



Wrangling Shape-Shifting Morpheeins to Tackle Disease and Approach Drug Discovery

Eileen K. Jaffe*

Fox Chase Cancer Center, Philadelphia, PA, United States

Homo-multimeric proteins that can come apart, change shape, and reassemble differently with functional consequences have been called morpheeins and/or transformers; these provide a largely unexplored context for understanding disease and developing allosteric therapeutics. This article describes such proteins within the context of protein structure dynamics, provides one detailed example related to an inborn error of metabolism and potential herbicide development, and describes the context for applying these ideas for understanding disease and designing bioactive molecules, such as therapeutics.

Keywords: protein structure dynamics, allostery, morpheein, fifth level of protein structure, drug discovery

OPEN ACCESS

Edited by:

Guang Hu,
Soochow University, China

Reviewed by:

Jian Zhang,
Shanghai Jiao Tong University, China
Claire Lesieur,
UMR 5005 Laboratoire Ampère
(Ampère), France
Athi N. Naganathan,
Indian Institute of Technology Madras,
India

*Correspondence:

Eileen K. Jaffe
Eileen.Jaffe@fccc.edu

Specialty section:

This article was submitted to
Biological Modelling and Simulation,
a section of the journal
Frontiers in Molecular Biosciences

Received: 13 July 2020

Accepted: 26 October 2020

Published: 27 November 2020

Citation:

Jaffe EK (2020) Wrangling
Shape-Shifting Morpheeins to Tackle
Disease and Approach Drug
Discovery.
Front. Mol. Biosci. 7:582966.
doi: 10.3389/fmolb.2020.582966

INTRODUCTION

A great number of medically relevant proteins are homo-multimers, some of which exist as an equilibrium of alternate assemblies that are both non-additive and functionally distinct. The phenomenon wherein protein homo-multimers can come apart, change shape while dissociated, and reassemble into an architecturally and functionally different assembly has been called the morpheein model of protein allostery (Jaffe, 2005; see Morpheein in Wikipedia¹). A key to this protein structure dynamic is that the required conformational change is *spatially forbidden* within the context of either assembly. Proteins with this capacity can be called morpheeins and the alternate assemblies can be called morpheein forms. The dynamic process of dissociation and association makes this mode of allostery distinct from the classic Monod-Wyman-Changeux and Koshland-Nemethy-Filmer models; it provides a conceptually distinct approach to understanding normal protein function, disease-associated protein dysfunction, drug action, and approaches to drug design. This article describes the morpheein model for allosteric regulation, provides a disease relevant example in the protein porphobilinogen synthase, and considers current and future research intended to capitalize on targeting quaternary structure shape shifting in many different proteins as a way to understand disease and develop therapies. Because there are so few well characterized examples, even the most comprehensive treatments of allosteric drug discovery do not address proteins that are established to sample a dynamic equilibrium of assemblies comprised of alternate protomer conformations whose interconversion is forbidden within the assemblies (Lu et al., 2019a,b).

MORPHEEINS WITHIN THE CONTEXT OF PROTEIN STRUCTURE DYNAMICS

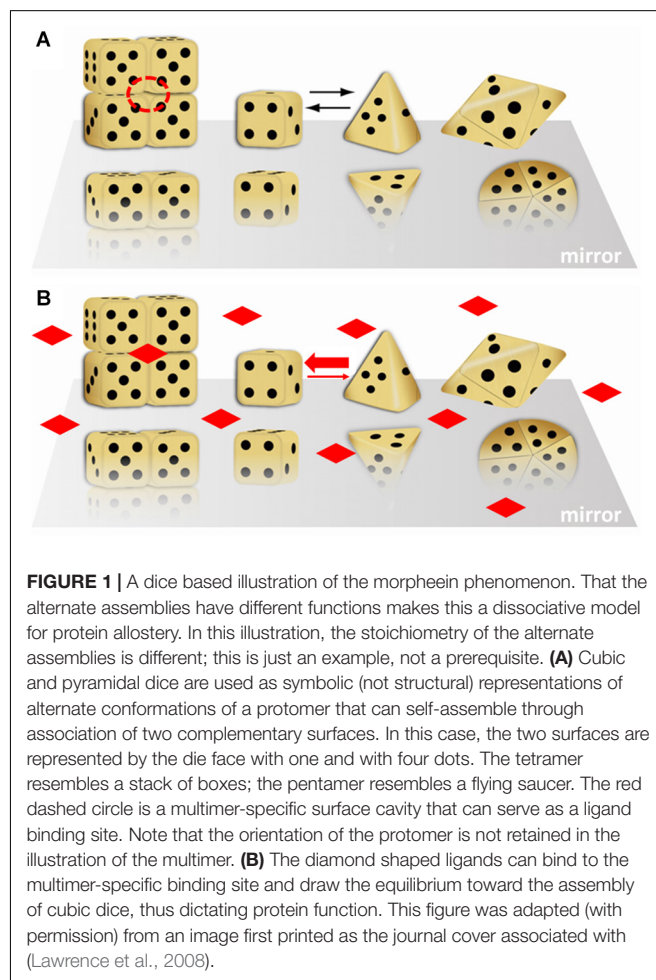
The existence of morpheeins is one of many protein structure dynamic phenomena that falls outside the classic one sequence – one structure – one function paradigm. Consequently, such discoveries

¹<http://en.wikipedia.org/wiki/Morpheein>

have been surprising (e.g., Breinig et al., 2003; Kashlan and Cooperman, 2003; Bornholdt et al., 2013), often serendipitous, and have been accompanied by the introduction of alternate related nomenclature such as transformers and metamorphic proteins (e.g., Murzin, 2008; Lopez-Pelegrin et al., 2014; Wasserman and Saphire, 2016; Dishman and Volkman, 2018). These refer to a continuum of quaternary structure dynamics which expand our view of protein structure beyond the level of primary, secondary, tertiary and quaternary. In the study of ribonucleotide reductase as a drug target, the investigator Aye has referred to going beyond quaternary structure as “breaking the fourth wall”; in the study of Ebola virology, the investigator Ollmann-Saphire has termed it the “fifth level of protein structure” (Wasserman and Saphire, 2016; Long et al., 2019). Herein, we use the term “fifth level of protein structure” to refer to equilibria of alternate assemblies comprised of alternate protomer conformations. This builds on the established concept that protein function is a consequence of an equilibrium of protein structures (Parisi et al., 2015). Both Aye’s and Ollmann-Saphire’s treatments highlight that normal protein structure dynamics can include architecturally distinct assemblies with alternate functions that are comprised of different protomer conformations. These assemblies exist as equilibria in the absence of chemical modification. The populations (e.g., mole fraction) of alternate morpheein forms responds to environmental factors (e.g., ionic strength, pH) and most significantly to ligand binding. These factors may govern the predominance of alternate morpheein forms in different cellular locations. Single amino acid substitutions that alter the mole fractions of alternate morpheein forms can cause disease (e.g., Jaffe and Stith, 2007). The morpheein model of protein allostery is a dissociative allosteric model most closely related to the equilibrium models of Nussinov (e.g., Kar et al., 2010) and Hilser (e.g., Motlagh et al., 2014), with the added dimension of quaternary structure. In the prototype morpheein described below, porphobilinogen synthase, the alternate functions are high activity (on) vs. low activity (off) (Jaffe and Stith, 2007; Jaffe and Lawrence, 2014; Jaffe, 2016, 2020). In the Ebola virus VP40 protein, the alternate functions are entirely separate activities, each one of which is essential for the viral life cycle (Bornholdt et al., 2013). Proteins that can moonlight (carry out unrelated functions, like VP40) were first discovered in the 1980’s (e.g., Gurney et al., 1986), and often arose from cloning the gene responsible for a biological function only to discover that the cloned protein sequence was already known to have a different function. A fascinating example is the protein originally identified as the glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase, which now has many documented functions, many of which can be targeted for drug discovery (e.g., Kopeckova et al., 2020; Lazarev et al., 2020). The known moonlighting proteins have recently been assembled by the investigator Jeffery into a MoonProt® database, which currently has ~400 listings (Mani et al., 2015; Chen et al., 2018). In most instances it remains to be determined if alternate moonlighting functions are associated with alterations at the fifth level of protein structure. A related fifth level phenomenon is the reversible filamentation of some enzymes, recently reviewed by Horton

