Organic & Biomolecular Chemistry

PAPER

View Article Online
View Journal | View Issue

Cite this: *Org. Biomol. Chem.*, 2019, **17**, 1506

Characterization of the flavin monooxygenase involved in biosynthesis of the antimalarial FR-900098†

Kim Nguyen,‡[a] Matthew A. DeSieno,‡[b,e] Brian Bae,§[a,e] Tyler W. Johannes,¶[b] Ryan E. Cobb,[b,e] Huimin Zhao*[a,b,c,d,e] and Satish K. Nair *[a,c,d,e]

The latter steps in this biosynthetic pathway for the antimalarial phosphonic acid FR-900098 include the installation of a hydroxamate onto 3-aminopropylphosphonate, which is catalyzed by the consecutive actions of an acetyltransferase and an amine hydroxylase. Here, we present the 1.6 Å resolution co-crystal structure and accompanying biochemical characterization of FrbG, which catalyzes the hydroxylation of aminopropylphosphonate. We show that FrbG is a flavin-dependent *N*-hydroxylating monooxygenase (NMO), which shares a similar overall structure with flavin-containing monooxygenases (FMOs). Notably, we also show that the cytidine-5′-monophosphate moiety of the substrate is a critical determinant of specificity, distinguishing FrbG from other FMOs in that the nucleotide cofactor-binding domain also serves in conferring substrate recognition. In the FrbG-FAD^{+}-NADPH co-crystal structure, the C4 of the NADPH nicotinamide is situated near the N5 of the FAD isoalloxazine, and is oriented with a distance and stereochemistry to facilitate hydride transfer.

Received 14th November 2018,
Accepted 21st January 2019

DOI: 10.1039/c8ob02840k

rsc.li/obc

Introduction

The rapid access to genomic information has spurred a renewed interest in the biosynthesis of natural products that contain phosphonate or phosphinate groups. Phosphonates are characterized by the presence of a hydrolytically inert C–P bond, and the biological activities of such compounds derive from their structural similarity to the corresponding labile phosphate ester. Studies by Seto and co-workers established that the thermodynamically unfavorable C–P bond formation reaction is catalyzed by the enzyme phosphoenolpyruvate mutase, and subsequent characterizations of phosphonate biosynthetic enzymes have revealed several unusual catalysts. Recent studies have resulted in the characterization of the biosynthetic clusters for several phosphonates, including the antibiotic fosfomycin, the herbicide phosphinothricin, the antifungal rhizocticin and the antimalarial candidate FR-900098 (compound **1** in Fig. 1a).[1]

FrbG was identified during the cloning and sequencing of the biosynthetic gene cluster for FR-900098 from *Streptomyces rubellomurinus*.[2] Initial sequence analysis of FrbG using bioinformatics tools failed to identify homologs in the protein database with significant sequence similarity. Further characterization of the enzymes, including whole cell feeding studies, established that FrbG functions as a *N*-hydroxylase during the biosynthesis of the hydroxamate moiety in FR-900098.[3] Upon purification, in the absence of exogenous compounds, the protein solution appears bright yellow, and a UV-vis spectroscopic analysis of the purified protein in solution produces an absorbance spectrum characteristic of a flavin-containing molecule (Fig. 1b), suggesting that FrbG is a flavoprotein monooxygenase.

Flavoprotein monooxygenases catalyze the regioselective monooxygenation on a wide variety of substrates including organic compounds, drugs, and xenobiotics. The flavin-containing monooxygenases (FMOs) and closely related microbial

