged in a bundle of α-helixes in antiparallel orientation (Goodwill et al., 1997; Fusetti et al., 1998). Only the structure of the N-terminal regulatory domain of PAH from rat has been determined (Kobe et al., 1999). The main core of the regulatory domain (residues 35–110 in human PAH) has the same $\beta\alpha\beta\beta\alpha\beta$ fold and topology as the ACT domain in 3PGDH (Figs. 1–3). The very N-terminus (up to residue 33) is shaped in an extended conformation that reaches and covers the active site; this sequence has been referred to as the autoregulatory sequence (A in Fig. 1a) (Kobe et al., 1999) and regulates the enzyme by intrasteric regulation (Kobe and Kemp, 1999). This N-terminal sequence is shorter in TPH1 and longer in human TH1–TH4 (the isoforms of TH generated by alternative splicing) and human TPH2, but the actual folding and function of these N-terminal segments in these other hydroxylases are not known. In PAH, the ACT domain has contacts with the catalytic domain of the adjacent subunit in one dimer, suggestive for a role of this regulatory domain both for the modulation of the activity of the catalytic domain in the same subunit and for the transmission of activating conformational changes.

4.2 Functional properties of PAH and regulation involving the ACT domain module

PAH catalyzes the hydroxylation of L-Phe to L-Tyr, the first step in the catabolic degradation of L-Phe, which in mammals takes place mainly in liver (Kaufman, 1993; Kappock and Caradonna, 1996; Flatmark and Stevens, 1999). The reaction mechanism includes two additional substrates, molecular oxygen and tetrahydrobiopterin (BH_4). Tetrameric mammalian PAH responds with positive cooperativity to increased concentration of the substrate (Parniak and Kaufman, 1981; Kaufman, 1993; Thórólfsson et al., 2002). The positive cooperativity of PAH in response to L-Phe concentration is believed to be of physiological significance as a mechanism controlling L-Phe homeostasis in blood (Kaufman, 1993; Scriver and Kaufman, 2001). It appears that the activation by L-Phe induces a conformational change that converts the enzyme to a high-affinity and high-activity state. However, the molecular basis for the cooperative activation

process is still not fully understood, partly due to the so far unsuccessful task to crystallize the tetrameric full-length enzyme (Flatmark and Stevens, 1999; Kobe et al., 1999).

Activation of the enzyme appears to require the removal of the N-terminal autoregulatory sequence away from the active site (Jennings et al., 2001; Wang et al., 2001; Thórólfsson et al., 2003). In addition, there are important residues in hinge regions and loops that are involved in the activation process. The sequence Gln31-Ala34 at the end of the autoregulatory sequence is in close proximity to the hinge Arg111-Thr117, which bridges the regulatory and catalytic domains and is likely to be involved in any domain movements leading to the activating conformational change (Kobe et al., 1999). Another hinge region is the Asp425-Gln429, which lies at the start of the tetramerization helix and is the only contact between the dimer of dimers that forms the tetrameric structure (Fusetti et al., 1998). Moreover, the area around the loop Arg68-Asp75 has important contacts both intra and inter-subunit, notably with Cys237 on the neighboring subunit in the dimer and the hinge Asp425-Gln429 (Thórólfsson et al., 2003).

The exact mechanism for this conformational change has long been debated. The most common interpretation of the cooperative substrate induced activation of PAH has been that it results from the binding of L-Phe to a putative allosteric/regulatory site, different from the catalytic site and located in the regulatory domain (Tourian, 1971; Shiman, 1980; Parniak and Kaufman, 1981; Kaufman, 1993; Kappock and Caradonna, 1996). An alternative explanation for the activation mechanism is that it results from the homotropic binding of L-Phe at the active site (Martínez et al., 1990; Martínez et al., 1993). We have recently presented further experimental evidence in favor of this proposed model of the L-Phe activation of PAH by binding to the catalytic site (Thórólfsson et al., 2002) and put forward a structural explanation of the activation mechanism using a combination of site-directed mutagenesis and molecular dynamic simulations (Thórólfsson et al., 2003): i) The binding of L-Phe to the active site would trigger the displacement of the autoregulatory sequence and consequently domain movements around the hinge region Arg111-Thr117; ii) this would further induce changes in the loop Arg68-Asp75, influencing the activation of the adjacent subunit in the dimer through interactions with Cys237 and neighboring residues; iii) the conformational changes would also result in changes in the orientation of the oligomerization domain through interactions between the loop Arg68-Asp75 and residues of the dimerization motif and/or the hinge Asp425-Gln429, as a way to transfer the cooperative activating conformational change from one dimer to the other dimer in the tetramer (Fig. 4). Recent mutational studies of hinge-bending regions have revealed further evidence that global activating conformational changes caused by L-Phe binding have the epicenter at the active site, and the changes are transmitted throughout the enzyme through hinge-bending motions (Stokka et al., 2004).

4.3 Functional properties of tyrosine hydroxylase (TH) and tryptophan hydroxylase (TPH)

TH catalyses the hydroxylation of L-Tyr to 3,4-dihydroxyphenylalanine (L-DOPA), the rate limiting step in the biosynthesis of catecholamines (dopamine, noradrenaline and adrenaline), functioning as hormones and neurotransmitters. The enzyme is not regulated by its amino acid substrate, but instead by phosphorylation at several serine residues located at the region N-terminal from the ACT domain, and by feedback inhibition by catecholamines (Flatmark and Stevens, 1999; Fitzpatrick, 2000). The catecholamines interact directly with the iron at the catalytic domain by bidentate coordination blocking the enzymatic activity (Andersson et al., 1988; Erlandsen et al., 1998).

