The first X-ray crystal structure of full-length mammalian phenylalanine hydroxylase

Emilia C. Arturo[1,2], Kushol Gupta[3], Annie Heroux[4], Penelope J. Cross[5], Emily J. Parker[5], Patrick J. Loll[2], Eileen K. Jaffe[1]

[1]Fox Chase Cancer Center (TUHS), Philadelphia, PA, [2]Drexel University College of Medicine, Philadelphia, PA, [3]Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA, [4]Brookhaven National Laboratory, Upton, NY, [5]University of Canterbury, Christchurch, New Zealand

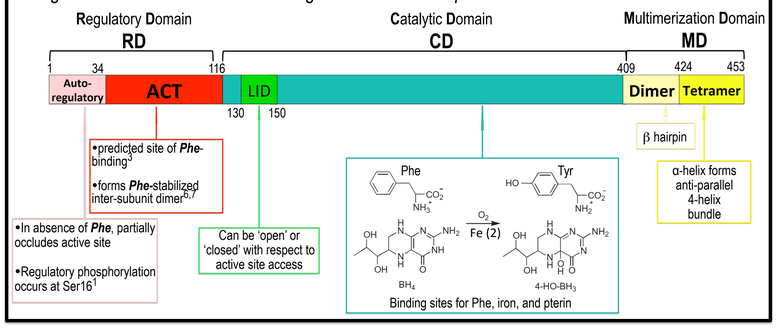
Abstract

Mammalian phenylalanine hydroxylase (PAH) is a non-heme iron-dependent mono-oxygenase that uses tetrahydrobiopterin (BH_4) to catalyze the rate-limiting step of phenylalanine catabolism, synthesizing tyrosine from phenylalanine (Phe, ***Phe***)[§]. At elevated concentrations, ***Phe*** allosterically activates PAH. Improper allosteric regulation and/or reduced enzymatic activity permit accumulation of neurotoxic levels of phenylalanine in the blood[1]; this is characteristic of one of the most common inborn errors of metabolism, phenylketonuria (PKU).[2] We previously proposed that allosteric regulation of PAH involves major inter-domain motions that interconvert multiple architecturally distinct tetramers that have low- or high-activity[3]; however, no full-length structure had yet been reported. This previously underappreciated fact--that PAH exists in a dynamic equilibrium of tetrameric assemblies in solution--has likely been a hurdle in obtaining diffraction-quality crystals of the full-length protein. By including an additional ion-exchange purification step that separates tetramers, crystallization was successful and we now report the first X-ray crystal structure of full-length rat PAH at 2.9 Å resolution. This structure supersedes a long-accepted composite homology model.[4] Furthermore, using small-angle X-ray scattering (SAXS), we confirm that this crystal structure is in a conformation consistent with the structure of a low-activity tetramer in solution. Comparison of this full-length structure with all other X-ray crystal structures of PAH—all are truncated—reveals features that likely play a role in tetramer interconversion, and in the regulation of active-site access. From the full-length structure we can now appreciate inter-domain interactions not previously reported, some of which may give new insight regarding the role of disease-associated single-residue substitutions. Ultimately, the goal is to discern the structures of the multiple tetramers available to PAH such that drugs can be rationally designed to bind to an allosteric ligand-binding site that stabilizes a high-activity form, or destabilizes a low-activity form. Such a drug could increase the activity of disease-associated variants whose oligomeric equilibrium is perturbed towards low-activity forms.

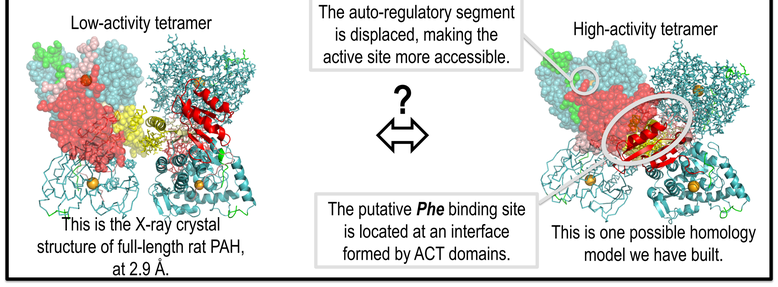
[§]Phe is both a substrate and an allosteric activator, therefore the distinction between the two Phe molecules, and their respective binding sites, will be emphasized by denoting the substrate in plain text ('Phe'), and the allosteric activator in bold italic text ('***Phe***').