(Park and Horton, 2019, 2020). Outstanding questions for many filament-forming proteins is whether they are morpheeins (with alternate protomer conformations), moonlighting proteins (with more than one function), or both. Two related enzymes, CTP synthase and IMP dehydrogenase are exemplars of this unknown. Each, separately and together, undergo changes in multimerization or filament formation in response to the state of the cell, but functional distinctions among these assemblies are yet unknown (Simonet et al., 2020).

Figure 1 illustrates the morpheein phenomenon using differently shaped dice to represent different conformations of the protomer. Monod first used dice assemblies to illustrate quaternary structures (Monod, 1965). **Figure 1** shows equilibration between two alternate conformations of the protomer, where one is represented by a cubic die and the other is represented by a pyramidal die. Although not obvious from the representative shapes, *the interconversion of these conformations does not require any substantial changes in the protein fold at the level of secondary or tertiary structure. The interconversion is spontaneous; it does not require any external input of energy.* It may involve small regions of order – disorder transition. For example, interconversion between the protomer conformations represented by the cube and pyramid could be a hinge motion



between two folded domains of each protomer that allows the die face with five dots to associate with the die face with six dots, burying these surfaces. Hinge motions allow protomer shape change without requiring a change in protein fold. In **Figure 1**, the higher order multimers form by association of the die face with one dot to the die face with four dots. The cubic die forms a symmetric tetramer; the pyramidal die forms a symmetric pentamer. It is easy to imagine how these two assemblies, though made up of chemically identical components, will interact with different cellular partners and potentially have different functions. Note that the surface of the tetramer contains a multimer-specific surface cavity that can serve as a ligand binding site (dashed circle in **Figure 1A**). The ligand could be a natural allosteric effector molecule, a drug, or another cellular entity (protein, nucleic acid, lipid, membrane surface). The pentamer does not have this same surface cavity and will not interact with the same ligand. In **Figure 1B**, addition of the imagined ligand causes stabilization of the tetramer, which will draw the structural equilibrium toward the tetramer and alter the protein's function to that of the tetramer. All of this happens in the absence of post-translational modifications or any other covalent changes to the protein. We note, however, that changes in protein sequence, post-translational modification, or the presence of purification tags can shift the position of the equilibrium (mole fraction of alternate morpheein forms) and enhance or inhibit allosteric ligand binding.

THE PROTOTYPE MORPHEEIN – PORPHOBILINOGEN SYNTHASE (PBGS)

PBGS Provides a Physiologic Relevance to the Morpheein Model of Allostery

The physiologic relevance of the morpheein model of allostery was first realized for the protein porphobilinogen synthase (PBGS), whose quaternary structure dynamic is illustrated in **Figure 2A** (Breinig et al., 2003; Selwood et al., 2008). Here I paint a broad picture of PBGS, highlighting key aspects of its fifth level of protein structure and refer the reader to recent reviews for more details (Jaffe and Lawrence, 2014; Jaffe, 2016). In the PBGS example, optimal enzyme activity requires controlled access to the enzyme active site, which is gated by the opening and closing of an active site lid. Each protomer has a complete active site, but securely closing the active site lid depends upon a network of molecular interactions that can only be achieved in the octamer (Jaffe, 2016). In the PBGS example, the alternate assemblies have different multimer-specific surface cavities that can be used for the development of bioactive molecules. The alternate assemblies also have different size, shape, and surface charge, which allows them to be separated by biochemical and biophysical methods such as native PAGE and ion exchange chromatography (Breinig et al., 2003).

The PBGS catalyzed reaction is essential for all organisms that rely on methanogenesis, photosynthesis, and/or respiration, thus covering every branch of cellular life. With all this evolutionary time to adapt to the organism's needs and to function in different

