Beyond active site residues: overall structural dynamics control catalysis in flavin-containing and heme-containing monooxygenases
Maximilian JLJ Furst, Filippo Fiorentini and Marco W Fraaije

Monooxygenases (MOs) face the challenging reaction of an organic target, oxygen and a cofactor – most commonly heme or flavin. To correctly choreograph the substrates spatially and temporally, MOs evolved a variety of strategies, which involve structural flexibility. Besides classical domain and loop movements, flavin-containing MOs feature conformational changes of their flavin prosthetic group and their nicotinamide cofactor. With similar mechanisms emerging in various subclasses, their generality and involvement in selectivity are intriguing questions. Cytochrome P450 MOs are often inherently plastic and large movements of individual segments throughout the entire structure occur. As these complicated and often unpredictable movements are largely responsible for substrate uptake, engineering strategies for these enzymes were mostly successful when randomly mutating residues across the entire structure.

Introduction
Aerobic life evolved to use O₂ as an electron acceptor in the respiratory chain and as a co-substrate to oxygenate organic compounds using enzymes such as monooxygenases (MOs). As the spin-forbidden reaction of triplet ground state O₂ with singlet organic compounds is very slow, enzymes lower the energy barrier by reductively activating oxygen. Unless the organic substrate provides the reducing power, this reaction requires a cofactor. Open-shell transition metals such as copper or iron can be deployed, and the latter is often complexed by a porphyrin scaffold — the heme cofactor. Alternatively, MOs use a purely organic flavin mononucleotide (FMN) or flavin adenine dinucleotide (FAD) cofactor. In the several hundred available MO structures, the two most frequently co-crystallized ligands are heme (43%) and FAD (14%), which are used by cytochrome P450 MOs (CYPs or P450s) and flavoprotein MOs, respectively. The traditional center of attention was the active site of the MOs, which provides the structural context for facilitating catalysis — electron transfer, O₂ activation, and oxygenation. However, if any static structure is insufficient in describing an enzyme’s mode of action, this is especially true with MOs due to their extremely dynamic nature (Figure 1). For a full understanding of the reaction of MOs, we need to look beyond the supposed catalytic center.

Flavoprotein monooxygenases
The isoalloxazine ring enables flavins to stabilize and shuttle between redox states. In flavoprotein MOs, oxygen is activated by the transfer of one electron from fully reduced flavin to O₂, followed by the coupling of the caged radical pair at the flavin’s C4a locus [1]. Characteristically, flavoprotein MOs stabilize the resulting catalytic (hydro)peroxylavin [2]. The electrons originate from a reduced nicotinamide cofactor – NAD(P)H – which can bind either transiently or permanently. The former is the case for the aromatic hydroxylases of class A flavoprotein MOs, where the nicotinamide cofactor dissociates immediately after reducing a mobile flavin [3,4] (Figure 1). These enzymes are quite narrow in substrate scope and ‘cautious’: before NAD(P)H is consumed, a potential substrate needs to be ‘proofread’ [5]. In contrast, NAD(P)H is consumed substrate-independently and bound in various conformations throughout the catalytic cycle in ‘bold’ class B flavoprotein MOs (Figure 1). These comprise N-hydroxylating MOs (NMOs), which are highly substrate-specific, heteroatom-oxygenating flavin-containing MOs (FMOs), and ketone to ester-transforming Baeyer-Villiger MOs (BVMOs), which often show relaxed substrate scopes.

Mobile flavins
For the prototype class A flavoprotein MO, p-hydroxybenzoate hydroxylase, a delicate dynamic interplay between the coenzyme NADPH and the prosthetic FAD cofactor, has been unraveled [6]. For reduction, the flavin of class A MOs swings toward NADPH into an ‘out’ position using the ribityl carbons as pivot points (Figure 2a). Next, NADP⁺ is released, FAD returns to the
‘in’ position [7], and the formed C4a-hydroperoxyflavin hydroxylates the substrate through electrophilic aromatic substitution. While this mechanism was elucidated decades ago [3,4], its clinical relevance was established recently, when a bacterial tetracycline MO that confers antibiotic resistance was shown to be efficiently inhibited by a substrate analogue, which locks FAD in the ‘out’ position [8**]. Furthermore, novel variations on the mobile flavin mechanism were discovered in two paralogous class A MOs converting the same multicyclic substrate to divergent products in a bifurcating metabolic pathway [9]. While one, RebC, substitutes a carboxyl group with a carbonyl group, the second, StaC, only decarboxylates. Apparently, RebC uses flavin mobility
for reduction before hydroxylation the substrate’s enol tautomer, while StaC’s mobile flavin accelerates the spontaneous decarboxylation of the keto tautomer via a steric and/or electrostatic clash. The same group also discovered that mobile flavins occur in N-hydroxylating MOs of class B [10].