PAH is a multimeric multi-domain non-heme iron-dependent mono-oxygenase that uses BH_4 to catalyze the hydroxylation of Phe to tyrosine.

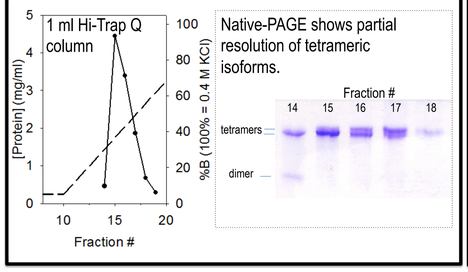
All figures herein are colored according to this domain map.



Our model[3] of allosteric regulation of mammalian PAH includes at least two architecturally distinct tetrameric assemblies in an dynamic equilibrium. The high-activity tetramer shown is a homology model based in part on the ACT-ACT interface found in the related, constitutively active aromatic amino acid hydroxylase, tyrosine hydroxylase[5]. The mechanism of interconversion is an active research pursuit.



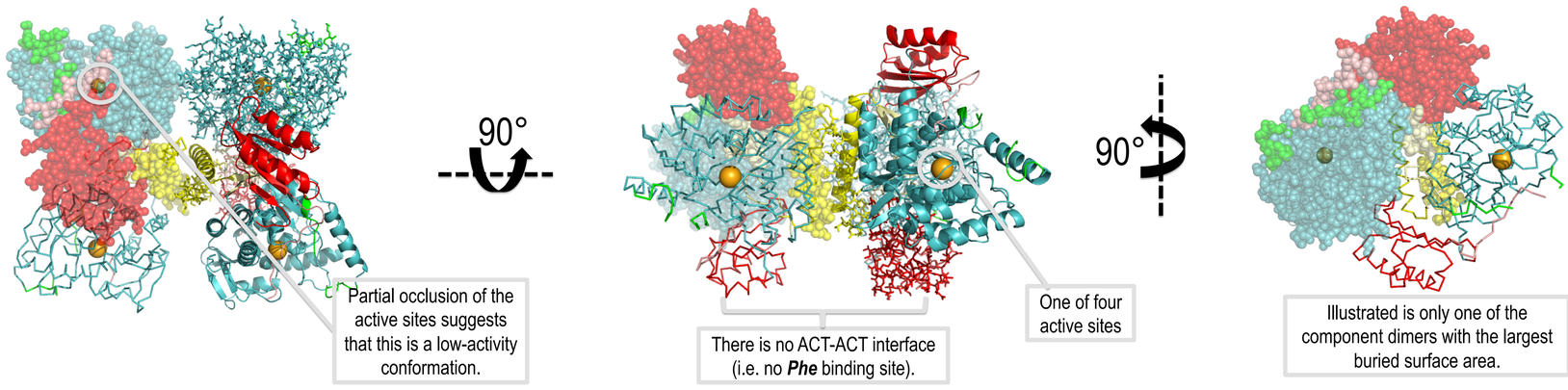
Solution of this X-ray crystal structure of full-length PAH was facilitated by purification of a single tetrameric assembly using ion-exchange chromatography applied to the phenyl-sepharose-purified protein.



