

# SCIENTIFIC REPORTS



OPEN

## Structural basis for ligand-dependent dimerization of phenylalanine hydroxylase regulatory domain

Received: 08 October 2015

Accepted: 08 March 2016

Published: 06 April 2016

Dipali Patel<sup>\*</sup>, Jolanta Kopec, Fiona Fitzpatrick<sup>†</sup>, Thomas J. McCorvie & Wyatt W. Yue

The multi-domain enzyme phenylalanine hydroxylase (PAH) catalyzes the hydroxylation of dietary L-phenylalanine (Phe) to L-tyrosine. Inherited mutations that result in PAH enzyme deficiency are the genetic cause of the autosomal recessive disorder phenylketonuria. Phe is the substrate for the PAH active site, but also an allosteric ligand that increases enzyme activity. Phe has been proposed to bind, in addition to the catalytic domain, a site at the PAH N-terminal regulatory domain (PAH-RD), to activate the enzyme via an unclear mechanism. Here we report the crystal structure of human PAH-RD bound with Phe at 1.8 Å resolution, revealing a homodimer of ACT folds with Phe bound at the dimer interface. This work delivers the structural evidence to support previous solution studies that a binding site exists in the RD for Phe, and that Phe binding results in dimerization of PAH-RD. Consistent with our structural observation, a disease-associated PAH mutant impaired in Phe binding disrupts the monomer:dimer equilibrium of PAH-RD. Our data therefore support an emerging model of PAH allosteric regulation, whereby Phe binds to PAH-RD and mediates the dimerization of regulatory modules that would bring about conformational changes to activate the enzyme.

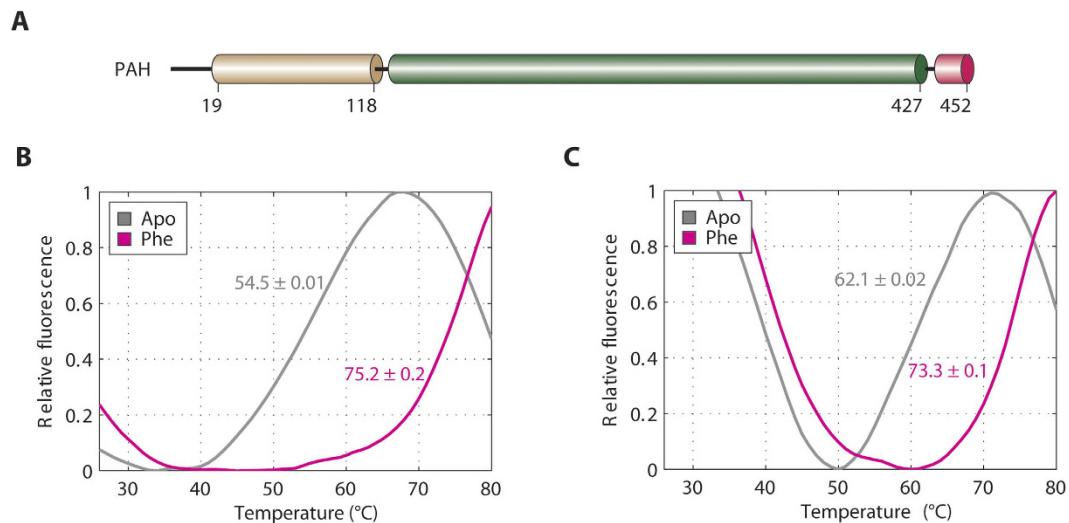
Phenylalanine hydroxylase (PAH; EC 1.14.16.1) catalyzes the first step of L-phenylalanine (Phe) degradation via its hydroxylation to L-tyrosine, and serves to metabolize excess Phe in the diet. Inherited PAH deficiency, caused by mutations on the human *pah* gene, results in the autosomal recessive disorder phenylketonuria (PKU; OMIM 261600), which occurs at an average incidence of 1 in 10,000 live births<sup>1</sup>. PKU leads to mental retardation caused by elevated Phe levels in the blood and brain if it is not treated with a lifelong dietary restriction of Phe. To date more than 800 disease-causing *pah* mutations have been reported (<http://www.biopku.org/pah>)<sup>2</sup>, of which ~60% represent single amino acid, missense changes.

PAH, together with tyrosine hydroxylase (TH, EC 1.14.16.2) and tryptophan hydroxylase (TPH1 and TPH2 isozymes, EC 1.14.16.4), constitute the family of aromatic amino acid hydroxylases (AAAH) that are tetrahydropterin (BH<sub>4</sub>)-dependent, non-haem Fe(II) monooxygenase enzymes<sup>3</sup>. Like other AAAHs, PAH adopts a multi-domain architecture comprised of an N-terminal regulatory domain (RD, ~100 aa), a central catalytic domain (CD, ~300 aa) and a C-terminal multimerization helix (MH, ~40 aa) (Fig. 1A)<sup>4</sup>. Crystal structures of rat PAH lacking the C-terminal MH (RD + CD fragment)<sup>5</sup>, and of human PAH lacking the N-terminal RD (CD + MH fragment)<sup>6</sup>, have provided biochemical insight into the enzyme mechanism. The CD is responsible for the iron-mediated incorporation of one atom of molecular oxygen into the amino acid substrate and the reducing substrate BH<sub>4</sub> to generate the hydroxylated products. The C-terminal MH mediates the formation of homo-tetramers that represent the native functional states<sup>3</sup>.

The role of the PAH-RD is less well defined. The enzymatic activity of PAH is known to be regulated at several levels<sup>7</sup>, including activation by substrate Phe<sup>8</sup>, inhibition by cofactor BH<sub>4</sub>, and also phosphorylation of Ser16<sup>9</sup>. Some, if not all, of these regulatory properties are likely mediated by the N-terminal RD, as deletion of the domain liberates the enzyme from activation and renders it constitutively active<sup>10,11</sup>. The inhibitory nature of the RD

Structural Genomics Consortium, Nuffield Department of Clinical Medicine, University of Oxford, UK OX3 7DQ.

<sup>\*</sup>Present address: Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, UK. <sup>†</sup>Present address: University of Cambridge, MRC Mitochondrial Biology Unit, Wellcome Trust/MRC Building, Hills Road, Cambridge, CB2 0XY, UK. Correspondence and requests for materials should be addressed to W.W.Y. (email: wyatt.yue@sgc.ox.ac.uk)



**Figure 1. Interaction between PAH-RD and Phenylalanine.** (A) Schematic of the regulatory domain (brown), catalytic domain (green) and tetramerization domain (pink) of the human PAH polypeptide. (B) DSF of the unliganded (grey line;  $T_m = 54.5^\circ\text{C} \pm 0.01$  SD) and Phe-bound (pink line;  $T_m = 75.2^\circ\text{C} \pm 0.2$  SD) hPAH-RD<sup>1-118</sup>. (C) DSF of the unliganded (grey line;  $T_m = 62.1^\circ\text{C} \pm 0.02$  SD) and Phe-bound (pink line;  $T_m = 73.3^\circ\text{C} \pm 0.1$  SD) hPAH-RD<sup>19-118</sup>.