TPH catalyses the hydroxylation of L-Trp to 5-hydroxytryptophan. This is the rate limiting step in the biosynthesis of 5-hydroxytryptamine (serotonin) and the first reaction in the synthesis of melatonin. TPH is the least known member of the AAAH family of enzymes, probably due to it being scarce and unstable (Cash, 1998; Martínez et al., 2001). Based on sequence comparisons with the other hydroxylases (Appendix 1) the ACT domain would correspond to residues 18–94 in human TPH1 and 66–142 in human TPH2. As for TH, very little is known about the role of the ACT domain in TPH, which also appears to be regulated by phosphorylation but not by its substrate or cofactor (Fitzpatrick, 2000; Martínez et al., 2001).

5 Common motifs and evolutionary insights

PAH of eukaryotic origin and the bacterial P-protein share two short sequence motifs in their regulatory (ACT) domains (see Fig. 1a and Appendix 1), the GAL and ESRP (Aravind and Koonin, 1999), which have been shown by mutational analysis to have important roles in binding of L-Phe and the consequent allosteric regulation of the P-protein (see above) (Pohnert et al., 1999). By

analogy, Gjetting et al. proposed that this motif also located the regulatory binding site for L-Phe in PAH, and found support from binding and mutational analysis on the isolated regulatory ACT domain (Gjetting et al., 2001). The protein utilized in the study by Gjetting et al. formed homodimers in solution and L-Phe was found to bind to these dimers. This behavior resembled that of the isolated regulatory ACT domains of the P-protein (see above) (Pohnert et al., 1999). Based on similarity to 3PGDH, for which a sandwich-like conformation of the ACT domains results in the formation of the amino acid binding site (Fig. 2) (Schuller et al., 1995; Chipman and Shaanan, 2001), formation of an L-Phe binding site at the dimer interface of the ACT domains of the P-protein may be envisioned. However, based on the crystal structures of truncated forms (Fusetti et al., 1998; Kobe et al., 1999) and the modeled full length tetrameric enzyme (Flatmark and Stevens, 1999; Thórólfsson et al., 2002), dimerization of the regulatory domain within the quaternary structure of PAH (Fig. 3b) does not seem possible. Nevertheless, as pointed out by both Pohnert et al. (1999) and Gjetting et al. (2001), the high conservation of the motifs GAL and ESRP between the P-protein and PAH is in fact strikingly (Figs. 1 and Appendix 1) which lead to the proposal that ''the common progenitor for the three vertebrate AAAH was formed by juxtaposition of