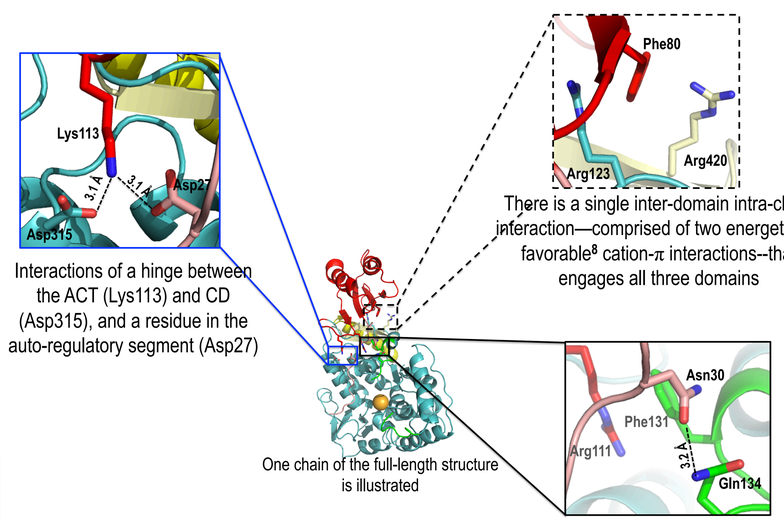
Statistics

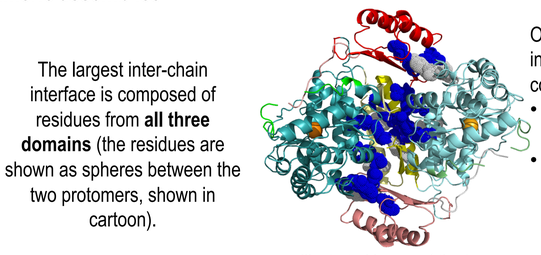
Data Collection	
Wavelength (Å)	1.1
Resolution range (Å)	38.1 - 2.9 (3.004 - 2.9)
Space group	C 1 2 1
Unit cell a, b, c (Å); β (°)	113.78, 89.16, 196.81; 104.53
Total reflections	73766 (5281)
Unique reflections	38661 (3231)
Multiplicity	1.9 (1.6)
Completeness (%)	91.01 (76.49)
Mean I/sigma(I)	9.06 (1.56)
Wilson B-factor	66.69
R-merge	0.067 (0.373)
R-meas	0.090
CC1/2	0.996 (0.833)
CC*	0.999 (0.953)
Refinement	
Reflections used for R-free	1930
R-work	0.234 (0.344)
R-free	0.305 (0.386)
No. of non-hydrogen atoms	13711
No. of protein atoms	13706
No. of ligands	4
water	1
Protein residues	1692
RMS(bonds)	0.01
RMS(angles)	1.41
Ramachandran favored (%)	95
Ramachandran outliers (%)	0.06
Clashscore	10.36
Average B-factor	71.5
macromolecules	71.5
ligands	59.1
solvent	24.2

The four monomers in the asymmetric unit (ASU) can be arranged as a tetramer that represents the *low-activity* tetrameric assembly of mammalian PAH



The full-length structure reveals intra (LEFT)**- and inter** (RIGHT)**-chain interactions (present in all four chains of the ASU) that are potentially important to the interconversion of tetrameric assemblies.**





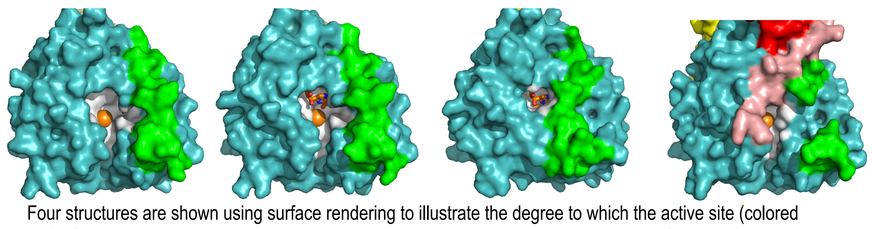
Of the 20 residues at the inter-chain interface of each of the two component dimers:

- 9 are disease-associated positions
- 3 of these (as single-residue substitutions) are among the most common PKU-associated PAH variants.[2]

Comparison of the full-length structure to CD-only PAH structures demonstrates:

A. that both the auto-regulatory segment and the active site lid modulate active site access;
B. disorder in the active site lid correlates with the presence of multiple domains and with active site occupancy; and
C. the backbone about the active site lid in the full-length structure is a hybrid of 'open' and 'closed' lid configurations; this may be the cause of disorder in residues 136-142 (a portion of the active site lid) in all RD-containing structures.

