JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

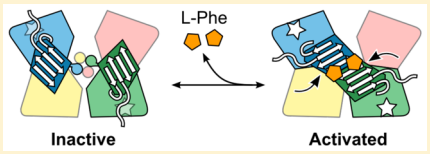
Domain Movements upon Activation of Phenylalanine Hydroxylase Characterized by Crystallography and Chromatography-Coupled Small-Angle X-ray Scattering

Steve P. Meisburger,[†] Alexander B. Taylor,[‡] Crystal A. Khan,[‡] Shengnan Zhang,[‡] Paul F. Fitzpatrick,*[,‡] and Nozomi Ando*[,†]

[†]Department of Chemistry, Princeton University, Princeton, New Jersey 08544, United States
[‡]Department of Biochemistry, University of Texas Health Science Center, San Antonio, Texas 78229, United States

Supporting Information

ABSTRACT: Mammalian phenylalanine hydroxylase (PheH) is an allosteric enzyme that catalyzes the first step in the catabolism of the amino acid phenylalanine. Following allosteric activation by high phenylalanine levels, the enzyme catalyzes the pterin-dependent conversion of phenylalanine to tyrosine. Inability to control elevated phenylalanine levels in the blood leads to increased risk of mental disabilities commonly associated with the inherited metabolic disorder, phenylketonuria. Although extensively studied, structural changes associated with allosteric activation in mammalian PheH have been elusive. Here, we examine the complex allosteric mechanisms of rat PheH using X-ray crystallography, isothermal titration calorimetry (ITC), and small-angle X-ray scattering (SAXS). We describe crystal structures of the preactivated state of the PheH tetramer depicting the regulatory domains docked against the catalytic domains and preventing substrate binding. Using SAXS, we further describe the domain movements involved in allosteric activation of PheH in solution and present the first demonstration of chromatography-coupled SAXS with Evolving Factor Analysis (EFA), a powerful method for separating scattering components in a model-independent way. Together, these results support a model for allostery in PheH in which phenylalanine stabilizes the dimerization of the regulatory domains and exposes the active site for substrate binding and other structural changes needed for activity.



INTRODUCTION

A member of the aromatic amino acid hydroxylase family, phenylalanine hydroxylase (PheH) is a nonheme iron-dependent enzyme that catalyzes the conversion of L-phenylalanine (L-Phe) to L-tyrosine using tetrahydrobiopterin (BH_4) as a cosubstrate.[1−4] In mammals, L-Phe is an essential amino acid that is acquired by dietary intake. PheH is highly expressed in the liver where it is inactive unless allosterically activated by high concentrations of L-Phe.[3,5−7] By clearing excess L-Phe, PheH also supplies tyrosine, a precursor for a number of catecholamine neurotransmitters such as dopamine.[2] Mutations in the human enzyme can have clinical consequences in the form of phenylketonuria (PKU), a relatively common genetic disorder characterized by an impaired ability to remove excess L-Phe, which can have severe and irreversible effects on cognitive development if untreated.[8] PheH thus plays a critical role in maintaining nonharmful levels of this essential amino acid in the blood.

Sequence alignments[1] and crystal structures solved in the absence of L-Phe[9,10] have revealed three functional domains in mammalian PheHs: an N-terminal regulatory domain, a central catalytic domain, and a C-terminal domain that stabilizes the tetrameric form (Figure 1A). The catalytic core is structurally homologous among all aromatic amino acid hydroxylases,[9,11−14] consistent with a common BH_4-dependent mechanism. The arrangement of subunits in the tetramer was initially established by a structure of a human construct that lacked the regulatory domain (PDB: 2PAH).[10] In this structure, the C-terminal domains bring the catalytic domains together by forming an antiparallel four-helical coiled-coil at the center of the tetramer (Figure S1A). The regulatory domain was visualized by a structure of a rat PheH construct lacking most of the C-terminal domain (PDB: 2PHM, Figure S2A).[9] The asymmetric unit of this structure contains a single copy, which forms a crystallographic dimer. Importantly, the regulatory domain was shown to consist of an N-terminal tail followed by a $\beta\alpha\beta\beta\alpha\beta$ fold that shares structural and sequence similarities with the ACT domain, a regulatory motif commonly found in enzymes involved in amino acid metabolism.[9,15−17]

PheH is known to be allosterically activated by its substrate.[5] However, the location of the allosteric binding site and the mechanism of activation have been controversial,[4−7,9,18,19] largely due to the lack of structural information. Crystal structures of PheH that include the regulatory domain have only been reported in the absence of L-Phe, when the enzyme is in the inactive conformation. In the dimeric PheH structure (2PHM),[9] the ACT domain (residues 34−118) binds near the dimer interface of the catalytic domains, placing the N-terminal tail (residues 19−33) over the active site.[9] Recently, a 2.9 Å

Received: February 11, 2016
Published: May 4, 2016

DOI: 10.1021/jacs.6b01563
J. Am. Chem. Soc. 2016, 138, 6506−6516

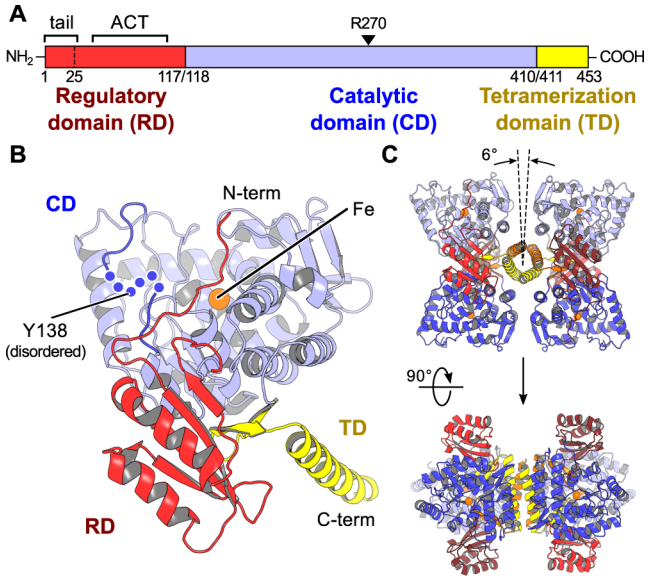


Figure 1. Crystal structure of tetrameric PheH in the absence of bound L-Phe. (A) Rat PheH contains multiple functional domains. (B) One chain of the tetrameric wt-PheH crystal structure determined in this study. The ACT motif of the regulatory domain (red) is bound to the catalytic domain (light blue) with residues 21−25 of the N-terminal tail extending over the active site, which contains the catalytic Fe. Tyr138 and its adjacent residues are disordered, rather than closed over the active site. (C) The tetramer is an asymmetrical dimer-of-dimers, suggesting flexibility within a hinge region of the tetramerization domain (yellow/orange). In the absence of bound L-Phe, there are two regulatory domains (red/maroon) on each face of the tetramer. The regulatory domains do not contact each other. Coordinates have been deposited in the PDB with accession codes 5FGJ (wt-PheH) and 5EGQ (PheH$^{\Delta 24\ R270K}$).

crystal structure of the full-length rat PheH was reported depicting the inactive tetramer (PDB: 5DEN).[20] The tetramer is arranged as a dimer-of-dimers, each resembling the dimeric structure (2PHM),[9] placing the regulatory domains on each side of the tetrameric assembly far removed from each other. This separation is unusual for ACT domains, which have been observed in other enzymes as oligomers, with the canonical form being a side-by-side dimer with an extended β sheet across the dimer interface[16] (Figure S3A). Structural changes have been shown to occur upon incubation with L-Phe by a variety of biophysical methods, including hydrogen/deuterium exchange mass spectrometry[21] and small-angle X-ray scattering (SAXS).[20] However, these data have been challenging to interpret, leaving the structure of PheH in the activated conformation open to speculation.

In the absence of a structure for the active PheH tetramer, current models for allosteric activation are largely divided on whether a large-scale movement of the ACT domains is thought to occur in response to L-Phe. In one set of models, allosteric activation by L-Phe does not require dimerization of the regulatory domains.[4,9,14,19,22] Instead, binding of L-Phe either at the active site[14,22] or another site[4,9,19] is thought to trigger local changes throughout the catalytic domain and thereby activate the enzyme. Computational docking studies on the 2PHM structure have suggested that L-Phe binds at the interface of the regulatory and catalytic domains, and hence the regulatory domains remain docked in both the inactive and activated states.[4,9,19]

By contrast, others have proposed that a large-scale movement of the regulatory domains is required to create an allosteric site for L-Phe.[7,18,23] In support of this model, a hydrogen/deuterium exchange study demonstrated that L-Phe alters the interaction between the regulatory and catalytic domains.[21] Jaffe et al.[18] were first to propose that the regulatory domains pivot to form dimers in a side-by-side arrangement similar to the ACT domain dimers in the structures of phosphoglycerate dehydrogenase (PDB ID: 1PSD)[24] and prephenate dehydratase (PDB ID: 2QMX)[25] (Figure S3A). Recent experiments on rat PheH support this dimerization model. First, the allosteric binding site for L-Phe was shown by mutagenesis and fluorescence quenching to be distinct from the active site.[26] In this study, an R270K mutation, which has been associated with PKU in humans,[27] was shown to abolish L-Phe binding in the active site in the full-length enzyme (K_m > 0.5 M), while retaining the ability to be activated by L-Phe. Furthermore, recent nuclear magnetic resonance (NMR) and analytical ultracentrifugation studies establish that a PheH construct containing only the regulatory domain undergoes a monomer−dimer transition in the presence of L-Phe[28,29] and that two molecules of L-Phe bind at a canonical ACT-domain dimer interface.[30] However, dimerization of the regulatory domains has not been visualized in the full-length enzyme.

