Activation of Phenylalanine Hydroxylase: Effect of Substitutions at Arg68 and Cys237†

Matthías Thórólfsson, Knut Teigen, and Aurora Martínez*

Department of Biochemistry and Molecular Biology, University of Bergen, Årstadveien 19, 5009-Bergen, Norway

Received January 8, 2003

ABSTRACT: Phenylalanine hydroxylase (PAH) is a multidomain tetrameric enzyme that displays positive cooperative substrate binding. This cooperative response is believed to be of physiological significance as a mechanism that controls L-Phe homeostasis in blood. The substrate induces an activating conformational change in the enzyme affecting the secondary, tertiary, and quaternary structures. Chemical modification and substitution with a negatively charged residue of Cys237 in human PAH (hPAH) also result in activation of the enzyme. As seen in the modeled structure of full-length hPAH, Cys237 is located in the catalytic domain close to residues in the oligomerization and regulatory domains of an adjacent subunit in the dimer, notably to Arg68. This residue is located in a prominent loop (68−75), which also has contacts with the dimerization motif from the same subunit. To investigate further the involvement of Cys237 and Arg68 in the activation of the enzyme, we have prepared mutants of hPAH at these positions, with substitutions of different charge and size. The mutations C237D, R68A, and C237A cause an increase of the basal activity and affinity for L-Phe, while the mutation C237R results in reduced affinity for the substrate and elimination of the positive cooperativity. The conformational changes induced by the mutations were studied by far-UV circular dichroism, fluorescence spectroscopy, and molecular dynamics simulations. All together, our results indicate that the activating mutations induce a series of conformational changes including both the displacement of the inhibitory N-terminal sequence (residues 19−33) that covers the active site and the domain movements around the hinge region Arg111−Thr117, in addition to the rearrangement of the loop 68−75. The same conformational changes appear to be involved in the activation of PAH induced by L-Phe.

The mechanisms of cooperativity and allosteric regulation in proteins have been intensively studied, and a great deal of information has been gathered throughout the past decades (*1*−*4*). However, few systems have been characterized in detail because of the scarcity of structural information on the different states of these proteins, which usually have complex multidomain and oligomeric structures. One of these proteins is the enzyme phenylalanine hydroxylase (PAH, phenylalanine 4-monooxygenase, EC 1.14.16.1),[1] which was first described in the late 1950s (for review see refs *5*−*7*). A series of recent structural reports (*8*−*12*) have provided important insights into the catalytic mechanism and the regulation of PAH. This enzyme catalyzes the first step in the catabolic degradation of L-phenylalanine (L-Phe), which in mammals takes place mainly in the liver (*5*−*7*). In the enzymatic reaction, a hydroxyl group is introduced into the para position on the phenyl group forming L-tyrosine (L-Tyr). The reaction mechanism includes two additional substrates, molecular oxygen and tetrahydrobiopterin (BH_4). The absence or severely reduced activity of the human PAH (hPAH) because of gene mutations results in phenylketonuria (PKU), which is the most frequent autosomal recessive inborn error of amino acid metabolism.

Tetrameric mammalian PAH is activated by preincubation with L-Phe (for review see ref *5*). In addition, the enzyme responds to increased concentration of the substrate showing positive cooperativity, as found both by steady-state kinetic measurements of the BH_4-dependent activity and by equilibrium binding analysis (*5, 13*). These regulatory properties have been mostly studied with the enzyme isolated from rat liver (*6*) and the recombinant rat and human enzymes (*14, 15*). 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 (*6, 16*). The activation by L-Phe induces a conformational change that converts the enzyme from a low-affinity and low-activity state to a high-affinity and high-activity state (the T- and R-states, respectively, according to the cooperative allosteric models (*17, 18*)). However, the molecular basis for the positive cooperativity and substrate activation of this enzyme remains unresolved, partly becasue of the so far unsuccessful task to crystallize the tetrameric full-length enzyme in the presence and the absence of L-Phe (*7, 10*).

As shown by deletion mutagenesis (*15*) and by the crystal structure (*8*−*10*), PAH has a three-domain structure. Each subunit contains (i) an N-terminal regulatory domain (up to Ser110) that contains an $\alpha-\beta$ sandwich core and an extended

† This work was supported by the Research Council of Norway.

* To whom correspondence should be addressed. Telephone: (47)-55586427. Fax: (47)55586400. E-mail: aurora.martinez@ibmb.uib.no.

[1] Abbreviations: BH_4, (6*R*)-L-*erythro*-5,6,7,8-tetrahydrobiopterin; CD, circular dichroism; hPAH, human phenylalanine hydroxylase; MD, molecular dynamics; PAH, phenylalanine hydroxylase; PKU, phenylketonuria; wt-hPAH, wild-type human phenylalanine hydroxylase.

10.1021/bi034021s CCC: $25.00
Published on Web 03/04/2003

N-terminal autoregulatory sequence (residues 19−33) that stretches over the active site of the catalytic domain; (ii) a catalytic domain (Asp111−Phe410), containing the active site iron and the substrate and BH_4 binding sites; and (iii) an oligomerization domain, with a dimerization motif (Ser411−Thr427) and a helical tetramerization motif at the C-terminus that promotes oligomerization through antiparallel helix−helix interactions (*9*). The displacement of the N-terminal autoregulatory sequence away from the active site is supposed to be an important activating conformational change that accompanies the transition from the T- to the R-state (*19*, *20*). In addition, the core of the regulatory domain has contacts with the catalytic domain of the neighboring subunit in one dimer (*10*), suggestive for a role of the 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. The residue Cys237 in the catalytic domain is positioned at this subunit interface. Modification of Cys237 by treatment with *N*-ethylmaleimide and 5,5′-dithiobis(2-nitrobenzoate) (DTNB) results in an activation of the enzyme comparable to that induced by L-Phe (*13*, *21*). In addition, we have previously shown that mutation of this residue to a negatively charged amino acid (C237D-hPAH) results in a form with higher basal activity than L-Phe-activated wt-hPAH (*22*). The closest residue to Cys237 in the regulatory domain on the adjacent subunit in the dimer is Arg68, which lies in a prominent loop (residues 68−75). This region of the regulatory domain also establishes contacts with the oligomerization domain, which includes both a dimerization motif (Ser411-Thr427) and a tetramerization helix at the C-terminus (*9*). To investigate the involvement of Cys237 and Arg68 in the activation of PAH, we have introduced substitutions of different charge and size at these positions, and we have studied the kinetic properties of the mutants. The conformational changes caused by the mutations have been studied by circular dichroism (CD) and fluorescence spectroscopy. Molecular dynamic (MD) simulations have also been used to analyze the structural effects of the mutations, and the results obtained have contributed to the interpretation of the experimental data and to the understanding of the activation mechanism of PAH.

EXPERIMENTAL PROCEDURES

Site-Directed Mutagenesis and Purification of Wild-Type hPAH and Mutants at Arg68 and Cys237. The plasmid containing the coding sequences for maltose binding protein (MBP), the factor Xa cleavage site (IEGR), and human PAH was prepared as described (*23*), and the mutations R68V, R68A, and C237A were introduced in the hPAH cDNA using the QuikChange site-directed mutagenesis kit (Stratagene). The primers for mutagenesis (provided by Eurogentec and M. W. G. Biotech) are shown in Table 1. The authenticity of the mutagenesis was verified by DNA sequencing. The mutant C237D-hPAH was prepared as described (*22*). The purification of the fusion proteins, cleavage by factor Xa (New England Biolabs), and further purification by high-performance size-exclusion chromatography on a HiLoad Superdex column (1.6. × 60 cm) was performed as described (*15*, *23*). Protein concentration was estimated by the absorbance at 280 nm, using the absorption coefficient A_{280} (1 mg mL^{-1} cm^{-1}) = 1.0 for the purified hPAH forms (*23*).