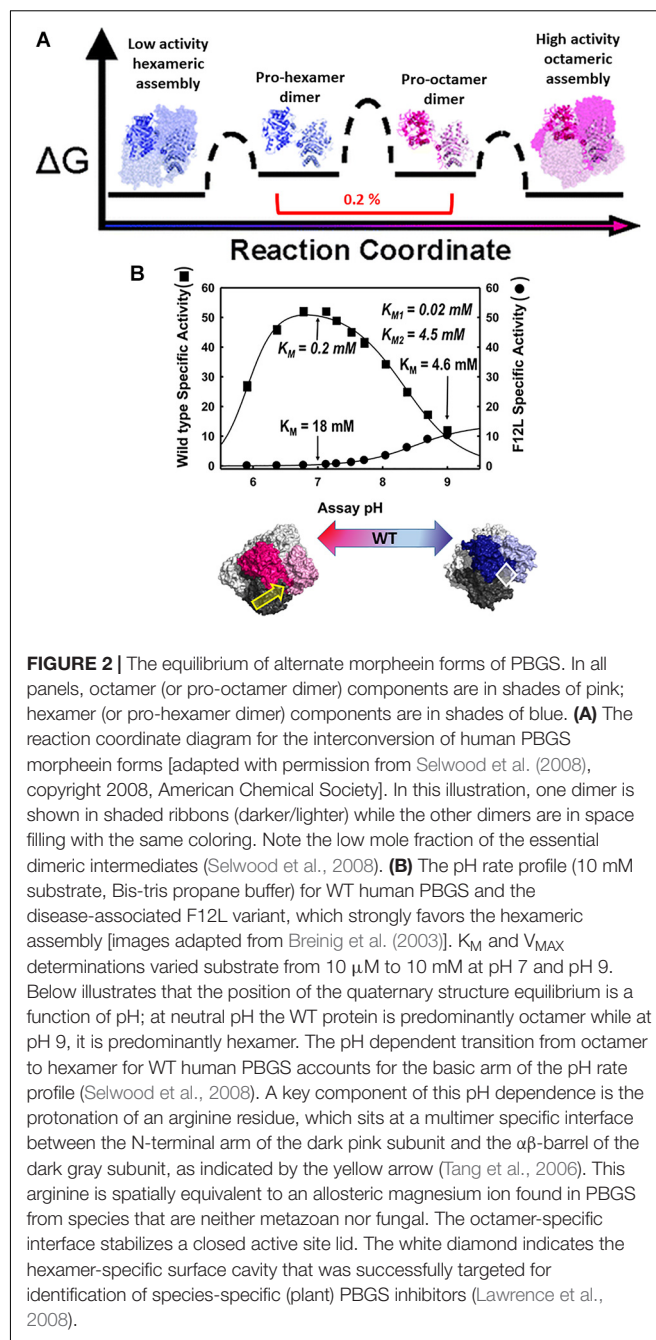


FIGURE 2 | The equilibrium of alternate morpheein forms of PBGS. In all panels, octamer (or pro-octamer dimer) components are in shades of pink; hexamer (or pro-hexamer dimer) components are in shades of blue. **(A)** The reaction coordinate diagram for the interconversion of human PBGS morpheein forms [adapted with permission from Selwood et al. (2008), copyright 2008, American Chemical Society]. In this illustration, one dimer is shown in shaded ribbons (darker/lighter) while the other dimers are in space filling with the same coloring. Note the low mole fraction of the essential dimeric intermediates (Selwood et al., 2008). **(B)** The pH rate profile (10 mM substrate, Bis-tris propane buffer) for WT human PBGS and the disease-associated F12L variant, which strongly favors the hexameric assembly [images adapted from Breinig et al. (2003)]. K_M and V_{MAX} determinations varied substrate from 10 μ M to 10 mM at pH 7 and pH 9. Below illustrates that the position of the quaternary structure equilibrium is a function of pH; at neutral pH the WT protein is predominantly octamer while at pH 9, it is predominantly hexamer. The pH dependent transition from octamer to hexamer for WT human PBGS accounts for the basic arm of the pH rate profile (Selwood et al., 2008). A key component of this pH dependence is the protonation of an arginine residue, which sits at a multimer specific interface between the N-terminal arm of the dark pink subunit and the α -barrel of the dark gray subunit, as indicated by the yellow arrow (Tang et al., 2006). This arginine is spatially equivalent to an allosteric magnesium ion found in PBGS from species that are neither metazoan nor fungal. The octamer-specific interface stabilizes a closed active site lid. The white diamond indicates the hexamer-specific surface cavity that was successfully targeted for identification of species-specific (plant) PBGS inhibitors (Lawrence et al., 2008).

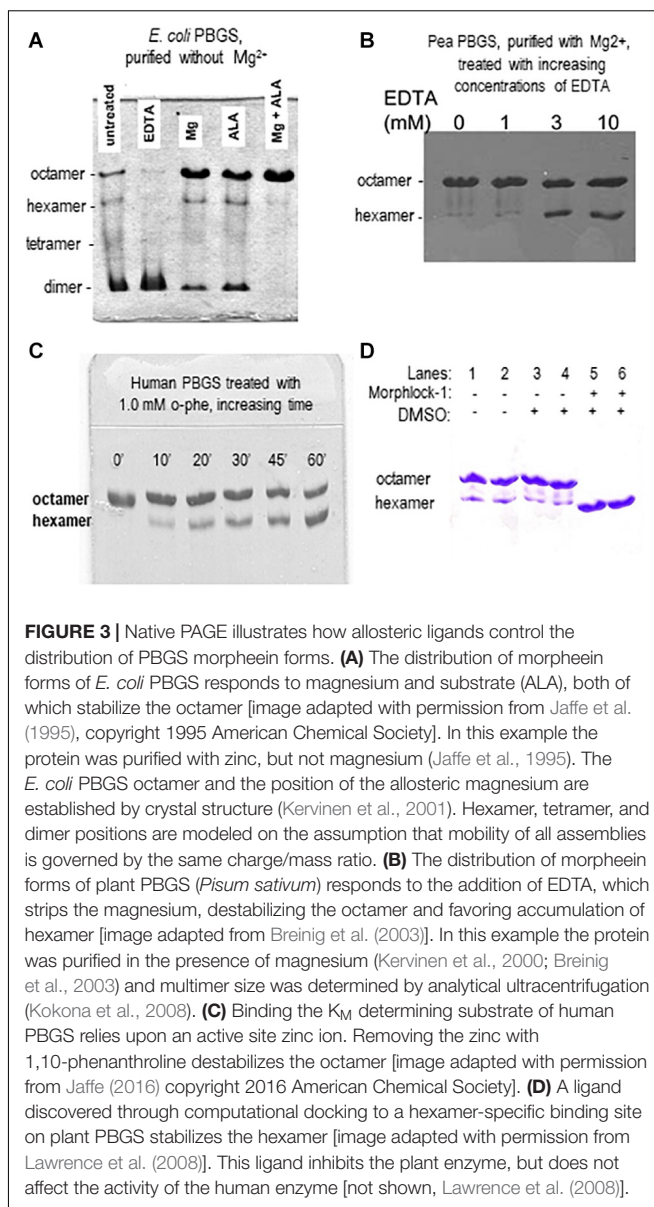
cellular environments (e.g., cytoplasm, chloroplast, apicoplast), factors governing the fifth level of PBGS protein structure are not evolutionarily conserved. Additionally, the amino acid composition of the targeted surface cavities is not conserved, unlike active site residues. This makes the allosteric regulation of PBGS a potential target for the development of antimicrobials and herbicides. In some species (e.g., plants, bacteria, archaea) the equilibrium position depends upon an allosteric magnesium binding at an interface only present in the octamer (see yellow arrow in **Figure 2B**). In PBGS from metazoa and fungi, which lack the magnesium binding site, in its place is the guanidinium

group of an arginine residue. In the human PBGS variant where a leucine is substituted for Phe12 (F12L), from which we obtained the crystal structure of the hexameric assembly (Breinig et al., 2003), this conserved arginine does not contact the neighboring N-terminal arm. In an apicoplast, PBGS has evolved to contain a C-terminal extension that prevents the hinge motion necessary to convert pro-octamer dimer to pro-hexamer dimer; this hinge motion is illustrated in **Figure 2A**. Thus, in apicoplast PBGS the equilibrium components are limited to octamer and pro-octamer dimer (Jaffe et al., 2011). This phylogenetic variation demonstrates that evolution of the fifth level of protein structure provides another opportunity for the adaptation of protein functional control.