Here, we examine rat PheH, which shares 92% sequence identity with the human enzyme. We present two crystal structures: the first depicts wt-PheH in the inactive conformation including all of its domains, and the second shows the effects of deletion of the autoinhibitory N-terminal tail and the R270K mutation. As crystals of the activated conformation of PheH have proven elusive, we describe the structural changes accompanying allosteric activation using two solution techniques, SAXS and isothermal titration calorimetry (ITC). To overcome the well-known tendency of PheH to form aggregates, particularly in the presence of L-Phe, we performed size exclusion chromatography (SEC) in-line with SAXS data collection. High quality SAXS profiles were obtained by applying a new analysis approach based on evolving factor analysis (EFA) to separate eluting species in a model-independent manner. Together, SAXS, crystallography, and ITC demonstrate that activation by L-Phe involves a large conformational change in the full-length PheH. This conformational change occurs even when active-site binding by L-Phe is effectively abolished by the R270K mutation. We propose an equilibrium model that accounts for the observed cooperativity of activation according to SAXS and stoichiometry of binding obtained from ITC. Finally, the SAXS data in the presence of L-Phe provide the first experimental constraints on the domain architecture of PheH in the activated conformation. These data are consistent with a ~90° rotation of the regulatory domains from their docked position in the inactive state to form canonical ACT domain dimers diagonally across the tetramer interface, releasing the N-terminal autoinhibitory tail from the active site.

RESULTS

Structure of the Inactive PheH Tetramer. To gain insight into the role of the N-terminal tail, we obtained crystal structures of two constructs: full-length rat PheH (wt-PheH) and a construct with the R270K active-site mutation and the first 24 residues of the N-terminal autoinhibitory tail removed (PheH$^{\Delta 24\ R270K}$). Crystals of the two constructs were obtained under highly similar conditions in the presence of 1 mM L-Phe.

DOI: 10.1021/jacs.6b01563
J. Am. Chem. Soc. 2016, 138, 6506−6516

Structures of wt-PheH and PheH$^{\Delta 24\ R270K}$ were solved to 3.6 and 2.5 Å resolution, respectively, using the truncated rat PheH structure (PDB: 2PHM)[9] as the search model for molecular replacement (Table S1). Consistent with the highly similar crystallization conditions, the two constructs form similar lattices ($P2_12_12_1$ space group), giving rise to structures that are superimposable (root-mean-square deviation (RMSD) of 1.1 Å for all C_α atoms, Figure S4A). The effect of the R270K mutation is localized to the loop containing Thr278; substitution of Arg270 with Lys270 not only eliminates ionic interactions that support binding of L-Phe in the active site,[26] but also introduces a new hydrogen bond with the backbone, which in turn appears to favor an alternate conformation of the Thr278 loop (Figure S4B,C). One active-site Fe is bound per monomer in our wt-PheH structure, whereas the Fe was lost during crystallization of PheH$^{\Delta 24\ R270K}$ (Figure S5). Although present in our crystallization conditions, electron density for L-Phe was not observed. More compellingly, we observe no L-Phe in any of the proposed allosteric sites, including the interface between the regulatory and catalytic domains,[4,7,9,13,18,19,23] and thus our structures represent the inactive conformation of the PheH tetramer.

As expected from the lack of bound L-Phe, our structures are highly similar to the recently reported 2.9 Å structure of inactive wt-PheH (PDB: 5DEN) that was obtained in a different space group (C121) in the absence of L-Phe (RMSD of 0.9 Å for all C_α atoms in the two wt-PheH structures).[20] Like the 5DEN structure, our structures resemble an asymmetric dimer of dimers that is held together by a central coiled coil formed by the tetramerization domains (Figure 1B,C). The asymmetric assembly is distinct from that observed in a previous tetrameric structure of a human PheH construct lacking the regulatory domain (PDB: 2PAH)[10] (Figure S1A,B). In the latter, the asymmetric unit consists of two chains with the catalytic domains extending from the coiled coil in opposing directions, and the complete tetramer is generated by 2-fold symmetry that results in symmetrically arranged dimers (Figure S1A). By contrast, the two dimers in our structures are arranged relative to each other with a ~6° angle, conferring a V-shaped asymmetry to the tetramer (Figures 1C and S1B). Additionally, the central coiled coil is rotated by ~15° relative to the two dimers (Figure S1B), which results from multiple backbone conformations of residues 424−427 in a previously identified hinge region of the tetramerization domain[10] (Figure S1A,B, right). Importantly, this hinge motion has now been observed in two different crystal forms, providing compelling support for the flexible nature of the PheH tetramer.

Within each dimeric unit of our structures, the arrangement of the regulatory and catalytic domains is superimposable with previous structures determined in the absence of L-Phe[9,20] (Figure S2B). Namely, residues 34−111 of each regulatory domain adopt $\beta\alpha\beta\beta\alpha\beta$ ACT-domain folds that are docked at the interface between the catalytic domains on either side of the dimer, with contacts being made by the β2 strand of the ACT motif on one chain and the catalytic domain of the other chain (Figure S3B). In our structures, the ACT domains are further stabilized in these positions by contacts with neighboring tetramers in the lattice. Notably, the ACT domains make no contact with each other within the tetramer (Figure 1C, red/maroon).

Residues 21−25 of the N-terminal tail in our wt-PheH structure are positioned across the active site (Figure 1B), partially occluding access,[31] whereas the first 20 residues (including the phosphorylation site Ser16) are disordered. The same residues are disordered in the previous dimeric structure[9] and the recent 5DEN structure[20] obtained in the absence of L-Phe. In our wt-PheH structure and in all previous structures that include the regulatory domain, residues 21−25 of the N-terminal tail are found above the BH_4 binding site, whereas the L-Phe binding site remains exposed. However, despite the presence of L-Phe in the crystallization solution, no L-Phe is observed in our structure of wt-PheH (Figure S5A), suggesting that a conformational change in the active site is required before L-Phe can bind.

Crystal structures of the catalytic domain that include BH_4, an L-Phe analogue, and Fe in the +2 oxidation state depict a large loop movement (residues 134−139) that brings Tyr138 from a surface position on the catalytic domain to a buried location in the active site, where it participates in a water-mediated hydrogen bond with BH_4.[14,32] We find that Tyr138 and its adjacent residues in this loop are disordered in both of our structures, rather than closed over the active site. A superposition of the closed-loop structure (PDB: 1MMK)[14] with our wt-PheH structure shows a possible steric clash between the N-terminal tail of the regulatory domain and the closed loop containing Tyr138, suggesting that the N-terminus may inhibit activity by preventing loop movement, rather than by blocking substrate access to the active site. However, the conformation of the Tyr138 loop is unaffected by deletion of the N-terminal residues in our structure of PheH$^{\Delta 24\ R270K}$, arguing against a steric mechanism involving only this loop. Instead, the proximity of the N-terminal tail to the BH_4 binding site that was previously observed in the PheH-BH_4 binary complex[32] suggests that its presence may perturb BH_4 binding,[31] which would in turn prevent the conformational changes that are needed to complete the active site pocket,[14] such as the closure of the Tyr138 loop and the loops containing Thr380 and Pro279. An additional clue pointing to BH_4's central role in allosteric activation is the observation of an ordered kinetic mechanism in which BH_4 binds before L-Phe,[33] despite the fact that L-Phe and BH_4 do not interact directly in the precatalytic ternary complex.[32] Thus, our structures support a specific order in which allosteric activation must occur: the N-terminal tail first unbinds from the active site to allow for ordered substrate binding, which is accompanied by the movement of the Tyr138 loop and compaction of the catalytic domain to a closed conformation that completes the active site.

Phenylalanine Induces a Cooperative Conformational Change. To examine the effect of L-Phe on the structure of the tetramer, we employed SAXS, a solution technique that is uniquely suited for the quantitative investigation of conformational changes associated with protein allostery.[34] To disentangle the potential effects of active site and allosteric site binding, we compared the scattering of three constructs: rat PheH (wt-PheH), the R270K mutant with compromised active site binding (PheH^{R270K}), and the same mutant with its N-terminal tail deleted (PheH$^{\Delta 24\ R270K}$). Samples were prepared at 25 μM (monomer concentration) and incubated at room temperature in the presence of varying concentrations of L-Phe. Scattering was measured simultaneously on two detectors and merged to produce scattering profiles (I vs q) over a wide q-range, where I is the scattering intensity, $q = 4\pi \sin(\theta)/\lambda$, 2θ is the scattering angle, and λ is the X-ray wavelength. The radius of gyration (R_g) was estimated by Guinier analysis and the

