

# Mechanism of molecular oxygen diffusion in a hypoxia sensing prolyl hydroxylase using multiscale simulation.

Carmen Domene,<sup>a,b,c,\*</sup> Christian Jorgensen,<sup>b</sup> and Christopher J. Schofield<sup>a,\*</sup>

<sup>a</sup> Chemistry Research Laboratory, Mansfield Road, University of Oxford, Oxford OX1 3TA, United Kingdom.

<sup>b</sup> Department of Chemistry, Britannia House, 7 Trinity Street, King's College London, London SE1 1DB, United Kingdom.

<sup>c</sup> Department of Chemistry, University of Bath, Claverton Down Bath BA2 7AY, United Kingdom

\* Corresponding authors: C.Domene@bath.ac.uk Tel: +44 (0)1225 386172  
christopher.schofield@chem.ox.ac.uk Tel: +44 (0)1865 275625

## Abstract

The chronic response of animals to hypoxia is mediated by the  $\alpha\beta$ -heterodimeric hypoxia-inducible transcription factors ( $\alpha,\beta$ -HIFs) which upregulate the expression of sets of genes that work to ameliorate the effects of limiting dioxygen. The HIF prolyl hydroxylase domain enzymes (PHDs) are Fe(II) and 2-oxoglutarate dependent oxygenases that act as hypoxia sensing components of the HIF system: prolyl-hydroxylation signals for dioxygen availability dependent HIF- $\alpha$  degradation via the ubiquitin proteasome system. The unusual kinetic properties of the PHDs, in particular a high  $K_m$  for dioxygen and slow reaction with dioxygen are proposed to enable their hypoxia sensing role. An understanding of how dioxygen is delivered to, and binds at, the active site of the PHDs is important for the development of a chemical understanding of the hypoxic response. We employed a combined multiscale approach involving classical atomistic equilibrium and non-equilibrium MD simulations combined with QM/MM trajectories to investigate dioxygen diffusion to, and binding at, the active site in the PHD2.Fe(II).2OG.HIF substrate complex; PHD2 is the most important of the three human PHDs. The transport of dioxygen to the active site is described; dioxygen transport follows a single well-defined hydrophobic tunnel, formed from both enzyme and substrate elements to reach the PHD2 active site. The results provide estimates for rate constants that define a diffusion-reaction model for dioxygen:PHD2 interactions; in combination with reported biophysical analyses they provide chemical insight into the basis of the slow reaction of PHD2 with dioxygen. They imply that the reversible binding of dioxygen is central to the hypoxia sensing capacity of the PHDs and that different PHD HIF- $\alpha$  substrate combinations might have different dioxygen sensitivity profiles. The extent of HIF- $\alpha$  substrate prolyl hydroxylation, which signals for subsequent HIF- $\alpha$  degradation is thus a manifestation of the equilibrium between dioxygen in bulk solution and dioxygen bound to the PHD2.Fe.2OG.HIF- $\alpha$  substrate complex.

## Introduction

The chronic response of animals, including humans, to hypoxia is substantially mediated by the  $\alpha\beta$ -heterodimeric hypoxia-inducible transcription factors ( $\alpha,\beta$ -HIFs), which upregulate the expression of a set of genes that work, in a context dependent manner, to ameliorate the effects of limiting molecular oxygen (dioxygen) availability at cellular, tissue and whole organism levels.<sup>1-2</sup> The key process by which the HIF system senses hypoxia involves *trans*-4-prolyl-hydroxylation of HIF- $\alpha$  subunits, of which there are three isoforms in humans (HIF1-3 $\alpha$ ), with HIF-1 $\alpha$  and HIF-2 $\alpha$  being extensively studied.<sup>1-2</sup> HIF-1 $\alpha$  and HIF-2 $\alpha$  hydroxylation occurs at two sites in an oxygen dependent degradation domain (ODD), i.e. the N-terminal ODD (NODD) and the C-terminal ODD (CODD). Human HIF- $\alpha$  hydroxylation is catalyzed by prolyl hydroxylase domain enzymes 1 to 3 (PHD1-3)<sup>3-4</sup> of which PHD2 is most conserved in animals and likely the most important PHD.<sup>1-2,5</sup> Prolyl-hydroxylation signals for HIF- $\alpha$  degradation via the ubiquitin proteasome system. The PHDs are Fe(II) and 2-oxoglutarate dependent oxygenases. When cellular dioxygen levels become limiting, PHD catalysis slows, HIF- $\alpha$  degradation slows, and HIF- $\alpha$  levels rise; HIF- $\alpha$  can then translocate to the nucleus and it can dimerize with HIF- $\beta$  and promote the transcription of HIF target genes.<sup>5</sup>

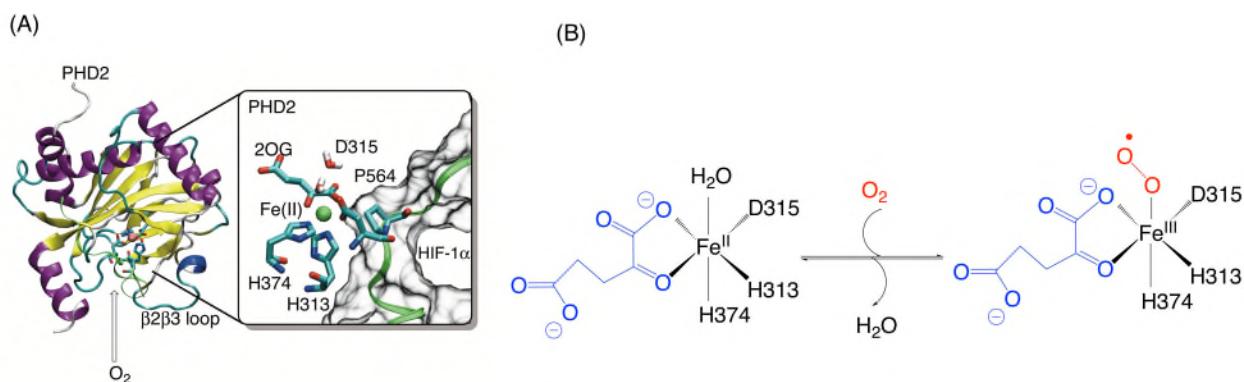
Despite their central sensing role in the hypoxic response, the underlying kinetic basis of the dioxygen-sensing function of the PHDs remains unclear. Given their sensing role, hydroxylation of HIF- $\alpha$  by the PHDs must be limited by dioxygen availability. However, little is known about the pathway or pathways that dioxygen might take to access the PHD active site. By contrast with some other 2-oxoglutarate oxygenases, experiments have demonstrated a high  $K_m$  for the PHDs<sup>6-7</sup> and that the reaction of a complex of PHD2, Fe(II), 2-oxoglutarate (2OG) and the C-terminal oxygen dependent degradation domain (CODD) of HIF- $\alpha$  with dioxygen to form hydroxylated CODD and succinate is substantially slower (~100-fold) than for other characterized 2OG oxygenases.<sup>8</sup>

The catalytic domain of PHD2 has been crystallized in several different forms, including without HIF- $\alpha$  substrate and in complex with Mn(II) (as a Fe(II) surrogate), its co-substrate 2OG (or N-oxalylglycine a close 2OG analogue), and the NODD and CODD sequences of its HIF-1 $\alpha$  substrate (Figure 1.A).<sup>9-12</sup> PHD2 has a distorted double-stranded  $\beta$ -helix (DSBH) core fold, which together with N-terminal elements is conserved in other 2OG oxygenases.<sup>13-14</sup> This

structural motif acts as a scaffold for the active site and surrounding loops, including the dynamic  $\beta 2\beta 3$  loop and C-terminal region.<sup>11-12</sup> The  $\beta 2\beta 3$  loop is important in substrate binding, which involves substantial conformational changes in the PHDs and which has been proposed to be involved in entry of dioxygen into the active site.<sup>12</sup>

As is the case with the PHDs, the Fe(II) centre in 2OG dependent hydroxylases is normally, but not always, coordinated by the side chains of three protein ligands, the His2-(Glu/Asp) triad (His-313, Asp-315 and His-374 in PHD2).<sup>13-14</sup> In the consensus ordered sequential mechanism of the 2OG oxygenases, 2OG binding is followed by that of substrate, and then dioxygen (Figure 1).<sup>15</sup> Substrate binding adjacent to the Fe(II) center is proposed to weaken binding of a coordinated water (often, but not always, observed by crystallography in enzyme.Fe(II).2OG.substrate complexes in the absence of dioxygen), thus enabling the ligation of dioxygen to Fe(II). In the case of PHD2, the Fe-ligating water appears to be particularly tightly bound. The reaction of dioxygen with Fe(II) comprises reduction of dioxygen concomitant with oxidation of Fe(II) to Fe(III) to give a metal associated superoxide species. Oxidative decarboxylation of 2OG gives carbon dioxide and a succinyl peroxide intermediate (at least in some cases<sup>15</sup>), which rearranges to give a reactive Fe(IV)=O species which is responsible for (prolyl) hydroxylation.<sup>8</sup>

Crystallographic analyses have suggested that during binding of a HIF- $\alpha$  ODD to PHD2, substantial conformational changes occur and that the PHD2.Fe(II).2OG. HIF- $\alpha$  complex likely has an unusually narrow entrance to its active site.<sup>9-12</sup> However, these factors alone are unlikely to be the cause of the slow reaction of PHD2 with dioxygen. It has been proposed that the rate limiting / unusually slow reaction of PHD2 with dioxygen reflects a need for dioxygen to displace the (unusually) tightly Fe(II)-coordinated water which is stabilized by hydrogen bonding and / or Fe-centred rearrangement during catalysis.<sup>6, 8, 11-12</sup>



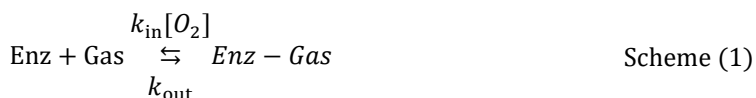
**Figure 1.** The catalytic domain of the hypoxia sensing prolyl hydroxylase PHD2. (A) Crystallographically derived view of PHD2 in complex with Fe(II), 2oxoglutarate (2OG) and the HIF-1 $\alpha$  C-terminal oxygen dependent degradation domain (CODD). The view was derived from PDB id 3HQR.<sup>10</sup> The entrance to the dioxygen diffusion path identified in this study is indicated by an arrow. The inset shows an enlarged representation of the PHD2 active site. The Fe(II) is shown as a green sphere and the residues that are part of the 2His-1carboxylate HXD...H triad are represented by the thicker lines. The HIF-1 $\alpha$  substrate is shown as a green ribbon with residue P564 (which is hydroxylated) highlighted. (B) Schematic outline of the studied diffusion-reaction model for PHD2.