Control of PBGS Morpheein Forms by pH and Ligand Binding

The human PBGS pH rate profile (**Figure 2B**) helped reveal a pH dependence to the quaternary structure equilibrium of wild type human PBGS (see the bottom panel of **Figure 2B**; Selwood et al., 2008). The crystal structures of the PBGS octamer (e.g., Jaffe, 2004; Jaffe et al., 2011; Mills-Davies et al., 2017) show that a non-reacting moiety of the K_M -determining substrate interacts with the active site lid, essentially closing the lid and allowing deprotonation of an active site lysine required for essential Schiff base formation. With an open active site, this deprotonation requires a high solvent pH, as is seen in the pH rate profile of the constitutively hexameric F12L variant. The crystal structures of the protomers in the human PBGS octamer vs. hexamer superpose remarkably well, but differ in two key ways. First is the backbone hinge between the $\alpha\beta$ -barrel and the N-terminal arm domains; this dictates assembly to octamer vs. hexamer. Second is the presence of an ordered active site lid, which is only present in one hemisphere of the octamer. Both structures contain active site ligands in a half-of-the-sites stoichiometry (one hemisphere) (PDB: 1E51, 1PV8). Only the ligand-containing octamer active sites contains atoms that derive from the K_M -determining substrate, which are securing a closed conformation of the active site lid through a network of bonds between the substrate's carboxyl group and basic residues on the lid.

Although the PBGS quaternary structure equilibrium is controlled by different factors in different branches of life, there are unifying characteristics in the pH rate profiles of alternate PBGS morpheein forms. Mammalian PBGS at neutral pH (see **Figure 2B**, bottom), and plant/bacterial PBGS with magnesium present are predominantly octameric and are documented to have K_M values in the range of $\sim 150 \mu\text{M}$ (Mitchell and Jaffe, 1992; Jaffe et al., 1995; Breinig et al., 2003), which is the range of the cellular substrate concentration. The isolated hexameric F12L variant (see **Figure 2B**), and the bacterial *Escherichia coli* PBGS without magnesium exhibit K_M values (at neutral pH) that are well above the physiological substrate concentration, in the range of 5–20 mM. Addition of magnesium to a magnesium-free *E. coli* PBGS sample has been shown to shift the quaternary structure equilibrium to octamer (**Figure 3A**) and reduce the K_M to that characteristic



of octamer (Mitchell and Jaffe, 1993; Jaffe et al., 1995). Similarly, treatment of a plant PBGS with EDTA causes a shift from octamer to hexamer (**Figure 3B**; Breinig et al., 2003). For human PBGS, this same phenomenon can be demonstrated by removal of a catalytic zinc ion that is essential for binding the K_M -determining substrate (**Figure 3C**; Jaffe, 2016). The high K_M values for PBGS that are not octameric derives from an inability to secure the closed active site lid, leaving the K_M -determining substrate loosely bound. What is striking about the pH rate profile of WT human PBGS is the shift in K_M values between pH 7 and pH 9, where the kinetic parameters resemble that of hexameric F12L. In fact, WT human PBGS at pH 9 exhibits a double hyperbolic kinetic behavior. At low substrate concentration, the observed reaction rate is dominated by the low K_M (high V_{MAX}) octameric component, which is at

a relatively low mole fraction. As the substrate concentration starts to approach the K_M of the hexamer, the observed enzyme catalyzed reaction rate become dominated by the high K_M (low V_{MAX}) hexameric component (which is at a high mole fraction). This double hyperbolic kinetic phenomenon, caused by a mixture of species with different kinetic constants follows a classic treatment of isozymes. The alternate morpheein forms of PBGS are not isozymes. Rather, they are a slowly exchanging mixture octamer and hexamer.

An Inborn Error of Metabolism Is Linked to Perturbation of an Equilibrium of Morpheein Forms

ALAD porphyria is a rare inborn error of metabolism caused by dysfunctional PBGS and inherited as a recessive disease (Maruno et al., 2001). ALAD, an abbreviation for amino-levulinic acid dehydratase, is an alternate name for PBGS; it remains in clinical use. There are eight known disease-associated variants and all patients are compound heterozygotes. Only one disease-associated variant alters an amino acid at the enzyme active site. When heterologously expressed and purified from *E. coli*, the disease-associated human PBGS variants all show an increased propensity to populate the hexameric assembly at neutral pH (Jaffe and Stith, 2007). This can be illustrated by the appearance of the double hyperbolic kinetic phenomenon at neutral pH (Figure 4), which is indicative of a mixture of octamer and hexamer. It is also seen by native PAGE and ion exchange chromatography, both of which separate the octamer from the hexamer, as confirmed by the established crystal structures of the wild-type and the constitutively hexameric F12L variant (Breinig et al., 2003).

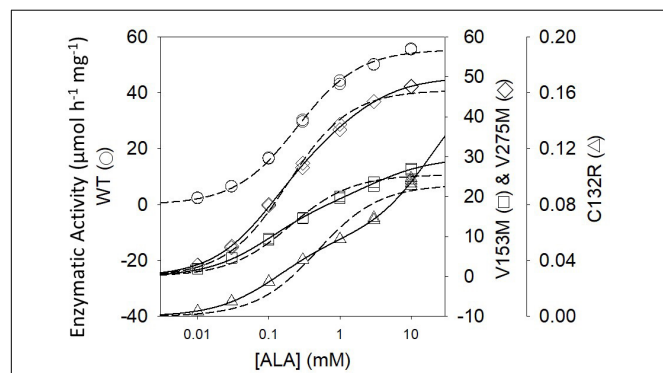


FIGURE 4 | The activity vs. substrate relationship for disease-associated human PBGS proteins indicate the presence of two morpheein forms at neutral pH. The dashed lines are fitted to the hyperbolic Michaelis-Menten equations, which is an excellent fit for the WT protein. The solid lines are double hyperbolic fits (the sum of two species with different K_M and V_{MAX} values) [image adapted with permission from Jaffe and Stith (2007)]. The kinetic behavior of each variant shows evidence for a mixture of low K_M and high K_M morpheein forms, corresponding to octamer and hexamer (Jaffe and Stith, 2007). The wild type protein shows this double hyperbolic behavior at pH 9 (see Figure 2B).

We have posited that the disequilibrium of alternate assemblies contributes to the complex phylogenetic patterns seen in the most common inborn error of amino acid metabolism, phenylketonuria, which is caused by dysfunctional phenylalanine hydroxylase (Jaffe et al., 2013; Jaffe, 2017). Although it is now established that phenylalanine hydroxylase can equilibrate between architecturally distinct high activity and low activity tetrameric assemblies, it is not yet established whether tetramer dissociation is a required component their interconversion (Jaffe, 2017).

WE COULD EASILY HAVE MISSED THE FIFTH LEVEL OF PROTEIN STRUCTURE WHEN STUDYING PBGS

The four levels of protein structure are introduced in every biochemistry text. Additionally, the relationship between sequence, structure, and function is the foundation of broad applications of bioinformatics, which drive much biomedical research. For the first 20 years that we studied PBGS, we were not looking for the fifth level of protein structure. All of our data prior to ~2003 was interpreted within the context that PBGS has one fixed quaternary architecture, which is an octamer (e.g., Wu et al., 1974), in possible equilibrium with tetramers or dimers comprised of the same protomer conformation. For example, our published model of the protein concentration dependence of a plant PBGS included only dimer, tetramer, and octamer (Kervinen et al., 2000). A few publications had suggested that PBGS is a hexamer (e.g., Stolz and Dornemann, 1996). But, we believed that one of these interpretations must be incorrect.