towards the CD is supported by the observation from the rat PAH structure of RD + CD that an N-terminal linker region of RD tails over to the catalytic domain and sterically blocks the active-site entrance<sup>5</sup>. While this manuscript was under review, the full-length structure of rat PAH in the absence of Phe has been reported<sup>12</sup>, revealing in molecular details the autoinhibitory nature of the N-terminal region. It has been suggested<sup>13,14</sup> that the binding of Phe to the PAH active site or CD triggers a conformational change to the enzyme that leads to its activation. There is to date no structural data for ligand binding to the PAH-RD. PAH-RD contains an ACT (Aspartate kinase, Chorismate mutase and TyrA) fold commonly found in enzymes for binding small molecule ligands<sup>15</sup>. Indirect evidence for a Phe binding site within PAH-RD has been accumulated by site-directed mutagenesis<sup>16</sup>, gel filtration and HSQC NMR<sup>17</sup>. More recently, binding of Phe to PAH-RD has been linked to its ability to homo-dimerize<sup>18</sup>, and to increase the enzyme activity<sup>19</sup>. In this study, we have determined the high resolution crystal structure of human PAH-RD bound with Phe. Our data provide the structural evidence for a Phe binding pocket at the subunit-subunit interface of a PAH-RD dimer, and demonstrate that PAH-RD dimerization depends on Phe binding. Furthermore, we show that dimerization can be affected by a disease-associated Phe-binding mutation.

## Results

**PAH-RD has a binding site for Phe.** The regulatory domain of PAH (PAH-RD) is located at the N-terminal 118 aa of the polypeptide (Fig. 1A). In the rat PAH structure of RD + CD<sup>5</sup>, aa 1-18 (same residue numbering as hPAH) was disordered, while aa 19-33 harbours an 'autoregulatory motif' for intramolecular active site gating<sup>5</sup>. To evaluate the binding of Phe to PAH-RD, we expressed hPAH-RD<sup>1-118</sup> and hPAH-RD<sup>19-118</sup> recombinantly in *E. coli*, and performed differential scanning fluorimetry (DSF). As-purified hPAH-RD<sup>1-118</sup> exhibits a melting temperature ( $T_m$ ) of  $54.5^\circ\text{C}$ , which was thermally stabilized to  $75.2^\circ\text{C}$  in the presence of Phe, an indicator of ligand binding to the domain (Fig. 1B). hPAH-RD<sup>19-118</sup> also shows a similar response to Phe ( $T_m$  shift from  $62.1^\circ\text{C}$  to  $73.3^\circ\text{C}$ ) (Fig. 1C), suggesting that the N-terminal 18 aa segment is dispensable for Phe binding. Our solution data therefore support the notion of a Phe binding site in the PAH-RD, besides its known binding to the active site in the CD<sup>20</sup>.

**Structure of Phe-bound PAH-RD is a side-by-side homodimer.** We next pursued a structural study of the Phe binding site of hPAH-RD. Following unsuccessful attempts to crystallize hPAH-RD<sup>1-118</sup> and hPAH-RD<sup>19-118</sup>, we performed *in situ* limited proteolysis<sup>21</sup>, with the objective of removing flexible regions to facilitate crystal lattice packing. Trypsin-treated hPAH-RD<sup>19-118</sup> yielded diffraction-quality crystals in the presence of Phe, allowing its structure determination at  $1.8\text{ \AA}$  resolution (Table 1). In the final model, the peptide regions Gly19-Gly33 and Asp112-Val118 are not present, which can be due to intrinsic disorder, or possible trypsinization, of the two regions. The structure of hPAH-RD is a homodimer (Fig. 2A), with two  $\beta\alpha\beta\beta\alpha\beta$  folds (one from each protomer; Fig. 2B) arranged side-by-side to form a continuous 8-stranded  $\beta$ -sheet, resulting in an internal 2-fold symmetry. The dimer interface is largely polar in nature, comprising main-chain/side-chain hydrogen bonds contributed from two amino acid regions Glu43-Leu48 and Asn61-Ser67, as well as from a bound Phe ligand (Supplementary Table 1). The PAH-RD side-by-side dimer closely resembles that of acetohydroxyacid synthase (AHAS) regulatory subunit (DALI Z-score 11.1, RMSD  $1.6\text{ \AA}$ , sequence identity 24%)<sup>22</sup>, but is topologically different from the regulatory domain of TH ( $4.9, 3.4\text{ \AA}$ , 13%)<sup>23</sup>, another member of the AAAH family (Fig. 2C).

Data Collection	
Space Group	P1
Wavelength (Å)	0.97623
a, b, c (Å)	36.14 37.41 64.08
$\alpha, \beta, \gamma$ (°)	98.94 103.83 94.79
Resolution range (Å)	34.81–1.80 (1.80–1.85)
Number of unique reflections*	28770 (2095)
$R_{\text{sym}}$ (%)*	5.6 (50.2)
$\langle I \rangle / \langle \sigma I \rangle$ *	12.0 (2.1)
Completeness (%)*	91.1 (90.5)
Multiplicity*	1.7 (1.6)
Refinement	
R factor (%)*	18.68
Free R factor (%)*	22.02
RMSD bond lengths (Å)	0.0090
RMSD bond angles (°)	1.4472
Ramachadran favoured	296 (100%)
Ramachandran disallowed	0 (0.00%)
PDB accession code	5FII

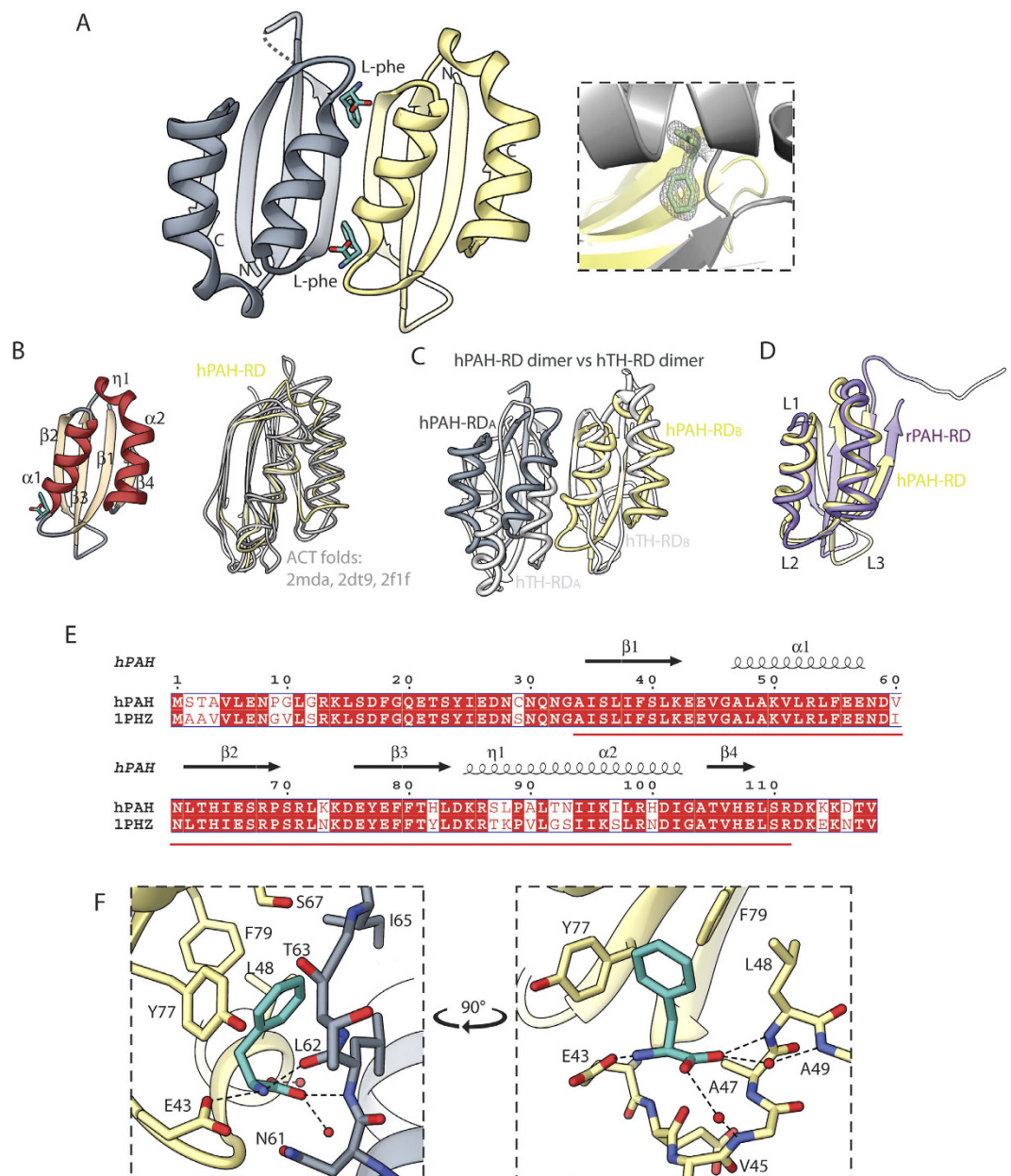
**Table 1. Summary of data collection and refinement statistics for the hPAH-RD structure.** †highest resolution shell shown in parentheses.