DOI: 10.1021/jacs.6b01563
J. Am. Chem. Soc. 2016, 138, 6506−6516

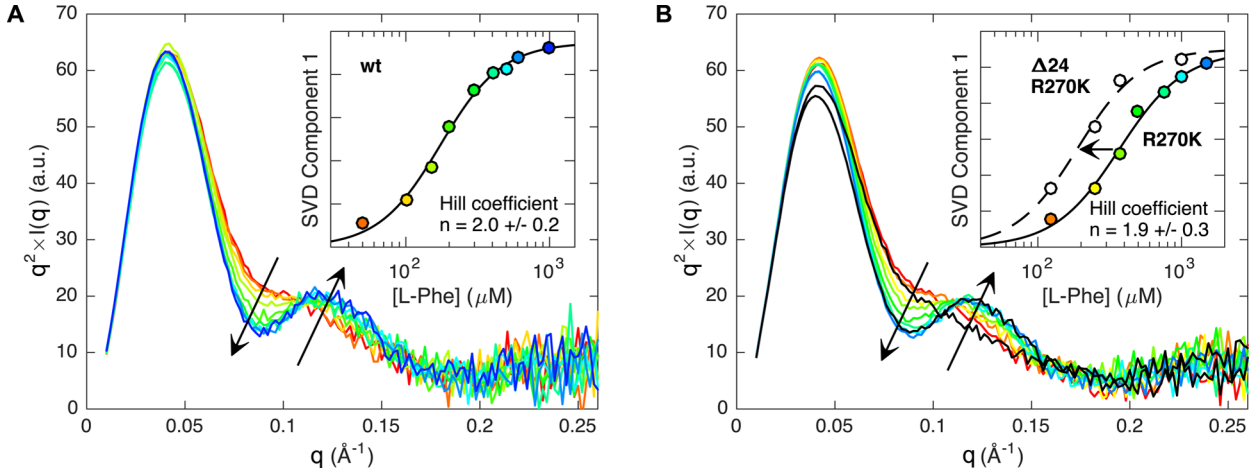


Figure 2. Addition of phenylalanine leads to a cooperative structural change in PheH. (A) Titration of 0−1 mM L-Phe to wt-PheH leads to a change in the SAXS profile in the mid-q region (red to blue). This cooperative change can be described by a Hill equation (eq 1) with an exponent of 2 (inset: solid line is the fit). (B) Titration of 0−1.5 mM L-Phe to PheH^{R270K} leads to a similar structural change (colored curves) also with a Hill coefficient of ∼2 (colored circles in inset: solid line is the fit). A similar titration was done with PheH$^{\Delta 24\ R270K}$ (black curves are at 0 and 1 mM L-Phe). The structural transition for PheH$^{\Delta 24\ R270K}$ is shifted toward lower [L-Phe] relative to PheH^{R270K} (open circles in inset: dashed line shows the fit to eq 1 with n = 2 fixed).

indirect Fourier transform method, and the mass was estimated by Porod analysis.[35,36]

First, we examined the scattering of wt-PheH, which is allosterically activated by L-Phe within the physiologically relevant range of 0−1 mM.[5] Over this concentration range, systematic changes are observed in the mid-q region (0.05−0.20 Å^{-1}) that are best visualized in a Kratky representation of the data (Iq^2 vs q) (Figure 2A). An iso-scattering point is observed at a q-value of 0.115 Å^{-1}, indicating that the data can be described as linear sums of interconverting species. These scattering profiles have similar forward scattering, $I(0)$, indicating that no major changes in oligomeric state occur in this range (Figure S6A, bottom set). However, in several profiles, slight upturns at low-q reveal the presence of aggregation. The amount of aggregation appears to vary at random among the identically prepared samples, suggesting that they represent aggregates that form irreversibly during normal handling, as documented previously.[18,37] Furthermore, the apparent molecular weights determined for these profiles are 10−25% larger than expected for tetrameric PheH (Figure S6B). Above 1 mM L-Phe, no further changes are observed in the mid-q region, but the forwarding scattering intensity and the apparent molecular weight increase dramatically (Figure S6, top set). Overall, we observe a conformational change consistent with domain rearrangement within tetrameric PheH that occurs between 0 and 1 mM L-Phe, while higher concentrations of L-Phe promote the formation of aggregates or high-order oligomers.

To determine the minimum number of structural states involved in the conformation change between 0 and 1 mM L-Phe, we applied singular value decomposition (SVD). SVD decomposes the scattering data into orthonormal components, known as singular vectors, that are ranked in order of significance by the singular values. Together, the singular values and singular vectors reconstruct the data set: $A = USV^T$, where A is a matrix whose columns are the data, U and V are orthonormal matrices whose columns are the left and right singular vectors respectively, and S contains the singular values along the diagonal in decreasing order. The columns of U correspond to basis states that are added together to produce each SAXS profile in A, while the columns of V show how the coefficients for each basis state vary between the samples. To prepare the data matrix A, difference profiles were produced by subtracting the scattering profile collected at 0 mM L-Phe from the remaining curves, and the difference profiles were weighted by the experimental error (Figure S7A). The matrix A was decomposed into singular vectors and values using the SVD function in MATLAB.

If the conformational change is two-state, we expect the difference-profile SVD to have one significant singular value with a corresponding right singular vector that is proportional to the change in the fraction of the final state. However, SVD of the wt-PheH data gives two major singular values, rather than one (Figure S7B). The first component's basis state corresponds to the dominant changes in scattering, which occur in the mid-q region (Figure S7, gray region), and the coefficients for this state vary systematically with [L-Phe] (Figure S7B, N = 1). The second component reports changes in the low-q region that correspond to the nonsystematic appearance of aggregates observed with Porod analysis (Figure S7B, N = 2). Thus, the filtering effect of SVD allows us to focus on the systematic change in structure with [L-Phe] by analysis of the first component in isolation. This component can be described as a two-state process using a Hill equation,

$$y([\text{L-Phe}]) = \frac{y_\infty([\text{L-Phe}]/[\text{L-Phe}]_{0.5})^n}{1 + ([\text{L-Phe}]/[\text{L-Phe}]_{0.5})^n} \quad (1)$$

with an exponent of n = 2.0 ± 0.2 and midpoint L-Phe concentration $[\text{L-Phe}]_{0.5}$ = 174 ± 9 μM (Figure 2A, inset), in agreement with measurements of allosteric activation by other methods.[37,38] Together, these results demonstrate that physiologically relevant levels of L-Phe cause a cooperative transition between two structurally distinct conformations.

Next, we examined the effect of L-Phe on a PheH mutant with compromised active-site binding, PheH^{R270K}. When plotted in the Kratky representation, the SAXS profiles determined at 0 and 1 mM L-Phe strongly resemble the corresponding SAXS profiles of wt-PheH, indicating that their conformations in solution are similar. Furthermore, the

DOI: 10.1021/jacs.6b01563
J. Am. Chem. Soc. 2016, 138, 6506−6516

distinctive [L-Phe]-dependent change in mid-q scattering that occurs in wt-PheH is also observed in PheH^{R270K} (Figure 2B, colored curves). When titration data are analyzed using SVD, an apparent two-state transition is observed with the same cooperativity as wt-PheH (n = 1.9 ± 0.3) (Figure 2B, colored circles in inset). The midpoint of the transition for the PheH^{R270K} mutant is shifted relative to the wild-type enzyme, occurring at a [L-Phe] of 370 ± 35 μM. A similar shift was previously observed by tryptophan fluorescence measurements of the same construct.[26] As a superposition of our PheH$^{\Delta 24\ R270K}$ and wt-PheH structures shows no significant change beyond the immediate vicinity of the mutation (Figure S4), the structural basis for this shift remains unclear. Crucially, however, the shift in $[\text{L-Phe}]_{0.5}$ due to the R270K mutation is small compared with its effect on the K_m for active site binding,[26] which exceeds the concentration of L-Phe seen to activate the enzyme by at least 2 orders of magnitude. Thus, the SAXS data show unambiguously that the conformational change associated with activation does not require binding of L-Phe to the active site. Rather, the conformational change is likely the direct result of L-Phe binding to one or more distinct allosteric sites.

Finally, the role of the N-terminal tail in allosteric activation was investigated by SAXS measurements of the PheH$^{\Delta 24\ R270K}$ construct as [L-Phe] was varied from 0 to 1 mM. The Kratky plots show that PheH$^{\Delta 24\ R270K}$ exhibits very similar conformational changes to both wt-PheH and PheH^{R270K} (Figure 2B, black curves). SVD analysis shows a clear shift in [L-Phe] required for activation, toward lower values relative to PheH with the R270K mutation alone. For this mutant, the number of SAXS profiles acquired as part of the titration was too few to determine the Hill coefficient uniquely, however the value of n = 2 obtained for wt-PheH and PheH^{R270K} provides a good fit with the PheH$^{\Delta 24\ R270K}$ data (Figure 2B, open circles in inset). The midpoint concentration estimated by fitting a Hill equation with n = 2 gives $[\text{L-Phe}]_{0.5}$ ~ 190 μM, which is 2-fold lower than the value obtained for the full-length R270K construct. Thus, the N-terminal tail likely stabilizes the inactive conformation relative to the active conformation, consistent with the proposal that the N-terminal tail unbinds from its observed location in the catalytic domain as part of the allosteric mechanism of activation.