Table 1: Primers Used for PCR Site-Directed Mutagenesis and DNA Sequencing[a]

mutant	sense	sequence (5′-3′)
R68V	forward	AACCACATTGAATCT**GTA**CCTTCTCGTTTA
R68V	reverse	TAAACGAGAAGG**TAC**AGATTCAATGTGGGT
R68A	forward	ACCCACATTGAATCT**GCA**CCTTCTCGTTTA
R68A	reverse	TAAACGAGAAGG**TGC**AGATTCAATGTGGGT
C237A	forward	CAATTCCTGCAGACT**GCC**ACTGGTTTCCGC
C237A	reverse	GCGGAAACCAGTG**GCA**GTCTGCAGGAATTG
C237R	forward	CAATTCCTGCAGACT**CGC**ACTGGTTTCCGC
C237R	reverse	GCGGAAACCAGTG**CGA**GTCGTCAGGAATTG

[a] Mismatched codons are shown in boldface.

Assay of Enzymatic Activity. The activity for the nonactivated (non L-Phe preincubated) and activated (L-Phe preincubated) hPAH enzymes was measured as described (*24*). The degree of activation (fold) refers to the ratio between the specific activities of the L-Phe-preincubated and the nonpreincubated samples. $S_{0.5}$ values for L-Phe were calculated for the enzyme preincubated with L-Phe at the indicated concentrations of L-Phe and 75 μM BH_4. The K_m values for BH_4 were obtained with variable concentration of BH_4 (0−500 μM) and 1 mM L-Phe. The kinetic parameters were calculated by nonlinear regression analysis of the experimental data using the Hill equation.

Circular Dichroism (CD). CD measurements were performed at 25 °C on a Jasco J-810 spectropolarimeter equipped with Jasco PTC-423S Peltier element for temperature control. Enzyme samples (5 μM subunit) were prepared in 20 mM sodium phosphate, 0.15 M KF, pH 7.0, in the presence and the absence of 1 mM L-Phe and placed in quartz cells with a path length of 1 mm. Buffer scans (with and without L-Phe) were routinely recorded and subtracted from the original spectra. Estimation of secondary structure elements was performed using the CDNN program that applies a back-propagation neural network model for the quantitative analysis of protein UV CD spectra (*25*).

Fluorescence Measurements. Measurements of intrinsic tryptophan fluorescence were performed at 25 °C in 20 mM Na-Hepes, 0.2 M NaCl at pH 7.0 with 1 μM PAH subunit. A Perkin-Elmer LS-50 luminescence spectrometer with a constant-temperature cell holder and 1-cm path length quartz cells was used. The excitation and emission slits were 3 and 5, respectively. All spectra were corrected for blank emission.

Molecular Modeling and Molecular Dynamics (MD) Simulations. A structural model of full-length human PAH was created by aligning the catalytic domains of the dimeric form of rat PAH (Ser19−Gln429) (*10*) and the tetrameric N-terminal truncated hPAH (Val118−Lys452) (*9*) and using this mosaic structure to create the complete tetrameric human isoform using the programs Swiss-PdbViewer (GlaxoSmith-Kline R&D) and Midas Plus (U. C. San Francisco) (*26*). The modeled structures of the mutants at Arg68 and Cys237 were prepared using the program WHAT IF (*27*). MD simulations were performed with the Discover module of the InsightII suite of programs (Accelrys), using the Amber force field (*28*), on a SGI Octane R10000 workstation. The simulated system comprised a dimeric form of the modeled full-length hPAH (19−452), consisting of 868 amino acids and 13 988 atoms. Although the cooperative unit of PAH is the dimer of dimers, which forms the tetrameric structure, MD simulations on the tetramer are unattainable so far because of the

large amount of atoms in this protein system. Simulations on the dimeric enzyme might be justified on the grounds that the isolated wt-dimer and dimeric truncated forms lacking the tetramerization helix dimer such as hPAH (Ser2−Arg428) do not show cooperativity but can, however, be activated by high concentrations (>5 mM) of L-Phe (*15*, *29*), indicating that certain conformational changes associated with activation by L-Phe also take place at the monomeric or dimeric levels in the absence of the dimer−dimer interactions (*15*, *19*). During energy minimization and MD simulations, part of the protein was kept rigid while residues 19−109, and all the amino acids having an atom within a radius of 10 Å from the residue 237, were allowed to move, giving a total of 1736 unrestrained atoms. Because of the large size of the protein system to be simulated, solvent water molecules were not included in the calculations, and a constant dielectric constant of 4 and a distance-dependent dielectric function were used. At these conditions, the rigid part of the macromolecule provides a realistic electrostatic environment for the part that moves, and this procedure has been found to satisfactorily reproduce the X-ray crystal structures of soluble proteins without including water molecules (*30*), and to maintain stability during MD simulations, notably for dimeric systems (*31*). The cutoff radius for nonbonded interactions was 12 Å, with a secondary cutoff radius of 15 Å. The molecular system was allowed to equilibrate at 300 K for the first 20 ps and was kept at this temperature for the rest of the simulation. The simulations were performed for 500 ps with a step length of 0.5 fs. Simulation of such a system on our SGI Octane single processor workstation took about 52 CPU hours for each simulation. Convergence criteria for molecular mechanical energy minimization of the protein were a maximum for the norm of the steepest descent energy gradient between successive steps of 1 kcal/mol. This gave a root-mean-square difference for the energy gradient of less than 0.01 kcal/mol for all the calculations. Coordinates were saved every 1 ps and used to calculate average structures at different time spans during simulations. The average structures were energy minimized until convergence. The calculation of solvent accessible surface of Trp120 in the different structures was performed using the program WHAT IF (*27*).

RESULTS

Kinetic Properties of the hPAH Mutants at Arg68 and Cys237. PAH is activated by preincubation with L-Phe, but while this activation results in a 20−30-fold increase of the BH_4-dependent activity of the rat enzyme, only a 2−3-fold activation has been measured for the human enzyme, which shows a higher level of basal activity prior to activation (*5*, *32*, *33*). This difference is probably related to the physiological responsiveness to blood concentration of L-Phe in different species. The apparent affinity for L-Phe ($S_{0.5}$) and the Hill coefficient for cooperativity ($h \approx 2$), which most probably reflects the cooperative function of the enzyme as a dimer of dimers, are nevertheless similar for rat and human PAH. The saturation curves for L-Phe and the steady-state kinetic parameters for the purified tetrameric forms of the mutants R68V-hPAH, R68A-hPAH, C237A-hPAH, C237R-hPAH, and C237D-hPAH (Figure 1 and Table 2) indicate that the mutations mainly affect the specific activity, the affinity for L-Phe, and the *h* for L-Phe binding. The affinity