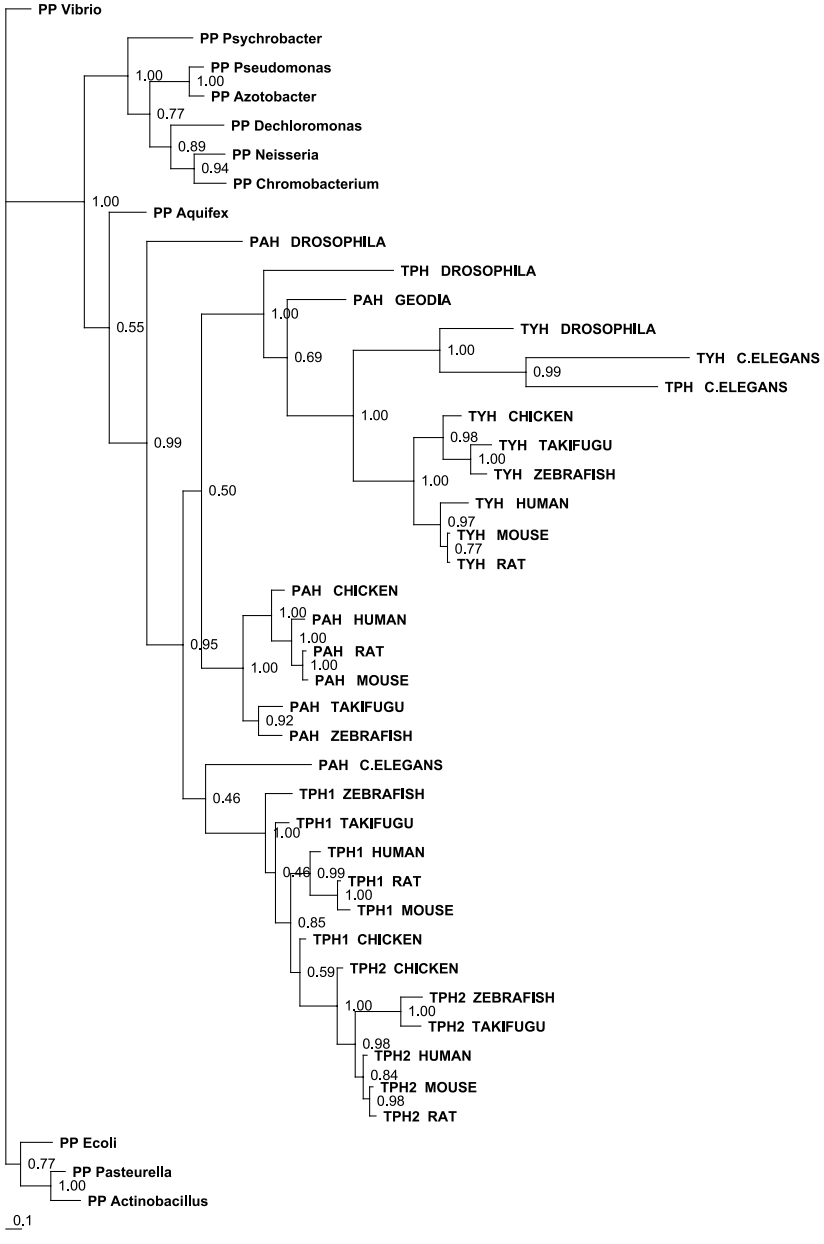


Fig. 5. The evolutionary relationships between the members of the AAAH family and P-proteins. The sequences where recruited by blasting the NCBI server (Altschul et al., 1997). A multiple sequence alignment was generated by T-Coffee (Notredame et al., 2000), and the tree was calculated using MrBayes (Ronquist and Huelsenbeck, 2003). MrBayes was run for 1500000 generations with 4 parallel chains. The tree is unrooted, and visualized with TreeView (Page, 1996). The underlying sequence alignment for the tree is shown in Appendix 1. TYH refers to TH

bacterial PAH with the regulatory domain of the P-protein, and that evolutionary changes in this domain have conferred at least some of the regulatory properties of the different hydroxylases'' (Gjetting et al., 2001). However, the motif has only slightly diversified in the other AAAH, although each and every one of the hydroxylases has their own conserved motif (see Appendix 1). As no regulatory role for their aromatic amino acid substrates is known for TH and TPH, we believe that the motifs GAL and ESRP may have another important, yet unknown, function.

To further investigate the relationship between the ACT domains of the AAAH and of the P-protein, we generated a phylogenetic tree using MrBayes (Fig. 5). In the obtained tree, where the posterior probabilities for nodes are over 69% in 86% of the cases, the AAAHs form a monophyletic group, and although the quality of the tree is low, PAH does not stand out to be the most common ancestor to the P-protein and no distinction can be made as to which one of the AAAH is more closely related to the P-protein. For the non-vertebrates the sequence for the ACT domain has diverged severely and one may ask if their function is conserved. Activation of L-Phe has not been found in PAH for lower eukaryotes e.g. *C. elegans* (Loer et al., 1999) and *D. melanogaster* (Bel et al., 1992) although at least the former contains the identical GAL and ESRP motifs (the motifs are GAL and ESRS in *D. melanogaster* PAH). Judging from the divergence in the part of the sequence corresponding to the ACT domain (Appendix 1) these proteins may function by different regulation mechanisms than their vertebrate homologs.

6 Concluding remarks

Several common features appear to characterize the allosteric regulation of 3PDGH and P-protein from *E. coli* by their amino acid effectors: the regulatory ACT domain is located at the C-terminal of the subunit sequences (Fig. 1a) and is involved in tetramerization of both enzymes; the enzymes operate at the committed step for the L-Phe (P-protein) and L-Ser (3PDGH) biosynthetic pathways; the regulatory amino acids (L-Phe and L-Ser) operate as feed-back heterotropic inhibitors (Fig. 1b) reducing the V_{max} for the reaction; the amino acids appear to shift the equilibrium from the more active to the less active conformation of the oligomer, the R (relaxed) and T (tense) conformations, respectively, in the allosteric nomenclature; at least for 3PDGH there is crystallographic evidence that the amino acid effector binds at the subunit interface created by two adjacent regulatory ACT subunits (Fig. 2) (Schuller et al., 1995), inducing both quaternary and tertiary inhibiting conformational changes that are communicated to the active site; for the P-protein there is also biochemical evidence that the binding of L-Phe at the interface between two regulatory subunits is essential for the correct allosteric regulation (Pohnert et al., 1999). Thus, truncation mutagenesis of the middle PDT domain compromises the L-Phe-induced oligomeric changes and leads to enzyme forms (containing the CM and regulatory domains) that exhibit allosteric activation of the CM activity by L-Phe, not the natural inhibition (Zhang et al., 2003).

On the other hand, the role of the ACT domain in PAH, as well as the characteristics of the allosteric regulation by the amino acid substrate in this enzyme, appear to be different to those encountered in 3PDGH and P-protein: the regulatory ACT domain is located at the N-terminal (Fig. 1a); the enzyme operates at the first step for the catabolic degradation of L-Phe; L-Phe is a positive activator of catalysis (Fig. 1b), mainly reducing the K_m for the binding of successive L-Phe molecules (Kaufman, 1993; Kappock and Caradonna, 1996); the amino acid shifts the equilibrium from the T to the R conformations; there is no crystallographic or clear biochemical evidence that in the full-length enzyme L-Phe binds at the regulatory ACT domain in addition to the catalytic site; in fact, according to the crystal structures of truncated forms of human (Fusetti et al., 1998) and rat PAH (Kobe et al., 1999), the enzyme does not contain binding interfaces implying two regulatory domains where L-Phe might bind in a similar way to L-Ser in 3PDGH (Figs. 2 and 3). Actually, the accumulated data on the cooperative activation by L-Phe (Kaufman, 1993; Kappock and Caradonna, 1996; Flatmark and Stevens, 1999) fit better with typical allosteric mechanisms found in other metabolic enzymes exhibiting positive cooperativity for their substrates than with mechanisms defending the binding of the substrate at two sites in the same subunit. Thus, in the common allosteric enzymes such as the prototypic aspartate transcarbamoylase the binding of the substrate at one active site increases the affinity for the substrate at the other active sites in the oligomer by homotropic effects – first described for the binding of dioxygen to hemoglobin-, while the other heterotropic effectors, different from the substrates, bind at the regulatory subunits (Perutz, 1990; Helmstaedt et al., 2001). So far no other allosteric physiologically relevant effectors that might regulate positively or negatively the activity of PAH, or the other AAAH, by binding at the ACT domain have been unquestionably put forward in the literature (Kappock and Caradonna, 1996).