Stoichiometry of L-Phe Binding to the Regulatory Domains. To investigate the role of the regulatory domains in the [L-Phe]-dependent structural change observed by SAXS, we prepared a construct of the isolated regulatory domain that lacks the first 24 residues of the flexible N-terminal tail (RDPheH$^{\Delta 24}$), corresponding to residues 25−117 in the full-length enzyme. This construct has been shown by analytical ultracentrifugation (AUC) to have a monomer−dimer equilibrium in solution, where titration of L-Phe favors the dimeric form.[28] To establish the stoichiometry and binding affinity of L-Phe, we performed isothermal titration calorimetry (ITC) measurements on RDPheH$^{\Delta 24}$. In this experiment, the enthalpy change (ΔH) has contributions from both dimerization and L-Phe binding. To determine binding stoichiometry, the concentration of RDPheH$^{\Delta 24}$ was chosen to be close to the dissociation constant (K_d) for L-Phe of 8 μM, as determined previously by AUC,[28] where approximately 80% of the domains exist as monomers in the absence of L-Phe. The time-dependent ITC traces were integrated to obtain the heat exchanged per mole of injected L-Phe. The resulting thermogram is well-described by a model for identical binding sites with a K_d of 15.2 ± 1.1 μM, ΔH of −10.6 ± 0.4 kcal mol^{-1}, and number of sites per monomer (n) of 0.9 ± 0.1 (Figure 3).

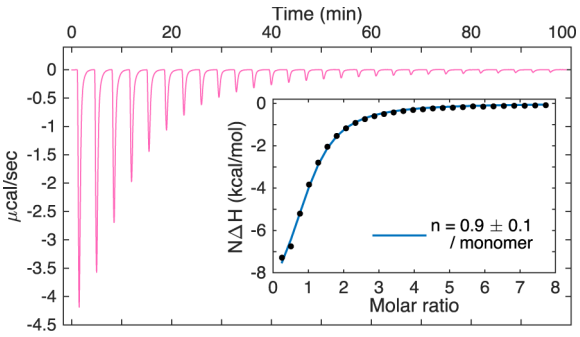


Figure 3. Stoichiometry of L-Phe binding determined by ITC. Titration of L-Phe to the isolated regulatory domain (RDPheH$^{\Delta 24}$) leads to heat changes that are consistent with one L-Phe per monomer, or two L-Phe per ACT dimer.

These results are in agreement with the values obtained by AUC[28] as well as a recent NMR study that shows that RDPheH$^{\Delta 24}$ forms a canonical side-by-side ACT domain dimer[30] similar to that observed in prephenate dehydratase[25] (Figure S3A), with the β2 strand of each monomer forming an extended β-sheet across the dimer interface. Two equivalent binding sites for L-Phe were identified in this study, each requiring residues from both chains in the ACT dimer to form.[30] Interestingly, in the crystal structures of the inactive PheH determined in the present study, the β2 strand makes contacts with the catalytic domain (Figure S3B). This observation strongly suggests that activation of the full-length PheH by L-Phe involves the undocking of the ACT domains from the catalytic domains followed by their dimerization to form two allosteric sites at the interface.

Chromatography-Coupled SAXS Yields Pure Scattering Components. As polydispersity can complicate structural interpretation of SAXS data, we applied an emergent technique known as size-exclusion chromatography-coupled SAXS (SEC-SAXS), where purification is performed in-line with X-ray data collection.[39,40] The SEC column was coupled directly to the SAXS flow cell and scattering data were collected continuously during elution (Figure 4A). Two data sets were acquired with 0 and 1 mM L-Phe in the running buffer. For each sample, ~600 μM wt-PheH (monomer concentration) was injected onto a pre-equilibrated column, and roughly 400 sequential profiles were collected at a flow rate of 0.075 mL/min. Profiles collected prior to the elution peak were averaged and used as a buffer blank to obtain the background-subtracted intensity (Figure 4B). Importantly, the scattering profiles before and after the main peak are superimposable, confirming that the buffer blank remains an excellent match during elution (Figure S8).

In the absence of L-Phe, we observe a single peak when the extrapolated forward scattering $I(0)$ is plotted as a function of frame number during elution (Figure 4C, dashed line). However, the apparent R_g decreases during elution, suggesting that the main peak is composed of multiple species whose elution profiles overlap (Figure S9A). Consistent with this result, SVD of the SEC-SAXS data set yields 3 significant singular values (Figure S10). Similar results were obtained for the SEC-SAXS data collected with 1 mM L-Phe. Again, we observe a single elution peak with varying R_g (Figure S9B).

DOI: 10.1021/jacs.6b01563
J. Am. Chem. Soc. 2016, 138, 6506−6516

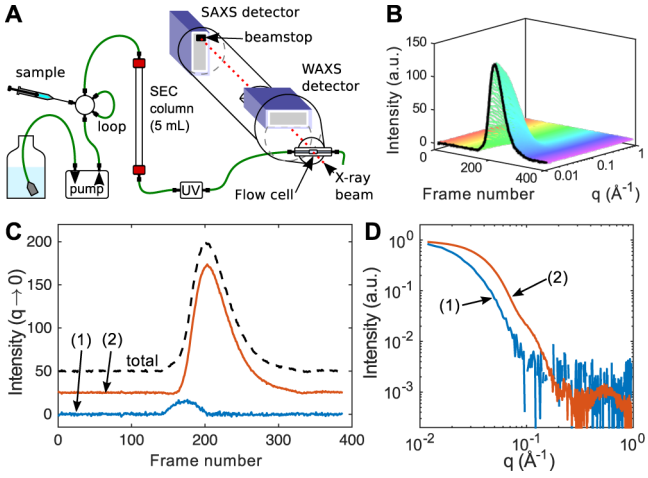


Figure 4. Separation of wt-PheH oligomers by SEC-SAXS. (A) Scattering images were collected continuously during elution. (B) The background-subtracted intensity for wt-PheH in 0 mM L-Phe is plotted as a function of number and momentum transfer, q. (C) SVD with EFA shows that the peak is composed of two overlapping components. The forward scattering intensities of each component are offset for clarity. (D) The normalized scattering profiles of components 1 and 2 were determined by EFA.

SVD of this data set also yields 3 significant values (Figure S11).

The problem of separating the scattering from overlapping peaks in SEC-SAXS data has been addressed using SVD;[41,42] however, previous efforts relied on models for the shape of either the elution peaks[41] or of the scattering profiles.[42] To analyze the data in a model-free way, we adapted a powerful variant of SVD known as evolving factor analysis (EFA) (Figure S12).[43] EFA allows for the identification of ranges within each elution peak where separate species elute from the abrupt changes in the number of significant singular values as scattering profiles are added or removed from the matrix (Figure S13–14). Once these peak ranges are identified, the right singular vectors can be rotated and converted into physically meaningful elution peaks. This rotation is determined automatically by requiring the concentration of each component outside its range to be zero; the same rotation can be applied to the left singular vectors to yield the corresponding scattering profiles in a model-independent manner (Figure S12–S14).

Using EFA, we resolve two overlapping peaks corresponding to protein scattering in both SEC-SAXS data sets, followed by a third species that corresponds to a buffer component (Figure 4C,D, Figure S13–S14). In the absence of L-Phe, a minor high-mass species elutes first (Figure 4C, blue), followed by the major species with an R_g of 40.3 ± 0.1 Å (Figure 4C, red). The apparent molecular weight of the major species was estimated to be ~210 kDa by Porod analysis, in good agreement with the expected mass for the tetramer (207 kDa). In the presence of 1 mM L-Phe, we again observe a minor high-mass species followed by a major species (R_g of 41.4 ± 0.1 Å, ~230 kDa) that is consistent with a tetramer (Figure S14). The R_g values we obtain for the inactive state and activated state are highly similar to previously reported values.[44]