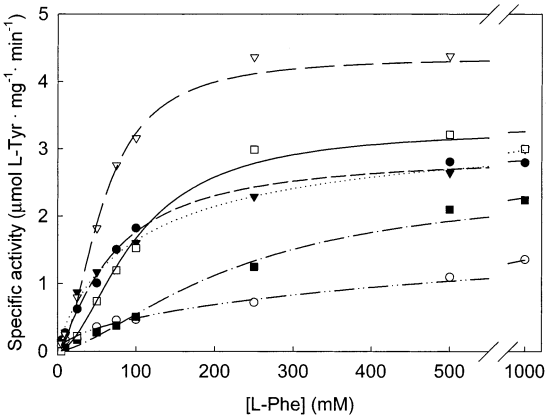

FIGURE 1: Effect of L-Phe concentration on the catalytic activity of wild-type and mutant forms of hPAH. Wild-type hPAH (□), R68V-hPAH (■), R68A-hPAH (●), C237A-hPAH (▼), C237R-hPAH (○), and C237D-hPAH (▽). The activity was measured at standard assay conditions (75 μM BH_4, 1 min assay, 25 °C) with 5 min preincubation of the enzymes at the corresponding concentrations of L-Phe at 25 °C.

for the cofactor BH_4 was only affected (2.5-fold increase) in C237R-hPAH (Table 2). The mutant R68V-hPAH showed a lower specific activity and lower affinity for L-Phe than wt-hPAH but was similarly activated by L-Phe with the same degree of positive cooperativity ($h = 1.9$) (Table 2). The other mutant at Arg68 (i.e., R68A-hPAH) was characterized by a high basal activity (in the absence of preincubation with L-Phe) and a high affinity for L-Phe. As a result, the activating effect of L-Phe (activation-fold) was lower for R68A-hPAH than for wt-hPAH, and the mutant also showed a reduced positive cooperativity for L-Phe ($h = 1.6$). The same properties characterized C237A-hPAH, although in this mutant the positive cooperativity for L-Phe was almost completely abolished (Figure 1 and Table 2). The insertion of a bulky positively charged residue at the same position (i.e., C237R-hPAH) resulted in a form of low basal activity that was not further activated by L-Phe and did not show positive cooperativity for the substrate. On the other hand, the incorporation of a negatively charged residue at position 237 results in an enzyme with very high basal activity and affinity for L-Phe, corresponding to an L-Phe activated form of the enzyme (*22*) (Figure 1, Table 2). C237R-hPAH thus seems to be a representative of the low-affinity, low-activity form of the enzyme (the tense T-form), while R68A-hPAH, C237A-hPAH, and C237D-hPAH only resemble an activated wt-hPAH (a relaxed R-form) since they still display some degree of positive cooperativity.

CD. As earlier found with the enzyme isolated from rat liver (*34*) and the recombinant rat enzyme (*35*), the far-UV CD spectrum of wt-hPAH shows two minima at 208 and 222 nm (Figure 2A), characteristic of proteins with high α-helical content. The amount of α-helical structure for wt-hPAH estimated from the CD spectrum (185−260 nm) was 30.1 ± 1.5%, a value that is in agreement with that obtained earlier by CD for rat PAH (*35*) and with the amount of secondary structure estimated from the crystal structure (*10*). The near-UV CD spectrum of hPAH is almost featureless and does not change upon L-Phe activation of the enzyme (data not shown). On the other hand, as previously reported by Kappock et al. (*35*) for the rat enzyme, an increase of the negative ellipticity in the far-UV region was observed

Table 2: Kinetic Properties of the wt-hPAH and the Mutant Forms at Residues 68 and 237

enzyme	specific activity[a] (μmol Tyr min^{-1} mg^{-1}) non L-Phe preincubated	specific activity[a] (μmol Tyr min^{-1} mg^{-1}) L-Phe preincubated	activation (fold)	$S_{0.5}$ (L-Phe) (μM)	K_m (BH_4) (μM)	*h* (L-Phe)
wt-hPAH	1.10 ± 0.01	3.25 ± 0.06	2.9	154 ± 6	35 ± 5	2.0 ± 0.2
R68V-hPAH	0.77 ± 0.05	2.14 ± 0.02	2.8	233 ± 41	37 ± 4	1.9 ± 0.4
R68A-hPAH	2.20 ± 0.03	2.94 ± 0.09	1.3	91 ± 10	37 ± 5	1.6 ± 0.2
C237A-hPAH	2.22 ± 0.20	2.97 ± 0.09	1.3	100 ± 12	34 ± 6	1.2 ± 0.3
C237R-hPAH	1.13 ± 0.03	1.08 ± 0.10	0.96	214 ± 39	14 ± 2	0.9 ± 0.1
C237D-hPAH[b]	4.19 ± 0.10	3.81 ± 0.09	0.94	91 ± 20	34 ± 3	1.7 ± 0.2

[a] Measured at standard conditions, 1 mM L-Phe and 75 μM BH_4. [b] From ref *22*.

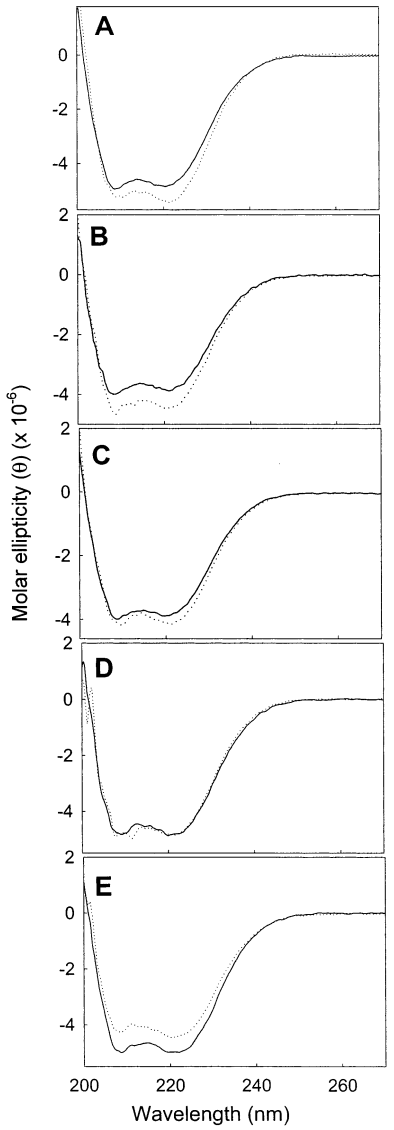

FIGURE 2: Effect of mutations at Arg68 and Cys237 on the far-UV CD spectra of hPAH. Spectra for the isolated enzymes in the absence (continuous line) and the presence (dotted line) of 1 mM L-Phe, taken in 20 mM sodium phosphate and 0.15 M KF, pH 7.0, 25 °C. Samples: wt-hPAH (A), R68V-hPAH (B), C237A-hPAH (C), C237D-hPAH (D), and C237R-hPAH (E). The concentration of hPAH was 5 μM subunit. [θ], molar ellipticity.

in the presence of L-Phe (Figure 2A). The high absorbance of L-Phe (added at 1 mM concentration) in the far UV hampered the acquisition of spectra at $\lambda < 200$ nm, which is required for accurate measurements of secondary structure content. From the change in ellipticity between 200 and 260 nm, we tentatively estimated a 3−4% increase in the apparent α-helical content of the L-Phe activated enzyme. The mutants at Arg68 and Cys237 revealed comparable far-UV CD spectra to that of wt-hPAH, and R68V-hPAH (Figure 2B) also showed a weak but significant increase in the negative ellipticity in the 205−230 nm range in the presence of L-Phe. The spectra of R68A-hPAH (data not shown), C237A-hPAH (Figure 2C), and C237D-hPAH (Figure 2D) were not significantly affected by the binding of L-Phe, while the substrate induces a significant reduction in the negative ellipticity of C237R-hPAH (Figure 2E).