The first PBGS crystal structure, published in 1997, showed the architecture of the octamer, establishing precedent, as crystal structures often do (Erskine et al., 1997). As described above, the third decade of our work with PBGS revealed the protein as the prototype morpheein. We would not have seen this but for a confluence factors. (1) With a focus on the evolution of metal ion usage, we had correlated sequence variations with *in vitro* behaviors looking at PBGS from mammals, bacteria, and plants (e.g., Jaffe, 2003). By the time we fully evaluated this relationship, the first crystal structure of a bacterial PBGS showed the location of the allosteric magnesium binding site (Frankenberg et al., 1999). It is present in the octamer and absent in the hexamer. (2) We had not used purification tags, which allowed ion exchange chromatography to reveal the separation of PBGS morpheein forms during purification. It was a surprise when the disease-associated PBGS variant F12L had a very different mobility on an ion exchange column relative to the charge-equivalent WT protein (Breinig et al., 2003). (3) Had we not chosen to study the F12L variant, we would not have had the stable hexamer as a reagent for comparative analysis. For decades we had been discarding the small amount of low activity hexameric mammalian PBGS that is present at neutral pH and separated during our purification of wild type proteins. When studying a human PBGS variant from which the catalytic zinc binding site had been removed, we consistently

saw two peaks on an ion exchange column, but could not discriminate these by circular dichroism or enzyme kinetics. By kinetic criteria they were the same, because addition of substrate converted the hexamer to the octamer during the long time-course assays for these very low activity variants. Substrate stabilization of the octameric assembly can be illustrated using 2-dimensional native PAGE during which the gel is incubated in assay mixture between the two dimensions (e.g., Jaffe et al., 1995; Lawrence et al., 2008; see also **Figure 3A**). (4) Had we not obtained the crystal structure of F12L, we would not have seen the structure of the hexamer. From this, and knowledge of the location of the allosteric magnesium from prior crystal structures, we realized the octamer \leftrightarrow hexamer equilibrium as the basis for allosteric regulation of PBGS. Our extensive work on *E. coli* PBGS (including the 1995 documentation of the native PAGE phenomenon shown in **Figure 3A**) could now be reinterpreted within the context of this new allosteric model. Having previously interpreted our data in terms of a classic primary, secondary, tertiary, and quaternary structure paradigm, we had overlooked hints that the PBGS octamer was part of an equilibrium of differently sized morpheein forms. For example, a ^{13}C NMR study of labeled substrate bound to *E. coli* PBGS (in the absence of magnesium) had unexpectedly yielded NMR line widths significantly narrower relative to an octameric mammalian PBGS (Jaffe and Markham, 1988; Mitchell and Jaffe, 1993), suggesting a faster rotational correlation time indicative of a smaller multimeric size.

FINDING PROTEINS THAT EXPERIENCE THE FIFTH LEVEL OF PROTEIN STRUCTURE

Since the fifth level of protein structure provides an additional way to manipulate protein function, it also provides an additional approach to allosteric drug discovery. Consequently, there is value in identifying proteins that behave as morpheeins. This goal remains difficult. The initial discovery of the morpheein character of PBGS was serendipitously based on *in vitro* protein behavior coupled with X-ray crystal structures showing architecturally different assemblies comprised of different protomer conformations, but not different folds. This discovery was not based on a bioinformatics approach. The PBGS example presented itself because there happened to be a naturally occurring variant, F12L, that sufficiently stabilized the hexameric assembly to obtain its crystal structure. Although the wild-type human PBGS was later realized to readily equilibrate between octamer and hexamer in a pH-dependent fashion (**Figure 2**), to date all conditions that have yielded diffraction quality crystals of the wild-type protein favored crystals comprised of octamer. Serendipity favored our studying F12L; we could not have predicted the effect of this mutation. Well-established computational approaches (e.g., the program FoldX; Schymkowitz et al., 2005) which analyze the overall thermodynamic effect of amino acid substitutions will fail to predict substitutions that shift the position of an equilibrium of morpheein forms. A similar story applies to the discovery of

alternate morpheein forms and alternate functions of the Ebola virus VP40 protein (Bornholdt et al., 2013).

Our initial approach to identifying proteins that use the fifth level of protein structure was to manually search the literature for proteins that have one or more characteristics that we had documented for PBGSs. This approach was especially challenging as most of these characteristics were not contained in the searchable abstract or keywords and much of the older literature was not yet available as text-searchable PDF documents. Some of the characteristics we focused on are listed in **Table 1**; the characteristics included in **Table 1** are each consistent with an equilibrium of alternate assemblies, but none are strictly diagnostic. Each of these behaviors can be otherwise attributed. Trevor Selwood's herculean efforts generated boxes of reprints and a list of putative morpheeins (Selwood and Jaffe, 2012). This list, originally posted on Wikipedia (Morpheein¹), to date has not been actively edited by the community. Failure to gain research support for using the literature to identify proteins that function

TABLE 1 | Protein behaviors that might indicate an equilibrium of morpheein forms.

Characteristic	Comments
SDS-pure protein separating into alternate bands on native PAGE can indicate multimers of alternate stoichiometry (or conformation).	Native PAGE can be routinely incorporated into the final stages of protein characterization.
SDS-pure protein that separates into alternate forms using ion exchange chromatography	This has been used to monitor the interconversion of alternate assemblies as a function of ligand.
The enzyme-kinetic phenomenon known as hysteresis.	This can be observed if the transition from a low activity form to a high activity form occurs during the assay [as we have shown for the R240A variant of PBGS (Tang et al., 2006)].
Double hyperbolic kinetics (the sum of two hyperbola with different kinetic constants).	This can indicate alternate morpheein forms with different kinetic constants. Seeing this may require using a broad range of substrate concentration (see Figure 4).
X-ray crystal structures of multi-domain proteins that cannot be superposed without clashes. This observation is often dismissed as an artifact of crystal packing.	Many multi-domain proteins do not produce diffraction quality crystals. A common approach is to truncate one or more domains. If overlaying the common elements of such structures causes domains to clash, this could be a sign of alternate assemblies. This was observed for alternate truncated constructs of HIV integrase (Andrake and Skalka, 2015).
Protein concentration dependent specific activity can indicate alternate activities associated with different multimeric stoichiometries.	This is seen for all PBGS that use only magnesium [e.g., (Petrovich et al., 1996; Kervinen et al., 2000; Shanmugam et al., 2010)].
Evidence for soluble protein multimers that dissociate along a hydrophilic protein-protein interface.	This is a difficult characteristic to search for as many crystal structure files are at insufficient resolution to position water molecules. PBGS crystal structures contain phylogenetically conserved water molecules at the subunit-subunit interfaces that dissociate upon formation of the dimeric morpheein forms (Selwood et al., 2008).

as morpheeins turned this to a back-burner approach. Yet, the relatively small number of established morpheeins precludes designing a bioinformatics method for their identification. Jeffrey has described the same dilemma for the development of a bioinformatics approach to identifying moonlighting proteins (Mani et al., 2015).