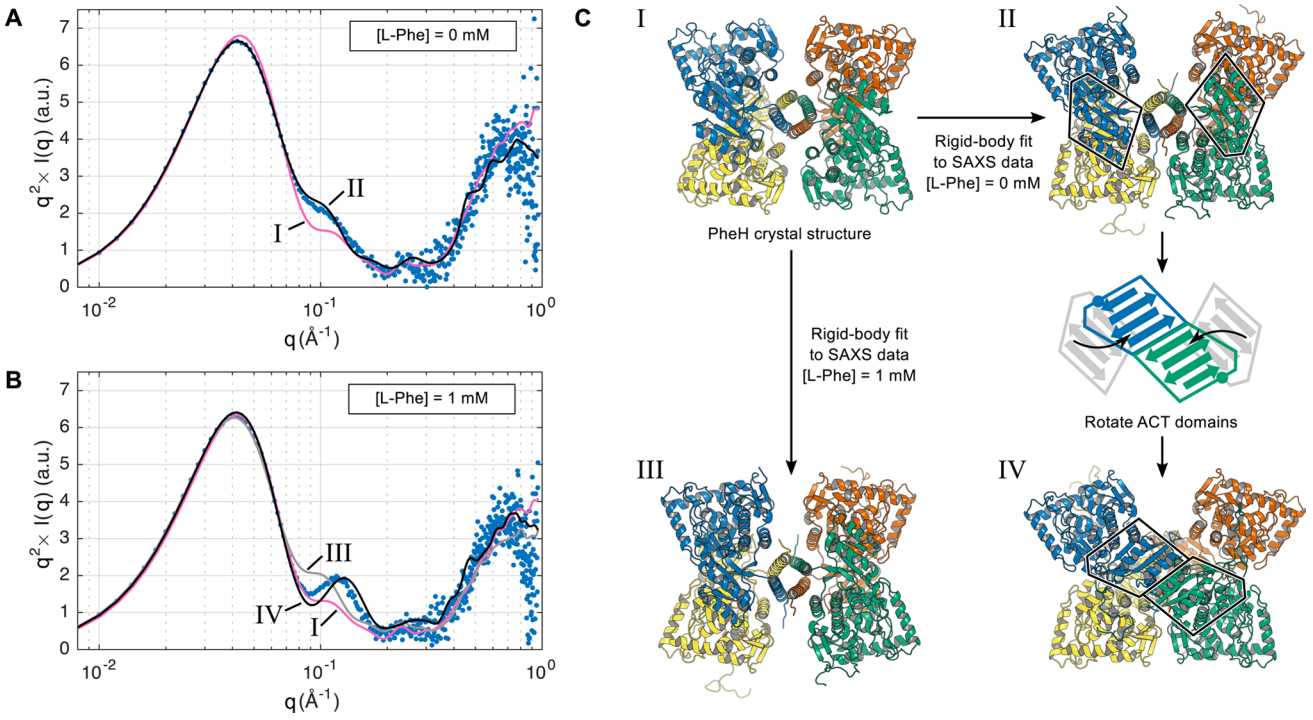


Figure 5. SAXS-based modeling of domain motion during activation of wt-PheH. (A) The EFA-separated scattering profile obtained at 0 mM L-Phe is shown as blue circles in Kratky representation (Iq^2 vs q). The crystal structure of PheH (model I, magenta line) agrees with the overall profile shape and is improved by rigid-body optimization of the dimer-of-dimer arrangement (model II, black line). (B) The EFA-separated scattering profile obtained in 1 mM L-Phe is shown as blue circles. An oscillatory feature at mid-q is not predicted by the crystal structure (model I, magenta line). The fit is not improved by rigid-body optimization with the regulatory domains docked to the catalytic domains (model III, gray line). However, the oscillatory feature is predicted when the regulatory domains are rotated to form an ACT domain dimer (model IV, black line). (C) Models I–IV of tetrameric PheH are shown as cartoons and colored by chain.

DOI: 10.1021/jacs.6b01563
J. Am. Chem. Soc. 2016, 138, 6506–6516

The minor high-mass species appearing in both data sets have similar average sizes, with R_g values of 66 ± 1 Å and 69 ± 1 Å at 0 and 1 mM L-Phe, respectively. Each has an apparent molecular weight of ~660 kDa according to Porod analysis. Although the consistent size and mass of this component in both data sets suggests a well-defined oligomeric state (~12 subunits), there is no known physiological role for PheH oligomers of greater than four subunits. The mass fraction of high-mass species in each sample was estimated from the areas under the EFA-derived elution peaks and the apparent molecular weight of each component. We find that the high-mass species accounts for 2% of the protein mass in the 0 mM L-Phe sample and increases to 4% in 1 mM L-Phe, consistent with previous reports of the greater tendency of PheH to aggregate in the presence of L-Phe.[18,37]

Structural Models of Inactive and Activated PheH in Solution. With the aggregation-free scattering profiles of PheH determined by EFA, the inactive and activated conformations in solution were examined. We began by testing whether the inactive conformation of PheH in the crystal is also observed in solution in the absence of L-Phe. The theoretical scattering profile of the wt-PheH crystal structure was calculated using CRYSOL[45] and fit to the EFA-separated scattering profile at 0 mM L-Phe (model I in Figure 5A,C and Figure S15A). The quality of the fit was assessed using the reduced χ ($\chi = \sqrt{\chi^2}$ in eq 3). We find that the scattering profile predicted from the crystal structure captures many of the essential features of the experimental scattering. However, it falls below the data at mid-q (0.064−0.15 Å^{-1}), giving a suboptimal fit (χ = 5.7). The fit is minimally affected when the missing N-terminal residues and disordered loops in the crystal structure are explicitly modeled (Figure S15A, red). In particular, the conformation of the N-terminal tail has an insignificant effect on the shape of the scattering profile at mid- and high-q (Figure S16), and thus cannot account for the structural differences we observe in the crystal and in solution.

Because scattering in the mid-q region can be sensitive to subunit arrangement,[34] we repeated the modeling, allowing for flexibility in the hinge region of the tetramerization domain (residues 424−427) that was observed to adopt multiple conformations in our crystal structures (Figure S1B, right). Restrained optimization of the tetrameric structure was performed in SASREF[45] treating the two dimeric units (residues 1−424) and the central coiled-coil of the tetramerization domain (residues 427−453) as separate rigid bodies. The best-fit model shows excellent agreement with the data (χ = 2.1) and has an overall conformation that is similar but with a more pronounced V-shaped asymmetry (~20°), leading to a slightly expanded tetramer relative to the crystal structure (model II in Figure 5A,C). This result is additionally supported by comparison of the pair-distance distribution functions, $P(r)$, calculated for the crystal structure and SAXS data. $P(r)$ for the PheH tetramer has a full width at half-maximum (FWHM) of 61 Å (Figure S15B, blue), slightly wider than the crystal structure's FWHM of 56 Å (Figure S15B, black/red). Thus, under the conditions of the SAXS experiment, the domain arrangement of inactive wt-PheH in solution appears similar to the crystal structure, while the absence of crystal lattice contacts appear to allow for an increase in the angle between the dimers.

Two classes of models for activated PheH have been proposed. In the first, the regulatory domains remain docked to the catalytic domains. In the second, the regulatory domains undock and form ACT domain dimers. The resolution of our SAXS data is sufficient to test whether the first, or the second, model is more likely to be correct. To test the first model, we performed rigid-body optimization of the crystal structure, as described above for the inactive conformation, to the EFA-separated scattering profile of wt-PheH in 1 mM L-Phe, again modeling a flexible hinge in the tetramerization domain. This optimization yields a poor fit to the scattering profile (χ = 5.8), especially in the mid-q region (model III in Figure 5B,C). Thus, flexibility in the tetramerization domain alone is not sufficient to explain the scattering profile of activated PheH.

To test whether dimerization of the regulatory domains can account for the change in the scattering of activated PheH, we first constructed a model that includes canonical ACT dimers as previously proposed.[18] Starting from the optimized SAXS model for 0 mM L-Phe, each ACT motif of the regulatory domain (residues 34−110) was reoriented to form canonical ACT dimers diagonally across the tetramer interface. Here, the ACT domain dimer of PDT (PDB: 2QMX)[25] was used as a template for aligning the ACT domains of PheH.[23,30] The missing N-termini and linker regions between the regulatory and catalytic domains (residues 111−117) were added using Modeller.[46] The theoretical scattering profile for this model shows the characteristic mid-q oscillation in the Kratky representation that is associated with activation by L-Phe in the SAXS data (model IV in Figure 5B,C). Importantly, the ACT domain dimer model shows significantly better agreement with the data (χ = 3.8) than the model in which the domains are docked, despite having no optimization of the domain positions to improve the fit. Thus, dimerization of the regulatory domains alone can account for most of the changes in the SAXS profile that characterize cooperative activation. In summary, the SAXS data are consistent with regulatory domain dimerization in the activated state and are inconsistent with the alternate model in which they remain docked to the catalytic domains during activation.

■ DISCUSSION

Mammalian PheH is responsible for keeping the concentration of L-Phe in the blood below potentially harmful levels. To accomplish this task, PheH switches between inactive and active conformations in response to substrate binding. Despite decades of study, the details of this allosteric mechanism, and the role of the ACT domain in particular, have been controversial due to a lack of structural data for the activated full-length enzyme. In this study, we used three techniques to gain insight into the structural basis for allosteric activation by L-Phe. Our crystal structures of the inactive PheH depict a flexible tetramer, in which the ACT domains are docked against the catalytic domains (Figure 1C). Using SAXS and ITC, we find that allosteric activation leads to a cooperative structural change that is consistent with the regulatory domains undocking and rotating to form a canonical ACT dimer with two allosteric sites for L-Phe. Importantly, the cooperativity is maintained even when L-Phe binding to the active site is prevented by an R270K mutation, thus establishing that binding of L-Phe to an allosteric site leads to the observed structural change. We further show with SAXS that unbinding of the autoinhibitory N-terminal sequence is a key step in allosteric activation. Together, our results suggest that rotation and subsequent dimerization of the ACT domains displaces the N-terminal tails that extend over the active sites in the inactive tetramer; this displacement allows substrate binding and further

DOI: 10.1021/jacs.6b01563
J. Am. Chem. Soc. 2016, 138, 6506−6516

structural changes in the catalytic domain, such as the closing of the Tyr138 loop,[14] that complete the active site.