Fluorescence Emission Spectra. hPAH contains three conserved tryptophan residues (120, 187, and 326). Activation by L-Phe has been shown to result in an increase of the intrinsic tryptophan fluorescence and a red shift of the emission maximun of both the rat (*36*) and the human PAH (*37*). L-Trp to L-Phe scanning mutagenesis has shown that Trp120 is largely responsible for the fluorescence properties of the enzyme (i.e., this residue accounts for about 61% of the total tryptophan fluorescence of hPAH). Also, the increase in fluorescence intensity and in emission maximum that accompanies activation of the enzyme by L-Phe have been found to be mainly due to changes in the emission of Trp120, which is supposed to exchange to a more polar environment, most probably the solvent (*37*). This residue is located close to a hinge region (Arg111−Thr117) between the catalytic and the regulatory domains from the same subunit (see below). In the absence of added L-Phe, the tryptophan fluorescence emission spectrum of the tetrameric wt-hPAH shows a maximum at about 337 nm (excitation at 295 nm), and on incubation with 1 mM L-Phe, a 10-nm red-shift and an increase in the quantum yield is observed (Figure 3A). The emission spectrum of R68V-hPAH was very similar to that of wt-hPAH, and incubation of the enzyme with L-Phe also resulted in a red shift of the emission maximum (Figure 3B). The emission maxima of the fluorescence spectra of R68A-hPAH (Figure 3C), C237A-hPAH (Figure 3D), and C237R-hPAH (Figure 3E) resembled that of the L-Phe activated wt-hPAH ($\lambda_{em\ max} = 347$ nm) and did not significantly change upon addition of the substrate. A decrease in the fluorescence intensity was observed in the emission spectrum of C237R-hPAH upon binding of L-Phe (Figure 3E). We have earlier reported that the mutant C237D-hPAH shows a red-shifted intrinsic tryptophan fluorescence spectrum that is not affected by preincubation with L-Phe (*22*).

Modeling and MD Simulations. Figure 4A shows a structural model of full-length hPAH in a dimeric form (*38*), including the position of Arg68, Cys237, and Trp120. The hinge regions 31−34, 111−117, 425−429, and the loop 68−

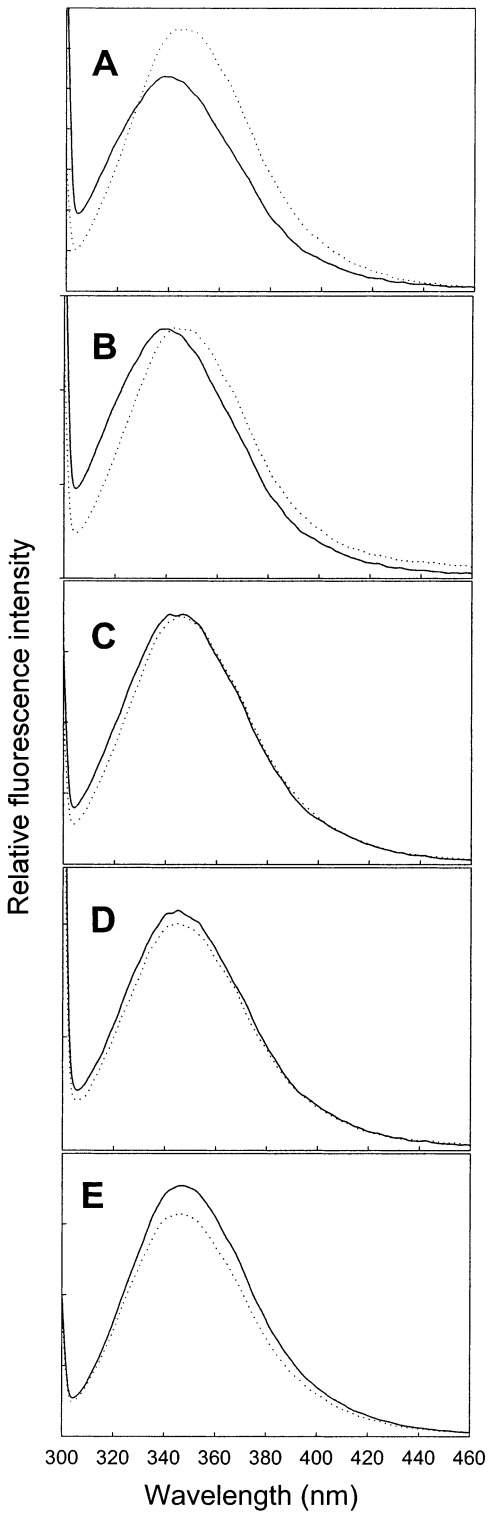

FIGURE 3: Fluorescence emission spectra of wild-type and mutant forms of hPAH. The enzyme samples (i.e., wt-hPAH (A), R68V-hPAH (B), R68A-hPAH (C), C237A-hPAH (D), and C237R-hPAH (E)) were diluted to a final concentration of about 60 μg/mL in 20 mM Na-Hepes, 0.2 M NaCl, pH 7.0. Fluorescence spectra were acquired in the absence (continuous line) and presence (dotted line) of 1 mM L-Phe. The excitation wavelength was 295 nm.

75 are also highlighted. In each subunit, the oligomerization domain establishes contacts with both the N-terminal and the catalytic domains. To obtain further understanding on the conformational changes effected in the mutant forms included in this study, the energy-minimized modeled structures of the selected enzyme forms wt-hPAH, C237A-hPAH, C237R-hPAH, and C237D-hPAH were subjected to 500-ps MD simulations. MD simulations were also performed on the wt-hPAH complexed with L-Phe at the active site according to the position of the substrate as determined by ^{1}H NMR and molecular modeling (*11*). No large deviations from the initial structure were measured through the MD simulations (i.e., the rmsd of the trajectory was $<$2.4 Å for all flexible Cα atoms for the four enzyme forms studied). However, some significant differences can be noticed in the structure of the mutants as compared with the wild-type form short after the start of the MD simulations. As seen in Figure 4B,C, showing the averaged structures at the time span 200−300 ps for the wt-hPAH and the representative mutant C237D-hPAH, changes were observed in the conformation of the core of the regulatory domain, especially in the loop 68−75, and in the N-terminal autoregulatory sequence (residues 19−33). This sequence stretches along the entrance of the active site in the wt-hPAH and is displaced from this position in the mutant C237D-hPAH providing an open active site. This movement of the autoregulatory sequence seems to be significant since this sequence acts as an internal inhibitor of the enzyme (*19*, *20*). As shown in Figure 5A, this displacement can be followed by measuring the distance between the Cα atom of Ser23 at the autoregulatory sequence and the Cα atom of Glu330, this being one of the coordinating residues to the catalytic non-heme iron at the active site. A displacement of the autoregulatory sequence can also be observed for wt-hPAH complexed with L-Phe and for the mutant C237A-hPAH (Figure 5A), although in this latter case the displacement occurs sporadically during the 500-ps simulation. After 420 ps of MD simulation, the initial closed conformation is recovered for C237D-hPAH but not for the wt-hPAH complexed with L-Phe (Figure 5A). It seems that C237A-hPAH and C237D-hPAH fluctuates between an open and a closed conformation of the active site. The unbound wt-hPAH and the C237R-hPAH mutant continued in a closed conformation during MD simulations, and all the forms studied show a rather stable structure (Figure 5A,B).