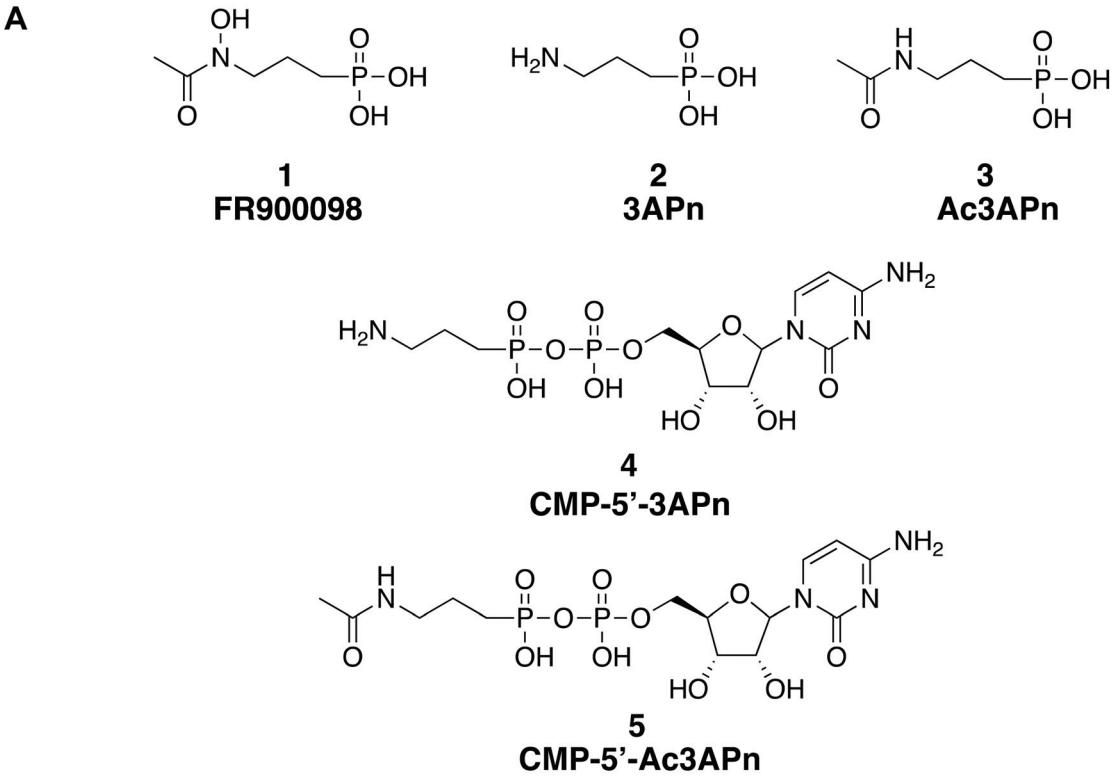
[a]*Department of Biochemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, IL 61801, USA*
[b]*Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, IL 61801, USA*
[c]*Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, IL 61801, USA*
[d]*Center for Biophysics and Quantitative Biology, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, IL 61801, USA*
[e]*Carl R. Woese, Institute for Genomic Biology, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, IL 61801, USA.*
E-mail: zhao5@illinois.edu, snair@illinois.edu; Fax: +(217) 333-5052, +(217) 244-5858; Tel: +(217) 333-2631, +(217) 333-0641

†Electronic supplementary information (ESI) available. See DOI: 10.1039/c8ob02840k

‡These authors contributed equally to this work.

§Current address: Rockefeller University, 600 York Avenue, New York, NY, 10021, USA.

¶Current address: The University of Tulsa, Department of Chemical Engineering, 800 South Tucker Drive, Tulsa, OK 74104, USA.