Omit Fo-Fc electron density map (Supplementary Fig. S1) revealed two Phe molecules bound to the hPAH-RD homodimer, at the interface of the two  $\beta\alpha\beta\beta\alpha\beta$  folds along the plane of the 2-fold axis (Fig. 2A, inset). Each Phe binding site (Fig. 2F) comprises residues from both protomers, including the ‘ExVxAL’ (aa 43–48) and ‘YxF’ (aa 77–79) sequence regions of one protomer, and the ‘NLTIXS’ (aa 61–67) region of the other. The residues involved in Phe-binding and dimer interaction are highly conserved among PAH orthologues, and also agree with a very recent NMR study on the Phe-binding region of rat PAH-RD<sup>24</sup>. To determine if Phe binding leads to conformational changes within the PAH-RD, we superimposed our Phe-bound hPAH-RD structure with that of unliganded PAH-RD, extracted from the rat PAH structure of RD + CD (83% sequence identity to hPAH; Fig. 2E). Significant C $\alpha$  differences (RMSD > 2 Å) between the two structures are found, and regions that undergo most conformational changes in the presence of Phe include (Fig. 2D): (i) loop L1 and the first turn of helix  $\alpha 1$  (aa 42–49), which brings in residues of the ‘ExVxAL’ motif for Phe binding, and (ii) the segment (aa 59–75) containing loop L2, strand  $\beta 2$  and loop L3, which is displaced outwards to ‘open up’ the Phe-binding site.

**Phe-dependent dimerization of PAH-RD in solution.** To provide solution evidence for the crystallographic PAH-RD homodimer, the oligomeric state of hPAH-RD<sup>1–118</sup> and hPAH-RD<sup>19–118</sup> was evaluated by size exclusion chromatography with multi-angle light scattering (SEC-MALS). In the presence of Phe, both hPAH-RD proteins yielded an experimental MW (26.3 kDa and 22.4 kDa; theoretical monomer 13.5 kDa and 11.7 kDa respectively) consistent with a predominant dimer (Fig. 3A,B). A low-resolution *ab initio* envelope, generated from small-angle x-ray scattering (SAXS) measurement of hPAH-RD<sup>19–118</sup> at multiple protein concentrations (to rule out concentration-dependent aggregation effect; see Experimental Procedure) (Fig. 3C), can accommodate one hPAH-RD dimer from the crystal structure (Fig. 3D). The larger volume for the *ab initio* model compared to our dimeric crystal structure is likely due to the tailing of the P(r) distribution curve (higher  $D_{\text{max}}$  region in Fig. 3C), and could be explained by (i) some aggregation present in the sample (also observed in the SEC-MALS data, Fig. 3B), and (ii) flexibility of the disordered N-terminal loop (aa 19–33) in solution. The latter is supported by a simulated P(r) distribution curve back-calculated from our PAH-RD structure encompassing aa 34–118 (Fig. 3C), which shows a smoother P(r) curve with a less pronounced tail region.

In the absence of Phe, both hPAH-RD<sup>1–118</sup> and hPAH-RD<sup>19–118</sup> proteins eluted as a mixed population of monomer (experimental MW 13.1 kDa and 11.3 kDa) and higher order aggregates, the latter indicated by the significant void volume peak at  $V_e = 8$  ml (Fig. 3A,B). While a previous analytical ultracentrifugation study using the rat PAH-RD protein indicated a monomer-dimer equilibrium in the absence of Phe<sup>18</sup>, we did not observe any discernible peak for an *apo* dimer in SEC-MALS. The aggregated nature of unliganded PAH-RD is supported further from SAXS profiles (Fig. 3C), where the Guinier approximation and pairwise distribution function P(r) indicate an amyloid-like structure, with larger values for radius of gyration ( $R_g$ ) and maximal intraparticle dimension ( $D_{\text{max}}$ ) as compared to the Phe-present sample (Fig. 3F). Furthermore, *ab initio* modelling of unliganded PAH-RD<sup>19–118</sup> (Fig. 3E) supports the presence of amyloid-like stacking which dominates the scattering signal, and is compatible with four PAH-RD dimers assembling with their  $\beta$ -sheets perpendicular to the long axis. Both SEC-MALS and SAXS data support the notion that in the absence of Phe, hPAH-RD is monomeric or prone to aggregation in solution, whereas the addition of Phe stabilizes the domain and reduces its aggregation tendency, likely through domain dimerization.