A detailed comparison of the MD average structures of wt-hPAH and the mutants at the atomic level is beyond justification when only short simulations have been performed in this work, but still it was interesting to note the specific orientation of the side chain of Arg68 depending on the mutation at 237 in the adjacent subunit and on the presence of L-Phe (Figure 6). An electrostatic favorable interaction between Asp237 and Arg68 can be foreseen for the activated mutant form C237D-hPAH (the distance between the O$^{\delta 1}$ in the carboxyl group of Asp237 and the N$^{\eta 1}$ of Arg68 was 3.0 Å in the averaged 200−300 ps structure). The opposite was observed for the residues at these positions in the C237R-hPAH mutant (the distance between the N$^{\eta 1}$ atoms in the adjacent arginines was 7.0 Å). These structural changes only slightly affect the backbone position at residues 68 and 237 for all enzyme forms studied, and a clear trend is only observed for wt-hPAH complexed with L-Phe in which a 1 Å approach between the Cα atoms of these residues can be observed after 300 ps simulation (Figure 5B). More noticeable effects are observed on the conformation of the loop Arg68−Asp75, including the interactions of Arg71 with residues from the dimerization

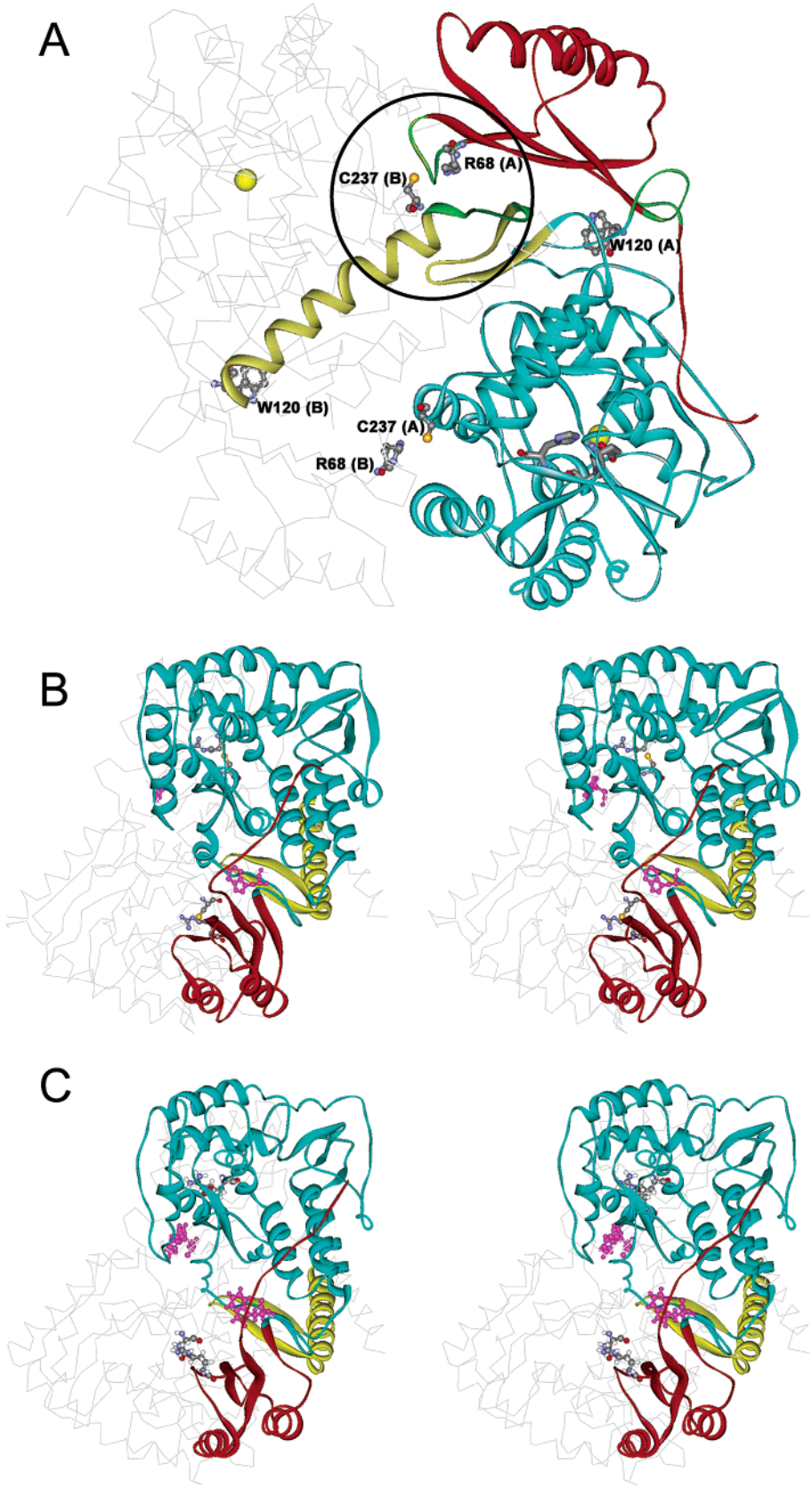

FIGURE 4: The modeled structure of the full-length hPAH subunit and results from MD simulations for wt-hPAH and C237D-hPAH. The N-terminal (red), catalytic (blue), and oligomerization (yellow) domains of subunit A are shown in ribbon representation. Subunit B is shown as a backbone wire. (A) Interdomain contacts in each subunit: the oligomerization domain establishes contacts with both the N-terminal domain and the catalytic domain. The region around residues 68 and 237 from adjacent subunits in the dimer is surrounded by a circle. The catalytic Fe is shown in yellow, and the coordinating residues His285, His290, and Glu330 are shown as sticks in subunit A. The residue exposed on activation (i.e., Trp120) and the residues mutated in this study (i.e., Arg68 and Cys237) are shown in ball-and-sticks representation. The hinge regions (see text) 31−34, 111−117, 425−429, and the loop 68−75 are shown in light green. (B and C) Averaged conformation (from 200 to 300 ps) from the MD simulations for wt-hPAH (B) and C237D-hPAH (C). Arg68, Cys237, and Trp120 (pink) are shown.