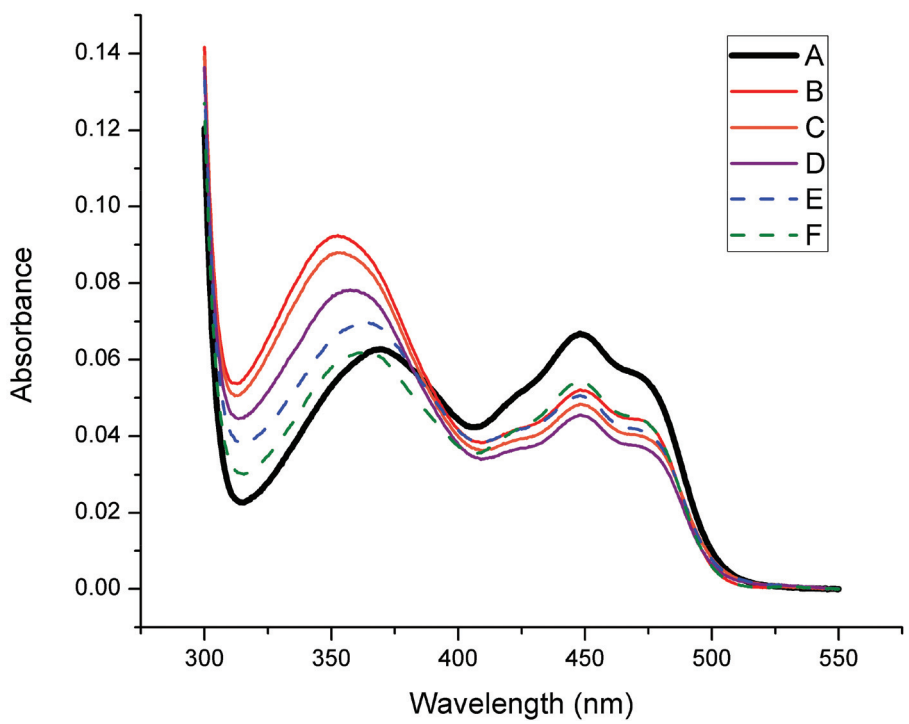


Fig. 1 (a) Chemical structures of (**1**) FR900098, (**2**) 3-aminopropylphosphonate, (**3**) *N*-acetyl-3-aminopropylphosphonate, (**4**) CMP-5′-3-aminopropylphosphonate, and (**5**) CMP-5′-*N*-acetyl-3-aminopropylphosphonate. Only compound **3** (CMP-5′-3-aminopropylphosphonate) is a substrate for FrbG. (b) Absorbance spectra of FrbG. Trace A represents the fully oxidized form of the enzyme (at a concentration of 1 μM) in the absence of NADPH. Trace B was obtained immediately after addition of equimolar NADPH under aerobic conditions. Traces C–F were recorded 1 min, 5 min, 10 min and 15 min after addition of the cofactor. The traces where the flavin is becoming more reduced are solid and where it is returning to the oxidized state at later times are dashed. These demonstrate the slow binding of NADPH by the enzyme and formation of the C4α-hydroperoxyflavin intermediate.

N-hydroxylating monooxygenases (NMOs), represent a subclass of these enzymes in which a single polypeptide can catalyze both monooxygenation and flavin reduction using an NADPH cofactor as an electron donor to reduce the FAD. Unlike other monooxygenases, FMOs and NMOs do not require the presence of the substrate for the reduction of the FAD prosthetic group by NADPH.[4] Instead, the enzyme-bound cofactor reacts with molecular oxygen to form an unstable C4α-hydroperoxyflavin intermediate, which is now primed to function as either an oxygenase (FMO) or *N*-hydroxylase (NMO).[5]

Prior biochemical and structural studies on FMOs suggests a model for catalysis in which the NADP(H) cofactor can participate in both flavin reduction and stabilization of the reactive C4α-hydroperoxyflavin against solvent-mediated decay to hydrogen peroxide.[6] Crystal structures of eukaryotic and bacterial FMOs,[6,7] along with that of the related phenylacetone Baeyer–Villiger monoxygenase,[8] reveal a conformation in which the $NADP^+$ cofactor is bound away from the flavin, in an orientation that would preclude hydride transfer, and in a conformation that is proposed to stabilize the C4α-hydroperoxyflavin. Consequently, the catalytic cycle of FMOs/NMOs is proposed to occur through conformational changes in the enzyme, involving a "pre-hydride transfer" conformation that promotes FAD reduction by NADPH, and is followed by a structural rearrangement to yield a "post-hydride transfer" conformation wherein the $NADP^+$ is now poised for stabilization of the intermediate.

We report here biochemical characterization that establishes FrbG as the flavin-dependent *N*-hydroxylating monooxygenase involved in FR-900098 biosynthesis. We also present the 1.6 Å resolution co-crystal structure of FrbG-FAD-NADPH complex. The structure of FrbG corresponds to a "pre-hydride transfer" conformation of an FMO family member, with the C4 of the nicotinamide and the N5 of the isoalloxazine in an orientation that would facilitate reduction of FAD by NADPH. Comparisons to structures of FMOs in the "post-hydride transfer" conformation,[6,7] in which the nicotinamide and flavin are distally situated, show structural reorganizations that likely occur following dioxygen reduction of the flavin. The structural data suggests a number of active site residues that may play a role in the catalysis, and site-specific mutagenesis at these residues is used to establish the function of these residues in the catalytic cycle. Notably, we demonstrate that conjugation of the substrate 3-aminopropylphosphonate to cytidine-5′-monophosphate (CMP) is essential for catalysis, suggesting that the nicotinamide-binding domain may also function in conferring substrate specificity.