Nevertheless, a key question remains as to how we can use what we have learned from PBGS to harness the fifth level of protein structure for drug discovery. In the PBGS example, the difference between the protomer that forms a hexamer and the protomer that forms an octamer is a hinge between two domains, without significantly altering the fold of these domains. This is also the case for alternate assemblies of the HIV integrase protein, where a hinge motion dictates formation of a core-core-dimer vs. a reaching dimer; and where it has been pointed out that small molecule stabilization of one or the other dimer, to prevent their interconversion, could yield a therapeutic (Bojja et al., 2013). For integrase, this promise has not yet materialized. In the case of VP40, alternate multimerization architectures also appear to arise from hinge motions between two domains, without significant refolding (Bornholdt et al., 2013). In the VP40 example, ligand association (RNA or membrane) dramatically stabilize one assembly relative to another allowing the protein to fulfill different essential function in the viral life cycle. Although not yet realized, drugs that prevent the interconversion of such assemblies could form the basis of a therapeutic. Research to facilitate this approach is ongoing (e.g., Buzon et al., 2020). In other cases wherein allosteric drugs could be imagined to work by stabilizing one of alternate assemblies (such as ribonucleotide reductase) the detailed molecular structures of alternate assemblies are only beginning to be revealed. Some of these are described in two recent reviews on the rich oligomeric and functional repertoire of both mammalian and bacterial ribonucleotide reductase enzymes (Thomas et al., 2019; Long et al., 2020).

IDENTIFYING ALLOSTERIC REGULATORS (E.G., THERAPEUTICS)

Using PBGS as an example, we have demonstrated how an equilibrium of morpheein forms can be manipulated in ways related to drugs. Realistically, ALAD porphyria is such a rare disorder that finding an octamer-stabilizing allosteric effector is, at best, an academic exercise. However, stabilization of a PBGS hexamer could form the basis for an antimicrobial or herbicide. To test this hypothesis, we targeted a hexamer-specific surface cavity on the model of a plant PBGS hexamer (*in silico* docking/*in vitro* testing) and found a hexamer-stabilizing inhibitor that did not affect human PBGS (Lawrence et al., 2008). **Figure 3D** illustrates the hexamer-stabilizing effect of the discovered compound (named morphlock-1). Unlike the enzyme active site, the multimer specific surface cavities is not phylogenetically conserved. We also used native PAGE to screen libraries of approved drugs as well as environmental contaminants (Lawrence et al., 2011, 2013). These studies revealed an explanation for some porphyria-promoting

drug side effects as well as the potential for environmental contributions to confound genotype/phenotype correlations. These demonstrations indicate that small molecule modulation of equilibria of morpheein forms is a viable approach to drug discovery. Nevertheless, a computational approach to identifying such modulators requires molecular resolution models and/or structures of alternate assemblies; these are not yet available for most putative morpheeins.

CONCLUSION AND FUTURE OUTLOOK

Proteins that can come apart, change shape, and reassemble differently with functional consequences provide expanded opportunities for understanding disease and designing therapeutics. This fifth level of protein structure provides another example where the one sequence/one structure/one function rule fails to provide the correct framework for data interpretation. Biophysical techniques that are becoming more widely available, (e.g., SEC-MALS, SEC-SAXS, cryo-EM, light scattering, to name just a few) are revealing the shape changing behavior of many disease-associated multimeric proteins. Other forms of microscopy are revealing changes in protein locations within cells. In some instances, these are associated with protein filamentation. All of these observations suggest the potential for harnessing the fifth level of protein structure for therapeutic advantage. As described, a few notable examples include ribonucleotide reductase (Thomas et al., 2019; Long et al., 2020), HIV integrase (Andrake and Skalka, 2015; Gupta et al., 2016), and Ebola VP40 protein (Bornholdt et al., 2013; Del Vecchio et al., 2018).

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

This work has been funded by grants from the National Institutes of Health 5R01 ES003654; 5R21 AI063324; 5R01 NS100081; and in part by the National Institutes of Health NCI Cancer Center Support Grant P30 CA006927. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health. Continuous support is acknowledged from the Fox Chase Cancer Center, now a part of the Temple University Health Systems.

ACKNOWLEDGMENTS

The author acknowledges many who have contributed to our deciphering the quaternary structure dynamics of PBGS, most notably Drs. Sabine Breinig, Sarah H. Lawrence, Trevor Selwood, and Lei Tang, and Ms. Linda Stith. Special acknowledgment is given to Emilia Arturo who has contributed to expanding our general appreciation for the fifth level of protein structure.