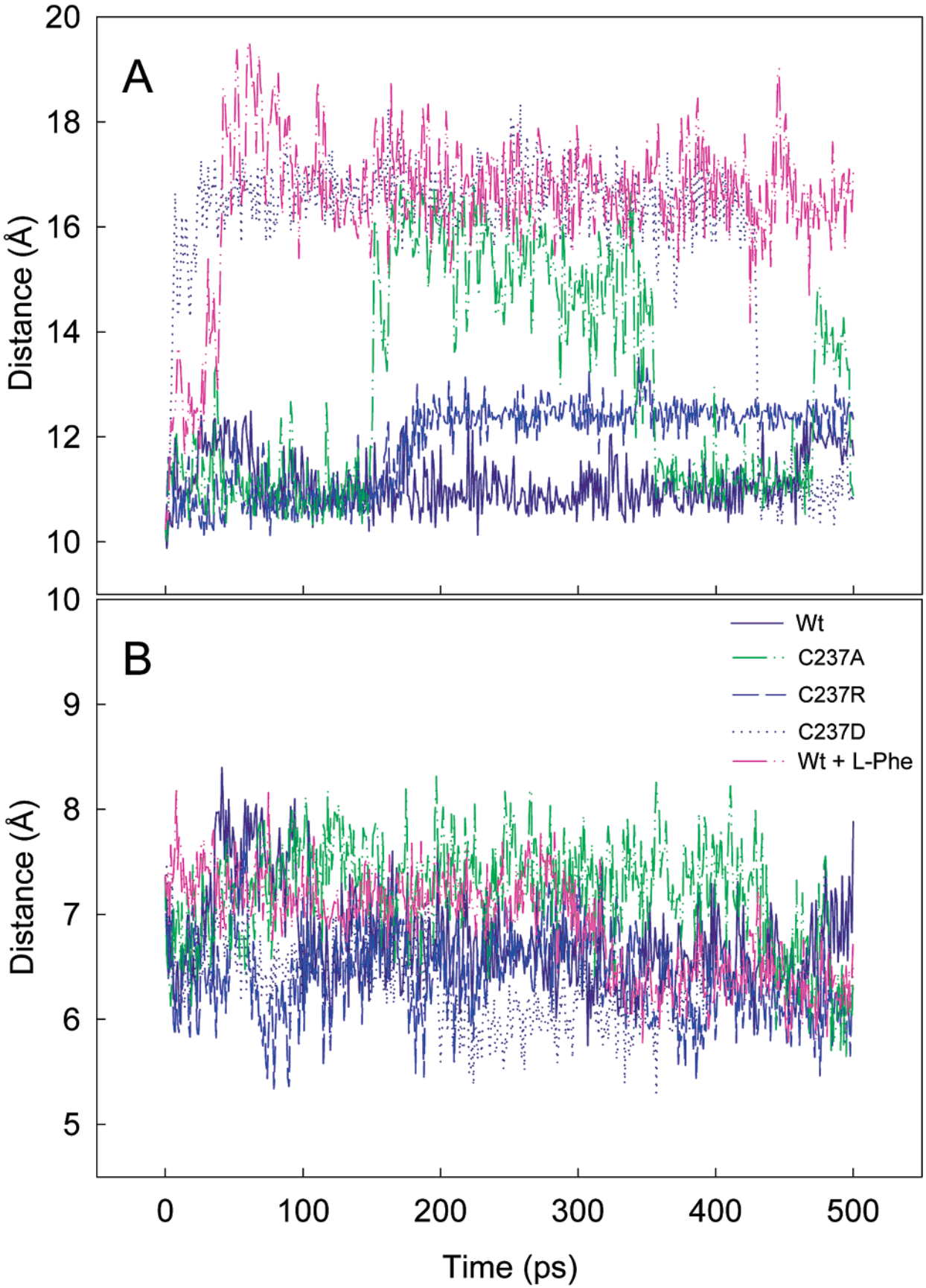

FIGURE 5: Comparison of the distances between the Cα atoms of Ser23 and Glu330 (A) and between the Cα atoms of Arg68 and the residue at position 237 (B) during MD simulations. The distances were measured in each snapshot taken every 2 ps for the wt-hPAH (—), C237D-hPAH (····), C237R-hPAH (- -), C237A-hPAH (-··-), and for wt-hPAH complexed with L-Phe at the active site (—··).

motif (Figure 6). Arg71 establishes ionic interactions with both Glu422 and Asp415, and these interactions are strengthened in the averaged 200−300 ps structure (corresponding to wt-hPAH and C237R-hPAH in the closed active site

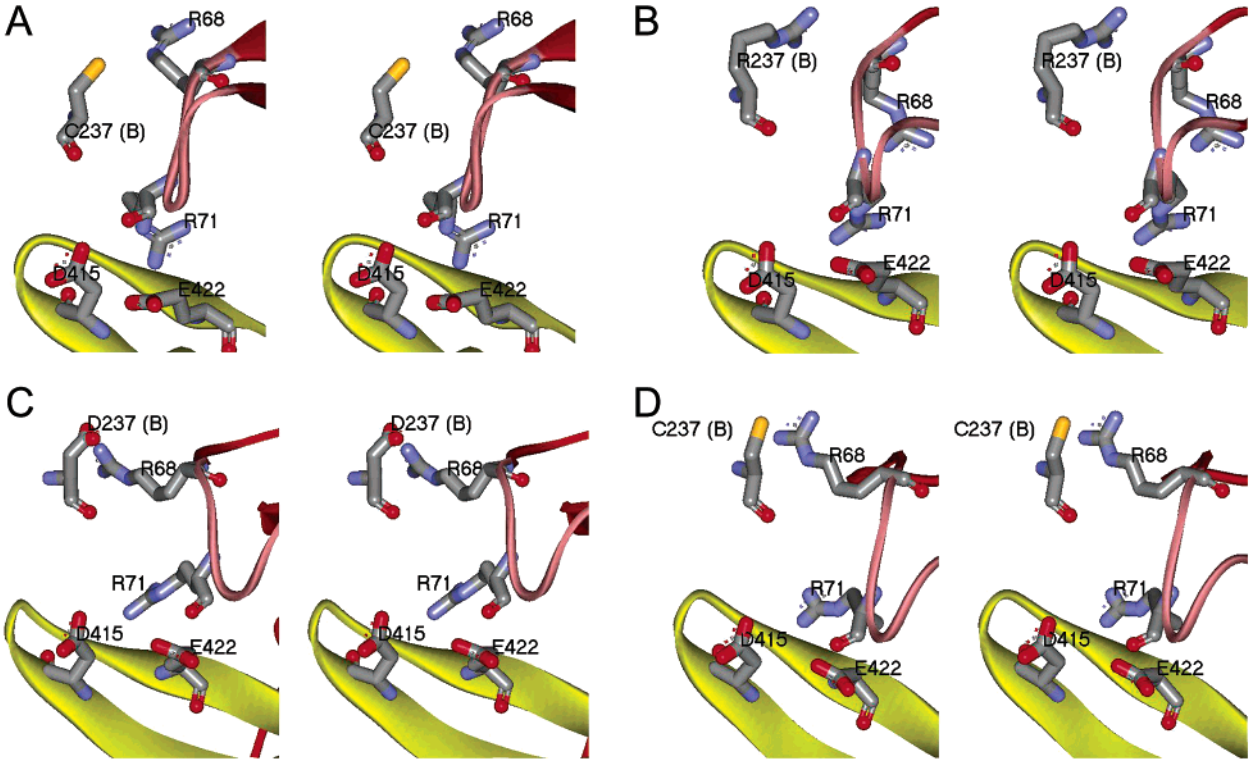

FIGURE 6: Close view showing the loop 68−75 and the two antiparallel β-strands forming the dimerization motif (residues 411−425) in the averaged conformation (200−300 ps) from the MD simulations. The loop is shown in light red and the dimerization motif in yellow for wt-hPAH (A), C237R-hPAH (B), C237D-hPAH (C), and wt-hPAH complexed with L-Phe at the active site (D). Selected residues are shown as sticks.

Table 3: Solvent Accessibility[a] of the Side Chain of Trp120 in the 200−300 ps Averaged Structure during Formation of the MD Simulations

enzyme	Å^2
wt-hPAH	0.7
Wt-hPAH + L-Phe	2.3
C237A-hPAH	1.7
C237R-hPAH	1.5
C237D-hPAH	3.0

[a] Measured as the accessible surface area, which is defined by all positions at the van der Waals surface of all atoms in the side chain where the center of a water molecule can be found.

conformation and to C237A-hPAH, C237D-hPAH, and wt-hPAH•L-Phe in the open conformation) of the activated forms but not in the other forms (Figure 6). Thus, the distances between $N^{\eta 2}$(Arg71) and $O^{\epsilon 1}$(Glu422) and $O^{\delta 2}$(Asp415) in wt-hPAH decreased from 3.9 and 3.8 Å, respectively, to about 3.4 and 3.1−3.2 Å, respectively, in the activated structures. The distance between $N^{\eta 2}$(Arg71) and $O^{\epsilon 1}$(Glu422) even increased to 4.4 Å in C237R-hPAH.