The joint specificity of the P-protein and PAH for L-Phe as a regulatory ligand and the conserved GAL and ESRP motifs in the sequences of their ACT domains have lead to the suggestion that these enzymes would have a closer relationship to each other than to any other ACT domain family member (Pohnert et al., 1999; Gjetting et al., 2001), but our preliminary phylogenetic analyses do not show a closer relationship between the P-protein with PAH than with the other AAAH (Fig. 5). Accordingly, this domain appears to participate in the allosteric regulation of the AAAH by the transmission of activating conformational changes (Thórólfsson et al., 2003) without providing a binding site for the regulatory amino acid substrate. In the case of TH and TPH the activating changes might be related to phosphorylation and interaction with 14.3.3 proteins (Fitzpatrick, 2000). In this context it also seems interesting that the 3PGDH of *Arabidopsis thaliana*, although it contains the ACT regulatory domain at the C-terminal, does not shown inhibition by L-Ser (Ho and Saito, 2001). The ACT domain could thus have been incorporated to these and other enzymes, as a handy, flexible, small module that provides allosteric regulation *via* transmission of finely tuned conformational changes, which are not necessarily initiated by ligand binding at the domain itself. In fact allostery appears to be associated to very specific interactions and subtle conformational changes in dynamic proteins (Gunasekaran et al., 2004). There are a number of studies showing that a single point mutation can turn an allosteric activation into inhibition and viceversa (Van Vliet et al., 1991; Muneyuki et al., 1997). In agreement with this notion, one of the residues in the catalytic domain interacting with the Arg (Arg68 in human PAH) in the ESRP motif at the regulatory domain is a strictly conserved Cys in all PAHs (C237 in human PAH), while all TPHs and THs sequences have an Arg (except golden hamster TPH that has a Cys). Interestingly, the mutant C237R of PAH is not activated by L-Phe (Thórólfsson et al., 2003). Another example on the fact that correct allosteric regulation requires a fine tuning of ligand induced changes in tertiary and quaternary structure is the change from negative to positive allosteric regulation of CM activity by L-Phe encountered upon deletion of the PDT domain in the *E. coli* P-protein (Zhang et al., 2003). Thus, the regulatory effect of the ACT domain appears to be set by its fine interactions with the partner domains.

Acknowledgements

We are very grateful to all members from our laboratory for critical discussions and to The Research Council of Norway for financial support.

References

Achouri Y, Rider MH, Schaftingen EV, Robbi M (1997) Cloning, sequencing and expression of rat liver 3-phosphoglycerate dehydrogenase. Biochem J 323(Pt 2): 365–370

Al-Rabiee R, Zhang Y, Grant GA (1996) The mechanism of velocity modulated allosteric regulation in D-3-phosphoglycerate dehydrogenase. Site-directed mutagenesis of effector binding site residues. J Biol Chem 271: 23235–23238

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389–3402

Anantharaman V, Koonin EV, Aravind L (2001) Regulatory potential, phyletic distribution and evolution of ancient, intracellular small-molecule-binding domains. J Mol Biol 307: 1271–1292

Andersson KK, Cox DD, Que L Jr, Flatmark T, Haavik J (1988) Resonance Raman studies on the blue-green-colored bovine adrenal tyrosine 3-monooxygenase (tyrosine hydroxylase). Evidence that the feedback inhibitors adrenaline and noradrenaline are coordinated to iron. J Biol Chem 263: 18621–18626

Aravind L, Koonin EV (1999) Gleaning non-trivial structural, functional and evolutionary information about proteins by iterative database searches. J Mol Biol 287: 1023–1040

Baldwin GS, McKenzie GH, Davidson BE (1981) The self-association of chorismate mutase/prephenate dehydratase from Escherichia coli K12. Arch Biochem Biophys 211: 76–85

Bel Y, Jacobson KB, Ferre J (1992) A comparative study of Drosophila phenylalanine hydroxylase with a natural and a synthetic tetrahydropterin as cofactor. Comp Biochem Physiol B 103: 557–562

Bell JK, Grant GA, Banaszak LJ (2004) Multiconformational states in phosphoglycerate dehydrogenase. Biochemistry 43: 3450–3458

Cash CD (1998) Why tryptophan hydroxylase is difficult to purify: a reactive oxygen-derived species-mediated phenomenon that may be implicated in human pathology. Gen Pharmacol 30: 569–574

Chipman DM, Shaanan B (2001) The ACT domain family. Curr Opin Struct Biol 11: 694–700

Erlandsen H, Flatmark T, Stevens RC, Hough E (1998) Crystallographic analysis of the human phenylalanine hydroxylase catalytic domain with bound catechol inhibitors at 2.0 A resolution. Biochemistry 37: 15638–15646

Fitzpatrick PF (2000) The aromatic amino acid hydroxylases. Adv Enzymol Relat Areas Mol Biol 74: 235–294

Flatmark T, Stevens RC (1999) Structural insight into the aromatic amino acid hydroxylases and their disease-related mutant forms. Chem Rev 99: 2137–2160