In this context of gas diffusion, diffusion pathways in non-heme and heme-dependent proteins, including deoxyhaemoglobin,<sup>16</sup> myoglobin,<sup>17</sup> FeFe hydrogenases,<sup>18-19</sup> globins,<sup>20</sup> and [Fe]-hydrogenase,<sup>21</sup> amongst others, have been studied using a variety of computational techniques, including locally enhanced sampling (LES),<sup>20, 22-25</sup> umbrella sampling,<sup>25</sup> and steered MD simulations.<sup>26-27</sup> In particular, migration of hydrogen, dioxygen, and carbon monoxide gases has been extensively studied in [NiFe]-hydrogenases,<sup>19, 26, 28-30</sup> with the results of Xe absorption studies being used to probe gas binding sites.<sup>29, 31</sup> The question of whether a single channel or multiple channels connects the enzyme active site with bulk solution has been considered.<sup>22, 25</sup> On the basis of previous work, identified dioxygen diffusion pathways have been proposed to share several common characteristics. In some cases, multiple pathways are proposed to converge to a predominantly hydrophobic main tunnel<sup>32</sup> which can accommodate clusters of gas molecules.<sup>25, 33, 22</sup> Compared to carbon monoxide, dioxygen diffusion has been reported to sample a greater volume of the same protein.<sup>29</sup>

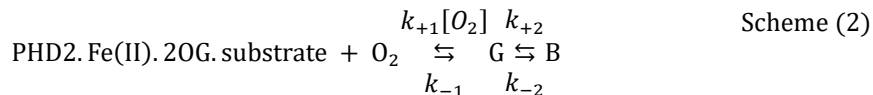
Our microscopic understanding of gas transport in oxygenases will be strengthened by computing diffusion rates and describing the pathways available to the gas to reach the catalytic center.<sup>19, 26</sup> Kinetic experiments typically yield an overall rate for ligand binding, but they usually do not provide information on whether it is the diffusion of the ligand into the active site

or the coupling of the ligand in the active site that determines the rate of the process. An understanding of how dioxygen is delivered to and binds at the active site of the 2OG dependent HIF hydroxylases is important for the development of a chemical understanding of the roles of individual enzymes in the hypoxic response under different conditions, in particular different dioxygen concentrations.

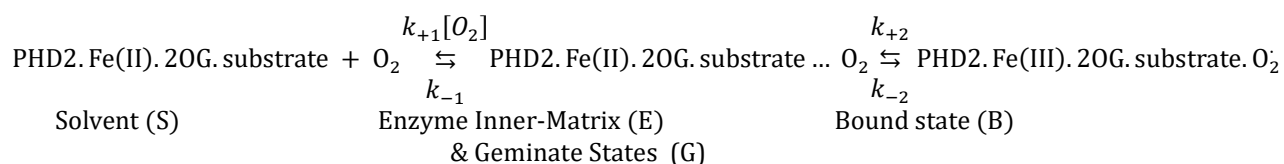
Classical molecular dynamics is a widely used computational technique that complements the knowledge gained from experimental structures providing dynamical information. The timescale accessible by MD simulations has rapidly increased in recent years, and some mechanistic and functional properties can now be estimated directly from MD trajectories.<sup>30, 34</sup> Nevertheless, force-field based MD methods alone are not sufficiently accurate in describing chemical binding to an active site especially if an active site involves a transition metal cluster with a complicated electronic structure. Classical MD simulation can only give a partial picture of the ligand binding process. We therefore adopted a general multiscale molecular simulation approach for the calculation of diffusion rates and determination of pathways by which dissolved dioxygen can reach the PHD2 active site. The method is not specific to this system and should be more generally applicable to the transport of small molecules in proteins.<sup>29</sup> A model, termed the diffusion-reaction model<sup>27</sup> as illustrated in Scheme (1), was adopted to describe the diffusion of dioxygen into PHD2:



This is a kinetic two-step model consisting of an initial step where the dioxygen diffuses followed by a metal-dioxygen molecule ligation (chemical reaction) step.<sup>29-30</sup> Within this framework, let us consider the binding of a dioxygen molecule to the PHD2.Fe(II).2OG.substrate active site, bound state B, via an intermediate state G, as illustrated in Scheme 2:



The mechanism can be further described as:



Scheme 3 illustrates a two-step diffusion process of the form: S to E to G to B. In the first step, dioxygen diffuses from the solvent (S) into the enzyme (E) inner-matrix. In the second step, dioxygen diffuses from the E state into the active site pocket, termed the geminate (G) state. In the G state, dioxygen is free to diffuse into the metal active site but it is not actually ligated to Fe(II). In the last step, dioxygen chemically binds to the metal center. It has postulated displacement of the water coordinating the iron center in the PHD2.Fe.2OG.substrate complex limits the rate of dioxygen binding.<sup>35</sup> The reaction of dioxygen with Fe(II) involves reduction of dioxygen to give a superoxide with concomitant oxidation of Fe(II) to Fe(III), resulting in the radical bound state (B) which likely reacts rapidly with 2OG.

To investigate the binding of dioxygen to PHD2, we employed a general multiscale molecular simulation approach to calculate the diffusion rates and examine the pathways by which dissolved dioxygen reaches the active site of the catalytic domain of the PHD2.Fe(II).2OG.substrate complex. The rate constants for the diffusive transitions between solvent and protein cavities and between the protein cavities to the different clusters were obtained from equilibrium and non-equilibrium MD simulations and the rate constants ( $k_{+2}$  and  $k_{-2}$ ) were obtained from density functional theory (DFT) calculations. The results provide insight into the chemical basis of unusually slow reaction of PHD2 with dioxygen and provide new mechanistic detail; they support the proposal that the reversible binding of dioxygen to the PHD proteins enables their hypoxia sensing capability.

## Materials and Methods

The initial configuration of the catalytic domain of the PHD2.Fe(II).2OG.HIF-1 $\alpha$  CODD fragment complex was taken from the crystal structure PDB id: 3HQR (PHD2 residues 186-421).<sup>10</sup> In the crystalline state, Mn(II) was substituted for Fe(II) and 2OG was substituted for N-oxalylglycine (NOG) to give a stable complex; the Mn(II) was replaced with Fe(II) and NOG was substituted for 2OG and the complex was solvated using the solvate plugin from VMD. The results predict that in solution, the overall structure of the PHD2.Fe(II).2OG.CODD complex (hereafter PHD2, except where specified) is similar to that observed for the analogous crystal

structure.<sup>10</sup> Then, 200 randomly chosen water molecules were replaced by dioxygen, which represents a concentration of ~600 mM. A relatively high concentration of dioxygen was chosen to obtain improved statistics for simulation of its diffusion process into PHD2. Interactions between gas molecules were assumed to have a negligible effect on the dynamics of gas diffusion. The protein was stable over the course of the simulation as reflected by the backbone root mean square deviation. Calculations were performed using NAMD 2.9,<sup>36</sup> employing the CHARMM force-field<sup>37</sup> for the protein together with the TIP3P water model.<sup>38</sup> Molecular oxygen and Fe(II) Lennard Jones parameters were taken from CHARMM.<sup>21,31</sup> After 500 steps of conjugate-gradient minimization with restraints on the protein, and 50 ps of simulation with backbone restraints, a 300 ns production run in the NPT ensemble was carried out.

The rate constant for transport of dioxygen from the bulk solvent (S) to the first inner-enzyme state (E) was estimated from unbiased MD simulations. The rate constant for the second transition (E) to (G),  $k_{GE}$ , was obtained from constant-force steered molecular dynamics (SMD) simulations using NAMD 2.9.<sup>36</sup> A single dioxygen molecule was used to sample the E to G transition in five independent sets of simulations. Forces of 140, 150, 160, 180, 200 kJ·mol<sup>-1</sup>·nm<sup>-1</sup> were applied to bias the trajectory of the dioxygen molecule along the pulling coordinate which corresponds to the vector connecting the center of mass (COM) of the dioxygen molecule and the Fe(II) ion. The time required for the transition between the two states was averaged over up to 100 trajectories with initial velocities drawn from a Boltzmann distribution at 300 K to yield the mean first passage time (MFPT),  $\tau_{G \leftarrow E}(F)$ , where  $F$  is the pulling force. The MFPT was extrapolated to zero force using a regression fit onto the Dudko-Hummer-Szabo (DHS) rate model for pulling. At a force  $F$  along the pulling coordinate and over a free-energy barrier  $\Delta G^\ddagger$ ; the parameter  $\nu$  describes two underlying free-energy surfaces (FES), a harmonic well FES with  $\nu = 1/2$  and a linear-polynomial FES DHS expression with  $\nu = 2/3$  that features a cusp-like barrier.<sup>39</sup> This predicts a relationship between the rate constant of diffusion and the pulling force (Equation 1), and by extension, the transition time (Equation 2).

$$k(F) = k_0 \left( 1 - \frac{\nu F x^\ddagger}{\Delta G^\ddagger} \right)^{1/\nu-1} e^{-\beta \Delta G^\ddagger \left( 1 - \left( 1 - \frac{\nu F x^\ddagger}{\Delta G^\ddagger} \right)^{1/\nu} \right)} \quad \text{Equation (1)}$$



$$\tau_{G \leftarrow E}(F) = 1 / k_{G \leftarrow E}(F) \quad \text{Equation (2)}$$

Where MFPTs were unavailable, they were extracted from an estimate from transition state theory (TST) using knowledge of the free energy barriers (Equation 3), and by extension, the MFPT was linked to the free-energy difference for a process (Equation 4).

$$\Delta G^\ddagger = \beta^{-1} \ln(\beta^{-1} \tau_{ji} / h) \quad \text{Equation (3)}$$

$$\Delta G = -\beta^{-1} \ln(k_{ij}/k_{ji}) = -\beta^{-1} \ln(\tau_{ij}/\tau_{ji}) \quad \text{Equation (4)}$$

The Adaptive Biasing Force (ABF) method<sup>40</sup> as implemented in NAMD 2.9<sup>36</sup> was used with the distance between Fe(II) and the center of mass of a dioxygen molecule as the collective variable ( $\xi$ ) to gain insight into the diffusion of dioxygen between the E and G states in PHD2. The range of  $\xi$  was chosen from 3 to 10 Å, in bins of 0.25 Å widths. The potential of mean force (PMF) was reconstructed by numerical integration of the re-assembled free-energy gradient  $dG/d\xi$ . Error analysis of the PMF was calculated using the Rodriguez-Gomez-Darve-Pohorille method<sup>41-42</sup> (see Supplementary Material).