On the basis of accumulating evidence in the literature,[21,26,28,30] as well as our SAXS and ITC data, we propose a minimal ligand-binding model for the allosteric regulation of PheH by L-Phe. The model assumes that the regulatory domains exist in equilibrium between undimerized and dimerized states, and equilibrium is shifted toward the dimerized state by the independent binding of L-Phe to two identical sites per ACT dimer (Figure 6). If each dimer is

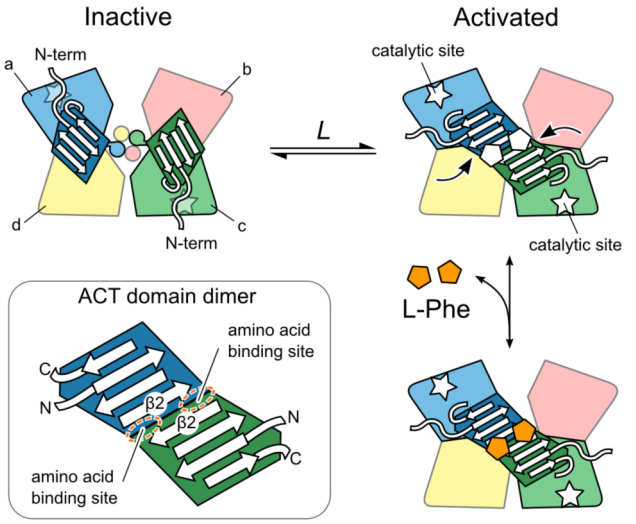


Figure 6. Model for allosteric activation of PheH. The regulatory domains are docked to the catalytic domains in the inactive tetramer. Dimerization of the regulatory domains creates two binding sites for L-Phe in a manner similar to other ACT dimers[16,24,25] (inset). The β strands of the $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$ ACT domain fold are drawn as arrows, while α helices are omitted for clarity.

independent of the other, the fraction of subunits in the active state depends on the equilibrium constant for undocking and dimerization of a regulatory domain pair (L) and the dissociation constant for L-Phe binding at each allosteric site (K_d),

$$f_{\text{active}} = \frac{L(1 + K_d^{-1}[\text{L-Phe}])^2}{1 + L(1 + K_d^{-1}[\text{L-Phe}])^2} \quad (2)$$

The limiting form of this expression when [L-Phe] $\gg K_d$ is a Hill equation with an exponent of 2 (eq 1), in excellent agreement with our SAXS titration data (Figure 2, insets) and prior biochemical studies.[37,38] The model provides an elegant explanation for the cooperativity observed in allosteric regulation of PheH: the Hill coefficient of 2 comes directly from the presence of 2 binding sites in each ACT domain dimer. The L-Phe concentration at half-saturation ($[\text{L-Phe}]_{0.5}$) occurs when $[\text{L-Phe}]_{0.5} = K_d(L^{-1/2} - 1) \sim K_d L^{-1/2}$. Combining the K_d obtained from ITC and $[\text{L-Phe}]_{0.5}$ obtained from SAXS, we estimate $L \sim 6.5 \times 10^{-3}$, in accord with previous estimates of <1% activity in unactivated PheH.[5] Furthermore, the dynamic nature of the regulatory domains in this equilibrium model also explains how the inactive tetramer could be crystallized even in the presence of L-Phe.

Previously, a morpheein model for allosteric regulation of PheH was proposed by Jaffe et al.[18] that also involves dimerization of the ACT domains, but differs fundamentally from the equilibrium model given here. In morpheein models generally, a conformational change important for regulation is prevented by steric constraints in a multimeric assembly, and cannot occur without dissociation.[18,47] Specifically in the case of PheH, the inactive tetramer is proposed to dissociate into a dimeric intermediate that changes into an active conformation *before* reassociating to form the activated tetramer. However, our results argue for a simpler, nondissociative pathway (Figure 6). First, we observe no involvement of a dimeric intermediate in our SAXS titration data; they are fully consistent with a two-state transition between tetrameric species. In addition, the cooperativity of activation (Hill coefficient of 2) is shown to result from the pair of binding sites formed when the ACT domains dimerize. This degree of cooperativity would not be observed in the morpheein model, because allosteric binding sites are not present in the dimeric intermediate, and therefore the association of dimers in the activated conformation would simultaneously form all four allosteric L-Phe sites. As a consequence, the morpheein model predicts a greater cooperativity (Hill coefficient of 4) than is observed. Finally, the morpheein model's supposed advantage is its ability to explain how all four ACT domains can rotate during activation without steric clashes. Yet, our SAXS data show that the PheH tetramer can adopt an expanded conformation in solution facilitated by the tetramerization domain hinge, and this hinge flexibility may facilitate the domain rearrangements required for activation. Consistent with this hypothesis, mutations in the hinge region of the tetramerization domain have been shown to affect allosteric activation (Figure S1).[48,49]

Finally, the structural data from this work and others[9,20,25,30] shed light onto the unique role of ACT domains in the allosteric regulation of L-Phe in nature. Conserved sequence motifs have been observed in the ACT domains of PheH and prephenate dehydratase (PDT),[23] an enzyme that catalyzes the conversion of prephenate to phenylpyruvate, a key step in the biosynthesis of L-Phe in the shikimate pathway of microorganisms and plants. With the structural data now available, we find that these highly conserved residues map to the interface of the dimerized ACT domains and the allosteric L-Phe binding sites (Figures S17–18). Specifically, a crystal structure of PDT from *Chlorobium tepidum* (PDB: 2QMX) depicts L-Phe bound to a side-by-side ACT-domain dimer (Figures S3A, S18A)[25] in the same manner as that observed in $\text{RDPheH}^{\Delta 24}$ by NMR.[30] Thus, a remarkable similarity is observed in the usage of the ACT motifs in these enzymes. Whereas this motif is involved in allosteric activation of PheH, PDT is allosterically inhibited by L-Phe, a downstream product in the shikimate pathway. Hence, although the shikimate pathway has been evolutionarily lost in animals, the regulatory mechanism used to sense and control L-Phe levels appears to have been retained.

■ CONCLUSION

The pterin-dependent aromatic amino acid hydroxylases play a critical role in cognitive and neurological health. Despite their importance, structural characterizations of this enzyme family have been challenging. Using a hybrid approach of crystallography and SAXS, we have established the unique role of the ACT domain in the allosteric activation of mammalian PheH. In particular, we demonstrate the power of combining SEC-SAXS and EFA for detecting even small populations of high-mass contaminants and producing robust structural information in a model-independent manner. With the approach described

DOI: 10.1021/jacs.6b01563
J. Am. Chem. Soc. 2016, 138, 6506–6516

here, we provide a new framework for interrogating the role of regulatory domains in other members of the aromatic amino acid hydroxylase family.

EXPERIMENTAL PROCEDURES

Protein Expression and Purification. The N-terminal 24-residue deletion mutant of the isolated regulatory domain of rat PheH ($RDPheH^{R270K}$) used for ITC measurements was expressed and purified as described previously.[28] The pERPHΔ24R270K expression vector for the enzyme with the N-terminal 24-residue deletion and the R270K mutation ($PheH^{\Delta24\ R270K}$) was generated by mutagenesis of the plasmid pERPH R270K[26] using the QuikChange mutagenesis protocol. Wt-PheH, $PheH^{\Delta24}$, and $PheH^{\Delta24\ R270K}$ were expressed and purified as previously described for the wild-type enzyme[50] with minor modifications.[26] Phenyl Sepharose fractions with <90% protein purity were further purified by a Q-Sepharose Fast Flow column (1.5 cm × 8 cm). The column was washed with 10 column volumes of 50 mM HEPES, 5% glycerol, 1 μM leupeptin and 1 μm pepstatin, pH 7.0, and eluted with a gradient of 0 to 0.3 M KCl. Fractions showing >95% protein purity by polyacrylamide gel electrophoresis were pooled and dialyzed against 50 mM HEPES, 15% glycerol, 1 μM leupeptin and 1 μM pepstatin A, pH 7.0. Each protein was concentrated to ∼200 μM using an Amicon stirred cell concentrator containing a PM 30 ultrafiltration membrane and stored at −80 °C. For SAXS experiments the enzyme was dialyzed into 50 mM HEPES, 5% glycerol, pH 7.0. Protein concentrations were determined using the calculated extinction coefficients at 280 nm.

Isothermal Titration Calorimetry (ITC). ITC experiments were conducted on a VP-ITC instrument (GE HealthCare-Microcal). $RDPheH^{\Delta24}$ (50 μM) in 50 mM sodium phosphate, 100 mM NaCl, pH 8.0 was loaded into the calorimeter cell, and the titration syringe was loaded with 1.8 mM L-Phe in 50 mM sodium phosphate, 100 mM NaCl, pH 8.0. Titrations were carried out using 28 10-μL injections at 4 min intervals at 25 °C. The data were fit with the equation for identical binding sites using Origin software (Origin Lab. Corporation, Northampton, MA, USA).

Protein Crystallization, Structure Determination and Refinement. Crystals of wt-PheH and $PheH^{\Delta24\ R270K}$ were grown at 25 °C in the UTHSCSA X-ray Crystallography Core Laboratory from commercial crystallization screen kits using a Phoenix crystallization robot (Art Robbins Instruments, Sunnyvale, CA). Crystals of wt-PheH were obtained from the Classics 2 Suite (Qiagen, Venlo, Netherlands) using the sitting drop vapor diffusion method with 200 nL 0.1 M Bis-Tris:HCl pH 5.5, 25% polyethylene glycol 3350, and 0.2 M magnesium chloride combined in a 1:1 ratio with 200 nL of 12 mg/mL wt-PheH in 30 mM HEPES pH 7.0 and 1 mM L-Phe. Likewise, crystals of $PheH^{\Delta24\ R270K}$ were obtained from the MCSG Suite (Microlytic, Woburn, MA) with 200 nL 0.1 M Bis-Tris:HCl pH 6.5, 25% polyethylene glycol 3350, and 0.2 M ammonium sulfate combined in a 1:1 ratio with 200 nL of 10 mg/mL $PheH^{24\ R270K}$ in 50 mM HEPES pH 7.0, 1 mM L-Phe and 5% glycerol. Crystals were mounted in undersized nylon loops with excess mother liquor removed by wicking and flash-cooled in liquid nitrogen prior to data collection. Diffraction data were measured at the Advanced Photon Source NE-CAT beamlines 24-ID-C and 24-ID-E and integrated and scaled using the program XDS.[51] Phases were obtained by the molecular replacement method implemented in PHASER[52] using the PheH coordinates in Protein Data Bank entry 2PHM[9] as the search model. Model coordinates were refined using the PHENIX program suite,[53] including simulated annealing with torsion angle dynamics, alternating with manual model adjustment using the program COOT.[54] Noncrystallographic symmetry (NCS) restraints were used in the refinement of wt-PheH. The wt-PheH crystal yielded significantly lower quality data compared to that of $PheH^{\Delta24\ R270K}$ in terms of the resolution limit and merging R-factors. Additionally, potential twinning was indicated by the Padilla-Yeates plot;[55] however, data processing and model refinement with the appropriate merohedral twin law in a lower symmetry monoclinic space group produced a statistically inferior model, so the highest symmetry space group ($P2_12_12_1$) was used. The refinement R-factors remained in the high 20% range, likely due to the relatively low data quality. Composite omit maps calculated using PHENIX and MolProbity statistics[56] were used to assess model quality. Data collection and refinement statistics are listed in Table S1. The refined coordinates and structure factors have been deposited in the Protein Data Bank under accession codes 5EGQ and 5FGJ.