Fusetti F, Erlandsen H, Flatmark T, Stevens RC (1998) Structure of tetrameric human phenylalanine hydroxylase and its implications for phenylketonuria. J Biol Chem 273: 16962–16967

Gjetting T, Petersen M, Guldberg P, Guttler F (2001) Missense mutations in the N-terminal domain of human phenylalanine hydroxylase interfere with binding of regulatory phenylalanine. Am J Hum Genet 68: 1353–1360

Goodwill KE, Sabatier C, Marks C, Raag R, Fitzpatrick PF, Stevens RC (1997) Crystal structure of tyrosine hydroxylase at 2.3 A and its implications for inherited neurodegenerative diseases. Nat Struct Biol 4: 578–585

Grant GA, Schuller DJ, Banaszak LJ (1996) A model for the regulation of D-3-phosphoglycerate dehydrogenase, a Vmax-type allosteric enzyme. Protein Sci 5: 34–41

Grant GA, Kim SJ, Xu XL, Hu Z (1999) The contribution of adjacent subunits to the active sites of D-3-phosphoglycerate dehydrogenase. J Biol Chem 274: 5357–5361

Grant GA, Hu Z, Xu XL (2001a) Amino acid residue mutations uncouple cooperative effects in Escherichia coli D-3-phosphoglycerate dehydrogenase. J Biol Chem 276: 17844–17850

Grant GA, Hu Z, Xu XL (2001b) Specific interactions at the regulatory domain-substrate binding domain interface influence the cooperativity of inhibition and effector binding in Escherichia coli D-3-phosphoglycerate dehydrogenase. J Biol Chem 276: 1078–1083

Grant GA, Hu Z, Xu XL (2002) Cofactor binding to Escherichia coli D-3-phosphoglycerate dehydrogenase induces multiple conformations which alter effector binding. J Biol Chem 277: 39548–39553

Grant GA, Hu Z, Xu XL (2003) Hybrid tetramers reveal elements of cooperativity in Escherichia coli D-3-phosphoglycerate dehydrogenase. J Biol Chem 278: 18170–18176

Grant GA, Xu XL, Hu Z (2004) Quantitative relationships of site to site interaction in Escherichia coli D-3-phosphoglycerate dehydrogenase revealed by asymmetric hybrid tetramers. J Biol Chem 279: 13452–13460

Gunasekaran K, Ma B, Nussinov R (2004) Is allostery an intrinsic property of all dynamic proteins? Proteins 57: 433–443

Helmstaedt K, Krappmann S, Braus GH (2001) Allosteric regulation of catalytic activity: Escherichia coli aspartate transcarbamoylase versus yeast chorismate mutase. Microbiol Mol Biol Rev 65: 404–421, table of contents

Herrmann KM, Weaver LM (1999) The Shikimate Pathway. Annu Rev Plant Physiol Plant Mol Biol 50: 473–503

Ho CL, Saito K (2001) Molecular biology of the plastidic phosphorylated serine biosynthetic pathway in Arabidopsis thaliana. Amino Acids 20: 243–259

Jennings IG, Teh T, Kobe B (2001) Essential role of the N-terminal autoregulatory sequence in the regulation of phenylalanine hydroxylase. FEBS Lett 488: 196–200

Kappock TJ, Caradonna JP (1996) The aromatic amino acid hydroxylases. Chem Rev 96: 2659–2756

Kaufman S (1993) The phenylalanine hydroxylating system. AdvEnzymol Relat Areas Mol Biol 67: 77–264

Kobe B, Kemp BE (1999) Active site-directed protein regulation. Nature 402: 373–376

Kobe B, Jennings IG, House CM, Michell BJ, Goodwill KE, Santarsiero BD, Stevens RC, Cotton RG, Kemp BE (1999) Structural basis of autoregulation of phenylalanine hydroxylase. Nat Struct Biol 6: 442–448

Lee AY, Karplus PA, Ganem B, Clardy J (1995) Atomic-Structure of the Buried Catalytic Pocket of Escherichia-Coli Chorismate Mutase. J Am Chem Soc 117: 3627–3628

Loer CM, Davidson B, McKerrow J (1999) A phenylalanine hydroxylase gene from the nematode C. elegans is expressed in the hypodermis. J Neurogenet 13: 157–180

Martínez A, Haavik J, Flatmark T (1990) Cooperative homotropic interaction of L-noradrenaline with the catalytic site of phenylalanine 4-monooxygenase. Eur J Biochem 193: 211–219

Martínez A, Ólafsdóttir S, Flatmark T (1993) The cooperative binding of phenylalanine to phenylalanine 4-monooxygenase studied by 1H-NMR paramagnetic relaxation. Changes in water accessibility to the iron at the active site upon substrate binding. Eur J Biochem 211: 259–266

Martínez A, Knappskog PM, Haavik J (2001) A structural approach into human tryptophan hydroxylase and its implications for the regulation of serotonin biosynthesis. Curr Med Chem 8: 1077–1091

Muneyuki E, Odaka M, Yoshida M (1997) A single mutation at the catalytic site of TF1-alpha3beta3gamma complex switches the kinetics of ATP hydrolysis from negative to positive cooperativity. FEBS Lett 413: 55–59

Notredame C, Higgins DG, Heringa J (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol 302: 205–217

Page RD (1996) TreeView: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 12: 357–358

Parniak MA, Kaufman S (1981) Rat liver phenylalanine hydroxylase. Activation by sulfhydryl modification. J Biol Chem 256: 6876–6882