To obtain an approximate estimate of the diffusion rate,  $k_{+1}$ , the individual rate constants for diffusive sub-steps were combined into a toy Markov chain. The states input into the Markov chain were obtained by clustering dioxygen densities from the equilibrium MD trajectories. All dioxygen molecules inside the protein complex were assigned to states, with a single state corresponding to dioxygen molecules in the solvent, defined as regions with zero density of protein atoms. The E state corresponds to the initial site of density for dioxygen in the inner-protein space, and the G state corresponds to the likely binding site of dioxygen to the Fe(II) center in the active site pocket, immediately prior to the ‘chemical reaction’.

A diffusion equation was employed to extract an estimate for  $k_{+1}$  from the equilibrium MD simulation statistics based on a simplified Markov chain estimate of the probability of residing in state  $i$ :<sup>43</sup>

$$\dot{p}_i(t) = \sum_j k_{ij} p_j(t) \quad \text{Equation (5)}$$

A three site Markov chain was built for the (S) to (E) to (G) transition. The rate constants  $k_{+1}$ ,  $k_{-1}$ , were obtained by a least-squares fit of the probability of occupying the active site geminate state G,  $p_G(t)$  (Equation 6) onto the phenomenological Equation 7. This is considered a valid approach for an assumed pseudo-first order kinetic regime at constant dioxygen concentrations.

$$p_G(t) = \alpha \exp(\lambda_1 t) + \beta \exp(\lambda_2 t) + \gamma \exp(\lambda_3 t) \quad \text{Equation (6)}$$

$$p_G(t) = \frac{k_{+1}[O_2]}{k_{+1}[O_2] + k_{-1}} \left[ 1 - \exp\left(-(k_{+1}[O_2] + k_{-1})t\right) \right] \quad \text{Equation (7)}$$

The TST estimated rate constant  $k_{+2}$  was obtained from QM/MM simulations describing the association of dioxygen to the active site Fe(II). Using the individual rate constants  $k_{+1}$ ,  $k_{-1}$ , and  $k_{+2}$ , the steady-state kinetics of the binding of dioxygen with the protein was approximated (Equation 8). It was assumed that (i) at the saturation limit and (ii) upon dioxygen binding to Fe(II) in the bound state B, limited recrossing occurs, such that the unbinding rate  $k_{-2}$  is smaller than  $k_{+2}$  ( $k_{+2} \gg k_{-2}$ ). This would hold true for PHD2 turnover of dioxygen that has entered into the catalytic B site.

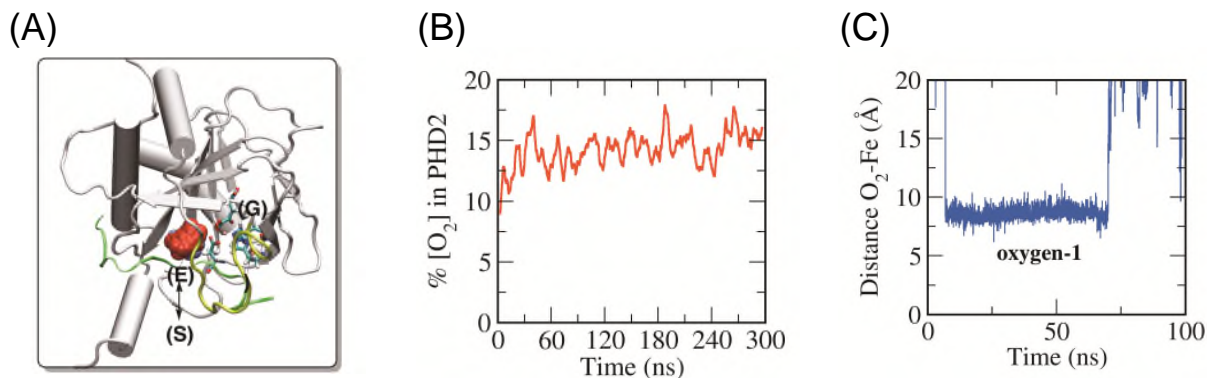
$$k_{in}^{ss} = \frac{k_{+1}k_{+2}}{k_{+1} + k_{+2}} \quad \text{Equation (8)}$$

In order to estimate  $k_{+2}$ , the ligation of dioxygen to high spin Fe(II) was simulated using *ab initio* QM/MM metadynamics,<sup>44</sup> as implemented in the DFT CP2K-QuickStep code<sup>45-46</sup> using Gaussian plane waves. The collective variable considered ( $\xi_a$ ) was the same as for MD simulations with ABF, that is, the center-of-mass distance of dioxygen to Fe(II). The bias along  $\xi$  was added in the form of a Gaussian function and the potential of mean force was reconstructed as the integral over all the Gaussian functions added.

## Results and Discussion

The diffusion of dioxygen into the PHD2 active site was observed to proceed via a single pathway as described in Figure 2A. Entry to the PHD2 active site occurs through a gate of width and height  $\sim 8.0$  Å that leads to a predominantly hydrophobic tunnel. A number of residues contribute in

delineating the path and rate of dioxygen entry into the PHD2 active site through the tunnel, including PHD2 Q239, I256 and W258. Notably, the substrate also contributes to the tunnel including HIF-1 $\alpha$  P564, Y565 and I566 residues. The fraction of dioxygen within the protein matrix was calculated as a percentage of the total 200 dioxygen molecules within a distance of 3.5 Å of the protein surface. On average, 12-16 water molecules were observed to populate the hydrophobic tunnel. On average only a single dioxygen molecule transition inside the hydrophobic tunnel, with a second dioxygen molecule residing at the entrance of the tunnel was consistently being observed, apparently poised for entry. The shape of the tunnel is illustrated in Figure 2.A. At the steady-state, 10-15% of dioxygen molecules reside on the surface of the PHD2.substrate complex (Figure 2.B). Dioxygen was observed to occupy the E state with an average  $10^2$  ns residence time within a distance of 8.0-9.0 Å from the metal center (Figure 2C). Dioxygen distributions were calculated using a 15 and 25 Å cut-offs and iso-contour surfaces of the dioxygen probability were obtained by discretizing the probability of finding any of the total 200 dioxygen molecules throughout a 300 ns trajectory. Contour calculations were done with the VMD Volmap feature.<sup>47</sup> Discretization of dioxygen states by iso-contour densities yielded 3, 9, and 19 states, respectively (See Supplementary Material). A single reactive trajectory was observed for PHD2, which was observed to be a transition between states S to E to G to B; of these only the S to E transition was detected in the classical MD simulation timescale. The E to G transition was simulated by SMD and the transition from G to B with QM/MM metadynamics.



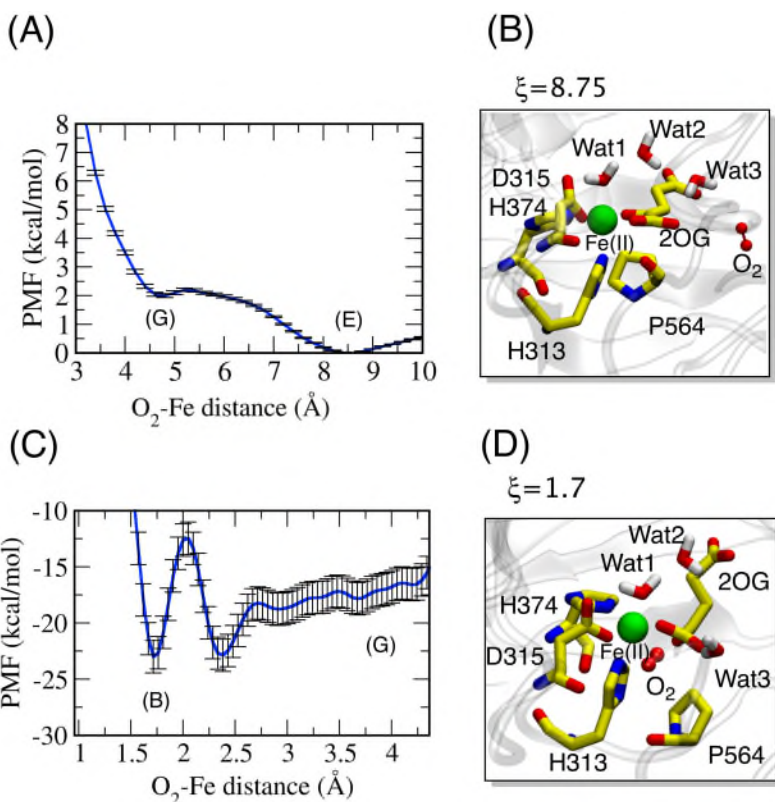
**Figure 2.** Results of unbiased MD simulations of dioxygen flooding into the PHD2.Fe(II).2OG.substrate (PHD2) complex active site. (A) Iso-density for the resulting dioxygen locations is in red; the productive pathway from S to E to G is indicated. Note that the dioxygen binding path is bordered by elements from both the enzyme and the substrate. (B) Percentage of dioxygen molecules within 3.5 Å of the protein surface as a function of simulation time. (C) Dioxygen-Fe(II) distance in the E state as a function of simulation time