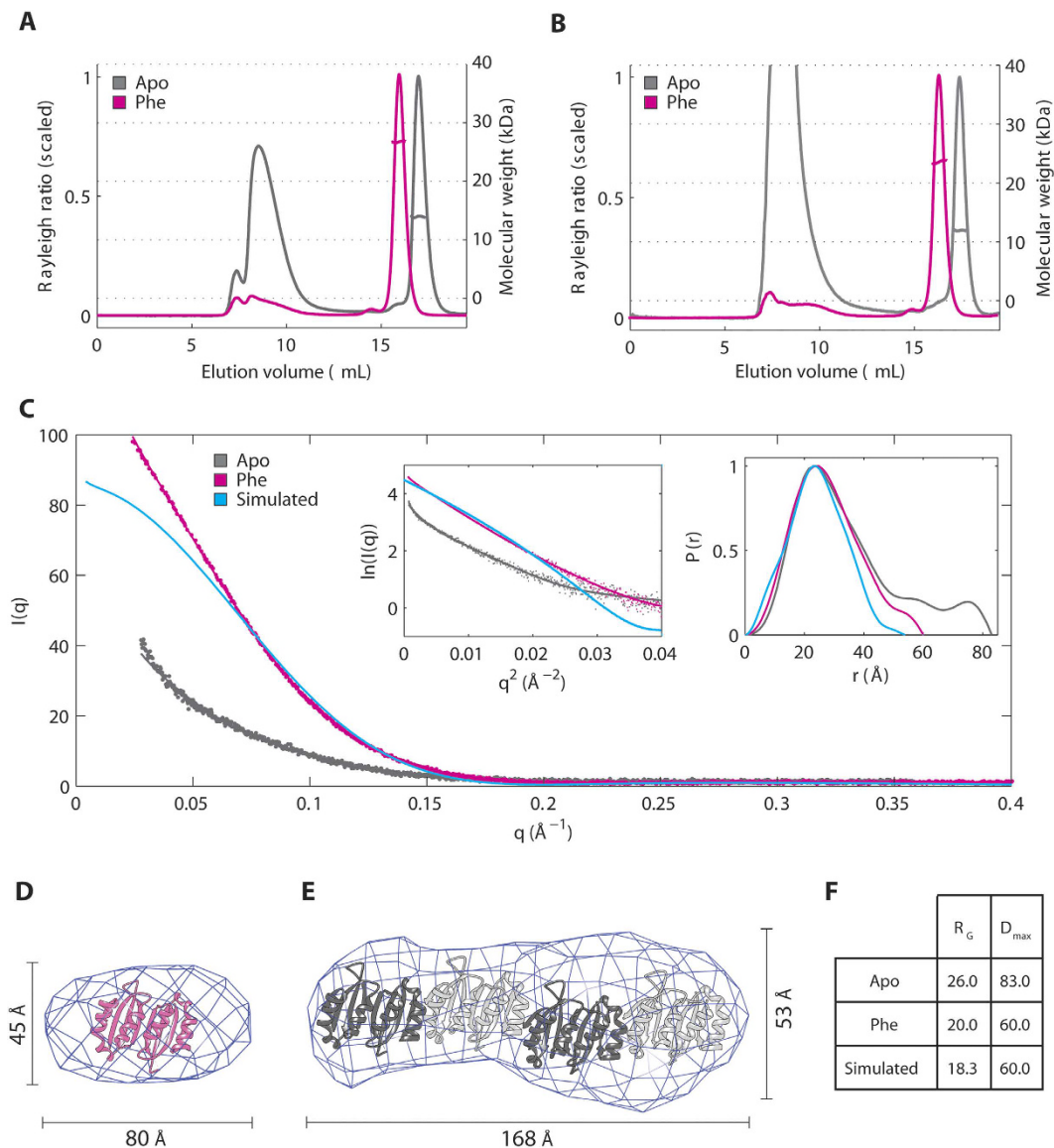
To verify that PAH-RD dimerization depends on its ability to bind Phe, we attempted to knock out Phe-binding residues by site-directed mutagenesis and study the oligomeric state of PAH mutants. To this end,



**Figure 2. Crystal structure of hPAH-RD.** (A) Ribbon representation of hPAH-RD dimeric structure, coloured grey and yellow for the two subunits. Phe is shown in sticks. (Inset)  $2F_o - F_c$  electron density map showing the bound Phe ligand. (B) *Left*: Topology of the ACT fold in hPAH-RD. *Right*: Structural superposition of hPAH-RD (yellow) with other ACT folds (grey) including PDB codes 2mda, 2dt9 and 2f1f. (C) Structural superposition of hPAH-RD dimer (grey and yellow subunits) with hTH-RD dimer (white subunits). The two subunits of a dimer are annotated as A and B. (D) Structural superposition of hPAH-RD (yellow) with the rat counterpart (rPAH) extracted from its RD + CD structure (purple). (E) Alignment of the RD sequences between hPAH (this study) and rat PAH (PDB code 1phz). Secondary structures from our hPAH-RD data are shown. Red line denotes the region covered in the hPAH-RD structure. (F) Phe binding site at the dimer interface of hPAH-RD<sup>19-118</sup>. The two subunits and Phe are colored as in A.

we selected 4 disease causing mutations, namely p.G46S (c.136G > A), p.T63P\_H64N (c.187\_190delACCCin-sCCCA), p.I65S (c.194T > G) and p.E76A (c.227A > C), based on their locations within or proximal to the Phe binding site revealed from the crystal structure. Three of them (p.G46S, p.T63P\_H64N and p.I65S) result in insoluble protein when overexpressed in *E. coli*, highlighting their inherent destabilization. The remaining mutant p.E76A was soluble, and the purified protein is significantly compromised in its stabilisation by Phe, as reflected by a much reduced  $T_m$  shift in DSF upon the addition of 1 mM Phe (from 49.8 to 53.0°C,  $\Delta T_m = 3.2^\circ\text{C}$ ), relative to its WT equivalent hPAH-RD<sup>1-118</sup> ( $\Delta T_m = 20.7^\circ\text{C}$ ) (Fig. 4A). SEC-MALS analysis shows that in the presence of Phe, p.E76A is not predominantly dimeric (as opposed to hPAH-RD<sup>1-118</sup> and hPAH-RD<sup>19-118</sup>), but instead exists as a monomer-dimer equilibrium as illustrated by the overlap seen in the monomer/dimer peaks – a feature not

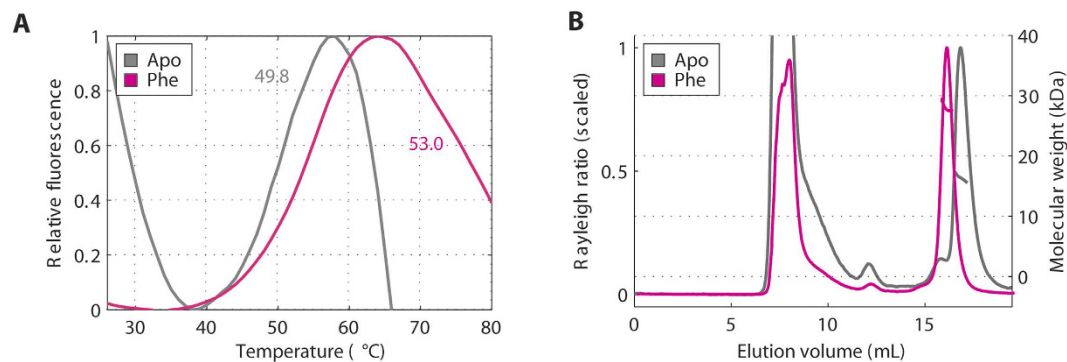




**Figure 3. Phenylalanine stabilizes the dimeric conformation of PAH-RD.** (A) SEC-MALS of unliganded (grey line) and Phe-bound (pink line) of hPAH-RD<sup>1-118</sup>. (B) SEC-MALS of unliganded (grey line) and Phe-bound (pink line) of hPAH-RD<sup>19-118</sup>. (C) SAXS profiles for hPAH-RD<sup>19-118</sup> are plotted for the unliganded (grey line), Phe-bound (pink line) and crystal-structure-simulated (blue line, calculated using CRYSOLO) data. Guinier plots (left) and real space  $P(r)$  distributions (right) are shown as inset. (D) *Ab initio* model of Phe-bound hPAH-RD<sup>19-118</sup> derived from experimental SAXS data (using a  $D_{max}$  estimate of 60 Å), superimposed with the crystallographic dimer of hPAH-RD. (E) *Ab initio* model of unliganded hPAH-RD<sup>19-118</sup> derived from experimental SAXS data (using a  $D_{max}$  estimate of 83 Å), revealing a rod-like shape as the dominant species, potentially accommodating four hPAH-RD crystallographic dimers as modelled. (F)  $R_G$  and  $D_{max}$  values calculated for the unliganded, Phe-bound and crystal-structure-simulated SAXS data of hPAH-RD.

present in the WT proteins. Furthermore, p.E76A also shows high-order aggregates in both Phe-present and Phe-absent samples (Fig. 4B), indicating its higher tendency for aggregation compared to wild-type, likely due to an insufficient capacity to bind Phe. Our data therefore illustrate a correlation between PAH-RD dimerization and its structural stability as well as ability to bind Phe.