Results and discussion

Biochemical analysis

Recombinant FrbG was heterologously produced in *E. coli* with a (cleavable) N-terminal polyhistidine fusion tag. Purification of protein through multiple chromatographic steps, specifically metal affinity, cation exchange, and size exclusion, yielded sample that was yellow in color indicative of a bound cofactor. Steady state kinetic analysis was performed at a fixed concentration of NADPH with a series of putative substrates selected on the basis of the proposed biosynthetic pathway of FR-900098. Based on the originally proposed pathway, FrbG was believed to *N*-hydroxylate either 3-aminopropylphosphonate (3APn; compound **2** in Fig. 1a) or *N*-acetyl-3-aminopropylphosphonate (Ac3APn; compound **3** in Fig. 1a).[2] However, we have recently demonstrated that the later steps in the pathway contain phosphonate intermediates that are conjugated to CMP,[3] and therefore, we also analyzed the CMP conjugates of 3APn and Ac3APn (compounds **4** and **5**, respectively, in Fig. 1a). FrbG demonstrated activity towards only one of these four substrates, CMP-5′-3APn (**4**) (Table 1 and Fig. S1 in ESI†).

Fully oxidized FrbG displays the expected UV absorption maxima at 265, 369 and 450 nm with a shoulder at 475 nm (Fig. 1b, trace A). Stabilization of the C4α-hydroperoxyflavin intermediate has been well documented for several other FMOs.[6,9–11] In these enzymes, mixing equimolar oxidized protein and NADPH in an aerated buffer solution lead to immediate formation of the highly stable intermediate, with a characteristic peak at ~360 nm.[9,12–14] Upon addition of NADPH to FrbG, an increase in absorbance and shift to 349 nm was recorded, instead of the expected peak (trace B). After 15 minutes, the spectrum resembled the C4α-hydroperoxyflavin intermediate (trace E). Also, the fully reduced form of FrbG could not be obtained, even in the presence of 10-fold excess of NADPH. In typical FMOs, the enzyme will become immediately reduced (seen as a shoulder instead of a peak at 450 nm). Due to the inherent NADPH oxidase activity of these enzymes, the spectrum will slowly revert to the spectrum of the oxidized flavin as hydrogen peroxide (H_2O_2) is released from the FMO.[6] In the case of FrbG, there is a slow decrease in the observed peak at 450 nm, instead of immediate formation of the fully reduced flavin (traces B–D). This peak reaches a minimum after approximately 5 minutes, followed by an increase in absorbance as the spectrum re-acquires the characteristic of oxidized flavin (traces E–F). This occurrence can be explained by the proposed "gatekeeper" role of $NADP^+$ in FMOs.[15] The normally tightly bound cofactor helps stabilize the C4α-hydroperoxyflavin intermediate, preventing rapid formation of H_2O_2 in the absence of substrate. The fully reduced flavin was unable to be obtained under anaerobic conditions as well with the addition of $NADP^+$. The inclusion of the oxidized cofactor inhibited the speed of flavin reduction compared to NADPH alone. These studies suggest alternate

Table 1 Kinetic parameters of FrbG

Substrate	k_{cat} (min^{-1}) (mean ± SD)	K_M (μM) (mean ± SD)	k_{cat}/K_M (mM^{-1} min^{-1})
CMP-5′-3APn (**4**)	5.9 ± 0.4	15.1 ± 1.5	391
CMP-5′-Ac3APn (**5**)			NR
3APn (**2**)			NR
Ac3APn (**3**)			NR

NR: no reaction.

binding between the substrate and the cofactor, making the C4α-hydroperoxyflavin intermediate more susceptible to NADPH oxidase activity compared to the case where $NADP^+$ is the final product to be released.

FrbG has relatively lower activity compared to other FMOs (k_{cat}/K_M of ~10^4 M^{-1} s^{-1} as compared to 10^5 to 10^6 M^{-1} s^{-1} for other characterized FMOs), which is likely the result of several factors. Enzymes in secondary metabolism tend to have lower rates compared to those within primary metabolism, so as to not divert key metabolites and resources away from critical reactions. This phenomenon is likely seen here with the susceptibility of FrbG to NADPH oxidase activity. Finally, the FrbG crystal structure, which will be discussed further, contained NADPH rather than the oxidized $NADP^+$. As a result, the equilibrium between the two must favor NADPH that would also lower the overall rate of the enzyme reaction.