To obtain estimates of mean first passage times (MFPT) for the transfer of dioxygen from state E to state G, five independent SMD simulations were run. The MFPT is the average time it takes for a gas molecule in state  $j$  to make a transition to state  $i$ . In order to obtain an MFPT converged estimate, each force set required up to 100 independent trajectories. Each trajectory was extended up to 10 ns. Only trajectories where the transfer was successful in bringing dioxygen within 5.0 Å of the metal center were considered. A regression of the SMD MFPT data fitted this data onto the logarithmic form of the Dudko-Hummer-Szabo (DHS) expression (Equation 5). This yielded MFPTs at zero pull force of 50 ns and 27 ns for the harmonic and linear polynomial models. From transition state theory, the SMD MFPT values translate to a  $\Delta G_{EG}$  of 2.4-2.5 kcal mol<sup>-1</sup> for the harmonic-well FES and the linear-polynomial FES, respectively (Table 2). The free energy barrier for dioxygen crossing was estimated from regression fit of the SMD MFPT data in the DHS expression, to give 5.0-7.0 kcal mol<sup>-1</sup> for the forward direction (entry into PHD2 state G), and 2.4-2.9 kcal mol<sup>-1</sup> for the reverse direction (exit from PHD2 state G). Importantly, these barrier heights illustrate that dioxygen entry into the active site is rate-limiting (compared to dioxygen exit). The extrapolated MFPT at zero force values, ranged from 27 to 50 ns, and it provides a bounded estimate of the residence time of dioxygen in the E state prior to transition into the G state.

Adaptive Biasing Force (ABF) simulations of dioxygen diffusion were employed in order to compare the E to G transition from SMD. The outputs from the ABF simulations reveal a transition pathway barrier of 2 kcal mol<sup>-1</sup> between the E state, located 8.8 Å from the metal ion, to the G state, which is located 4.4 Å from the metal ion (Figure 3.A). This is in agreement with the SMD pulling distance histograms where the average O<sub>2</sub>-Fe(II) arrival distance is 4.05-4.17 Å, with the conformation depicted in Figure 3B. An error analysis was performed using a bin of width 0.25 Å consisting of a correlated time-series of 175,000 force samples. The statistical variance of the instantaneous force time series value was 78 N, and the result of applying the Flyvbjerg-Petersen block renormalization algorithm to estimate the variance of the average force was found to be 0.01 N. The correlation length was calculated to be  $\tau=0.38$ ;  $\sigma(\Delta G) = 0.003$  kcal mol<sup>-1</sup> in the bin corresponding to  $\xi=8.75$  Å. Subsequently, this error analysis was propagated along the rest of the bins, such that the errors ranged between  $\pm 0.003$  and  $\pm 0.050$  kcal mol<sup>-1</sup>.

The cosubstrate 2-oxoglutarate (2OG) coordinates the metal ion in a bidentate mode in PHD2 as shown in Figure 1. Although our work principally focuses on the passage of dioxygen to the PHD2 active site rather than its chemical mechanism, in our simulations, it is notable that upon arrival of dioxygen to the G state, 2OG alternates between a bidentate and a monodentate coordination (Figure 3B, D). Comparison between the bidentate and monodentate configurations suggests that this change in complexation corresponds to a lower-energy monodentate state; the bidentate to monodentate energy difference is  $\Delta\Delta G = +0.6 \text{ kcal mol}^{-1}$ . A monodentate state has been suggested as an intermediate step in the catalytic cycles of some 2OG oxygenases.<sup>10, 48-49</sup> In 2OG complexed crystal structures, the coordination position of the 2OG C-1 carboxylate has been observed to adopt two geometries, either being *trans* to His-313 or *trans* to His-374 (using PHD2 numbering).<sup>9-10</sup> The latter has been observed in PHD2 structures (as in Figure 3B)<sup>10</sup>, creating a mechanistic problem, i.e. that there is not a vacant coordination position for dioxygen to bind adjacent to the substrate at the active site.<sup>10</sup> This could be solved by a rearrangement subsequent to dioxygen reaction with the Fe(II),<sup>50</sup> or as implied by our work on PHD2, by a change in complexation prior to dioxygen reaction with the active site metal ion.

The incorporation of an oxygen atom from atmospheric dioxygen into an organic substrate as occurs during PHD catalysis requires the reaction of triplet spin state dioxygen with a singlet spin state organic substrate. Although such reactions are thermodynamically favorable, they are kinetically slow because of the spin mismatch.<sup>51</sup> To overcome the kinetic barrier for incorporating an oxygen atom into its HIF- $\alpha$  substrate, PHD2, like most other 2OG dependent family oxygenases, uses mononuclear non-heme ferrous iron and 2OG.<sup>9</sup> Starting from a conformation from the converged ABF simulation of the E to G transition, where dioxygen is located in the G state ( $\xi = 4.3 \text{ \AA}$ ), an *ab initio* (QM/MM) metadynamics simulation was performed to estimate the free energy for the reaction of dioxygen and Fe(II) to give an Fe(III) superoxide complex or state B. The calculation started with 2OG complexed to Fe(II) via both its C-1 carboxylate and ketone oxygen atoms, as observed in the PHD2 crystal structures and those of most other 2OG oxygenases.<sup>13-15</sup>



**Figure 3.** Binding of dioxygen to the active site metal center. (A) Dioxygen diffusion potential of mean force (PMF) as calculated with ABF simulations for dioxygen entry into the G state with the collective variable ( $\xi$ ) set as the distance between the center of mass of dioxygen and the metal center. (B) Representative conformation of the active site as taken at  $\xi = 8.75$  Å, in which dioxygen is located in the E state. (C) Potential of mean force of dioxygen binding from QM/MM metadynamics simulations along  $\xi$  which is the distance from the center of mass of  $O_2$  to Fe(II). (D) Representative conformation of the PHD2 active site at  $\xi = 1.7$  Å, in which dioxygen is bound (state B).

At the beginning of the *ab initio* trajectory, a transition from bidentate 2OG chelation to an intermediate monodentate state was observed, consistent with the ABF calculation results described above. End-on binding of dioxygen to the so created open coordination site on the Fe(II) then occurs (as in Figure 3D) with the resulting potential of mean force shown in Figure 3C. Dioxygen binding occurs at a tilt angle conformation ( $\sim 120^\circ$ , Figure 3D) with the closest oxygen atom 1.7 Å from Fe(III). Dioxygen binding requires overcoming a barrier to cross from the precursor state G to the bound state B of 11 kcal/mol. The energy difference between states G and B is  $\sim -7.35$  kcal/mol. This value is considered to be an estimate of the binding affinity of dioxygen to the Fe metallo-center. This value is similar to those reported for analogous dioxygen

binding to [Fe-Fe] centres in hydrogenases (-4.4 to -7.0 kcal/mol<sup>52</sup> and -7.1 kcal/mol<sup>34</sup>) and hemerythrin (-5.2 kcal/mol)<sup>53</sup> from DFT calculations.

The absolute values obtained for the potential of mean force (Figure 3B, D) are very different for the ABF ( $\Delta G_{GE} = 2.0$  kcal/mol) and QM/MM metadynamics ( $\Delta G_{BG} \approx -7.35$  kcal/mol) methods, as are the maximum barrier heights (ABF,  $\Delta G^\ddagger = 2.2$  kcal/mol; QM/MM metadynamics,  $\Delta G^\ddagger = 11.0$  kcal/mol). The origin of such differences can, at least in part, be ascribed to methodological and pathway differences. The E  $\rightarrow$  G transition is diffusive in origin, and is a low-energy transition within the inner protein matrix, while the G  $\rightarrow$  B transition involves the reaction of dioxygen onto the coordination sphere of Fe(II), which is a significantly more costly energetic process. Similar values have been reported for a G  $\rightarrow$  B type barrier transition for O<sub>2</sub> binding to [Fe-Fe] hydrogenases (12.9 kcal/mol)<sup>34</sup> and CO<sub>2</sub> binding to a [NiFe<sub>3</sub>S<sub>4</sub>] cluster metallo-center (11.9 kcal/mol) in the carbon monoxide dehydrogenase/acetyl-CoA synthase (CODH/ACS) complex using DFT.<sup>54</sup>

Values for the second-order rate constant ( $k_{+1}$ ) for the diffusion of dioxygen from the bulk solvent to the active site and the first-order rate constant for diffusion out of the protein active site ( $k_{-1}$ ) were estimated to be  $1 \times 10^2 \text{ s}^{-1} \text{ m M}^{-1}$ , and  $6 \times 10^6 \text{ s}^{-1}$ , respectively (Table 1). Using a transition state equation, an estimate for  $k_{+2}$  based on the QM/MM potential of mean force for dioxygen binding was calculated (Figure 3.C). A barrier of 11 kcal mol<sup>-1</sup> results in  $k_{+2} = 5.9 \times 10^4 \text{ s}^{-1}$ . Based on these rate constants ( $k_{+1}$ ,  $k_{-1}$  and  $k_{+2}$ ) and using the steady-state approximation, the resultant on-rate constant  $k_{ss}^{in}$  is estimated to be  $0.98 \text{ s}^{-1} \text{ m M}^{-1}$ . This rate constant spans the transitions of the entire dioxygen binding process, *i.e.*, from S to E to G to B (Scheme 3).