A.



Four structures are shown using surface rendering to illustrate the degree to which the active site (colored WHITE) is accessible—the active site lid can be 'open' or 'closed' where resolved in the CD-only structures, or partially disordered in all—including this full-length structure—the multi-domain structures.

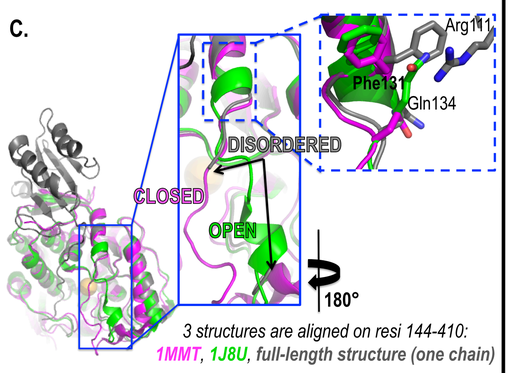
B.

domains present	PDB id	year	resolution (Å)	occupancy of ligand binding sites at the active site: Fe	Phe*	pterin*	organism h=human r=rat
CD	1PAH	1997	2.0	Fe(3)	none	none	h
	3PAH	1998	2.0	Fe(3)	none	XDE	h
	4PAH	1998	2.0	Fe(3)	none	LNR	h
	6PAH	1998	2.2	Fe(3)	DAH	none	h
	5PAH	1998	2.1	Fe(3)	none	LDP	h
	1DMW	2000	2.0	Fe(3)	none	HBI	h
	1J8U	2001	1.5	Fe(2)	none	H4B	h
	1J8T	2001	1.7	Fe(2)	none	none	h
	1KW0	2002	2.5	Fe(3)	TIH	H4B	h
	1LRM	2002	2.1	Fe(2)	none	HBI	h
	1MMK	2003	2.0	Fe(2)	TIH	H4B	h
	1MMT	2003	2.0	Fe(2)	NLE	H4B	h
	1TDW	2004	2.1	Fe(3)	none	none	h
	1TG2	2004	2.2	Fe(3)	none	H2B	h
	4ANP	2012	2.1	Fe(3)	**	**	h
RD + CD	1PHZ #	1999	2.2	Fe(3)	none	none	r
	2PHM	1999	2.6	Fe(3)	none	none	r
CD + MD	2PAH	1998	3.1	Fe(3)	none	none	h
RD + CD + MD		2015	2.9	Fe(3)	none	none	r

colored by active site lid configuration: DISORDERED, 'OPEN', or 'CLOSED'
* Chemical Component Dictionary code; **unusual binding of chaperone at active site;
# phosphorylated at Ser16; *italics*: single-residue substituted variant

Superposition of all 'CLOSED' structures results in a rmsd of ~0.2 Å, and superposition of all 'OPEN' structures results in a rmsd of ~0.3 Å.

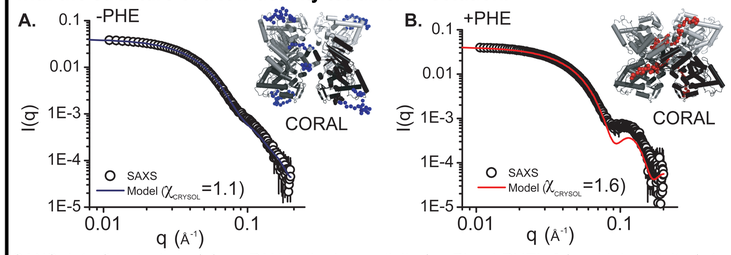
C.