Overall structure

In order to gain molecular insights into the mechanism of catalysis, we solved the structure of full-length FrbG to a resolution of 1.6 Å (Fig. 2a). Despite the fact that no exogenous ligands were added to the crystallization buffers, the crystals had a very yellowish appearance, and clear electron density can be observed for both NADPH and FAD in both experimental electron density maps and difference maps calculated with phases from the final refined model (Fig. 3a). The non-planar electron density observed for the nicotinamide ring is consistent with the presence of the reduced form of the cofactor. In the crystal, there is one protein-NADPH-FAD complex per asymmetric unit, consistent with the monomeric state observed for the enzyme in solution (Fig. S2†). The structure has been refined to a final free *R* factor of 22.1% (see Table 2 for relevant crystallographic statistics).

FrbG is composed of two structural domains (Fig. 2a and b). Residues 167–340 form a small (NADPH binding) domain and residues 1–166 along with 341–430 form a larger single (FAD binding) domain. Each of the two motifs contains a β-α-β motif characteristic of mononucleotide binding sequences. The two domains are situated close to one another such that a single channel that extends from one domain to the other is nestled within the core of the protein. The region where the prosthetic group, FAD, and NADPH interact is located at the interface of the larger and smaller domains. The NADPH cofactor is sandwiched between the *re* side of the flavin and the carboxy edge of the central parallel β-sheet and is enveloped predominantly by the smaller domain of FrbG.

A DALI structure-based comparison against the Protein Data Bank using the intact structure of FrbG identifies very few structural homologs, and is likely the consequence of various insertions and structural rearrangements between the two domains that distinguishes FrbG from other FMOs. Bioinformatics analysis using each of the isolated domains identifies the FAD-binding domain of the thioredoxin reductase TrxR from *Helicobacter pylori*[16] as a structural homolog of the larger FAD-binding domain (RMS deviation of 3.3 Å over 271 Cα atoms with 13% sequence identity) and the NADPH-binding domains of the FMO from *Schizosaccharomyces pombe* (RMS deviation of 3.2 Å over 287 Cα atoms with 18% sequence identity) and *Methylophaga* sp. SK1 (RMS deviation of 3.3 Å over 285 Cα atoms with 14% sequence identity). Although the overall fold of FrbG is similar to those of other FMOs, the enzyme is distinguished by insertions of several secondary structural elements (*vide infra* and Fig. S3†).

FAD binding domain

The nucleotide-binding motif GGGPAG forms the loop connecting the carboxy end of strand β1 to helix α1 and stabilizes binding of the adenosine diphosphate portion of FAD (Fig. 3b). The glycine-rich region of the motif positions the central part of FAD close to the protein framework. The amino end of helix α1, specifically the nitrogen of Ala-14, makes a hydrogen bond with the pyrophosphate group O1P of FAD. The adenine nucleotide is stabilized by hydrogen bonding contacts to the main chain nitrogen atoms of Ala-41 and Val-131 and the carbonyl oxygen of Val-131. The prosthetic group is well solvated in the channel with water molecules mediating protein contacts. A well-localized water molecule further anchors the flavin pyrophosphate group to the GGGPAG motif. This bound water molecule hydrogen bonds with the nitrogen atoms of Gly-12 and Gly-15, the carbonyl oxygen of Gly-10, and O1P of the flavin phosphate. An oxygen atom of the ribitol group along with an oxygen atom and several nitrogen atoms from the isoalloxazine ring of the flavin nucleotide interact with the large domain *via* water molecules. In the enzyme-FAD-NADPH complex, the isoalloxazine ring of FAD is stacked with the nicotinamide of $NADP^+$ (Fig. 3c). The oxygen atoms, O2 and O4, of the isoalloxazine ring form hydrogen bonds with the main chain nitrogens of Val-412 and Asn-58, respectively. Although the flavin nucleotide extends toward the interaction domain to engage in stacking interactions with the nicotinamide, the prosthetic group remains predominantly enclosed in the larger domain of FrbG surrounded by a network of water molecules.