**Table 1.** Experimental and calculated phenomenological rate constants.  $k_{+1}$  and  $k_{-1}$  were calculated from classical potential calculations and  $k_{+2}$  was estimated from the QM/MM metadynamics potential of mean force. Combining  $k_{+1}$ ,  $k_{-1}$  and  $k_{+2}$  and using the steady-state approximation, the resultant on-rate constant  $k_{ss}^{in}$  was estimated.

Rate constant	Value
$k_{+1} (10^2 \text{ s}^{-1} \text{ mM}^{-1})$	1
$k_{-1} (10^6 \text{ s}^{-1})$	6
$k_{+2} (10^4 \text{ s}^{-1})$	5.9
$k_{in}^{ss} (\text{s}^{-1} \text{ mM}^{-1})$	0.98
$k_{in} (\text{s}^{-1} \text{ mM}^{-1})$ experimental <sup>55-56</sup>	0.12-1.2

The findings from SMD, ABF and QM/MM metadynamics are summarized in Table 2. Notable methodological differences exist between these approaches. Firstly, ABF and QM/MM metadynamics yields the potential of mean force along the collective variable  $\xi$ , which is at odds with SMD. Secondly, the barrier height estimate from SMD simulations is obtained as a regression parameter from the fit of the MFPT data to the DHS expression; disparities between the ABF potential of mean force barrier and the SMD regression parameter barrier are observed.<sup>39, 57</sup> Table 2 compares the raw output data from these simulations, together with transition state theory interpretations of the data, but is not intended to make direct comparisons of the data from the different simulation methods. The data from two transitions, E to G, and G to B are presented in Table 2. From SMD, for the E to G transition, unbiased transition times were in the range of 27-50 ns, and rapid (0.44-0.7 ns) reverse transition of dioxygen out of the protein (G to E) was observed. For the ABF simulation describing the reversible E to G transition, a rate estimate was calculated with transition state theory by inputting the barrier height free energy into Equation 3. Reconciling distinct values from different methodologies is not simple. TST estimates for the rate of the initial diffusive entry of dioxygen into PHD2 are of the order of  $10^7 \text{ s}^{-1}$  from SMD and  $10^{10} \text{ s}^{-1}$  from ABF trajectories. The B to G transition, sampled using QM/MM metadynamics, was estimated to be  $\sim 10^4 \text{ s}^{-1}$ .

**Table 2.** Comparison of Mean First Path Time (MFPT  $\tau_0$ ), rate constant ( $k$ ), free-energy barrier ( $\Delta G^\ddagger$ ) and free energy difference ( $\Delta G$ ) between dioxygen when residing in states E or G obtained for SMD and ABF. The error for the MFPT estimates is reported as  $\pm \sigma$  for values of harmonic vs. linear polynomial fit.

Method	DHS $v$	Direction	MFPT $\tau_0$ (ns) $\pm \sigma$	$k$ ( $10^6 \text{ s}^{-1}$ )	$\Delta G^\ddagger$ (kcal/mol)	$\Delta G$ (kcal/mol)
ABF	-	$E \rightleftharpoons G$	$769 \pm 0$ (TST)	$13000 \pm 0$ (TST)	$2.2 \pm 0.1$	$2.0 \pm 0.1$
SMD Harmonic	$\frac{1}{2}$	$E \rightarrow G$	$50 \pm 1$	$20 \pm 9$	$7 \pm 6$	2.5
		$G \rightarrow E$	$0.7 \pm 0.1$	$1429 \pm 535$	$2.4 \pm 4.0$	
SMD Linear poly	$\frac{2}{3}$	$E \rightarrow G$	$27 \pm 11$	$37 \pm 9$	$7 \pm 6$	2.4
		$G \rightarrow E$	$0.4 \pm 0.1$	$2500 \pm 535$	$2.9 \pm 4.0$	
QM/MM	-	$G \rightleftharpoons B$	$16666 \pm 0$ (TST)	$0.06 \pm 0.00$ (TST)	$11 \pm 3$	$-7 \pm 3$

Characterization of the diffusion pathways of gases into proteins has attracted considerable attention due to the importance of gas transport in the functions of transport proteins and enzymes (e.g. hydrogenases or oxygenases), and the misfunction of such proteins in disease.<sup>16, 18, 20-22, 24-25, 31, 33, 58</sup> Diffusion pathways for both molecular hydrogen and dioxygen are reported to



occur through hydrophobic protein channels, and their binding to the interior of proteins is sometimes rate-limited at the final transition into active sites, which often contain cofactors such as metal ions.<sup>26-27, 29-30</sup>

There is evidence that in some circumstances, factors other than dioxygen availability can affect PHD activity and the roles of the PHDs in HIF- $\alpha$  degradation, e.g. Fe(II) / 2OG / availability or the presence of regulatory proteins / inhibitors.<sup>5</sup> However, there is a strong body of evidence supporting the assignment of their prime roles, at least in healthy circumstances, as hypoxia sensors, in a manner relating to limitation of their cellular activity by dioxygen availability. Given the proposed central role of the PHDs in human hypoxic sensing, including from a medicinal chemistry perspective (PHD inhibitors have recently been approved for clinical use in treatment of anemia in kidney disease<sup>59</sup>), there is considerable interest in understanding, how dioxygen binds to them. The results of experimental dioxygen consumption assays with PHD2 have estimated the thermodynamic binding constant of dioxygen in the range of 0.065-0.240 mM,<sup>55</sup> corresponding to 4.9 to 5.7 kcal mol<sup>-1</sup>, i.e. they are indicative of strong dioxygen binding. Strikingly, reported kinetic assays reveal dioxygen turnover to be unusually slow for PHD2,<sup>8</sup> with rate estimates ranging between 0.12 and 1.2 s<sup>-1</sup>mM<sup>-1</sup>. Coupled with their conserved ability to efficiently bind, and form a stable complex with, Fe(II) and 2OG,<sup>60</sup> this property is proposed to be important in enabling the hypoxia sensing function of PHD2 and, by implication other PHDs,<sup>55-56</sup> which are likely present in all animals.<sup>61-62</sup>

We employed a combined multiscale approach involving 1.8  $\mu$ s of classical MD simulation trajectories comprising equilibrium, non-equilibrium MD together with QM/MM metadynamics calculations to investigate dioxygen diffusion and binding to the active site in the PHD2.Fe(II).2OG.substrate complex and provide estimates for the rate constants that define the diffusion-reaction model for molecular oxygen binding. The results provide clear theoretical support for a slow turnover-rate for dioxygen by PHD2. Using a multistate formalism, an estimate for the steady-state rate  $k_{in}$  (Scheme 1) was obtained based on probabilistic interpretation of equilibrium MD and biased MD simulations. Unbiased estimates were calculated for  $k_{+1}$  and  $k_{-1}$  using a phenomenological equation that unweighs the effect of dioxygen concentration from the estimates (Schemes 2, 3). The rate constant  $k_{+2}$  for the process was calculated from transition state theory based on a barrier from QM/MM metadynamics

calculations. Comparison of  $k_{+2}$  with  $k_{+1}$  suggests that  $k_{+1}$  is the rate-limiting step in the binding / reaction of dioxygen with PHD2 (Scheme 3). Notably, comparison of  $k_{+1}$  from MD simulation and  $k_{+2}$  from QM/MM metadynamics simulations shows that the rate of chemical interaction between dioxygen and the active site Fe ( $k_{+2} = 5.9 \times 10^4 \text{ s}^{-1}$ ) is about three orders of magnitude faster than the rate of diffusion of dioxygen to the active site ( $k_{+1} = 0.81 \times 10^2 \text{ s}^{-1} \text{ mM}^{-1}$ ). The overall on-rate binding constant ( $k_{\text{in}} = 0.49 \text{ s}^{-1} \text{ mM}^{-1}$ ) is about two orders of magnitude slower than diffusion. Thus, our results imply a PHD2 oxygen sensing mechanism consisting of a rate-limiting diffusive entry followed by a faster ‘chemical binding’ step. By combining these individual rate processes, an expression for the overall rate of dioxygen reaction with the PHD2.Fe.2OG.substrate complex was derived; its value is in agreement with experimental estimates,<sup>8, 56</sup> i.e. supports the proposal of the unusually slow reaction of the PHDs with dioxygen.

## Conclusions

The overall results imply that the reversible binding of dioxygen is central to the hypoxia sensing capacity of PHD2. HIF- $\alpha$  substrate prolyl hydroxylation, which signals for subsequent HIF- $\alpha$  degradation, is thus, at last principally, a manifestation of the equilibrium between dioxygen in solution and that bound to the PHD2.Fe.2OG.substrate complex. The importance of reversible binding of the sensed dioxygen molecule to the PHDs suggests that heme<sup>63</sup> and non heme based dioxygen / hypoxia sensing systems are perhaps more similar than might sometimes have been perceived.