**Dimerization of other AAAH regulatory domains does not depend on ligand binding.** Among the AAAH family of enzymes, only PAH is shown to be allosterically activated by its own amino acid substrate, while the others (TH, TPH1, TPH2) are not, despite the presence of a regulatory domain in their N-terminal regions<sup>4</sup> (Supplementary Fig. S2A). The NMR structure of the RD of human TH further shows an ACT-fold homodimer in the unliganded form<sup>23</sup>. We therefore isolated the RDs of human TH and TPH1 by recombinant expression, and explored if their oligomeric status is influenced by amino acid binding. In DSE, hTH-RD and hTPH1-RD are not thermally stabilized by L-Tyrosine and L-Tryptophan respectively, suggesting that they do not



**Figure 4. A phenylalanine binding mutation disrupts the propensity of hPAH-RD to homodimerize.** (A) DSF of the unliganded (grey line;  $T_m = 49.8^\circ\text{C} \pm 0.2$  SD) and Phe-bound (pink line;  $T_m = 53.0^\circ\text{C} \pm 0.03$  SD) hPAH-RD p.E76A. (B) SEC-MALS of unliganded (grey line) and Phe-bound (pink line) of PAH-RD p.E76A.

bind their cognate amino acid substrate (Supplementary Fig. S2B). Furthermore, SEC-MALS analysis of hTH-RD reveals a homodimer even in the absence of amino acid ligand (Supplementary Fig. S2C), confirming that its dimerization is not a consequence of L-Tyrosine binding.

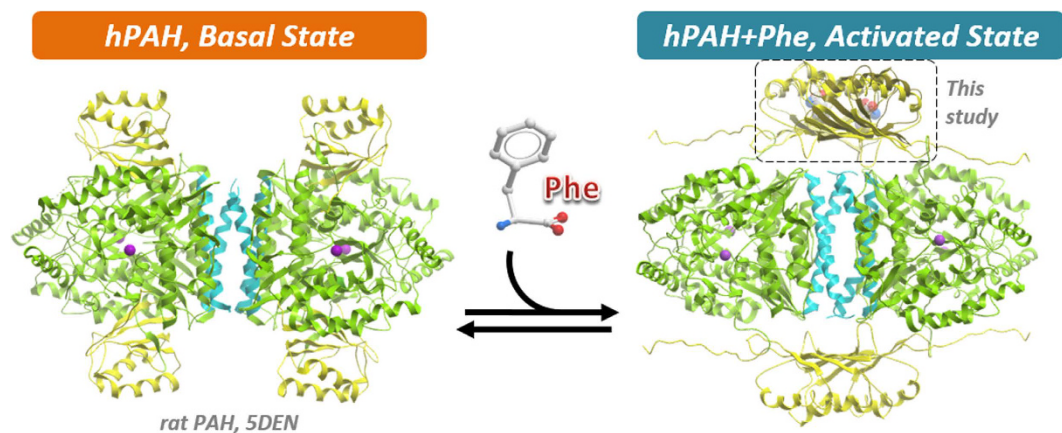
## Discussion

It has long been suggested that phenylalanine serves as a homotropic ligand to phenylalanine hydroxylase, *i.e.* both as a substrate to the catabolic enzyme and as a regulator to its hydroxylation activity. The homotropic effect could be explained by Phe binding to one single site/region for both catalytic and regulatory functions, as proposed in some studies<sup>13,14</sup>, or by Phe binding to two distinct sites as proposed by others<sup>8</sup>. Much progress has since been made on biochemical and structural grounds to delineate the multi-domain organisation of PAH<sup>3</sup>, revealing the rationale for an N-terminal regulatory domain in gating substrate access to the C-terminal catalytic core<sup>5,12</sup>, and thereby modulating its enzyme activity. Recent solution studies have supported such an allosteric mechanism of PAH, involving a Phe-stabilized ACT fold during enzyme activation<sup>18,19,24</sup>. This study presents the structural evidence for Phe binding to PAH-RD as a ligand of the ACT fold, and provides the first high-resolution atomic view for such an allosteric binding site.

The key finding of our study is that the binding of Phe to PAH-RD restructures the ACT fold, forming the necessary inter-subunit dimer interface and the resultant Phe allosteric site. The Phe-bound hPAH-RD structure reported here is in agreement with examples of ACT-containing metabolic enzymes<sup>25,26</sup>, where their ACT folds function in concert, as dimers or higher-order oligomers stabilized by the binding of their cognate regulatory ligand (often an amino acid, as in PAH)<sup>15</sup>. To reconcile the difference with the recently reported rat PAH full-length structure, which did not reveal ACT:ACT interactions in the unliganded form<sup>12</sup>, we propose that the rat full-length PAH structure in the absence of Phe and our human PAH-RD structure in the presence of Phe could represent the unliganded ‘active-site blocking mode’ and the Phe-bound ‘multimerization mode’ of the domain, respectively. The two conformations therefore could exist in an equilibrium driven by the allosteric Phe ligand. Disturbed allostery is a key observation of PKU-associated mutations found in all three domains of PAH<sup>27</sup>, and our data gives the first structural viewpoint for this phenomenon.

It is therefore of no surprise that our PAH-RD dimer neatly confirms the proposed ACT interface that is integral to an active PAH tetramer model<sup>28</sup> (Fig. 5). In this model, Jaffe *et al.* posited that a  $90^\circ$  rotation would be required to translate the PAH-RD from the active-site blocking mode into the multimerization mode, as a necessary path for the conversion of low-activity PAH dimers/tetramers into high-activity PAH tetramers. Such rotation of the PAH-RD would likely sequester the N-terminal auto-regulatory motif away from its inhibitory interaction at the active site<sup>12</sup>. Indeed there is precedence for long-range conformational movement in other ACT-domain enzymes, such as the DAH7PS enzyme where a  $\sim 110^\circ$  reorientation of its RD, driven by Tyr binding, is part of the allosteric gating mechanism<sup>26,29</sup>. The inter-domain plasticity of PAH has also been supported by indirect solution studies, including limited proteolysis<sup>30</sup>, hydrogen/deuterium exchange<sup>31</sup> and more recently, SAXS<sup>12</sup>. Additionally, the proposed model, coupled with our structure, explains the increase in volume<sup>32</sup> and shift towards tetramers<sup>33</sup> observed in the presence of Phe, making it highly likely that such a conformational change does occur in full-length PAH, as part of an allosteric activation mechanism in response to Phe binding. We are cognizant that our study involves an isolated PAH-RD that may not be fully extrapolated to its behaviour in the entire protein, therefore further biophysical and structural studies in the context of the full-length PAH tetramer are warranted to decipher the activating mechanism and underlying conformational changes of the PAH multi-domain enzyme.