Perutz MF (1990) Mechanisms of cooperativity and allosteric regulation in proteins. Cambridge University Press, Cambridge, New York, pp x, 101, 108 of plates

Pind S, Slominski E, Mauthe J, Pearlman K, Swoboda KJ, Wilkins JA, Sauder P, Natowicz MR (2002) V490M, a common mutation in 3-phosphoglycerate dehydrogenase deficiency, causes enzyme deficiency by decreasing the yield of mature enzyme. J Biol Chem 277: 7136–7143

Pohnert G, Zhang S, Husain A, Wilson DB, Ganem B (1999) Regulation of phenylalanine biosynthesis. Studies on the mechanism of phenylalanine binding and feedback inhibition in the Escherichia coli P-protein. Biochemistry 38: 12212–12217

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574

Schuller DJ, Grant GA, Banaszak LJ (1995) The allosteric ligand site in the Vmax-type cooperative enzyme phosphoglycerate dehydrogenase. Nat Struct Biol 2: 69–76

Scriver CR, Kaufman S (2001) Hyperphenylalaninemia:phenylalanine hydroxylase defciency. In: Scriver CR, Beaudet AL, Valle D, Sly WS (eds) The metabolic and molecular bases of inherited disease, 8th edn. McGraw-Hill, New York, pp 1667–1724

Shiman R (1980) Relationship between the substrate activation site and catalytic site of phenylalanine hydroxylase. J Biol Chem 255: 10029–10032

Shindyalov IN, Bourne PE (1998) Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. Protein Eng 11: 739–747

Stokka AJ, Carvalho RN, Barroso JF, Flatmark T (2004) Probing the role of crystallographically defined/predicted hinge-bending regions in the substrate-induced global conformational transition and catalytic activation of human phenylalanine hydroxylase by single-site mutagenesis. J Biol Chem 279: 26571–26580

Thórólfsson M, Ibarra-Molero B, Fojan P, Petersen SB, Sanchez-Ruiz JM, Martínez A (2002) L-Phenylalanine binding and domain organization in human phenylalanine hydroxylase: a differential scanning calorimetry study. Biochemistry 41: 7573–7585

Thórólfsson M, Teigen K, Martínez A (2003) Activation of phenylalanine hydroxylase: effect of substitutions at Arg68 and Cys237. Biochemistry 42: 3419–3428

Tourian A (1971) Activation of phenylalanine hydroxylase by phenylalanine. Biochim Biophys Acta 242: 345–354

Van Vliet F, Xi XG, De Staercke C, de Wannemaeker B, Jacobs A, Cherfils J, Ladjimi MM, Herve G, Cunin R (1991) Heterotropic interactions in aspartate transcarbamoylase: turning allosteric ATP activation into inhibition as a consequence of a single tyrosine to phenylalanine mutation. Proc Natl Acad Sci USA 88: 9180–9183

Walther DJ, Peter JU, Bashammakh S, Hortnagl H, Voits M, Fink H, Bader M (2003) Synthesis of serotonin by a second tryptophan hydroxylase isoform. Science 299: 76

Wang GA, Gu P, Kaufman S (2001) Mutagenesis of the regulatory domain of phenylalanine hydroxylase. Proc Natl Acad Sci USA 98: 1537–1542

Zhang S, Pohnert G, Kongsaeree P, Wilson DB, Clardy J, Ganem B (1998) Chorismate mutase-prephenate dehydratase from Escherichia coli. Study of catalytic and regulatory domains using genetically engineered proteins. J Biol Chem 273: 6248–6253

Zhang S, Wilson DB, Ganem B (2003) An engineered chorismate mutase with allosteric regulation. Bioorg Med Chem 11: 3109–3114

Authors' address: Aurora Martínez, Professor, Department of Biomedicine, University of Bergen, Jonas Lies vei 91, 5009 Bergen, Norway, Fax: (47)55586360, E-mail: aurora.martinez@biomed.uib.no