The results identify a slow rate of dioxygen entry involves diffusion through a single wide channel lined by hydrophobic residues. The dioxygen access channel is formed by elements from both the enzyme and the HIF- $\alpha$  substrate (Figure 2). Given that there are three human PHDs and three HIF- $\alpha$  isoforms, the results thus raise the possibility that different PHD HIF- $\alpha$  substrate combinations might have different dioxygen sensitivity profiles; this is something that could be explored computationally in the future. In this regard, it is notable that multiple non HIF- $\alpha$  substrates for the PHDs have been reported.<sup>64</sup> Although a recent report has not provided support for any of these acting as substrates for isolated PHDs,<sup>64</sup> it is possible that some of the proteins reported to interact with the PHDs modulate PHD sensitivity to dioxygen by altering dioxygen

transport to the active sites of the PHDs. It is also of interest to explore details of the binding of dioxygen to factor inhibiting HIF (FIH).<sup>65</sup> FIH is a HIF- $\alpha$  asparaginyl hydroxylase, the activity of which reduces the interaction between HIF and transcriptionally enhancing histone acetyl transferases.<sup>65</sup> Notably, FIH is less sensitive than the PHDs to limiting dioxygen availability, as observed both in cells and in isolated form.<sup>66-68</sup> FIH also accepts multiple non HIF- $\alpha$  substrates, mostly ankyrin repeat domain proteins, as shown by work from several groups.<sup>69-70</sup> Interestingly, crystallographic studies reveal that, unlike PHD2, 2OG binds to FIH in a manner such that a vacant coordination site is adjacent to the substrate and that the FIH active site is apparently more open than that of the PHDs,<sup>71</sup> properties that are consistent with the apparently more efficient binding of dioxygen to FIH compared to the PHDs.<sup>72</sup>

The results set the stage for a dissection of the full PHD2 reaction that should provide new insight into the mechanism of dioxygen activation by the 2OG-dependent dioxygenases. In this regard, the change in bidentate to monodentate coordination by 2OG that enables dioxygen reaction with the Fe is particularly intriguing. Although further work is required to define the full mechanism of PHD2, this observation suggests that dioxygen can bind without displacement of the apparently tightly bound water from the Fe. Alternative mechanisms to enable dioxygen binding in the coordination site adjacent to the substrate are possible, e.g. movement of the 2OG C-1 carboxylate to the position occupied by the ligating water in the PHD2.Fe(II).2OG.substrate complex, or rearrangement of an intermediate (ferryl flip<sup>50</sup>); it is also near certain that there is mechanistic variation between different 2OG oxygenases. There is also precedent for dioxygen binding to a water ligated Fe(II) in catalysis by isopenicillin N synthase, which does not employ 2OG but which is closely related to the 2OG oxygenases from a structural / mechanistic perspective.<sup>73-74</sup> The maintenance of an Fe-ligated water during dioxygen reaction provides a route for exchange of oxygen from dioxygen with that from water during hydroxylation, as evidenced by less than stoichiometric incorporation of <sup>18</sup>O label from <sup>18</sup>O<sub>2</sub> in some cases. This occurs with some, but not all, 2OG oxygenases; by contrast there is consistently high level of incorporation of one <sup>18</sup>O label from <sup>18</sup>O<sub>2</sub> not succinate.<sup>75</sup>

Fe(II) and 2OG-dependent oxygenases and structurally / mechanistically related enzymes form a large superfamily that catalyzes a very range of reactions, including hydroxylations, halogenations, ring closures, desaturations, epimerizations, ring expansions, and epoxidations.<sup>76</sup>

In humans, our 60-70 2OG oxygenases have roles in collagen biosynthesis, epidermal growth factor domain like protein hydroxylation, fatty acid metabolism, DNA and RNA repair / modification, demethylation related to epigenetic regulation, modification / regulation of the translation machinery, and hypoxia sensing.<sup>51</sup> Strikingly, despite the variations in substrate and subfamily type of 2OG oxygenase employed in biology, the rate of reaction of the PHD2:Fe(II):2OG:HIF- $\alpha$  complex with dioxygen is much slower than for most, if not all, other human 2OG oxygenases that have been studied. However, in some cases, e.g. microbial antibiotic biosynthesis, 2OG oxygenases operate in a very efficient manner, at least within cells, if not always in isolated form. Understanding the molecular factors governing dioxygen transport may be useful in understanding these differences. Our identification of a clearly defined path for dioxygen binding to the PHD2 active site, raises the question as to whether this is always the case. Linked with this is the question of whether consequently arising distinct kinetic properties with respect to the sensed element are a hallmark of enzyme sensors. It will be important to investigate this question within the context of comparing the properties of very efficient family members, e.g. the oxygenases of penicillin and cephalosporin biosynthesis, with (potential) sensing enzymes, including other 2OG dependent protein/ nucleic acid oxygenases.

Studies on dioxygen binding to PHD homologues in early animals and microbes are also of interest, since these do not have as far as is known HIF- $\alpha$  homologue substrates. In the case of the social amoeba, *Dictyostelium discoideum*, and the protozoan parasite, *Toxoplasma gondii*, PHD homologues are proposed to act as a hypoxia sensors via dioxygen limited prolyl hydroxylation of the S-Phase Kinase Associated Protein 1 (Skp) protein so regulating transcription.<sup>77</sup> In *Pseudomonas spp.*, a PHD homologue (PPHD) catalyzes the hydroxylation of elongation factor thermally unstable (EF-Tu), a key element of the translational machinery in bacteria. Interestingly, unlike the HIF- $\alpha$  ODDs which are largely unstructured in solution, EF-Tu is a well-ordered protein, that undergoes a very substantial conformational change during induced fit binding to PPHD to give a stable complex.<sup>78</sup> Investigating the details of dioxygen binding to sensing and non-sensing oxygenases in organisms from environments with different oxygen availabilities, may provide general molecular insight into the evolution of aerobic metabolic systems.

From a practical perspective, the results on dioxygen binding to oxygenases are of interest with respect to optimizing the use of 2OG oxygenases as biocatalysts and in identifying more subtle and selective inhibitors of them than are presently available. PHD inhibitors for the treatment of anemia in chronic kidney disease have recently been approved for clinical use.<sup>59</sup> However, these compounds block activity by complexing with active site Fe(II) and competing with 2OG.<sup>59</sup> The results presented here suggest that identifying compounds that modulate dioxygen by binding to one or more the PHD.Fe.2OG.substrate complexes should be possible. Such inhibitors should be safer because they will not completely block PHD activity, so reducing the chance of overdose. There is also the possibility of PHD / HIF- $\alpha$  isoform substrate selective inhibition by this approach via compound binding to a specific PHD / HIF- $\alpha$  combinations.

## ASSOCIATED CONTENT

Details of force field parameterization for 2OG, computational methods and protocols. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

*Corresponding Authors:* Carmen Domene [C.Dmene@bath.ac.uk](mailto:C.Dmene@bath.ac.uk) & Christopher J. Schofield [christopher.schofield@chem.ox.ac.uk](mailto:christopher.schofield@chem.ox.ac.uk)

## ACKNOWLEDGMENTS

CD acknowledges the use of ARCHER at the UK National Supercomputing Service (<http://www.archer.ac.uk>) facility through the PRACE initiative, EPSRC RAP calls and the SimBiosim consortium. C.J. thanks King's College London for the award of a Graduate Teaching Assistant studentship. C.J.S. thanks the Wellcome Trust, Cancer Research UK, and the BBSRC for funding.

## REFERENCES

1. Kaelin, W. G.; Ratcliffe, P. J., Oxygen sensing by metazoans: The central role of the hif hydroxylase pathway. *Molecular cell* **2008**, 30 (4), 393-402.
2. Mole, D. R.; Ratcliffe, P. J., Cellular oxygen sensing in health and disease. *Pediatr Nephrol* **2008**, 23, 681-694.
3. Bruick, R. K.; McKnight, S. L., A conserved family of prolyl-4-hydroxylases that modify hif. *Science* **2001**, 294 (5545), 1337-1340.
4. Epstein, A. C.; Gleadle, J. M.; McNeill, L. A.; Hewitson, K. S.; O'Rourke, J.; Mole, D. R.; Mukherji, M.; Metzen, E.; Wilson, M. I.; Dhanda, A.; Tian, Y. M.; Masson, N.; Hamilton, D. L.; Jaakkola, P.; Barstead, R.; Hodgkin, J.; Maxwell, P. H.; Pugh, C. W.; Schofield, C. J.; Ratcliffe, P. J., C. Elegans egl-9 and mammalian homologs define a family of dioxygenases that regulate hif by prolyl hydroxylation. *Cell* **2001**, 107 (1), 43-54.