An early indication for a conformational change in NMOs was the proposed allosteric regulation [11] of t-ornithine MO (SidA) by t-arginine [12]. However, the regulation is likely rather a competitive inhibition, as structures later revealed t-arginine to bind at the same position in SidA [13] as t-ornithine in a homologous NMO (PvdA) [14]. Eventually, structures of another homolog (KtzI) showed FAD to undergo conformational changes [10]. While the swing of the flavin in class A MOs occurs nearly in the plane of the isoalloxazine ring, KtzI’s flavin pivots largely at the ribityl C1 and rotates out of the plane (Figure 2b). As this trajectory clashes with the nicotinamide riboside, it might represent an NADP+ ejection mechanism. In the resting state, the oxidized flavin is probably in an equilibrium between ‘in’ and ‘out’. No hydride transfer orientation was observed, but reduced flavin was always ‘in’ and the hydroperoxyflavin likely retains this position. A distorted nicotinamide in crystals of PvdA trapped with the product suggested an initial destabilization of NADP+ [14], which then would be ejected by the moving flavin.

Mobile nicotinamide cofactors
As they bind NADP stably [2], class B MOs are often crystallized in complex with both cofactors. Several orientations of NADP can be observed in available structures. With varying degrees of confidence, these have been attributed to the dual role of the cofactor over the course of the catalytic cycle: reduction of the flavin and stabilization of the (hydro)peroxyflavin [2]. As the two roles require different orientations and no structure appropriate for hydride transfer is known, a ‘sliding mechanism’ has been proposed [15] (Figure 3a). Accordingly, NADPH reduces the flavin while sliding over the isoalloxazine into its fixed and commonly observed ‘stabilization’ position. Various structures appear to show the positions sampled on the way: stacked above the flavin in steroid MO (STMO, PDB IDs 4AOS), and an intermediate position in one crystal form of cyclohexalone MO (CHMO, PDB ID 3GWF). Problematically, however, the model conflicts with experiments showing that NADPH’s pro-R hydride reduces the flavin, which is incompatible with the anti conformation of the flavin-stacked NADPH observed in the before-mentioned structures. Although the stereoselectivity can be altered by active site mutagenesis, it is conserved throughout the class B MOs [16]. Two exceptions in the PDB display a more suitable syn conformation: cadaverine MO (PDB ID 508R [17]) where unfortunately the NADP was modeled on diffuse electron density and its validity is doubtful; and a mutant of a bacterial trimethyl-amine MO (TMM, PDB ID 5IQ4 [18]), where the electron density of the nicotinamide suffered from low occupancy (Figure 3b–c).

When NADP+ is in its ‘usual’ position, a hydrogen bond from the amide oxygen crucially stabilizes the N5 hydrogen of the reduced flavin [18] and the subsequently forming peroxyflavin [19]. Additionally, the ribose 2’ hydroxyl group hydrogen bonds to the reaction intermediate in BMOs, and donates its proton to form the hydroperoxyflavin in FMOs/NMOs [20]. By a flip of the amide, the amine can also interact with the N5 of the oxidized flavin after product formation in a retained overall conformation of NADP+. The distinction is difficult, as the orientation of
NADP and protein mobility. (a) Cut-open surface of PAMO (1W4X) with FAD- (yellow), NADP- (blue) and helical domains (orange). An ‘L’ marks a moving BVMO loop with a conserved tryptophan (light grey), which can be folded in (2YL, white cartoon) when NADP$^+$ is present or form a β-hairpin (3UOZ, dark grey) in a homolog. The inset magnifies the flavin (yellow carbons) and the various positions found in class B MOs of NADP’s nicotinamide ring. ‘N1’ marks the apparent ‘sliding’ movement by overlaying STMO (4AOS, green carbons), CHMO (5GWF, cyan carbons), and PAMO (2YL, blue carbons). ‘N2’ marks an apparent rotation via a half-rotated (TMM, 5GSN, dark violet carbons and mFMO, 2XLR, violet carbons) to a fully rotated form in CHMO (3UCL, pink carbons). (b-e) Electron densities ($e = 1$) of structures with controversial NADP$^+$ modeling: (b) cadaverine MO (6OBR) and (c) the TMM Y207S mutant (5G04) are modeled with NADP in a hydride transfer-suitable syn conformation, but suffer from poor electron density at the nicotinamide end. (d) CHMO (3UCL) and (e) TMM (5GSN) with half-rotated, and fully rotated NADP$^+$, respectively, where additional density connected to NADP was modeled as substrate molecules.