In addition to the changes in the rearrangement of the N-terminal autoregulatory sequence and the loop/dimerization domain, we also investigated the environment of Trp120. Table 3 summarizes the solvent accessibility of the side chain of Trp120 during MD simulations in the average structure at the time span of 200−300 ps. It is interesting to note that the activated enzyme forms, exhibiting open-active site conformations in the course of the MD simulations (Figure 5A), show a red-shifted fluorescence spectra and a more exposed Trp120 (Figure 3 and Table 3). However, the nonactivatable C237R-hPAH is also characterized by a similar fluorescence spectrum (Figure 3E), although the MD structures indicated a more limited exposure of Trp120 in this mutant (Table 3).

DISCUSSION

The cooperative activation of PAH by L-Phe has been interpreted to result from binding of the substrate either to (i) a putative allosteric/regulatory site, which is distinct from the catalytic site (*29*, *39*−*41*) or (ii) to the active site near the iron (*42*, *43*). A recent differential scanning calorimetry study on the binding of L-Phe to hPAH has demonstrated that there are no binding sites for L-Phe in the regulatory domain, indicating that the activation of the enzyme by L-Phe is the result of its homotropic cooperative binding to the active sites (*38*). The regulatory domain appears to be involved in cooperativity through its interactions with the catalytic and oligomerization domains via the prominent loop Arg68−Asp75, which interacts with both the catalytic domain in the opposite subunit and the oligomerization domain of the same subunit. Positive cooperativity for the binding of L-Phe is only manifested by the full-length tetrameric enzyme and not by dimeric or truncated forms (*5*, *15*), although dimeric forms as the isolated wt-dimer and the truncated hPAH (Ser2−Arg428) also can be activated by high concentrations (>5 mM) of L-Phe, and their intrinsic fluorescence spectrum is similarly affected by activation as the full-length tetrameric forms (*15*, *19*, *29*).

The subunits are organized in the tetramer as a dimer of dimers, and dimer−dimer contact only takes place through the tetramerization motifs of the four subunits that are arranged in a bundle of α-helices in antiparallel orientation (*9*). The tetramerization motif is thus an important region for the transfer of cooperative conformational changes.

Accordingly, the mutant T427P, which most probably results in a different conformation and/or orientation of the C-terminal helix, had decreased affinity and enzymatic activity and lacked positive cooperativity for L-Phe, proving the contribution of the flexible hinge region Asp425−Gln429 at the start of the tetramerization motif to the cooperative substrate binding (*44*). This hinge region establishes several contacts with the loop (residues 68−75) of the same subunit, and together with some residues around Cys237 from the adjacent subunit, it might constitute a structural control unit important in the modulation of intra- and intersubunit domain−domain interactions and for the transmission of activating conformational changes (circled region in Figure 4A). Hence, chemical modification of Cys237 by treatment with *N*-ethylmaleimide or DTNB (*13*, *21*) or the substitution of this cysteine with a negatively charged residue by site-directed mutagenesis (*22*) results in activation of the enzyme. Arg68 is also located at this interdomain contacting unit, and as seen in this work, certain mutations at these positions have a large effect on the activation state of the enzyme. As earlier found with C237D-hPAH (*22*), both mutants R68A-hPAH and C237A-hPAH resemble an R-form in that they show high basal activity and affinity for L-Phe. On the other hand, C237R-hPAH shows lower affinity for L-Phe than wt-hPAH and cannot be activated by L-Phe; this form is thus a good representative for a T-form of the enzyme. The mutant R68V-hPAH, while still maintaining the ability to be activated by L-Phe, shows lower affinity for L-Phe than wt-hPAH.

MD simulations are important tools for understanding internal motions in biological macromolecules. When they are used to reveal conformational changes as a function of time (i.e., the actual dynamics), long simulations (up to 10 ns) are required (*45*). Short simulations of the type applied in this study are typically used to refine structures and sample part of the configuration space for the macromolecule, still providing important structural and functional information. Our results from MD simulations show that wt-hPAH is kept in a closed conformation and that the binding of L-Phe at the active site induces the open site conformation (Figure 5A). In agreement, the wild-type form requires preincubation with L-Phe to be activated. The mutants C237D-hPAH and C237A-hPAH seem to switch between conformations with open and closed active sites. This might indicate that both conformations coexist in equilibrium in purified preparations of these mutants, in agreement with the finding that these mutants still show some degree of positive cooperativity (Table 2). These mutants show a fluorescence emission in the absence of L-Phe characteristic of a solvent exposed Trp120. This residue is located at the end of a hinge region (111−117) between the catalytic and the regulatory domains and is covered by the start of the N-terminal autoregulatory sequence in the unactivated wt-hPAH. MD simulations in fact show evidence of a higher degree of solvent exposure of Trp120 for the activated forms (Table 3), in agreement with the movement of the domains around the hinge region 111−117 and the rearrangement of the autoregulatory sequence.

Ionic types of interactions seem to be important when analyzing the activating effect of the mutation C237D-hPAH, in which the MD simulated structures show the formation of a salt bridge between Asp237 and Arg68 from adjacent subunits (Figure 6C). In the wild type form, the formation of a hydrogen bonding interaction between residues Cys237 and Arg68 seems to be important, but not essential, in activation of the enzyme. Accordingly, when a better hydrogen bond acceptor as a serine residue is placed at position 237 instead of the cysteine, the resulting mutant form of hPAH is partially activated (*22*), while the R68V-hPAH mutant shows a lower affinity for L-Phe and a lower degree of basal activity (Table 2). Also, a packing effect because of intersubunit area conservation could be the cause for the reorientation of the side chain at Arg68 in the C237A-hPAH mutant (data not shown), which could partially mimic the activating effect of the R237D mutation. An ionic facilitated activation is also supported by the fact that the mutant C237R-hPAH is not activatable.

The most remarkable conformational effects induced by the activating mutations and the binding of L-Phe at the active site seem to be rearrangements at the loop 68−75, notably involving Arg71, which appears to be an important residue in the structural control unit (see above). The reinforcement of the electrostatic interactions between Arg71 and both Asp415 and 422 at the dimerization motif observed in the activated forms indicates that activation of PAH may affect the hydrogen bonding network in all residues at the structural control unit, and finally, alter the orientation of the oligomerization motif. Our results also indicate that the mutations at residue 237 (and the mutation at Arg68 in R68A-hPAH) by affecting the topology of the loop Arg68−Asp75 also affect the orientation of the autoregulatory sequence and the domain orientation around the hinge Arg111−Thr117, and ultimately, the solvent exposure of Trp120. A similar series of events can be envisioned as being triggered by L-Phe binding to the active site, proceeding as follows: (i) the displacement of the autoregulatory sequence and domain movements around the hinge region 111−117; (ii) changes at the loop 68−75 and neighboring residues at the adjacent subunit in the dimer; and (iii) changes in the orientation of the oligomerization domain as a way to transfer the cooperative activating conformational change from one dimer to the other dimer in the tetramer. Events i and ii might take place both in tetrameric and in dimeric forms of the enzyme, the latter being activated by high concentrations of L-Phe (*15*, *29*), while event iii, which would be directly related to cooperative activation, would require the participation of the tetrameric form and the switch between the R- and T-form quaternary structures. An indication of the activation-induced reorganization of the bundle of α-helices in the center of the tetramer might be the increase of negative ellipticity of the far-UV CD spectrum of the wild-type enzyme observed in the presence of L-Phe (ref *35* and this paper). This series of events fit well with the proposal that allosteric activation of multimeric proteins most probably occur through a combination of sequential conformational changes according to the KNF model and a concerted change of quaternary structure as foreseen in the MWC model (*3*).