Another notable finding is the observation that the regulatory domain of other AAAH enzymes (e.g. tyrosine hydroxylase and tryptophan hydroxylase) can also form the structural homodimer, albeit without the need to bind the cognate amino acid substrate. This corroborates with the NMR structure of TH-RD<sup>23</sup> revealing an unliganded ACT dimer evolved with a distinct, ‘staggered’ interface that renders the canonical small molecule binding site no longer present. We posit that dimerization is a common mechanism for the ACT fold, which could be facilitated for some enzymes by ligand binding, but for others by alternative, post-translational triggers that



**Figure 5. Proposed model of PAH activation by Phenylalanine.** In the basal, unliganded state, as represented by the full-length rat PAH structure (PDB code 5DEN), the four RDs of a PAH tetramer do not interact with each other. Each RD (yellow) interacts with its corresponding CD domain (green) via the N-terminal loop region, resulting in steric hindrance of the CD active site. Phe binding relieves the steric hindrance due to homodimerization of RD, supported in this study, hence allowing unhindered access of substrates into the active sites.

remain poorly understood. It is also possible that the dependence (or not) of the ACT domain towards binding its cognate substrate is dictated by structural adaptation of the ACT dimer for either the 'staggered' (non-binding) or 'side-by-side' (binding) dimer interface. Structural characterization of the TPH-RD, not shown to respond to amino acid ligands, would help validate the hypothesis.

Hundreds of *pah* inherited mutations are the molecular cause of PKU, the most common inborn error of amino acid metabolism. PKU is often referred to as a conformational disorder<sup>34,35</sup>, in which protein destabilization, misfolding and aggregation are molecular hallmarks associated with disease<sup>36–39</sup>. Our solution data demonstrate the tendency of the N-terminal PAH-RD to form higher-order amyloid-like aggregates in the unliganded state, a phenomenon shown to be ameliorated by Phe stabilization *in vitro*. This supports previous claims that misfolding-induced aggregation is implicated in the *in vivo* disease mechanism<sup>34</sup>. A novel therapeutic approach, involving the use of small molecule target-specific 'pharmacological chaperones' (PCs) to stabilize aggregation-prone mutant proteins, is emerging as potential treatment for conformational disorders<sup>35,40</sup>. Relevant to this, PAH constitutes one of several multi-domain metabolic enzymes where a native regulatory ligand exists to activate the target via an allosteric site. The natural cofactor BH<sub>4</sub> currently serves as a PC therapy for PKU, though it is not successful in treating all PKU disease alleles. Our study provides a rationale for the structure-guided drug design of PAH, using the PAH-RD as a target for small molecule activation. Rather than developing a PC molecule directed to the PAH active site, a small molecule that mimics Phe in stabilising the PAH-RD dimerization interface could activate the mutant enzyme above the threshold sufficient for relieving disease phenotypes, and prove useful as an alternative PC therapy, especially for those PKU mutations that are not BH<sub>4</sub> responsive. The recent report of Phe-like molecules that bind PAH could potentially be acting in such a manner<sup>41</sup>, although caution will be needed to avoid cross-reactivity with the active site. Lessons could therefore be learnt from this and other examples of metabolic enzymes (e.g. cystathionine  $\beta$ -synthase, the cause of homocystinuria<sup>42</sup>; porphobilinogen synthase, the cause of ALAD porphyria<sup>43</sup>), for the design of novel PC molecules that target the allosteric domains as a stabilization strategy.

## Methods

**Expression and Purification of PAH, TH and TPH regulatory domains.** A DNA fragment encoding the PAH regulatory domain (residues 1–118 or 19–118; GenBank number 4557819) was subcloned into the pNIC28-Bsa4 (GenBank Accession No. EF198106) vector incorporating an N-terminal TEV-cleavable His<sub>6</sub>-tag. The resulting plasmid was transformed into *E. coli* BL21(DE3), cultured in Terrific Broth at 37 °C until OD<sub>600</sub> ~1.5, and induced with 0.5 mM IPTG for overnight growth at 18 °C. Cells were harvested and homogenized in buffer A (50 mM HEPES pH 7.5, 500 mM NaCl, 5 mM imidazole, 0.5 mM TCEP, EDTA-free protease inhibitor). Insoluble debris was removed by further centrifugation. Proteins were purified by passing cell extracts through a 1 ml HisTrap column pre-equilibrated with buffer A, and eluted with buffer B (buffer A + 250 mM Imidazole). Eluted fractions were treated with TEV protease at a protein:protease ratio of 1:20, and incubated overnight at 4 °C in order to cleave off the His<sub>6</sub>-tag. The tag-cleaved protein was applied onto a 1 ml HisTrap column, pre-equilibrated with GF buffer (50 mM HEPES, pH 7.5, 300 mM NaCl, 5% glycerol and 0.5 mM TCEP). The flow-through sample was applied onto a HiLoad 16/60 Superdex 75 column pre-equilibrated with GF buffer. The regulatory domains of human TH (residues 65–160; GenBank number 88900503) and TPH1 (residues 1–100; GenBank number 4759248) were subcloned into the pFB-Lic-Bse vector incorporating an N-terminal TEV-cleavable His<sub>6</sub>-tag, overexpressed in insect sf9 cells, and purified as above.

**Crystallization and structure determination of PAH-RD.** Purified hPAH-RD was concentrated to 13.5 mg/ml, where L-Phenylalanine (Phe) was added to a final concentration of 10 mM. The sample was treated with trypsin at a protein:trypsin mass ratio of 1:100 immediately prior to crystallization set up. Crystals were grown by

vapour diffusion in sitting drops at 20 °C. A sitting drop consisting of 100 nl trypsin-treated protein and 200 nl well solution was equilibrated against well solution containing 25% PEG 3350, 0.20 M NaCl and 0.1 M BIS-Tris pH 5.5. The crystals were mounted directly from the drop using 25% ethylene glycol as a cryoprotectant and flash-cooled in liquid nitrogen. Diffraction data was collected at the Diamond Light Source and processed with the CCP4 program suite<sup>44</sup>. The structure of hPAH-RD was solved by molecular replacement using the program PHASER<sup>45</sup> with the *Chlorobium tepidum* prephenate dehydratase structure (PDB code 2QMX) as search model. Iterative cycles of restrained refinement and manual model building were performed using COOT<sup>46</sup> and PHENIX<sup>47</sup>.

**Differential scanning fluorimetry.** hPAH-RD was assayed for shifts in melting temperature caused by the presence of Phe in a 96-well PCR plate using an Mx3005p RT-PCR machine (Stratagene). Each reaction well (20 µl) consists of protein (10 µM in a buffer containing 50 mM Hepes, pH 7.5, 50 mM NaCl), SYPRO-Orange (Invitrogen) diluted 1,000×, and 1 mM Phe. Fluorescence intensities were measured from 25 °C to 96 °C with a ramp rate of 1 °C/min, as described<sup>48</sup>. The temperature shifts,  $\Delta T_m^{obs}$ , for each ligand were determined<sup>48</sup>. The experiment was repeated three times and an error was calculated based on standard deviation.