Structural and biochemical studies of FMO from *S. pombe* identified an Asn-91 residue located near the isoalloxazine ring of FAD that is believed to be directly involved in the catalytic mechanism.[7] This asparagine is thought to be crucial in supplying molecular oxygen to the isoalloxazine ring and stabilizing the resultant C4α-hydroperoxyflavin intermediate. FrbG has Asn-58 located at an analogous position near the isoalloxazine ring and the Nδ2 of Asn-58 is 2.98 Å from a water molecule. This same water molecule is 3.80 and 2.96 Å from C4α and N5 of FAD, respectively. Initial refinement of *S. pombe* FMO had identified a water molecule in a comparable location that was later modeled with a dioxygen molecule.[7] Dioxygen molecules near enzyme active sites have been observed with naphthalene dioxygenase and cytochromes P450 as well.[17,18] Mutation of Asn-58 to Ala in FrbG demonstrated its importance in the catalytic cycle of the enzyme. The FrbG-N58A variants exhibited robust NADPH oxidase activity, but activity with the CMP-5′-3APn substrate was compromised relative to

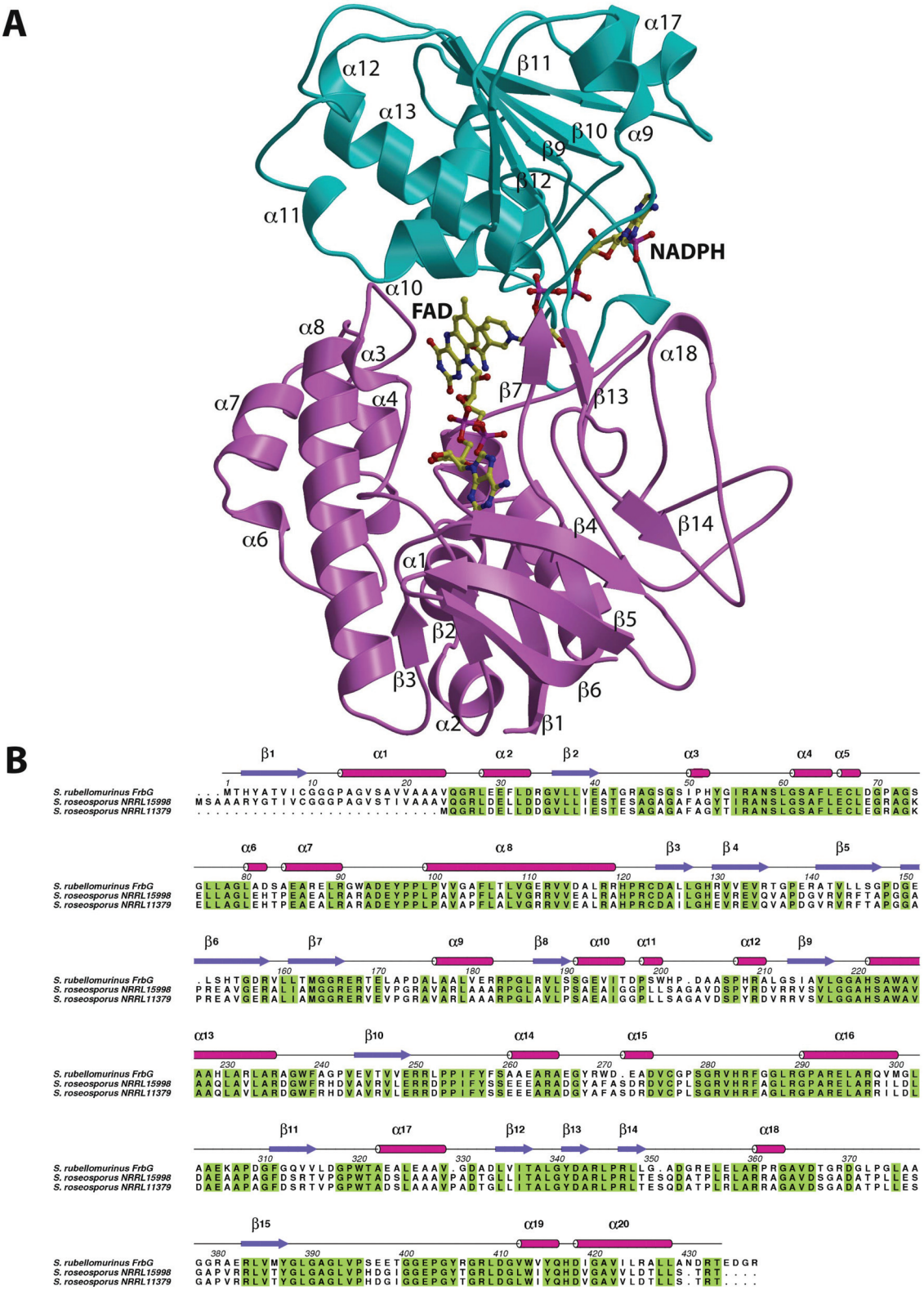


Fig. 2 (a) Ribbon diagram derived from the 1.6 Å resolution structure of FrbG with the $NADP^+$ binding domain colored in cyan and the FAD binding domain colored in magenta. The coordinates for each of the ligands are shown as stick figures. (b) Structure-based sequence alignment of FrbG with other flavin dependent monooxygenases. Secondary structure elements demarcated in the alignment correspond to those noted in the ribbon diagram.