However, this assignment is controversial, as it stands in contrast to previous ligand positions and there is a noticeable connection to the density of the nicotinamide riboside (Figure 3d–e). It can, therefore, hardly be excluded that the origin is an alternative conformation of NADP, rather than a ligand. Further research should clarify the substrate position and whether the rotated cofactor is a general mechanism of the enzyme class. This may contribute to solving two remaining puzzles: the structural basis for the different mechanisms and reactivities, and the cause of the vast discrepancy in substrate specificity.

**Mobility of loops and domains in flavoprotein MOs**

Substrate acceptance is an intensely researched enzyme trait with biotechnological relevance, and protein flexibility was identified as ‘perhaps the single most important mechanism’ to achieve promiscuity [24]. The most flexible protein structures are loops and unsurprisingly, this structural element differs most among otherwise similar flavoprotein MOs.
In BVMOs, a long omega loop (where start and end are close and act as a hinge [25]) appears crucial for function and was called ‘control loop’ [26]. If visible, the loop folds on top of the Rossmann fold-bound NADP, thereby often trapping the cofactor in the crystal structure (Figure 3a). SAXS experiments indicate that NADP* exposure favors this folded state, which also coincides with ‘closed’ enzyme conformations. In ‘open’ conformations, not only the disordered loop may be unstructured, but also a wide swing into the solvent (deemed a crystallization artefact) was seen in phenylacetone MO (PAMO, PDB ID 1W4X [27]), and 2-oxo-Δ^3-4,5,5-trimethylcyclopentenylacetyl-coenzyme A MO, where the loop adopts a structured β-hairpin (e.g. PDB ID 3UOZ [28]) (Figure 3a). A central role in loop reorganization is assumed for a conserved tryptophan (Figure 3a), which is an active site residue if the loop is folded and whose removal drastically reduces enzyme activity [15]. The loop may also act as an ‘atomic switch’ [15,26] that connects the active site and the BVMO signature motif [29], a strictly conserved stretch at the edge of the NADP domain, inexplicably far from the active site. A histidine in this motif adopts varying conformations and can form contacts with the linker region, which in turn is connected to the control loop [15]. The importance and ability of the linker for long-range effects became also apparent when mutations in this region drastically altered enzymatic activity [30]. Considering that the SAXS results were not fully explainable by loop movements, these data collectively suggested that larger movements of the domains could occur. Domain rotations of up to 6° [15,31] were already observed, but the extent might have been artificially hindered by crystal packing [26]. A drastic domain rotation of 30° has been observed for an NMO, NbtG [32], but it is unknown whether other NMOs, let alone other class B families can sample this conformation as well. More distantly related enzymes with the same domain architecture are able to rotate by even 67° [33], and some members of class A flavoprotein MOs can cover their active site with a flexible ‘lid’ domain [34]. Future discoveries on such mechanisms in class B MOs can be expected, and these may be key in understanding their varying selectivities. It might also allow to explain the profound allostERIC effects of active site-remote mutations [35], and the surprisingly mild effects of removal of residues that (seemingly) form the active site [36*].

Cytochrome P450s

Referred to as ‘nature’s blowtorch’ [37], the iron-oxo species forming in the core of cytochrome P450s MOs (P450s) are endowed with the oxidative power to catalyze various reactions: besides performing dealkylations, heteroatom oxidations and epoxidations, P450s hydroxylate non-activated C–H and C–C bonds in substrates of diverse size, functional group composition, and polarity [38]. Similar to class A flavoprotein MOs, the catalytic mechanism is initiated by substrate binding, which causes a separate or translationally fused reductase to shuttle NAD(P)H-derived electrons to the heme (Figure 1). Dioxygen binds to the one electron-reduced ferrous heme and the second electron creates the ferric peroxy complex, which matures to the catalytically active ‘Compound I’. Despite amino acid sequence differences of up to 90%, all P450s share a common fold with identical topology and conserved secondary structural elements. The question arises, how such a highly conserved architecture can sustain the observed immense variety in catalyzed reactions. Clearly, the P450 fold evolved early as a safe platform for an inherently dangerous reaction – the activation of molecular oxygen – and as a versatile scaffold. As such, the variability of P450 reactions cannot be attributed to the composition and capacity of the active site but is rather a result of the concerted and dynamic action of the whole enzyme. A large body of research spanning both selective prokaryotic and highly promiscuous eukaryotic P450s demonstrates the essential role of plasticity in the selection of suitable substrates and their delivery to the heme.