**Small Angle X-ray Scattering.** The hPAH-RD sample was purified by size-exclusion chromatography prior to sample preparation to remove any large aggregate species. hPAH-RD was prepared at 10 mg/ml with a final Phe concentration of 10 mM for the Phe-bound sample. Data collection was performed in batch mode, where samples were subjected to SAXS by flowing sample through an in-vacuum quartz capillary of 1.6 mm diameter. Data were collected using a Pilatus2M detector (Dectris, CH) at a sample-detector distance of 3914 mm and a wavelength of  $\lambda = 1 \text{ \AA}$ . The range of momentum transfer of  $0.1 < s < 5 \text{ nm}^{-1}$  was covered ( $s = 4\pi\sin\theta/\lambda$ , where  $\theta$  is the scattering angle). Comparison of eighteen 10 s exposures was performed. Radiation damage was monitored for the batch mode by monitoring changes in radius of gyration in each frame, and no significant changes were observed. Samples were measured at three concentrations of 2.5 mg/ml, 5 mg/ml and 10 mg/ml for both unligand and liganded samples, to eliminate any concentration dependent aggregation during the SAXS experiment. The raw scattering curves for all concentrations were identical in profile, albeit poorer in the signal-to-noise ratio upon decreasing protein concentrations. The data were radially averaged and the scattering of the buffer was subtracted. The forward scattering  $I(0)$ , radius of gyration  $R_g$ , pair distribution of the particle  $p(r)$ , and maximum dimension  $D_{max}$  were analyzed using ScATTER and the ATSAS suite of programs<sup>49</sup>.

**Size Exclusion Chromatography – Multiple Angle Light Scattering.** SEC-MALS analysis was performed at 20 °C using an analytical Superdex 75 column (GE Healthcare) and a Shimadzu (Kyoto, Japan) chromatography system, connected in-line to a Heleos8 + multi-angle light scattering detector and an Optilab T-rEX refractive index (RI) detector (Wyatt Technologies, Goleta, CA). hPAH-RD was used at a final concentration of 3 mg/ml and Phe was used at a final concentration of 10 mM, in a buffer containing 10 mM Tris-HCl, pH 7.5, 300 mM NaCl, 2 mM  $\beta$ -mercaptoethanol. Samples measured with Phe were used in a buffer containing 10 mM Tris-HCl, pH 7.5, 300 mM NaCl, 2 mM  $\beta$ -mercaptoethanol, 10 mM Phe. Samples were injected in this system, and the resulting MALS, RI and UV traces processed in ASTRA 6 (Wyatt Technologies). On-column protein concentration was calculated from the differential RI, assuming  $dn/dc$  of 0.1850 ml/g. SEC-MALS data were fit with a two-state model to derive association parameters.

## References

- Konecki, D. S. & Lichter-Konecki, U. The phenylketonuria locus: current knowledge about alleles and mutations of the phenylalanine hydroxylase gene in various populations. *Hum. Genet.* **87**, 377–388 (1991).
- Wettstein, S. *et al.* Linking genotypes database with locus-specific database and genotype-phenotype correlation in phenylketonuria. *Eur. J. Hum. Genet.* **23**, 302–309 (2015).
- Fitzpatrick, P. F. Mechanism of aromatic amino acid hydroxylation. *Biochemistry* **42**, 14083–14091 (2003).
- Grenett, H. E., Ledley, F. D., Reed, L. L. & Woo, S. L. Full-length cDNA for rabbit tryptophan hydroxylase: functional domains and evolution of aromatic amino acid hydroxylases. *Proc. Natl. Acad. Sci. USA* **84**, 5530–5534 (1987).
- Kobe, B. *et al.* Structural basis of autoregulation of phenylalanine hydroxylase. *Nat. Struct. Biol.* **6**, 442–448 (1999).
- Fuseti, F., Erlandsen, H., Flatmark, T. & Stevens, R. C. Structure of tetrameric human phenylalanine hydroxylase and its implications for phenylketonuria. *J. Biol. Chem.* **273**, 16962–16967 (1998).
- Fitzpatrick, P. F. Allosteric regulation of phenylalanine hydroxylase. *Arch. Biochem. Biophys.* **519**, 194–201 (2012).
- Shiman, R. & Gray, D. W. Substrate activation of phenylalanine hydroxylase. A kinetic characterization. *J. Biol. Chem.* **255**, 4793–4800 (1980).
- Doskeland, A. P., Martinez, A., Knappskog, P. M. & Flatmark, T. Phosphorylation of recombinant human phenylalanine hydroxylase: effect on catalytic activity, substrate activation and protection against non-specific cleavage of the fusion protein by restriction protease. *Biochem. J.* **313** (Pt 2), 409–414 (1996).
- Daubner, S. C., Hillas, P. J. & Fitzpatrick, P. F. Expression and characterization of the catalytic domain of human phenylalanine hydroxylase. *Arch. Biochem. Biophys.* **348**, 295–302 (1997).
- Knappskog, P. M., Flatmark, T., Aarden, J. M., Haavik, J. & Martinez, A. Structure/function relationships in human phenylalanine hydroxylase. Effect of terminal deletions on the oligomerization, activation and cooperativity of substrate binding to the enzyme. *Eur. J. Biochem.* **242**, 813–821 (1996).
- Arturo, E. C. *et al.* First structure of full-length mammalian phenylalanine hydroxylase reveals the architecture of an autoinhibited tetramer. *Proc. Natl. Acad. Sci. USA* doi: 10.1073/pnas.1516967113 (2016).
- Thorolfsson, M. *et al.* L-phenylalanine binding and domain organization in human phenylalanine hydroxylase: a differential scanning calorimetry study. *Biochemistry* **41**, 7573–7585 (2002).
- Thorolfsson, M., Teigen, K. & Martinez, A. Activation of phenylalanine hydroxylase: effect of substitutions at Arg68 and Cys237. *Biochemistry* **42**, 3419–3428 (2003).
- Lang, E. J., Cross, P. J., Mittelstadt, G., Jameson, G. B. & Parker, E. J. Allosteric ACTION: the varied ACT domains regulating enzymes of amino-acid metabolism. *Curr. Opin. Struct. Biol.* **29**, 102–111 (2014).
- Gjetting, T., Petersen, M., Guldberg, P. & Guttler, F. Missense mutations in the N-terminal domain of human phenylalanine hydroxylase interfere with binding of regulatory phenylalanine. *Am. J. Hum. Genet.* **68**, 1353–1360 (2001).