REFERENCES

- Andrake, M. D., and Skalka, A. M. (2015). Retroviral integrase: then and now. *Annu. Rev. Virol.* 2, 241–264. doi: 10.1146/annurev-virology-100114-055043
- Bojja, R. S., Andrade, M. D., Merkel, G., Weigand, S., Dunbrack, R. L. Jr., and Skalka, A. M. (2013). Architecture and assembly of HIV integrase multimers in the absence of DNA substrates. *J. Biol. Chem.* 288, 7373–7386. doi: 10.1074/jbc.m112.434431
- Bornholdt, Z. A., Noda, T., Abelson, D. M., Halfmann, P., Wood, M. R., Kawaoka, Y., et al. (2013). Structural rearrangement of ebola virus VP40 begets multiple functions in the virus life cycle. *Cell* 154, 763–774. doi: 10.1016/j.cell.2013.07.015
- Breinig, S., Kervinen, J., Stith, L., Wasson, A. S., Fairman, R., Wlodawer, A., et al. (2003). Control of tetrapyrrole biosynthesis by alternate quaternary forms of porphobilinogen synthase. *Nat. Struct. Biol.* 10, 757–763. doi: 10.1038/nsb963
- Buzon, P., Ruiz-Sanz, J., Martinez, J. C., and Luque, I. (2020). Stability, conformational plasticity, oligomerization behaviour and equilibrium unfolding intermediates of the Ebola virus matrix protein VP40. *J. Biomol. Struct. Dyn.* 38, 4289–4303. doi: 10.1080/07391102.2019.1671226
- Chen, C., Zabad, S., Liu, H., Wang, W., and Jeffery, C. (2018). MoonProt 2.0: an expansion and update of the moonlighting proteins database. *Nucleic Acids Res.* 46, D640–D644.
- Del Vecchio, K., Frick, C. T., Gc, J. B., Oda, S. I., Gerstman, B. S., Saphire, E. O., et al. (2018). A cationic, C-terminal patch and structural rearrangements in Ebola virus matrix VP40 protein control its interactions with phosphatidylserine. *J. Biol. Chem.* 293, 3335–3349. doi: 10.1074/jbc.m117.816280
- Dishman, A. F., and Volkman, B. F. (2018). Unfolding the mysteries of protein metamorphosis. *ACS Chem. Biol.* 13, 1438–1446. doi: 10.1021/acscchembio.8b00276
- Erskine, P. T., Senior, N., Awan, S., Lambert, R., Lewis, G., Tickle, L. J., et al. (1997). X-ray structure of 5-aminolaevulinatase dehydratase, a hybrid aldolase. *Nat. Struct. Biol.* 4, 1025–1031. doi: 10.1038/nsb1297-1025
- Frankenberg, N., Erskine, P. T., Cooper, J. B., Shoolingin-Jordan, P. M., Jahn, D., and Heinz, D. W. (1999). High resolution crystal structure of a Mg²⁺-dependent porphobilinogen synthase. *J. Mol. Biol.* 289, 591–602. doi: 10.1006/jmbi.1999.2808
- Gupta, K., Turkki, V., Sherrill-Mix, S., Hwang, Y., Eilers, G., Taylor, L., et al. (2016). Structural basis for inhibitor-induced aggregation of HIV integrase. *PLoS Biol.* 14:e1002584. doi: 10.1371/journal.pbio.1002584
- Gurney, M. E., Heinrich, S., Lee, M. R., and Yin, H. S. (1986). Molecular cloning and expression of neuroleukin, a neurotrophic factor for spinal and sensory neurons. *Science* 234, 566–574. doi: 10.1126/science.3764429
- Jaffe, E. K. (2003). An unusual phylogenetic variation in the metal ion binding sites of porphobilinogen synthase. *Chem. Biol.* 10, 25–34. doi: 10.1016/s1074-5521(02)00296-x
- Jaffe, E. K. (2004). The porphobilinogen synthase catalyzed reaction mechanism. *Bioorg. Chem.* 32, 316–325. doi: 10.1016/j.bioorg.2004.05.010
- Jaffe, E. K. (2005). Morpheeins—a new structural paradigm for allosteric regulation. *Trends Biochem. Sci.* 30, 490–497. doi: 10.1016/j.tibs.2005.07.003
- Jaffe, E. K. (2016). The remarkable character of porphobilinogen synthase. *Acc. Chem. Res.* 49, 2509–2517. doi: 10.1021/acs.accounts.6b00414
- Jaffe, E. K. (2017). New protein structures provide an updated understanding of phenylketonuria. *Mol. Genet. Metab.* 121, 289–296. doi: 10.1016/j.ymgme.2017.06.005
- Jaffe, E. K. (2020). “Chapter three - porphobilinogen synthase: an equilibrium of different assemblies in human health, in progress,” in *Molecular Biology and Translational Science*, eds J. Giraldo, and F. Ciruela (Cambridge, MA: Academic Press), 85–104. doi: 10.1016/bs.pmbts.2019.11.003
- Jaffe, E. K., Ali, S., Mitchell, L. W., Taylor, K. M., Volin, M., and Markham, G. D. (1995). Characterization of the role of the stimulatory magnesium of *Escherichia coli* porphobilinogen synthase. *Biochemistry* 34, 244–251. doi: 10.1021/bi00001a029
- Jaffe, E. K., and Lawrence, S. H. (2014). “The dance of porphobilinogen synthase in the control of tetrapyrrole biosynthesis,” in *Handbook of Porphyrin Science*, ed. G. C. Ferreira (Toh Tuck: World Scientific Publishing Co. Pte. Ltd), 79–128. doi: 10.1142/9789814407755_0002
- Jaffe, E. K., and Markham, G. D. (1988). C-13 Nmr-Studies of methylene and methine carbons of substrate-bound to a 280000-Dalton protein, porphobilinogen synthase. *Biochemistry* 27, 4475–4481. doi: 10.1021/bi00412a039
- Jaffe, E. K., Shanmugam, D., Gardberg, A., Dieterich, S., Sankaran, B., Stewart, L. J., et al. (2011). Crystal structure of *Toxoplasma gondii* porphobilinogen synthase: insights on octameric structure and porphobilinogen formation. *J. Biol. Chem.* 286, 15298–15307. doi: 10.1074/jbc.m111.226225
- Jaffe, E. K., and Stith, L. (2007). ALAD porphyria is a conformational disease. *Am. J. Hum. Genet.* 80, 329–337. doi: 10.1086/511444
- Jaffe, E. K., Stith, L., Lawrence, S. H., Andrade, M., and Dunbrack, R. L. Jr. (2013). A new model for allosteric regulation of phenylalanine hydroxylase: implications for disease and therapeutics. *Arch. Biochem. Biophys.* 530, 73–82. doi: 10.1016/j.abb.2012.12.017
- Kar, G., Keskin, O., Gursory, A., and Nussinov, R. (2010). Allostery and population shift in drug discovery. *Curr. Opin. Pharmacol.* 10, 715–722. doi: 10.1016/j.coph.2010.09.002
- Kashlan, O. B., and Cooperman, B. S. (2003). Comprehensive model for allosteric regulation of mammalian ribonucleotide reductase: refinements and consequences. *Biochemistry* 42, 1696–1706. doi: 10.1021/bi020634d
- Kervinen, J., Dunbrack, R. L. Jr., Litwin, S., Martins, J., Scarrow, R. C., Volin, M., et al. (2000). Porphobilinogen synthase from pea: expression from an artificial gene, kinetic characterization, and novel implications for subunit interactions. *Biochemistry* 39, 9018–9029. doi: 10.1021/bi000620c
- Kervinen, J., Jaffe, E. K., Stauffer, F., Neier, R., Wlodawer, A., and Zdanov, A. (2001). Mechanistic basis for suicide inactivation of porphobilinogen synthase by 4,7-dioxosebacic acid, an inhibitor that shows dramatic species selectivity. *Biochemistry* 40, 8227–8236. doi: 10.1021/bi010656k
- Kokona, B., Rigotti, D. J., Wasson, A. S., Lawrence, S. H., Jaffe, E. K., and Fairman, R. (2008). Probing the oligomeric assemblies of pea porphobilinogen synthase by analytical ultracentrifugation. *Biochemistry* 47, 10649–10656. doi: 10.1021/bi801128d
- Kopeckova, M., Pavkova, I., and Stulik, J. (2020). Diverse localization and protein binding abilities of Glyceraldehyde-3-Phosphate dehydrogenase in pathogenic bacteria: the key to its multifunctionality? *Front. Cell. Infect. Microbiol.* 10:89.
- Lawrence, S. H., Ramirez, U. D., Tang, L., Fazliyev, F., Kundrat, L., Markham, G. D., et al. (2008). Shape shifting leads to small-molecule allosteric drug discovery. *Chem. Biol.* 15, 586–596. doi: 10.1016/j.chembiol.2008.04.012
- Lawrence, S. H., Selwood, T., and Jaffe, E. K. (2011). Diverse clinical compounds alter the quaternary structure and inhibit the activity of an essential enzyme. *Chem. Med. Chem.* 6, 1067–1073. doi: 10.1002/cmdc.201100009
- Lawrence, S. H., Selwood, T., and Jaffe, E. K. (2013). Environmental contaminants perturb fragile protein assemblies and inhibit normal protein function. *Curr. Chem. Biol.* 7, 196–206. doi: 10.2174/2212796811307020011
- Lazarev, V. F., Guzova, I. V., and Margulis, B. A. (2020). Glyceraldehyde-3-phosphate dehydrogenase is a multifaceted therapeutic target. *Pharmaceutics* 12:416. doi: 10.3390/pharmaceutics12050416
- Long, M. J. C., Hall-Beauvais, A. Van, and Aye, Y. (2020). The more the merrier: how homo-oligomerization alters the interactome and function of ribonucleotide reductase. *Curr. Opin. Chem. Biol.* 54, 10–18. doi: 10.1016/j.cbpa.2019.09.003
- Long, M. J. C., Hnedzko, D., Kim, B. K., and Aye, Y. (2019). Breaking the fourth wall: modulating quaternary associations for protein regulation and drug discovery. *ChemBiochem* 20, 1091–1104. doi: 10.1002/cbic.201800716
- Lopez-Pelegrin, M., Cerda-Costa, N., Cintas-Pedrola, A., Herranz-Trillo, F., Peinado, J. R., et al. (2014). Multiple stable conformations account for reversible concentration-dependent oligomerization and autoinhibition of a metamorphic metalloproteinase. *Angew. Chem. Int. Ed. Engl.* 53, 10624–10630. doi: 10.1002/anie.201405727
- Lu, S., He, X., Ni, D., and Zhang, J. (2019a). Allosteric modulator discovery: from serendipity to structure-based design. *J. Med. Chem.* 62, 6405–6421. doi: 10.1021/acs.jmedchem.8b01749
- Lu, S., Shen, Q., and Zhang, J. (2019b). Allosteric methods and their applications: facilitating the discovery of allosteric drugs and the investigation of allosteric mechanisms. *Acc. Chem. Res.* 52, 492–500. doi: 10.1021/acs.accounts.8b00570
- Mani, M., Chen, C., Amblee, V., Liu, H., Mathur, T., Zwicke, G., et al. (2015). MoonProt: a database for proteins that are known to moonlight. *Nucleic Acids Res.* 43, D277–D282.
- Maruno, M., Furiyama, K., Akagi, R., Horie, Y., Meguro, K., Garbaczewski, L., et al. (2001). Highly heterogeneous nature of delta-aminolevulinatase dehydratase