A full understanding of the activation process for PAH must await the until now unattainable task of solving high-resolution structures of the enzyme in the resting and activated states. Nevertheless, the combination of selected site-directed mutagenesis, kinetic and spectroscopic investigations, and MD simulations, even when short timed, have

provided valuable insights on the activation of this allosteric enzyme.

ACKNOWLEDGMENT

We thank Randi M. Svebak for expert technical assistance, Per M. Knappskog for help in the preparation of bacterial strains expressing the mutants of human phenylalanine hydroxylase, and Rebecca Wade for valuable comments on the MD simulations.

REFERENCES

1. Perutz, M. F. (1989) *Q. Rev. Biophys. 22*, 139−237.
2. Koudelka, G. B. (2000) *Curr. Biol. 10*, R704−707.
3. Koshland, D. E., Jr., and Hamadani, K. (2002) *J. Biol. Chem. 277*, 46841−46844.
4. Kantrowitz, E. R., and Lipscomb, W. N. (1990) *Trends Biochem. Sci. 15*, 53−59.
5. Kaufman, S. (1993) *Adv. Enzymol. Relat. Areas Mol. Biol. 67*, 77−264.
6. Kappock, T. J., and Caradonna, J. P. (1996) *Chem. Rev. 96*, 2659−2756.
7. Flatmark, T., and Stevens, R. C. (1999) *Chem. Rev. 99*, 2137−2160.
8. Erlandsen, H., Fusetti, F., Martínez, A., Hough, E., Flatmark, T., and Stevens, R. C. (1997) *Nat. Struct. Biol. 4*, 995−1000.
9. Fusetti, F., Erlandsen, H., Flatmark, T., and Stevens, R. C. (1998) *J. Biol. Chem. 273*, 16962−16967.
10. Kobe, B., Jennings, I. G., House, C. M., Michell, B. J., Goodwill, K. E., Santarsiero, B. D., Stevens, R. C., Cotton, R. G., and Kemp, B. E. (1999) *Nat. Struct. Biol. 6*, 442−448.
11. Teigen, K., Frøystein, N. Å., and Martínez, A. (1999) *J. Mol. Biol. 294*, 807−823.
12. Andersen, O. A., Flatmark, T., and Hough, E. (2002) *J. Mol. Biol. 320*, 1095−1108.
13. Parniak, M. A., and Kaufman, S. (1981) *J. Biol. Chem. 256*, 6876−6882.
14. Citron, B. A., Davis, M. D., and Kaufman, S. (1992) *Protein Expr. Purif. 3*, 93−100.
15. Knappskog, P. M., Flatmark, T., Aarden, J. M., Haavik, J., and Martínez, A. (1996) *Eur. J. Biochem. 242*, 813−821.
16. Kaufman, S., Citron, B. A., Davis, M., and Milstien, S. (1993) *Adv. Exp. Med. Biol. 338*, 97−102.
17. Monod, J., Wyman, J., and Changeux, J. P. (1965) *J. Mol. Biol. 12*, 88−118.
18. Koshland, D. E., Jr., Nemethy, G., and Filmer, D. (1966) *Biochemistry 5*, 365−385.
19. Jennings, I. G., Teh, T., and Kobe, B. (2001) *FEBS Lett. 488*, 196−200.
20. Wang, G. A., Gu, P., and Kaufman, S. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 1537−1542.
21. Gibbs, B. S., and Benkovic, S. J. (1991) *Biochemistry 30*, 6795−6802.
22. Knappskog, P. M., and Martinez, A. (1997) *FEBS Lett. 409*, 7−11.
23. Martínez, A., Knappskog, P. M., Olafsdottir, S., Døskeland, A. P., Eiken, H. G., Svebak, R. M., Bozzini, M., Apold, J., and Flatmark, T. (1995) *Biochem. J. 306*, 589−597.
24. Miranda, F. F., Teigen, K., Thorolfsson, M., Svebak, R. M., Knappskog, P. M., Flatmark, T., and Martinez, A. (2002) *J. Biol. Chem. 277*, 40937−40943.
25. Bohm, G., Muhr, R., and Jaenicke, R. (1992) *Protein Eng. 5*, 191−195.
26. Ferrin, T. E., Hunag, C. C., Jarvis, L. E., and Langridge, R. (1988) *J. Mol. Graphics 6*, 13−27.
27. Vriend, G. (1990) *J. Mol. Graphics 8*, 52−56.
28. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) *J. Am. Chem. Soc. 117*, 5179−5197.
29. Parniak, M. A. (1987) in *Unconjugated pterins and related biogenic amines* (Curtius, H.-C., and Blau, N., Eds.) pp 327−337, Walter de Gruyter & Co., Berlin.
30. Christensen, I. T., and Jorgensen, F. S. (1997) *J. Biomol. Struct. Dyn. 15*, 473−488.
31. Kumar, A., Shankar, S., and Kothekar, V. (2001) *J. Biomol. Struct. Dyn. 19*, 449−458.
32. Døskeland, A. P., Martínez, A., Knappskog, P. M., and Flatmark, T. (1996) *Biochem. J. 313*, 409−414.
33. Olafsdottir, S., and Martínez, A. (1999) *J. Biol. Chem. 274*, 6280−6284.
34. Abita, J. P., Parniak, M., and Kaufman, S. (1984) *J. Biol. Chem. 259*, 14560−14566.
35. Kappock, T. J., Harkins, P. C., Friedenberg, S., and Caradonna, J. P. (1995) *J. Biol. Chem. 270*, 30532−30544.
36. Phillips, R. S., Parniak, M. A., and Kaufman, S. (1984) *Biochemistry 23*, 3836−3842.
37. Knappskog, P. M., and Haavik, J. (1995) *Biochemistry 34*, 11790−11799.
38. Thórólfsson, M., Ibarra-Molero, B., Fojan, P., Petersen, S. B., Sanchez-Ruiz, J. M., and Martínez, A. (2002) *Biochemistry 41*, 7573−7585.
39. Tourian, A. (1971) *Biochim. Biophys. Acta 242*, 345−354.
40. Shiman, R., Xia, T., Hill, M. A., and Gray, D. W. (1994) *J. Biol. Chem. 269*, 24647−24656.
41. Gjetting, T., Petersen, M., Guldberg, P., and Guttler, F. (2001) *Am. J. Hum. Genet. 68*, 1353−1360.
42. Martínez, A., Haavik, J., and Flatmark, T. (1990) *Eur. J. Biochem. 193*, 211−219.
43. Martínez, A., Olafsdottir, S., and Flatmark, T. (1993) *Eur. J. Biochem. 211*, 259−266.
44. Bjørgo, E., de Carvalho, R. M., and Flatmark, T. (2001) *Eur. J. Biochem. 268*, 997−1005.
45. Karplus, M., and McCammon, J. A. (2002) *Nat. Struct. Biol. 9*, 646−652.

BI034021S