Questions concerning P450 flexibility involved in substrate binding have already been raised after the first crystal structure. In P450cam, the camphor substrate is effectively sealed from the outside, implying a structural plasticity that enables the protein to open for substrates to enter and products to leave [39]. Subsequent crystal and NMR structures as well as molecular dynamics simulations have since then confirmed how an impressive degree of flexibility in P450s facilitates a stepwise adaptation of the enzyme to the substrate in order to lead it to the active site.

Binding mechanisms in P450s

Work on CYP3A4, a human P450 involved in xenobiotic metabolism, supported an induced fit substrate binding mechanism. The enzyme structure in complex with midazolam hints at substrate-induced, global structural readjustments, with concurrent reshaping of the active site. In particular, a conformational switch of two helices (the F–G segment) and long-range residue movements transmitting from remote areas (the D, E, H, and I helices) triggered a collapse of the active site cavity and ligand immobilization. Productive substrate positioning can occur at two overlapping binding sites near the I helix, and a substrate concentration-dependent collapse or widening of the catalytic cavity determines the reaction’s regioselectivity [40]. Structural investigations of the prokaryotic OleP in complex with a macrolactone are also consistent with an induced-fit binding, whereby a cascade of interactions responsible for substrate-induced conformational changes was proposed [41]. Some P450s, however, were shown to explore an incessant motion between different conformations regardless of the presence of substrates. The ligand-free structures of the erythromycin-converting P450 EryK suggest the presence of a
heterogeneous conformational ensemble between an open and a closed state [42].

Notably, the conformational changes occurring upon substrate recognition can show striking similarities between very distant representatives. P450cam and MycG are only 29% identical on sequence level and act on the structurally diverse substrates camphor and mycinamicin IV, respectively. Using a combination of NMR structural studies, site-directed mutagenesis and functional assays, several regions far from the active site of P450cam were demonstrated to be critical to ensure efficient recognition and orientation of the substrate into the catalytic center. Many of the same secondary structural features in MycG are perturbed upon substrate binding. The most-affected residues were subsequently found to be functionally important and lie in a conical region roughly anti-symmetric with the triangular shape of the P450 molecule [43*].

P450s’ substrate selection via tailored plasticity

With twelve entries deposited in the protein data bank, CYP2B enzymes show one of the highest degree of plasticity among crystallographically characterized P450s — about one third of the protein is accounted for by five plastic regions (PRs). Comparison of PR2 and PR4 allowed to distinguish four distinct conformations: ‘open’ to allow substrate access, ‘closed’ and ‘expanded’ upon binding of small and large ligands to CYP2B4, respectively, and an ‘intermediate’ form induced by and molded to the inhibitor 1-biphenyl-4-methyl-1H-imidazole (1-PBI) (Figure 4a) [44]. As catalysis involves subtle, concerted conformational changes spanning a large part of the enzyme, allosteric effects are frequently observed and sometimes drastic. In CYP2Bs, mutations of residues remote from the active site caused not only a switch in selectivity for some substrates, but also profound functional changes affecting the enzyme’s catalytic rates and inhibition [45]. Interestingly, mutations targeting active site residues produced much smaller changes [46]. In CYP2B1, equally distant mutations enhanced the metabolism of several substrates including the anticancer prodrugs cyclophosphamide and ifosfamide [47]. Similarly, the enhanced activity of a rat CYP1A1 mutant toward a dibenzo-p-dioxin toxin is triggered by a more efficient binding of the substrate in the active site even though the mutation is over 25Å away [48]. In this scenario, it is not surprising how most of the single nucleotide polymorphisms (SNPs) that make CYP2B6 highly polymorphic and, accordingly, differently active in the metabolism of a variety of drugs lie far from the active site of the enzyme [49]. Another demonstration of how the creation of a new activity passes mostly through mutations in flexible regions involved in substrate recognition [50] is the engineering of P450-BM3 toward a propane monoxygenase [51] where only a fraction of the mutations was located in the active site (Figure 4b).