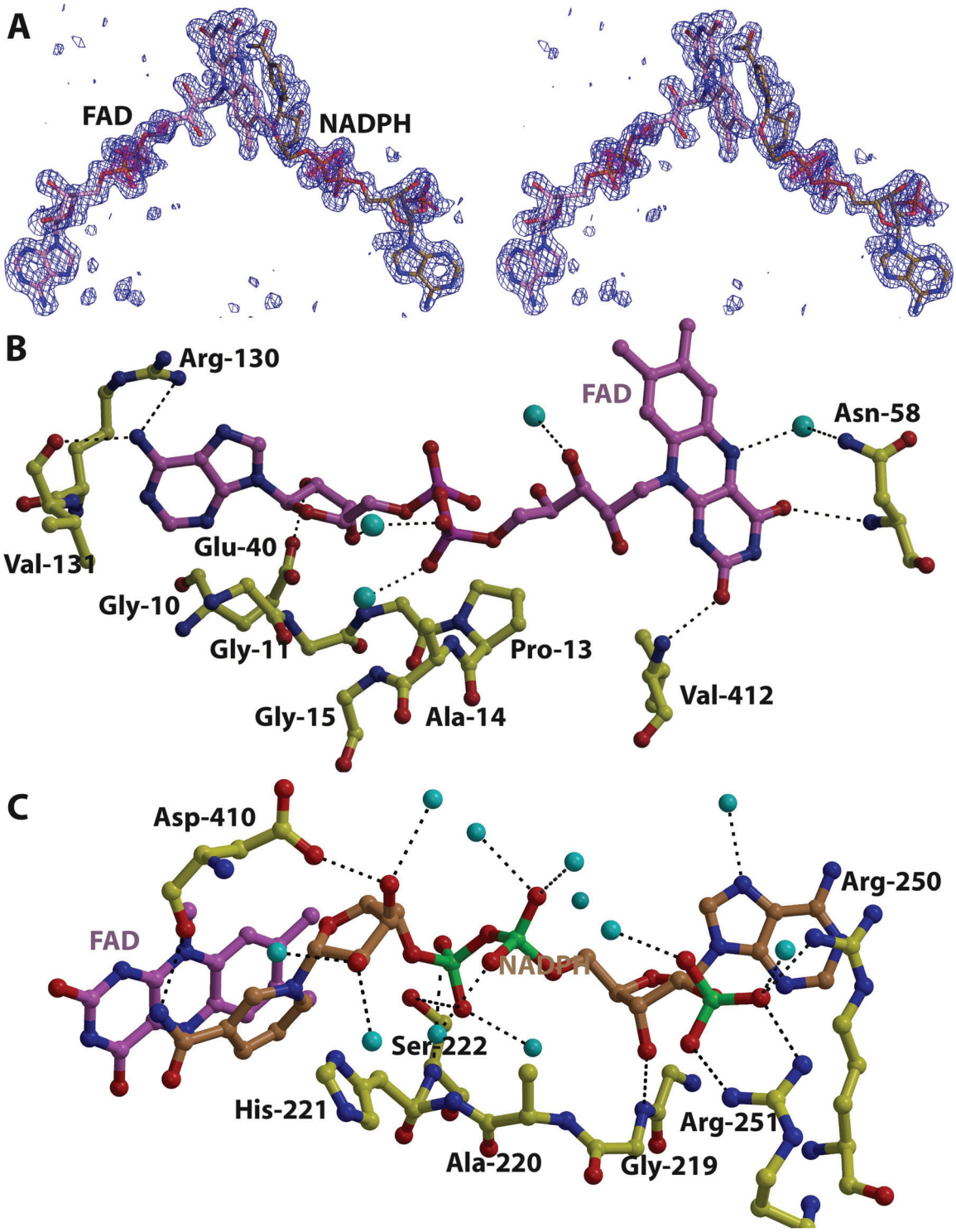


Fig. 3 (a) Electron density maps calculated using Fourier coefficients $F_{obs} - F_{calc}$ with phases derived from the refined 1.6 Å resolution structure of FrbG calculated with the coordinates for $NADP^+$ and FAD omitted prior to one round of crystallographic refinement. The map is contoured at 3σ (blue mesh) and 8σ (red mesh) and the final refined coordinates for each ligand are superimposed. (b) Protein residues involved in the binding of the $NADP^+$ are shown in yellow, the $NADP^+$ cofactor is shown in orange and the isoalloxazine ring of FAD is shown in purple. (c) Protein residues (yellow) involved in interactions with FAD (purple). Solvent molecules within the first hydration shell are shown as cyan spheres.

Table 2 Data collection, phasing and refinement statistics

	Hi-resolution	SeMet FrbG
Data collection		
Space group	*C*2	*C*2
Unit cell dimensions		
a, *b*, *c* (Å)	121.3, 64.7, 69.2	121.3, 64.6, 69.1
α, *β*, *γ* (°)	90.0, 124.7, 90.0	90.0, 124.8, 90.0
Resolution (Å)	50–1.6 (1.66–1.6)[a]	50–1.8 (1.86–1.8)[a]
R_{sym} [b] (%)	5.8 (27.1)	6.9 (20.6)
$I/\sigma(I)$	42.1 (4.8)	37.6 (8.5)
Completeness (%)	97.2 (88.4)	97.7 (96.3)
Redundancy	7.5 (5.9)	5.9 (5.6)
$\|F_H\|/e$ (anomalous)		1.27
FOM/DM FOM[c]		0.376/0.76
Refinement		
Resolution (Å)	25.0–1.6	50–1.8
Number of reflections	53 510	37 762
R_{work}/R_{free} [d]	19.3%/22.1%	18.7%/21.7%
Number of atoms		
Protein	3122	3156
Solvent	487	568
NADP	48	48
FAD	53	53
Average *B* value		
Protein	19.8	20.8
Solvent	31.9	35.3
NADP	15.9	16.4
FAD	11.8	11.5
R.m.s deviations		
Bond lengths (Å)	1.64	1.35
Bond angles (°)	0.006	0.007

[a] Highest resolution shell is shown in parenthesis. [b] $R_{sym} = \sum(|I_i - \langle I\rangle|)/\sum|I_i$ where I_i = intensity of the *i*th reflection and $\langle I\rangle$ = mean intensity. [c] Mean figure of merit before and after density modification (DM). [d] *R*-Factor = $\sum(|F_{obs}| - k|F_{calc}|)/\sum|F_{obs}|$ and *R*-free is the *R* value for a test set of reflections consisting of a random 5% of the diffraction data not used in refinement.