- (ALAD) deficiencies in ALAD porphyria. *Blood* 97, 2972–2978. doi: 10.1182/blood.v97.10.2972
- Mills-Davies, N., Butler, D., Norton, E., Thompson, D., Sarwar, M., Guo, J., et al. (2017). Structural studies of substrate and product complexes of 5-aminolaevulinic acid dehydratase from humans, *Escherichia coli* and the hyperthermophile *Pyrobaculum calidifontis*. *Acta Crystallogr. D. Struct. Biol.* 73(Pt 1), 9–21.
- Mitchell, L. W., and Jaffe, E. K. (1992). *Escherichia-Coli* porphobilinogen synthase is a Zn(II) metalloenzyme. *FASEB J.* 6, A459–A459.
- Mitchell, L. W., and Jaffe, E. K. (1993). Porphobilinogen synthase from *Escherichia-coli* is a Zn(II) metalloenzyme stimulated by Mg(II). *Arch. Biochem. Biophys.* 300, 169–177. doi: 10.1006/abbi.1993.1024
- Monod, J. (1965). [Reflections on the relationship between the structure and function of globular proteins]. *Annee Biol.* 59, 231–240.
- Motlagh, H. N., Wrabl, J. O., Li, J., and Hilser, V. J. (2014). The ensemble nature of allostery. *Nature* 508, 331–339. doi: 10.1038/nature13001
- Murzin, A. G. (2008). Metamorphic proteins. 320, 1725–1726. doi: 10.1126/science.1158868
- Parisi, G., Zea, D. J., Monzon, A. M., and Marino-Buslje, C. (2015). Conformational diversity and the emergence of sequence signatures during evolution. *Curr. Opin. Struct. Biol.* 32, 58–65. doi: 10.1016/j.sbi.2015.02.005
- Park, C. K., and Horton, N. C. (2019). Structures, functions, and mechanisms of filament forming enzymes: a renaissance of enzyme filamentation. *Biophys. Rev.* 11, 927–994. doi: 10.1007/s12551-019-00602-6
- Park, C. K., and Horton, N. C. (2020). Novel insights into filament-forming enzymes. *Nat. Rev. Mol. Cell Biol.* 21, 1–2. doi: 10.1038/s41580-019-0188-1
- Petrovich, R. M., Litwin, S., and Jaffe, E. K. (1996). Bradyrhizobium japonicum porphobilinogen synthase uses two Mg(II) and monovalent cations. *J. Biol. Chem.* 271, 8692–8699. doi: 10.1074/jbc.271.15.8692
- Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., and Serrano, L. (2005). The FoldX web server: an online force field. *Nucleic Acids Res.* 33, W382–W388.
- Selwood, T., and Jaffe, E. K. (2012). Dynamic dissociating homo-oligomers and the control of protein function. *Arch Biochem. Biophys.* 519, 131–143. doi: 10.1016/j.abb.2011.11.020
- Selwood, T., Tang, L., Lawrence, S. H., Anokhina, Y., and Jaffe, E. K. (2008). Kinetics and thermodynamics of the interchange of the morphein forms of human porphobilinogen synthase. *Biochemistry* 47, 3245–3257. doi: 10.1021/bi702113z
- Shanmugam, D., Wu, B., Ramirez, U., Jaffe, E. K., and Roos, D. S. (2010). Plastid-associated porphobilinogen synthase from *Toxoplasma gondii*: kinetic and structural properties validate therapeutic potential. *J. Biol. Chem.* 285, 22122–22131. doi: 10.1074/jbc.m110.107243
- Simonet, J. C., Burrell, A. L., Kollman, J. M., and Peterson, J. R. (2020). Freedom of assembly: metabolic enzymes come together. *Mol. Biol. Cell* 31, 1201–1205. doi: 10.1091/mbc.e18-10-0675
- Stolz, M., and Dornemann, D. (1996). Purification, metal cofactor, N-terminal sequence and subunit composition of a 5-aminolevulinic acid dehydratase from the unicellular green alga *Scenedesmus obliquus*, mutant C-2A'. *Eur. J. Biochem.* 236, 600–608. doi: 10.1111/j.1432-1033.1996.00600.x
- Tang, L., Breinig, S., Stith, L., Mischel, A., Tannir, J., Kokona, B., et al. (2006). Single amino acid mutations alter the distribution of human porphobilinogen synthase quaternary structure isoforms (morpheins). *J. Biol. Chem.* 281, 6682–6690. doi: 10.1074/jbc.m511134200
- Thomas, W. C., Brooks, F. III, Burnim, A. A., Bacik, J., Stubbe, J., Kaelber, J. T., et al. (2019). Convergent allostery in ribonucleotide reductase. *Nat. Commun.* 10:2653.
- Wasserman, H., and Saphire, E. O. (2016). More than meets the eye: hidden structures in the proteome. *Annu. Rev. Virol.* 3, 373–386. doi: 10.1146/annurev-virology-100114-054923
- Wu, W. H., Shemin, D., Richards, K. E., and Williams, R. C. (1974). The quaternary structure of delta-aminolevulinic acid dehydratase from bovine liver. *Proc. Natl. Acad. Sci. U.S.A.* 71, 1767–1770. doi: 10.1073/pnas.71.5.1767
- Conflict of Interest:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Jaffe. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.