3 structures are aligned on resi 144-410: 1MMT, 1J8U, full-length structure (one chain)

The Arg111-Phe131-Gln134 network may select for the 'closed' configuration of the active site lid (*Arg111 is not resolved in any of the CD-only structures*)

SAXS confirms that the crystal structure of the full-length protein is consistent with the structure of a low-activity tetramer in solution.



SAXS data (black circles) from PAH in the low-activity (no Phe, '-PHE', **A.**) or high-activity (+1mM Phe, '+PHE', **B.**) tetrameric state (fraction 15 from Hi-Trap Q column), showing the recorded scattering intensity as a function of q ($q=4\pi\sin\theta/\lambda$, where 2θ is the scattering angle) as a log-log plot. **A.** The black line is the fit corresponding to the SAXS-refined model of the 2.9 Å crystal structure corrected for the disordered regions modeled as beads (χ = 1.1)[9]; the modeled residues are shown as blue spheres in the inset. **B.** The red line shows the SAXS-refined model of the high-activity homology model with missing linkers modeled as beads (red, χ = 1.6).[9] *SAXS data was collected at the Australian Synchrotron using an inline SEC-SAXS configuration.*

Outstanding questions

- We have solved one full-length structure of PAH, and have identified several inter-domain interactions that may be important for stabilizing the low-activity tetramer. *Are these interactions broken in tetramer interconversion/PAH activation?*
- We have identified residues 130-150 as a putative active site lid, but *we do not yet know what determines the configuration of the lid.*
- The full-length structure of PAH allows us to rationalize the means by which disease-associated missense variants cause PAH dysfunction. *Are any common variants deficient in the ability to interconvert quaternary structure assemblies, thereby being improperly allosterically regulated by* Phe/stabilized by ***Phe****?*
- We have built a homology model for the high-activity tetramer that resembles the tetrameric species in the +PHE SAXS experiment, but the scattering profile shows additional features. *What is the structure of the high-activity form?*

References

1. Fitzpatrick PF (2012) Allosteric regulation of phenylalanine hydroxylase. *Archives of biochemistry and biophysics* 519(2):194-201.
2. The PAHdb: http://www.pahdb.mcgill.ca/ Accessed July 1, 2015.
3. Jaffe EK, Stith L, Lawrence SH, Andrake M, & Dunbrack RL, Jr. (2013) A new model for allosteric regulation of phenylalanine hydroxylase: implications for disease and therapeutics. *Archives of biochemistry and biophysics* 530(2):73-82.
4. Erlandsen H & Stevens RC (1999) The structural basis of phenylketonuria. *Molecular genetics and metabolism* 68(2):103-125.
5. Zhang S, Huang T, Ilangovan U, Hinck AP, & Fitzpatrick PF (2014) The solution structure of the regulatory domain of tyrosine hydroxylase. *Journal of molecular biology* 426(7):1483-1497.
6. Zhang S, Roberts KM, & Fitzpatrick PF (2014) Phenylalanine binding is linked to dimerization of the regulatory domain of phenylalanine hydroxylase. *Biochemistry* 53(42):6625-6627.
7. Roberts KM, Khan CA, Hinck CS, & Fitzpatrick PF (2014) Activation of phenylalanine hydroxylase by phenylalanine does not require binding in the active site. *Biochemistry* 53(49):7846-7853.
8. Gallivan JP & Dougherty DA (1999) Cation-pi interactions in structural biology. *Proceedings of the National Academy of Sciences of the United States of America* 96(17):9459-9464.
9. Petoukhov MV ,Franke D, Shkumatov AV, Tria G, Kikhney AG, Gajda M, Gorba C, Mertens HD, Konarev PV, and Svergun DI (2012) New developments in the program package for small-angle scattering data analysis. *J Appl Crystallogr* **45**, 342-350

Acknowledgements

Linda Stith contributed significantly towards cloning, expression, and purification of the protein used for crystallization. Dr. Ursula Ramirez and Thomas Scary optimized crystallization conditions. Dr. Mark Andrake and the modeling facility at the Fox Chase Cancer Center are acknowledged for their help in the construction of the homology models of the high-activity tetramer.