Small-Angle X-ray Scattering. X-ray scattering measurements of wt-PheH were performed at the G1 Station of the Cornell High Energy Synchrotron Source (CHESS) using 11.75 keV X-rays with a flux of 10^{11} photons per second at a beam size of 250 × 480 μm^2. Small-angle and wide-angle X-ray scattering (SAXS/WAXS) images were collected simultaneously on two photon-counting detectors (Pilatus 100K) at sample-to-detector distances of 1.47 and 0.42 m, respectively. The SAXS detector covered a q-range of 0.01 to 0.336 $Å^{-1}$, and the WAXS detector covered a q-range of 0.338 to 0.960 $Å^{-1}$, where q is the momentum transfer, defined as $q = 4\pi \sin(\theta)/\lambda$, where λ is the X-ray wavelength and 2θ is the scattering angle. Samples were held in a flow cell at room temperature and oscillated during exposure to minimize radiation damage. Images were integrated and normalized by the incident X-ray intensity as measured by an N_2-filled ion chamber located after the beam-defining slits. SAXS data for $PheH^{\Delta24\ R270K}$ and $PheH^{R270K}$ were collected at the CHESS G1 Station with an X-ray energy of 9.83 keV, SAXS detector at 1.49 m covering the q-range of 0.01 to 0.262 $Å^{-1}$, and WAXS detector at 0.42 m covering the q-range of 0.264 to 0.77 $Å^{-1}$. Samples were oscillated in the flow cell at 4 °C during data collection, and scattering profiles were normalized by the photocurrent recorded from a PIN diode in the beamstop. Data were processed following established protocols.[57] Analysis of structural parameters, SVD, and EFA were performed using custom code written in MATLAB (The MathWorks, Inc., Natick, MA).

The SAXS titration data for wt-PheH were produced from samples prepared at 25 μM (monomer concentration) in 50 mM HEPES pH 7.0, 5% (v/v) glycerol and variable amounts of L-Phe. Samples of $PheH^{\Delta24\ R270K}$ and $PheH^{R270K}$ were prepared in the same way; however, the buffer additionally contained 100 mM NaCl for stability. Background subtraction was performed with buffers containing identical concentrations of L-Phe. Immediately before data collection, protein samples were incubated at room temperature to allow for activation and then centrifuged for ∼2 min. To generate difference profiles for SVD, the scattering profile collected at 0 mM L-Phe was subtracted from the remaining curves. The unequal distribution of noise in the SAXS data was corrected by dividing the difference profiles by the average experimental uncertainty at each q-value (Figure S7A).

Samples of wt-PheH for SEC-SAXS were passed continuously through an in vacuo X-ray sample cell via an in-line size exclusion column (GE Superdex 200 5/15GL) operated by a room-temperature GE Äkta Purifier using a flow rate of 0.075 mL min^{-1}. The column was pre-equilibrated with running buffer, consisting of 50 mM HEPES pH 7.0, 5% (v/v) glycerol with 0 or 1 mM L-Phe. In anticipation of a ∼10-fold dilution in the size exclusion column, protein samples were prepared at ∼600 μM PheH (monomer concentration) with 0 or 1 mM L-Phe and injected into a 50 μL loop. Approximately 400 eight-second exposures were collected per sample, and 140 buffer profiles preceding the elution peaks were averaged and used for background subtraction. Previous SEC-SAXS studies have contended with drift in the background scattering during elution, possibly due to the accumulation of material on the X-ray windows or instability of the normalizing detector.[41,42] Therefore, we assessed the quality of the buffer blank by comparing the first and last profiles within each data set; they superimpose within the experimental noise, signifying that the buffer blank was a good match for the scattering during elution (Figure S8).

Although EFA has been used successfully in chromatography for several decades, its application to chromatography-coupled SAXS has not been described. Methods and associated software for analyzing SEC-SAXS data using EFA were developed (Figure S12). First, the buffer-subtracted SAXS profiles collected during elution were

DOI: 10.1021/jacs.6b01563
J. Am. Chem. Soc. 2016, 138, 6506−6516

uncertainty-weighted and stored as sequential columns in a matrix, A. Then, the range of each potentially overlapping peak in the data was determined using EFA.[43] Because sudden jumps in the rank of A occur when new species elute, the start of each component's peak was identified by plotting the singular values of the first n columns of A while n was varied (Figure S13−S14, solid lines in top panels). Similarly, sudden decreases in the rank of A that signify the end of a peak were found from the evolution of the singular values of A with the first n columns removed (Figure S13−S14, dotted lines in top panels). The values of n where jumps in the rank of A occur were found graphically, and pairs of these points from the forward and reverse EFA plots defined the peak windows according to the first-in, first-out principle. Once the peak windows were determined, the basis vectors were rotated to zero the contribution of each component outside its peak window. Finally, the SAXS profiles for each component and their estimated uncertainties were calculated (Figure S13−S14).

SAXS Models of PheH. To predict SAXS profiles from atomic models, disordered residues missing in the crystal structures of PheH were added using the program Modeller.[46] Theoretical scattering profiles were calculated in the program CRYSOL[45] with 50 spherical harmonics and 501 points between 0 and 1 Å^{-1}. The solvent density was modified from the default value of 0.334 electrons/Å^3 (water) to 0.340 electrons/Å^3 to account for 5% glycerol present in the buffer. The overall scale factor, c, and solvation parameters were obtained by fitting to the experimental profile. In the fitting step, CRYSOL minimizes the χ^2 for the theoretical profile and the data points with corresponding errors, defined as

$$\chi^2 = \frac{1}{N_p}\sum_{i=1}^{N_p}\left(\frac{I_e(q_i) - cI_{\mathrm{calc}}(q_i)}{\sigma(q_i)}\right)^2 \quad (3)$$

To model alternate conformations of the tetramer, we used the program SASREF (version 6.0),[58] which performs rigid-body optimization of molecular models to SAXS data with distance constraints. Because alternate conformations of the disordered N-termini minimally affect the shape of the scattering profile (Figure S16), they were held fixed during rigid-body optimization. Three rigid bodies were defined: the two dimers (residues 1−424), and the coiled-coil of tetramerization domain α-helices (residues 427−453). To model the hinge, a distance restraint was imposed between residues 424 and 427 based on the maximum distance of 8.6 Å between the C_α atoms of these residues in the crystal structure. The spherical harmonic expansion of each rigid body was calculated using CRYSOL with 20 spherical harmonics, the maximum allowed by SASREF, and the optimization was performed using the SAXS data. Due to the limited resolution of the calculation, the data were truncated at q = 0.3 Å^{-1}. After the optimization was complete, we performed a CRYSOL fit to the entire q-range to obtain meaningful χ values, as described above. We note that similar results were obtained with the program CORAL,[58] which allows for explicit modeling of flexible linkers. However, a meaningful quality of fit could only be obtained with SASREF; because it is able to utilize a greater number of spherical harmonics than CORAL, it can accurately calculate the scattering over the wide q-range of our SAXS data.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b01563.

Table S1 and Figures S1−18. (PDF)
Crystallographic data for 5EGQ. (CIF)
Crystallographic data for 5FGJ. (CIF)

■ AUTHOR INFORMATION

Corresponding Authors

*fitzpatrickp@uthscsa.edu
*nozomi.ando@princeton.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors are grateful to Dr. Katherine Davis, Saad Imran, Max Watkins, Will Thomas, Brian Chang, Emily Adler, and Dr. Richard Gillilan (CHESS) for assistance with SAXS data collection. We thank Dr. Kenneth M. Roberts for assistance in the crystallization of wild-type PheH, Saad Imran for assistance in the preparation of mutant PheH samples for SAXS, and Max Watkins for critical reading of this manuscript. This work was supported by National Health Institutes (NIH) grants GM100008 to NA and GM098140 to PFF, Welch Foundation grant AQ-1245 to PFF, and start-up funds from Princeton University to NA. Support for the X-ray Crystallography Core Laboratory and the Institutional Mass Spectrometry Laboratory by the UTHSCSA Executive Research Committee and the Cancer Therapy Research Center is gratefully acknowledged. Crystallography was conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P41 GM103403). The Pilatus 6M detector on 24-ID-C beamline is funded by a NIH-ORIP HEI grant (S10 RR029205). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. SAXS was conducted at the Cornell High Energy Synchrotron Source (CHESS), which is supported by the National Science Foundation award DMR-1332208, using the MacCHESS facility, which is supported by the NIH National Institute of General Medical Sciences (NIGMS) award GM103485.