Appendix 1. T-Coffee alignment used to generate the free shown in Fig. 5. The secondary structure (E = beta strand, H = alpha helix) are given for the rat PAH structure, PDB 1PHZ

```
                          EEEEEEEE       HHHHHHHHHH        EEEEE         EEEEEE
PP_Vibrio           -----KTTLIMSTSQ---EAGSLVETLLVLQRYGINMTKLESRPIMGNPW---EEMFYVD
PP_Ecoli            -----KTTLLMATGQ---QAGALVEALLVLRNHNLIMTRLESRPIHGNPW---EEMFYLD
PP_Pasteurella      -----KTLLLMTTSQ---QAGALVDALFVFKKHQINMTKLESRPIYGKPW---EEMFYLE
PP_Actinobacillus   -----KTLLLMTTSQ---QAGALADALMVFKQHQIRMTKLESRPIYGKPW---EEMFYIE
PP_Aquifex          -----KTSILFGVKD---EPGALYKALEVFYKHGINLTKIESRPSKKKAW---DYVFFVD
PP_Dechloromonas    -----KTSLIMSAPN---RTGALHELLLPLSTAGVSMCRLESRPARNALW---EYVFYVD
PP_Neisseria        -----KTSLAVSAPN---RAGAVASLLQPLTESGISMTKFESRPSKSVLW---EYLFFID
PP_Chromobacterium  -----KTSIVVSAPN---RPGAVHQLLEPVAANGVSMSKFESRPSRAGLW---DYVFFID
PP_Pseudomonas      -----KTSIIVSMRN---KPGALHELLVPFHNNGIDLTRIETRPSRSGKW---TYVFFID
PP_Azotobacter      -----KTSIIVSMNN---KPGALHALLMPFHENGIDLTRIETRPSRSGKW---TYVFFID
PP_Psychrobacter    -----KTSIMVSAHD---KAGALIEILKPLSYHGVSMTSIETRPERPNKW---AYVFFID
PAH___DROSOPHILA    ------TCLLFSPKDSSLSSGALANILAIFKKHDINLVHIESRSSLRVP----GYEFFVE
PAH___GEODIA        ---SANLSLIFSLKD---EQGSLVTSLKPFQDMGINMTHLESRPSKSNPG--SEYDFYVD
PAH___HUMAN         CNQNGAISLIFSLKE---EVGALAKVLRLFEENDVNLTHIESRPSRLKKD---EYEFFTH
PAH___RAT           SNQNGAISLIFSLKE---EVGALAKVLRLFEENDINLTHIESRPSRLNKD---EYEFFTY
PAH___MOUSE         SNQNGAVSLIFSLKE---EVGALAKVLRLFEENEINLTHIESRPSRLNKD---EYEFFTY
PAH___CHICKEN       -NKNGVISLIFSLKE---EVGALAKVLRTFEEKGINLTHIESRPSRLNKD---EYEFFIN
PAH___TAKIFUGU      ----EVISCIFSLKQ---EVGPLTKALKLFEEKGINLTHIESRPSRTNTG---QYEFFIN
PAH___ZEBRAFISH     -NKSGVVSCIFSLKE---EVGALVKALKLFEEKGINLTHIESRPSRKNNE---EYEFFIS
TPH1__RAT           SSERGRVTLIFSLKN---EVGGLIKALKIFQENHVNLLHIESRKSKRRNS---EFEIFVD
TPH1__MOUSE         -SERGRVTLIFSLEN---EVGGLIKVLKIFQENHVSLLHIESRKSKQRNS---EFEIFVD
TPH1__HUMAN         --ERGRATLIFSLKN---EVGGLIKALKIFQEKHVNLLHIESRKSKRRNS---EFEIFVD
TPH1__CHICKEN       --ERGRTAIIFSLKN---EVGGLVKALKLFQEKHVNLVHIESRKSKRRNS---EFEIFVD
TPH1__TAKIFUGU      -SEKARATIIFSLKN---EVGGLVKALKLFQENHVNLVHIESRKSKRRNS---EFEIFVD
TPH1__ZEBRAFISH     -SETGRAAVVFSLKN---EVGGLVKALKLFQENHVNLVHIESRKSKRRNS---EFEIFVD
TPH2__MOUSE         -TESSKTAVVFSLKN---EVGGLVKALRLFQEKHVNMLHIESRRSRRRSS---EVEIFVD
TPH2__RAT           -SESSKTAVVFSLKN---EVGGLVRALRLFQEKHVNMLHIESRRSRRRSS---EVEIFVD
TPH2__HUMAN         -TESGKTAVVFSLKN---EVGGLVKALRLFQEKRVNMVHIESRKSRRRSS---EVEIFVD
TPH2__CHICKEN       -SEGGKTAVVFSLKN---EVGGLVKALRLFQEKHVSMVHIESRKSKRRNS---EVEIFVD
TPH2__ZEBRAFISH     ----GKLAVIFSLKN---EVGFLVKALRLFQEKHVNLAHIESRRSKRLTN---EIEIYAE
TPH2__TAKIFUGU      -----KTAVIFSLKN---EVGCLVKALRLFQEKRVNLRHIESRASRRLTD---EVEIFAD
PAH___C.ELEGANS     ----GKTTIVFTLRE---KAGALAETLKLFQAHDVNLSHIESRPSKTHEG---CYEVLVE
TYH___CHICKEN       -DGRAMLNLFFMLKGA--KTSPLSRALKVFETFEAKIHHLETRLSRKPREGTAELEYFVR
TYH___TAKIFUGU      -DGRALLNIFFTLRNT--KTPALSRTLKVFETFEAKIHHLETRPCPKLKDSLEGLEYFVR
TYH___ZEBRAFISH     -DGKALISLFFTLRCS--KSPALSRTLKVFETFEAKIHHLETRPSRKPKDGLEDLEYYVQ
TYH___MOUSE         -DGNAVLNLLFSLRGT--KPSSLSRALKVFETFEAKIHHLETRPAQRPLAGSPHLEYFVR
TYH___RAT           -DGNAVLNLLFSLRGT--KPSSLSRAVKVFETFEAKIHHLETRPAQRPLAGSPHLEYFVR
TYH___HUMAN         -EGKAVLNLLFSPRAT--KPSALSRAVKVFETFEAKIHHLETRPAQRPRAGGPHLEYFVR
TYH___C.ELEGANS     ------FSIILTSTDP-TLSNFVSDILQNMSSAKVQICHVETRGNEASHDVLLACKA-TK
TYH___DROSOPHILA    -AAMQSAALVVRLKE---GISSLGRILKAIETFHGTVQHVESRQSRVEG---VDHDVLIK
TPH___DROSOPHILA    -SPGERISIIFTLRN---QVGNLARALQVFQELGINVLHLELSPLEMATN---QADVLVD