The role of dynamics of the overall P450 fold is also well exemplified by the long-range effects of putidaredoxin (Pdx) binding to the proximal face of P450cam, which influences motions on the opposite side of the protein. The open/close motion of the F/G helical region is
coupled to a movement of the C helix, which directly contacts Pdx. The Pdx-induced changes in the F/G helical region are instrumental to carry out the enzymatic activity: it triggers free an important aspartate involved in the proton delivery network required for $O_2$ activation [52*]. Even the entrance of molecular oxygen into the active site is tuned by protein dynamics. Simulations of the protein backbone dynamics of P450-BM3 revealed the transient nature of some channels, with subchannels forming and merging and $O_2$ molecules hopping in between [53,54*].

The full understanding of P450s catalysis is pivotal for exploiting their selectivity in industrial processes and designing tailored inhibitors for drug metabolism. The joint participation of remote, flexible elements can represent a complication, as their influence on specificity and catalytic activity may be difficult to predict. This explains why directed evolution approaches with this enzyme family have been much more successful than rational approaches focused on active-site engineering. A picture emerges where the active site of P450s are reduced to a mere accessory role. A recent structural characterization of different members of CYP153s illustrates this. Among these homologs, all active site residues are conserved, but the enzymes display varying hydroxylation activities with alkanes, fatty acids, and heterocyclic compounds. The comparison of five crystal structures allowed to plot out the regions which exhibited the most pronounced sequence variabilities and conformational changes. In this manner, it was possible to identify the B/C-loop, the F, G, and H helices and the F/G-loop to be responsible for substrate recognition and binding [55*].

Conclusions
While flavin-dependent MOs compensate their subdomain’s intrinsic rigidity by linker and loop movements and/or cofactor mobility, P450s counterbalance the heme cofactor’s inflexibility by widely dispersed mobile regions involved in substrate binding. The structural and mechanistic complexity found in flavoprotein MOs reflects the complex catalytic duty of efficiently coordinating three substrates by the same active site in a timely regulated fashion. A complete understanding of the reaction mechanism relies on future discoveries, specifically with regard to hydride transfer and substrate selectivity differences. When considering P450s, novel features of their mechanisms have emerged from various P450 subfamilies. For both monooxygenase classes, it has become clear that structural dynamics play an important role in their catalytic functioning. Besides better understanding their molecular functioning, new insights will hopefully clarify vast discrepancies in substrate acceptance and fuel the design of enzyme engineering strategies. Clearly, such rational approaches need to take all steps and loci involved in enzyme catalysis into consideration, rather than focusing solely on the chemical step thought to occur in a static active site.

Conflict of interest statement
Nothing declared.

Acknowledgements
The research for this work has received funding from the European Union (EU) project ROBOX (grant agreement no 635734) under EU’s Horizon 2020 Programme Research and Innovation actions H2020-LEIT BIO-2014-1.

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:
• of special interest
• of outstanding interest


This study performed the structural characterization of several pathogenic resistance-conferring bacterial tetracycline destructases in complex with tetracyclines. With observed variations in ligand binding and flavin conformation, the authors realized that some tetracyclines lock FAD in one orientation, thus laying the basis for a new class of tetracycline deconstructase inhibitors.

Catalysis and regulation


The surprisingly small influence of the simultaneous removal of active-site or substrate tunnel residues in a thermostable MO led to the proposition that substrate promiscuity can stem from an ‘excluding rather than binding’ mechanism.


Using two very distantly related P450s acting on diverse substrates, this study demonstrates how some conformational changes involved in substrate binding seem to be conserved among the entire P450 class.


47. Kumar S, Chen CS, Waxman DJ, Halpert JR: Directed evolution of mammalian cytochrome P450 2B1: mutations outside of the active site enhance the metabolism of several substrates


Using structural and computational methods, this study investigates the long-range effects exhibited by P450cam’s redox partner, putidaredoxin. Upon binding, an internally propagated movement frees a crucial residue necessary for proton delivery, and the enzyme is found in an intermediate state resembling the ‘open’ form, but with a highly ordered active site.


54. Ebert MC, Guzman Espinola J, Lamoureux G, Pelletier JN:

Using a combination of steered molecular dynamics simulations and docking, catalytically relevant steps along the binding trajectory of a fatty acid substrate to P450-BM3 are revealed and experimentally verified.

55. Fiorentini F, Hatzl A-M, Schmidt S, Savino S, Glieder A, Mattevi A:
• The extreme structural plasticity in the CYP153 subfamily of P450s directs development of designer hydroxylases. Biochemistry 2018, 57:6701-6714.

In a combined structural and biochemical study, the authors show how variations in the mobile regions involved in substrate uptake (and not the active-site composition) are the sole elements responsible for selectivity differences in CYP153 enzymes.