17. Li, J., Ilangovan, U., Daubner, S. C., Hinck, A. P. & Fitzpatrick, P. F. Direct evidence for a phenylalanine site in the regulatory domain of phenylalanine hydroxylase. *Arch. Biochem. Biophys.* **505**, 250–255 (2011).
18. Zhang, S., Roberts, K. M. & Fitzpatrick, P. F. Phenylalanine binding is linked to dimerization of the regulatory domain of phenylalanine hydroxylase. *Biochemistry* **53**, 6625–6627 (2014).
19. Roberts, K. M., Khan, C. A., Hinck, C. S. & Fitzpatrick, P. F. Activation of phenylalanine hydroxylase by phenylalanine does not require binding in the active site. *Biochemistry* **53**, 7846–7853 (2014).
20. Andersen, O. A., Stokka, A. J., Flatmark, T. & Hough, E. 2.0 Å resolution crystal structures of the ternary complexes of human phenylalanine hydroxylase catalytic domain with tetrahydrobiopterin and 3-(2-thienyl)-L-alanine or L-norleucine: substrate specificity and molecular motions related to substrate binding. *J. Mol. Biol.* **333**, 747–757 (2003).
21. Wernimont, A. & Edwards, A. *In situ* proteolysis to generate crystals for structure determination: an update. *Plos One* **4**, e5094 (2009).
22. Kaplun, A. *et al.* Structure of the regulatory subunit of acetohydroxyacid synthase isozyme III from *Escherichia coli*. *J. Mol. Biol.* **357**, 951–963 (2006).
23. Zhang, S., Huang, T., Ilangovan, U., Hinck, A. P. & Fitzpatrick, P. F. The solution structure of the regulatory domain of tyrosine hydroxylase. *J. Mol. Biol.* **426**, 1483–1497 (2014).
24. Zhang, S. & Fitzpatrick, P. F. Identification of the Allosteric Site for Phenylalanine in Rat Phenylalanine Hydroxylase. *J. Biol. Chem.* doi: 10.1074/jbc.M115.709998 (2016).
25. Schuller, D. J., Grant, G. A. & Banaszak, L. J. The allosteric ligand site in the V<sub>max</sub>-type cooperative enzyme phosphoglycerate dehydrogenase. *Nat. Struct. Biol.* **2**, 69–76 (1995).
26. Cross, P. J., Dobson, R. C., Patchett, M. L. & Parker, E. J. Tyrosine latching of a regulatory gate affords allosteric control of aromatic amino acid biosynthesis. *J. Biol. Chem.* **286**, 10216–10224 (2011).
27. Gersting, S. W. *et al.* Loss of function in phenylketonuria is caused by impaired molecular motions and conformational instability. *Am. J. Hum. Genet.* **83**, 5–17 (2008).
28. Jaffe, E. K., Stith, L., Lawrence, S. H., Andrade, M. & Dunbrack, R. L. Jr. A new model for allosteric regulation of phenylalanine hydroxylase: implications for disease and therapeutics. *Arch. Biochem. Biophys.* **530**, 73–82 (2013).
29. Cross, P. J., Allison, T. M., Dobson, R. C., Jameson, G. B. & Parker, E. J. Engineering allosteric control to an unregulated enzyme by transfer of a regulatory domain. *Proc. Natl. Acad. Sci. USA* **110**, 2111–2116 (2013).
30. Phillips, R. S., Iwaki, M. & Kaufman, S. Ligand effects on the limited proteolysis of phenylalanine hydroxylase: evidence for multiple conformational states. *Biochem. Biophys. Res. Commun.* **110**, 919–925 (1983).
31. Li, J., Dangott, L. J. & Fitzpatrick, P. F. Regulation of phenylalanine hydroxylase: conformational changes upon phenylalanine binding detected by hydrogen/deuterium exchange and mass spectrometry. *Biochemistry* **49**, 3327–3335 (2010).
32. Kappock, T. J., Harkins, P. C., Friedenber, S. & Caradonna, J. P. Spectroscopic and kinetic properties of unphosphorylated rat hepatic phenylalanine hydroxylase expressed in *Escherichia coli*. Comparison of resting and activated states. *J. Biol. Chem.* **270**, 30532–30544 (1995).
33. Martinez, A. *et al.* Expression of recombinant human phenylalanine hydroxylase as fusion protein in *Escherichia coli* circumvents proteolytic degradation by host cell proteases. Isolation and characterization of the wild-type enzyme. *Biochem. J.* **306** (Pt 2), 589–597 (1995).
34. Muntau, A. C. & Gersting, S. W. Phenylketonuria as a model for protein misfolding diseases and for the development of next generation orphan drugs for patients with inborn errors of metabolism. *J. Inher. Metab. Dis.* **33**, 649–658 (2010).
35. Underhaug, J., Aubi, O. & Martinez, A. Phenylalanine hydroxylase misfolding and pharmacological chaperones. *Curr. Top. Med. Chem.* **12**, 2534–2545 (2012).
36. Leandro, J., Simonsen, N., Saraste, J., Leandro, P. & Flatmark, T. Phenylketonuria as a protein misfolding disease: The mutation pG46S in phenylalanine hydroxylase promotes self-association and fibril formation. *Biochim. Biophys. Acta.* **1812**, 106–120 (2011).
37. Leandro, J., Saraste, J., Leandro, P. & Flatmark, T. The G46S-hPAH mutant protein: a model to study the rescue of aggregation-prone PKU mutations by chaperones. *Mol. Genet. Metab.* **104**, Suppl, S40–44 (2011).
38. Pey, A. L., Stricher, F., Serrano, L. & Martinez, A. Predicted effects of missense mutations on native-state stability account for phenotypic outcome in phenylketonuria, a paradigm of misfolding diseases. *Am. J. Hum. Genet.* **81**, 1006–1024 (2007).
39. Erlandsen, H. *et al.* Correction of kinetic and stability defects by tetrahydrobiopterin in phenylketonuria patients with certain phenylalanine hydroxylase mutations. *Proc. Natl. Acad. Sci. USA* **101**, 16903–16908 (2004).
40. Muntau, A. C., Leandro, J., Staudigl, M., Mayer, F. & Gersting, S. W. Innovative strategies to treat protein misfolding in inborn errors of metabolism: pharmacological chaperones and proteostasis regulators. *J. Inher. Metab. Dis.* **37**, 505–523 (2014).
41. Montalbano, F. *et al.* Phenylalanine iminoboronates as new phenylalanine hydroxylase modulators. *RSC Adv* **4**, 61022–61027 (2014).
42. McCorvie, T. J. *et al.* Inter-domain communication of human cystathionine beta-synthase: structural basis of S-adenosyl-L-methionine activation. *J. Biol. Chem.* **289**, 36018–36030 (2014).
43. Jaffe, E. K. & Lawrence, S. H. Allostery and the dynamic oligomerization of porphobilinogen synthase. *Arch. Biochem. Biophys.* **519**, 144–153 (2012).
44. CCP4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763 (1994).
45. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. Likelihood-enhanced fast translation functions. *Acta Crystallogr. D Biol. Crystallogr.* **61**, 458–464 (2005).
46. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
47. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
48. Niesen, F. H., Berglund, H. & Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2**, 2212–2221 (2007).
49. Petoukhov, M. V. *et al.* New developments in the program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **45**, 342–350 (2012).

## Acknowledgements

We thank Diamond Light Source for access to beamlines I03 and B21. We are grateful to Robert Rambo and James Douth for help with SAXS data collection and processing. The Structural Genomics Consortium is a registered charity (number 1097737) that receives funds from AbbVie, Boehringer Ingelheim, the Canada Foundation for Innovation, the Canadian Institutes for Health Research, Genome Canada, GlaxoSmithKline, Janssen, Lilly Canada, the Novartis Research Foundation, the Ontario Ministry of Economic Development and Innovation, Pfizer, Takeda, and the Wellcome Trust (092809/Z/10/Z).

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Patel, D. *et al.* Structural basis for ligand-dependent dimerization of phenylalanine hydroxylase regulatory domain. *Sci. Rep.* **6**, 23748; doi: 10.1038/srep23748 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>