TPH___C.ELEGANS     EEGVQILTIIVKSSRVSEDISKMIANL----PDHTRIKHLETRDSQDGSS--KTMDVLLE
                                .           :   :          :  .*
                          HHHHHHHHHHHHHH                        EEEE----E
PP_Vibrio           LEAHLGSTEMQQALQELTKIT--------------------KHLKVLGCYPSENIKP--
PP_Ecoli            IQANLESAEMQKALKELGEIT--------------------RSMKVLGCYPSENVVPV-
PP_Pasteurella      IEANIHNPETQQALDELKQYS--------------------HYLKVLGCYPSEIIKPV-
PP_Actinobacillus   LQANIHSENTQQALKALENVT--------------------SYIKVLGCYPSEIIEPV-
PP_Aquifex          LEGHKEEERVEKALKELKEKT--------------------QFLKVLGSYPKALLQE--
PP_Dechloromonas    IEGHRDEPAIKAALEKLAGYA--------------------AYLKILGSYPVAVY----
PP_Neisseria        IEGHRRDAQIQTALERLGERA--------------------SFVKAIGSYPTAVL----
PP_Chromobacterium  LEGHRQDERVLKALAGLGERT--------------------SFVKVLGSYPTAVL----
PP_Pseudomonas      FVGHHKEPLIKDVLEKIGQEA--------------------VALKVLGSYPKAVL----
PP_Azotobacter      FVGHRHDPLVKSVLERLSQEV--------------------VALKVLGSYPKAVL----
PP_Psychrobacter    INGHIADPNVSAAIADIRPLV--------------------KDLRVLGSYPKAVL----
PAH___DROSOPHILA    ADGKSG--ALGKAIEDVKEQC--------------------SYFNII----SRDYK---
PAH___GEODIA        CVCPPD--KKEDLLSSLRANS--------------LTVNILSRDPGE-DEVPWFPRK
PAH___HUMAN         LD-KRSLPALTNIIKILRHDI--------------------GATVHEL----SRDKKK--
PAH___RAT           LD-KRTKPVLGSIIKSLRNDI--------------------GATVHEL----SRDKEK--
PAH___MOUSE         LDK-RSKPVLGSIIKSLRNDI--------------------GATVHEL----SRDKEK--
PAH___CHICKEN       LEG-KNVPALDKIIKSLRNDI--------------------GVTVHEL----SRTKKK--
PAH___TAKIFUGU      VDTSCS-QALDEVIDSLRTEI--------------------SGHVHEL----SRNKEK--
PAH___ZEBRAFISH     VD-QASSKVLDEIVDGLRTQI--------------------RGQVHEL----SRNKQK--
TPH1__RAT           CDINR--EQLNDIFPLLKSHT--------------------TVLSVDS-PDQLPEK--
TPH1__MOUSE         CDISR--EQLNDIFPLLKSH---------------------ATVLSVDS-PDQLTAK--
TPH1__HUMAN         CDTNR--EQLNDIFHLLKSHT----------------------NVLSVTPPDNFTMK-
TPH1__CHICKEN       CDSNR--EQLNEIFQLLKSH---------------------VSIVSMNPTEHFN---
TPH1__TAKIFUGU      CDSNH--EQLNEIIQLLQKN---------------------VSVLDMEPLD------
TPH1__ZEBRAFISH     CDSNR--EQLHEIIQLLRKHV--------------------NVVEMDA-PDNRLPEE-
TPH2__MOUSE         CECGK--TEFNELIQLLKFQT--------------------TIVTLN--PPESI----
TPH2__RAT           CECGK--TEFNELIQLLKFQT--------------------TIVTLN--PPDNI----
TPH2__HUMAN         CECGK--TEFNELIQLLKFQT--------------------TIVTLN--PPENI----
TPH2__CHICKEN       CDCSK--KEFNELIQLLKFQT--------------------NIVSLNPPENI----
TPH2__ZEBRAFISH     CNCTK--KEFNELVQHLKDHV--------------------NIVSYNT-PQHV-----
TPH2__TAKIFUGU      CSCGP--KEFNELLEHLKDHV--------------------NIISFNT-P--------
PAH___C.ELEGANS     FAEAEDHRKIEGVIEHFQQKA--------------------EKKVLVQDWNTKNK-----
TYH___CHICKEN       CEVHSS--DLNTFISSIKRVA--------------------EDVRTT-KEDKFHWFPRK
TYH___TAKIFUGU      CELHLS--DVSTLIGSLKRNA--------------------EDVKTTKEV-KFHWFPKK
TYH___ZEBRAFISH     CEVHLS--DVSTLVSSLKRSA--------------------EDVKTTKEV-KFHWFPRK
TYH___MOUSE         FEVPSG--DLAALLSSVRRVS--------------------DDVRSARE-DKVPWFPRK
TYH___RAT           FEVPSG--DLAALLSSVRRVS--------------------DDVRSARE-DKVPWFPRK
TYH___HUMAN         LEVRRG--DLAALLSGVRQVS--------------------EDVRSPAG-PKVPWFPRK
TYH___C.ELEGANS     NQLIHSAELLTQNHVALTKFS-----------------IFAKKLSDEKNQSQIWFPRH
TYH___DROSOPHILA    LDMTRG--NLLQLIRSLRQSG--------------SFSSMNLMADNNLNVKAPWFPKH
TPH___DROSOPHILA    VECDQ--RRLDQVVKMLNREV-ASVNYTSVNTQGLARAPSLSACSSFDFGDMVWFPRK
TPH   C.ELEGANS     IELFHYGKQ--EAMDLMRLNGLDVHEVSSTIRPTAIKEQYTEPGSDDATTGSEWFPKS
```