5. Schofield, C. J.; Ratcliffe, P. J., Oxygen sensing by hif hydroxylases. *Nature Rev Mol Cell Biol* **2004**, *5* (5), 343-354.
6. Flagg, S. C.; Giri, N.; Pektas, S.; Maroney, M. J.; Knapp, M. J., Inverse solvent isotope effects demonstrate slow aquo release from hypoxia inducible factor-prolyl hydroxylase (phd2). *Biochemistry* **2012**, *51* (33), 6654-6666.
7. Hirsilä, M.; Koivunen, P.; Günzler, V.; Kivirikko, K. I.; Myllyharju, J., Characterization of the human prolyl 4-hydroxylases that modify the hypoxia-inducible factor. *J Biol Chem* **2003**, *278* (33), 30772-30780.
8. Flashman, E.; Hoffart, L. M.; Hamed, R. B.; Bollinger Jr, J. M.; Krebs, C.; Schofield, C. J., Evidence for the slow reaction of hypoxia-inducible factor prolyl hydroxylase 2 with oxygen: Prolyl hydroxylase 2 reaction with oxygen. *FEBS Journal* **2010**, *277* (19), 4089-4099.
9. McDonough, M. A.; Li, V.; Flashman, E.; Chowdhury, R.; Mohr, C.; Liénard, B. M. R.; Zondlo, J.; Oldham, N. J.; Clifton, I. J.; Lewis, J.; McNeill, L. A.; Kurzeja, R. J. M.; Hewitson, K. S.; Yang, E.; Jordan, S.; Syed, R. S.; Schofield, C. J., Cellular oxygen sensing: Crystal structure of hypoxia-inducible factor prolyl hydroxylase (phd2). *PNAS* **2006**, *103* (26), 9814-9819.
10. Chowdhury, R.; McDonough, M. A.; Mecnović, J.; Loenarz, C.; Flashman, E.; Hewitson, K. S.; Domene, C.; Schofield, C. J., Structural basis for binding of hypoxia-inducible factor to the oxygen-sensing prolyl hydroxylases. *Structure* **2009**, *17* (7), 981-989.
11. Chowdhury, R.; Leung, I. K. H.; Tian, Y.-M.; Abboud, M. I.; Ge, W.; Domene, C.; Cantrelle, F.-X.; Landrieu, I.; Hardy, A. P.; Pugh, C. W.; Ratcliffe, P. J.; Claridge, T. D. W.; Schofield, C. J., Structural basis for oxygen degradation domain selectivity of the hif prolyl hydroxylases. *Nat. Commun.* **2016**, *7* (1), 12673.
12. Flashman, E.; Bagg, E. A. L.; Chowdhury, R.; Mecnovic, J.; Loenarz, C.; McDonough, M. A.; Hewitson, K. S.; Schofield, C. J., Kinetic rationale for selectivity toward n- and c-terminal oxygen-dependent degradation domain substrates mediated by a loop region of hypoxia-inducible factor prolyl hydroxylases. *J Biol Chem* **2008**, *283* (7), 3808-3815.
13. Aik, W.; McDonough, M. A.; Thalhammer, A.; Chowdhury, R.; Schofield, C. J., Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases. *Cur Opin Struct Biol* **2012**, *22* (6), 691-700.
14. McDonough, M. A.; Loenarz, C.; Chowdhury, R.; Clifton, I. J.; Schofield, C. J., Structural studies on human 2-oxoglutarate dependent oxygenases. *Cur Opin Struct Biol* **2010**, *20* (6), 659-672.
15. Hausinger, R. P., Fe(ii)/ $\alpha$ -ketoglutarate-dependent hydroxylases and related enzymes. *Crit Rev Biochem Mol* **2004**, *39* (1), 21-68.
16. Shadrina, M. S.; English, A. M.; Peslherbe, G. H., Effective simulations of gas diffusion through kinetically accessible tunnels in multisubunit proteins: O<sub>2</sub> pathways and escape routes in t-state deoxyhemoglobin. *J Am Chem Soc* **2012**, *134* (27), 11177-11184.
17. Shadrina, M. S.; English, A. M.; Peslherbe, G. H., Benchmarking rapid tles simulations of gas diffusion in proteins: Mapping o<sub>2</sub> migration and escape in myoglobin as a case study. *J Chem Theory Comput* **2016**, *12* (4), 2038-2046.
18. Lautier, T.; Ezanno, P.; Baffert, C.; Fourmond, V.; Cournac, L.; Fontecilla-Camps, J. C.; Soucaille, P.; Bertrand, P.; Meynial-Salles, I.; Léger, C., The quest for a functional substrate access tunnel in fefe hydrogenase. *Faraday Discuss* **2011**, *148*.
19. Kubas, A.; Orain, C.; De Sancho, D.; Saujet, L.; Sensi, M.; Gauquelin, C.; Meynial-Salles, I.; Soucaille, P.; Bottin, H.; Baffert, C., Mechanism of o<sub>2</sub> diffusion and reduction in fefe hydrogenases. *Nature Chem* **2016**.
20. Cohen, J.; Schulten, K., O<sub>2</sub> migration pathways are not conserved across proteins of a similar fold. *Biophys J* **2007**, *93* (10), 3591-3600.
21. Cohen, J.; Kim, K.; Posewitz, M.; Ghirardi, M. L.; Schulten, K.; Seibert, M.; King, P., Molecular dynamics and experimental investigation of h<sub>2</sub> and o<sub>2</sub> diffusion in [fe]-hydrogenase. *Biochem Soc Trans* **2005**, *33* (1), 80-2.
22. Baron, R.; Riley, C.; Chenprakhon, P.; Thotsaporn, K.; Winter, R. T.; Alfieri, A.; Forneris, F.; van Berkel, W. J. H.; Chaiyen, P.; Fraaije, M. W.; Mattevi, A.; McCammon, J. A., Multiple pathways guide oxygen diffusion into flavoenzyme active sites. *PNAS* **2009**, *106* (26), 10603-10608.
23. Cohen, J.; Kim, K.; King, P.; Seibert, M.; Schulten, K., Finding gas diffusion pathways in proteins: Application to O<sub>2</sub> and H<sub>2</sub> transport in cpi [fefe]-hydrogenase and the role of packing defects. *Structure* **2005**, *13* (9), 1321-1329.

24. Orlowski, S.; Nowak, W., Locally enhanced sampling molecular dynamics study of the dioxygen transport in human cytoglobin. *J Mol Model* **2007**, *13* (6-7), 715-723.
25. Topin, J.; Diharce, J.; Fiorucci, S.; Antonczak, S.; Golebiowski, J., O<sub>2</sub> migration rates in [nife] hydrogenases. A joint approach combining free-energy calculations and kinetic modeling. *J Phys Chem B* **2014**, *118* (3), 676-681.
26. Wang, P.-h.; Best, R. B.; Blumberger, J., Multiscale simulation reveals multiple pathways for h<sub>2</sub> and o<sub>2</sub> transport in a [nife]-hydrogenase. *JACS* **2011**, *133* (10), 3548-3556.
27. Wang, P.-h.; Bruschi, M.; De Gioia, L.; Blumberger, J., Uncovering a dynamically formed substrate access tunnel in carbon monoxide dehydrogenase/acetyl-coa synthase. *JACS* **2013**, *135* (25), 9493-9502.
28. Teixeira, V. H.; Baptista, A. M.; Soares, C. M., Pathways of h<sub>2</sub> toward the active site of [nife]-hydrogenase. *Biophys J* **2006**, *91* (6), 2035-2045.
29. Wang, P.-h.; Best, R. B.; Blumberger, J., A microscopic model for gas diffusion dynamics in a [nife]-hydrogenase. *Phys Chem Chem Phys* **2011**, *13* (17).
30. Wang, P. h.; Blumberger, J., Mechanistic insight into the blocking of co diffusion in [nife]-hydrogenase mutants through multiscale simulation. *PNAS* **2012**, *109* (17), 6399-6404.
31. Cohen, J.; Arkhipov, A.; Braun, R.; Schulten, K., Imaging the migration pathways for o<sub>2</sub>, co, no, and xe inside myoglobin. *Biophys J* **2006**, *91* (5), 1844-1857.
32. Mahinthichaichan, P.; Gennis, R. B.; Tajkhorshid, E., All the o<sub>2</sub> consumed by thermus thermophilus cytochrome ba<sub>3</sub> is delivered to the active site through a long, open hydrophobic tunnel with entrances within the lipid bilayer. *Biochemistry* **2016**, *55* (8), 1265-1278.
33. Liebgott, P.-P.; Leroux, F.; Burlat, B.; Dementin, S.; Baffert, C.; Lautier, T.; Fourmond, V.; Ceccaldi, P.; Cavazza, C.; Meynial-Salles, I.; Soucaille, P.; Fontecilla-Camps, J. C.; Guigliarelli, B.; Bertrand, P.; Rousset, M.; Léger, C., Relating diffusion along the substrate tunnel and oxygen sensitivity in hydrogenase. *Nat Chem Biol* **2010**, *6* (1), 63-70.
34. Kubas, A.; De Sancho, D.; Best, R. B.; Blumberger, J., Aerobic damage to [fefe]-hydrogenases: Activation barriers for the chemical attachment of o<sub>2</sub>. *Angewandte Chemie* **2014**, *53* (16), 4081-4.
35. Chowdhury, R.; Hardy, A.; Schofield, C. J., The human oxygen sensing machinery and its manipulation. *Chem Soc Rev* **2008**, *37* (7).
36. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K., Scalable molecular dynamics with namd. *J Comput Chem* **2005**, *26* (16), 1781-1802.
37. Brooks, B.; Brooks, C.; MacKerell, A.; Nilsson, L.; Petrella, R.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S., Charmm: The biomolecular simulation program. *J Comput Chem* **2009**, *30* (10), 1545-1614.
38. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L., Comparison of simple potential functions for simulating liquid water. *J Chem Phys* **1983**, *79* (2), 926-935.
39. Dudko, O.; Hummer, G.; Szabo, A., Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys Rev Lett* **2006**, *96* (10).
40. Darve, E.; Pohorille, A., Calculating free energies using average force. *J Chem Phys* **2001**, *115* (20).
41. Darve, E.; Rodríguez-Gómez, D.; Pohorille, A., Adaptive biasing force method for scalar and vector free energy calculations. *J Chem Phys* **2008**, *128*.
42. Rodríguez-Gomez, D.; Darve, E.; Pohorille, A., Assessing the efficiency of free energy calculation methods. *J Chem Phys* **2004**, *120* (8).
43. Buchete, N.-V.; Hummer, G., Coarse master equations for peptide folding dynamics. *J Phys Chem B* **2008**, *112* (19), 6057-6069.
44. Laio, A.; Parrinello, M., Escaping free-energy minima. *PNAS* **2002**, *99* (20), 12562-12566.
45. VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J., Quickstep: Fast and accurate density functional calculations using a mixed gaussian and plane waves approach. *Comput Phys Commun* **2005**, *167* (2), 103-128.
46. Krack, M.; Parrinello, M., Quickstep: Make the atoms dance. *High performance computing in chemistry* **2004**, *25*, 29.
47. Humphrey, W.; Dalke, A.; Schulten, K., Vmd: Visual molecular dynamics. *J Mol Graphics* **1996**, *14* (1), 33-38.