■ REFERENCES

(1) Fitzpatrick, P. F. *Annu. Rev. Biochem.* **1999**, *68*, 355−381.
(2) Fitzpatrick, P. F. *Biochemistry* **2003**, *42*, 14083−14091.
(3) Kappock, T. J.; Caradonna, J. P. *Chem. Rev.* **1996**, *96*, 2659−2756.
(4) Flydal, M. I.; Martinez, A. *IUBMB Life* **2013**, *65*, 341−349.
(5) Shiman, R.; Gray, D. *J. Biol. Chem.* **1980**, *255*, 4793−4800.
(6) Fitzpatrick, P. F. *Arch. Biochem. Biophys.* **2012**, *519*, 194−201.
(7) Fitzpatrick, P. F. *Curr. Opin. Struct. Biol.* **2015**, *35*, 1−6.
(8) Williams, R. A.; Mamotte, C. D. S.; Burnett, J. R. *Clin. Biochem. Rev.* **2008**, *29*, 31−41.
(9) Kobe, B.; Jennings, I. G.; House, C. M.; Michell, B. J.; Goodwill, K. E.; Santarsiero, B. D.; Stevens, R. C.; Cotton, R. G.; Kemp, B. E. *Nat. Struct. Biol.* **1999**, *6*, 442−448.
(10) Fusetti, F.; Erlandsen, H.; Flatmark, T.; Stevens, R. C. *J. Biol. Chem.* **1998**, *273*, 16962−16967.
(11) Goodwill, K. E.; Sabatier, C.; Marks, C.; Raag, R.; Fitzpatrick, P. F.; Stevens, R. C. *Nat. Struct. Biol.* **1997**, *4*, 578−585.
(12) Wang, L.; Erlandsen, H.; Haavik, J.; Knappskog, P. M.; Stevens, R. C. *Biochemistry* **2002**, *41*, 12569−12574.
(13) Ronau, J. A.; Paul, L. N.; Fuchs, J. E.; Corn, I. R.; Wagner, K. T.; Liedl, K. R.; Abu-Omar, M. M.; Das, C. *Eur. Biophys. J.* **2013**, *42*, 691−708.
(14) Andersen, O. A.; Stokka, A. J.; Flatmark, T.; Hough, E. *J. Mol. Biol.* **2003**, *333*, 747−757.
(15) Aravind, L.; Koonin, E. V. *J. Mol. Biol.* **1999**, *287*, 1023−1040.
(16) Lang, E. J.; Cross, P. J.; Mittelstädt, G.; Jameson, G. B.; Parker, E. J. *Curr. Opin. Struct. Biol.* **2014**, *29*, 102−111.
(17) Zhang, S.; Huang, T.; Ilangovan, U.; Hinck, A. P.; Fitzpatrick, P. F. *J. Mol. Biol.* **2014**, *426*, 1483−1497.

DOI: 10.1021/jacs.6b01563
J. Am. Chem. Soc. 2016, 138, 6506−6516

(18) Jaffe, E. K.; Stith, L.; Lawrence, S. H.; Andrake, M.; Dunbrack, R. L. *Arch. Biochem. Biophys.* **2013**, *530*, 73−82.
(19) Carluccio, C.; Fraternali, F.; Salvatore, F.; Fornili, A.; Zagari, A. *J. Biomol. Struct. Dyn.* **2016**, *34*, 497−507.
(20) Arturo, E. C.; Gupta, K.; Héroux, A.; Stith, L.; Cross, P. J.; Parker, E. J.; Loll, P. J.; Jaffe, E. K. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 2394−2399.
(21) Li, J.; Dangott, L. J.; Fitzpatrick, P. F. *Biochemistry* **2010**, *49*, 3327−3335.
(22) Thórólfsson, M.; Ibarra-Molero, B.; Fojan, P.; Petersen, S. B.; Sanchez-Ruiz, J. M.; Martínez, A. *Biochemistry* **2002**, *41*, 7573−7585.
(23) Gjetting, T.; Petersen, M.; Guldberg, P.; Güttler, F. *Am. J. Hum. Genet.* **2001**, *68*, 1353−1360.
(24) Schuller, D. J.; Grant, G. A.; Banaszak, L. J. *Nat. Struct. Biol.* **1995**, *2*, 69−76.
(25) Tan, K.; Li, H.; Zhang, R.; Gu, M.; Clancy, S. T.; Joachimiak, A. *J. Struct. Biol.* **2008**, *162*, 94−107.
(26) Roberts, K. M.; Khan, C. A.; Hinck, C. S.; Fitzpatrick, P. F. *Biochemistry* **2014**, *53*, 7846−7853.
(27) Pey, A. L.; Stricher, F.; Serrano, L.; Martinez, A. *Am. J. Hum. Genet.* **2007**, *81*, 1006−1024.
(28) Zhang, S.; Roberts, K. M.; Fitzpatrick, P. F. *Biochemistry* **2014**, *53*, 6625−6627.
(29) Zhang, S.; Hinck, A. P.; Fitzpatrick, P. F. *Biochemistry* **2015**, *54*, 5167−5174.
(30) Zhang, S.; Fitzpatrick, P. F. *J. Biol. Chem.* **2016**, *291*, 7418−7425.
(31) Jennings, I. G.; Teh, T.; Kobe, B. *FEBS Lett.* **2001**, *488*, 196−200.
(32) Andersen, O. A.; Flatmark, T.; Hough, E. *J. Mol. Biol.* **2002**, *320*, 1095−1108.
(33) Roberts, K. M.; Pavon, J. A.; Fitzpatrick, P. F. *Biochemistry* **2013**, *52*, 1062−1073.
(34) Ando, N.; Brignole, E. J.; Zimanyi, C. M.; Funk, M. A.; Yokoyama, K.; Asturias, F. J.; Stubbe, J.; Drennan, C. L. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 21046−21051.
(35) Glatter, O.; Kratky, O. *Small Angle X-ray Scattering*; Academic Press: London, 1982.
(36) Petoukhov, M. V.; Franke, D.; Shkumatov, A. V.; Tria, G.; Kikhney, A. G.; Gajda, M.; Gorba, C.; Mertens, H. D.; Konarev, P. V.; Svergun, D. I. *J. Appl. Crystallogr.* **2012**, *45*, 342−350.
(37) Kappock, T. J.; Harkins, P. C.; Friedenberg, S.; Caradonna, J. P. *J. Biol. Chem.* **1995**, *270*, 30532−30544.
(38) Knappskog, P. M.; Flatmark, T.; Aarden, J. M.; Haavik, J.; Martínez, A. *Eur. J. Biochem.* **1996**, *242*, 813−821.
(39) Mathew, E.; Mirza, A.; Menhart, N. *J. Synchrotron Radiat.* **2004**, *11*, 314−318.
(40) Berthaud, A.; Manzi, J.; Pérez, J.; Mangenot, S. *J. Am. Chem. Soc.* **2012**, *134*, 10080−10088.
(41) Brookes, E.; Pérez, J.; Cardinali, B.; Profumo, A.; Vachette, P.; Rocco, M. *J. Appl. Crystallogr.* **2013**, *46*, 1823−1833.
(42) Malaby, A. W.; Chakravarthy, S.; Irving, T. C.; Kathuria, S. V.; Bilsel, O.; Lambright, D. G. *J. Appl. Crystallogr.* **2015**, *48*, 1102−1113.
(43) Maeder, M. *Anal. Chem.* **1987**, *59*, 527−530.
(44) Loeb, K. E.; Westre, T. E.; Kappock, T. J.; Mitic, N.; Glasfeld, E.; Caradonna, J. P.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. *J. Am. Chem. Soc.* **1997**, *119*, 1901−1915.
(45) Svergun, D.; Barberato, C.; Koch, M. *J. Appl. Crystallogr.* **1995**, *28*, 768−773.
(46) Fiser, A.; Do, R. K. G.; Šali, A. *Protein Sci.* **2000**, *9*, 1753−1773.
(47) Jaffe, E. K. *Trends Biochem. Sci.* **2005**, *30*, 490−497.
(48) Bjørgo, E.; de Carvalho, R. M. N.; Flatmark, T. *Eur. J. Biochem.* **2001**, *268*, 997−1005.
(49) Stokka, A. J.; Carvalho, R. N.; Barroso, J. F.; Flatmark, T. *J. Biol. Chem.* **2004**, *279*, 26571−26580.
(50) Daubner, S. C.; Hillas, P. J.; Fitzpatrick, P. F. *Biochemistry* **1997**, *36*, 11574−11582.
(51) Kabsch, W. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 125−132.
(52) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J. Appl. Crystallogr.* **2007**, *40*, 658−674.
(53) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 213−221.
(54) Emsley, P.; Cowtan, K. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126−2132.
(55) Padilla, J. E.; Yeates, T. O. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2003**, *59*, 1124−1130.
(56) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 12−21.
(57) Skou, S.; Gillilan, R. E.; Ando, N. *Nat. Protoc.* **2014**, *9*, 1727.
(58) Petoukhov, M. V.; Svergun, D. I. *Biophys. J.* **2005**, *89*, 1237−1250.

DOI: 10.1021/jacs.6b01563
J. Am. Chem. Soc. 2016, 138, 6506−6516