48. Ehrismann, D.; Flashman, E.; Genn, D. N.; Mathioudakis, N.; Hewitson, K. S.; Ratcliffe, P. J.; Schofield, C. J., Studies on the activity of the hypoxia-inducible-factor hydroxylases using an oxygen consumption assay. *Biochem J* **2007**, *401* (1), 227-34.
49. Flashman, E.; Hoffart, L. M.; Hamed, R. B.; Bollinger, J. M., Jr.; Krebs, C.; Schofield, C. J., Evidence for the slow reaction of hypoxia-inducible factor prolyl hydroxylase 2 with oxygen. *FEBS* **2010**, *277* (19), 4089-99.
50. Zhang, Z.; Ren, J.-s.; Harlos, K.; McKinnon, C. H.; Clifton, I. J.; Schofield, C. J., Crystal structure of a clavamate synthase- $\text{Fe(II)}$ -2-oxoglutarate-substrate- $\text{no}$  complex: Evidence for metal centred rearrangements. *FEBS Lett* **2002**, *517* (1-3), 7-12.
51. Martinez, S.; Hausinger, R. P., Catalytic mechanisms of  $\text{Fe(II)}$ - and 2-oxoglutarate-dependent oxygenases. *J Biol Chem* **2015**, *290* (34), 20702-20711.
52. Hong, G.; Pachter, R., Inhibition of biocatalysis in  $[\text{Fe-Fe}]$  hydrogenase by oxygen: Molecular dynamics and density functional theory calculations. *ACS Chem Biol* **2012**, *7* (7), 1268-1275.
53. Wirstam, M.; Lippard, S. J.; Friesner, R. A., Reversible dioxygen binding to hemerythrin. *J Am Chem Soc* **2003**, *125* (13), 3980-3987.
54. Wang, P. H.; Bruschi, M.; De Gioia, L.; Blumberger, J., Uncovering a dynamically formed substrate access tunnel in carbon monoxide dehydrogenase/acetyl-coa synthase. *J Am Chem Soc* **2013**, *135* (25), 9493-502.
55. Ehrismann, D.; Flashman, E.; Genn, David N.; Mathioudakis, N.; Hewitson, Kirsty S.; Ratcliffe, Peter J.; Schofield, Christopher J., Studies on the activity of the hypoxia-inducible-factor hydroxylases using an oxygen consumption assay. *Biochem J* **2007**, *401* (1), 227.
56. Tarhonskaya, H.; Hardy, A. P.; Howe, E. A.; Loik, N. D.; Kramer, H. B.; McCullagh, J. S.; Schofield, C. J.; Flashman, E., Kinetic investigations of the role of factor inhibiting hypoxia-inducible factor (fih) as an oxygen sensor. *J Biol Chem* **2015**, *290* (32), 19726-19742.
57. Dudko, O. K.; Hummer, G.; Szabo, A., Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *PNAS* **2008**, *105* (41), 15755-15760.
58. van Lun, M.; Hub, J. S.; van der Spoel, D.; Andersson, I.,  $\text{CO}_2$  and  $\text{O}_2$  distribution in rubisco suggests the small subunit functions as a  $\text{CO}_2$  reservoir. *J Am Chem Soc* **2014**, *136* (8), 3165-3171.
59. Yeh, T.-L.; Leissing, Thomas M.; Abboud, M. I.; Thinnies, C. C.; Atasoylu, O.; Holt-Martyn, J. P.; Zhang, D.; Tumber, A.; Lippl, K.; Lohans, C. T.; Leung, I. K. H.; Morcrette, H.; Clifton, I. J.; Claridge, T. D. W.; Kawamura, A.; Flashman, E.; Lu, X.; Ratcliffe, P. J.; Chowdhury, R.; Pugh, C. W.; Schofield, C. J., Molecular and cellular mechanisms of hif prolyl hydroxylase inhibitors in clinical trials. *Chem Sci* **2017**, *8* (11), 7651-7668.
60. McNeill, L. A.; Flashman, E.; Buck, M. R. G.; Hewitson, K. S.; Clifton, I. J.; Jeschke, G.; Claridge, T. D. W.; Ehrismann, D.; Oldham, N. J.; Schofield, C. J., Hypoxia-inducible factor prolyl hydroxylase 2 has a high affinity for ferrous iron and 2-oxoglutarate. *Mol Biosyst* **2005**, *1* (4), 321-324.
61. Loenarz, C.; Coleman, M. L.; Boleininger, A.; Schierwater, B.; Holland, P. W. H.; Ratcliffe, P. J.; Schofield, C. J., The hypoxia-inducible transcription factor pathway regulates oxygen sensing in the simplest animal, trichoplax adhaerens. *Embo Rep* **2011**, *12* (1), 63-70.
62. K., L.; A., B.; M.A., M.; M.I., A.; H., T.; R., C.; C., L.; C.J., S., Born to sense: Biophysical analyses of the oxygen sensing prolyl hydroxylase from the simplest animal trichoplax adhaerens. *Hypoxia* **2018**, *6*, 57-71.
63. José López-Barneo; Ricardo Pardal, a.; Ortega-Sáenz, P., Cellular mechanism of oxygen sensing. *Ann Rev Physiol* **2001**, *63* (1), 259-287.
64. Cockman, M. E.; Lippl, K.; Tian, Y.-M.; Pegg, H. B.; Figg, W. D. J.; Abboud, M. I.; Heilig, R.; Fischer, R.; Myllyharju, J.; Schofield, C. J.; Ratcliffe, P. J., Lack of activity of recombinant hif prolyl hydroxylases (phds) on reported non-hif substrates. *eLife* **2019**, *8*, e46490.
65. Hewitson, K. S.; McNeill, L. A.; Riordan, M. V.; Tian, Y.-M.; Bullock, A. N.; Welford, R. W.; Elkins, J. M.; Oldham, N. J.; Bhattacharya, S.; Gleadle, J. M.; Ratcliffe, P. J.; Pugh, C. W.; Schofield, C. J., Hypoxia-inducible factor (hif) asparagine hydroxylase is identical to factor inhibiting hif (fih) and is related to the cupin structural family. *J Biol Chem* **2002**, *277* (29), 26351-26355.
66. Koivunen, P.; Hirsilä, M.; Günzler, V.; Kivirikko, K. I.; Myllyharju, J., Catalytic properties of the asparaginyl hydroxylase (fih) in the oxygen sensing pathway are distinct from those of its prolyl 4-hydroxylases. *J Biol Chem* **2004**, *279* (11), 9899-9904.



67. Tian, Y.-M.; Yeoh, K. K.; Lee, M. K.; Eriksson, T.; Kessler, B. M.; Kramer, H. B.; Edelman, M. J.; Willam, C.; Pugh, C. W.; Schofield, C. J.; Ratcliffe, P. J., Differential sensitivity of hypoxia inducible factor hydroxylation sites to hypoxia and hydroxylase inhibitors. *J Biol Chem* **2011**, *286* (15), 13041-13051.
68. Hangasky, J. A.; Saban, E.; Knapp, M. J., Inverse solvent isotope effects arising from substrate triggering in the factor inhibiting hypoxia inducible factor. *Biochemistry* **2013**, *52* (9), 1594-1602.
69. Hardy, A. P.; Prokes, I.; Kelly, L.; Campbell, I. D.; Schofield, C. J., Asparaginyl  $\beta$ -hydroxylation of proteins containing ankyrin repeat domains influences their stability and function. *J Mol Biol* **2009**, *392* (4), 994-1006.
70. Cockman, M. E.; Lancaster, D. E.; Stolze, I. P.; Hewitson, K. S.; McDonough, M. A.; Coleman, M. L.; Coles, C. H.; Yu, X.; Hay, R. T.; Ley, S. C.; Pugh, C. W.; Oldham, N. J.; Masson, N.; Schofield, C. J.; Ratcliffe, P. J., Posttranslational hydroxylation of ankyrin repeats in ikb proteins by the hypoxia-inducible factor (hif) asparaginyl hydroxylase, factor inhibiting hif (fih). *PNAS* **2006**, *103* (40), 14767-14772.
71. Elkins, J. M.; Hewitson, K. S.; McNeill, L. A.; Seibel, J. F.; Schlemminger, I.; Pugh, C. W.; Ratcliffe, P. J.; Schofield, C. J., Structure of factor-inhibiting hypoxia-inducible factor (hif) reveals mechanism of oxidative modification of hif-1 $\alpha$ . *J Biol Chem* **2003**, *278* (3), 1802-1806.
72. Tarhonskaya, H.; Hardy, A. P.; Howe, E. A.; Loik, N. D.; Kramer, H. B.; McCullagh, J. S. O.; Schofield, C. J.; Flashman, E., Kinetic investigations of the role of factor inhibiting hypoxia-inducible factor (fih) as an oxygen sensor. *J Biol Chem* **2015**, *290* (32), 19726-19742.
73. Schofield, C. J.; Baldwin, J. E.; Byford, M. F.; Clifton, I.; Hajdu, J.; Hensgens, C.; Roach, P., Proteins of the penicillin biosynthesis pathway. *Curr Opin Struc Bio* **1997**, *7* (6), 857-864.
74. Roach, P. L.; Clifton, I. J.; Hensgens, C. M. H.; Shibata, N.; Schofield, C. J.; Hajdu, J.; Baldwin, J. E., Structure of isopenicillin synthase complexed with substrate and the mechanism of penicillin formation. *Nature* **1997**, *387* (6635), 827-830.
75. Welford, R. W. D.; Kirkpatrick, J. M.; McNeill, L. A.; Puri, M.; Oldham, N. J.; Schofield, C. J., Incorporation of oxygen into the succinate co-product of iron(ii) and 2-oxoglutarate dependent oxygenases from bacteria, plants and humans. *FEBS Lett* **2005**, *579* (23), 5170-5174.
76. Loenarz, C.; Schofield, C. J., Physiological and biochemical aspects of hydroxylations and demethylations catalyzed by human 2-oxoglutarate oxygenases. *Trends Biocheml Sci* **2011**, *36* (1), 7-18.
77. Xu, Y.; Brown, K. M.; Wang, Z. A.; van der Wel, H.; Teygong, C.; Zhang, D.; Blader, I. J.; West, C. M., The skp1 protein from toxoplasma is modified by a cytoplasmic prolyl 4-hydroxylase associated with oxygen sensing in the social amoeba dictyostelium. *J Biol Chem* **2012**, *287* (30), 25098-25110.
78. Scotti, J. S.; Leung, I. K. H.; Ge, W.; Bentley, M. A.; Paps, J.; Kramer, H. B.; Lee, J.; Aik, W.; Choi, H.; Paulsen, S. M.; Bowman, L. A. H.; Loik, N. D.; Horita, S.; Ho, C.-h.; Kershaw, N. J.; Tang, C. M.; Claridge, T. D. W.; Preston, G. M.; McDonough, M. A.; Schofield, C. J., Human oxygen sensing may have origins in prokaryotic elongation factor tu prolyl-hydroxylation. *PNAS* **2014**, *111* (37), 13331-13336.

For Table of Contents Only