the wild-type (Fig. S4†). These data suggest that NADPP oxidation is uncoupled to substrate hydroxylation in this variant. One plausible explanation for this observation is that FrbG-N58A may not be able to sufficiently stabilize the C4α-hydroperoxyflavin intermediate to carry out the hydroxylation of the substrate.

NADP(H) binding domain

The NADPH cofactor binds on the *re* side of the flavin and is situated closely to the second nucleotide-binding motif, GGAHS (Fig. 3c). The nitrogen atoms of Gly-219 hydrogen bond to the 3′-OH of ribose of the adenine group and the main chain nitrogen atoms of His-221 and Ser-222 anchor the nicotinamide pyrophosphates *via* hydrogen bonding. This glycine-rich region, which shares structural homology to the FAD nucleotide-binding motif, forms a short loop connecting the carboxy end of strand β5 to helix α5. A side chain nitrogen of residue Arg-250 hydrogen bonds to the phosphate present on the 2′ carbon of the adenosine ribose of NADPH. The remaining side chain nitrogens of Arg-250 engage in stacking interactions with the adenine nucleotide of NADPH. In addition, numerous water molecules are bound within the channel and make hydrogen-bonding contacts with the protein as well as the cofactor. Water molecules, along with the side chain oxygen of Asp-410, stabilize the nicotinamide ribose as the reduced nicotinamide extends toward the interaction domain to stack with the isoalloxazine ring of the flavin moiety of FAD.

NADPH is bound at the carboxy edge of the central parallel β-sheet and is enveloped predominantly by the smaller domain of FrbG. Binding of NADPH in the smaller domain buries 101 Å^2 of surface area, in contrast with the 237 Å^2 of surface area buried upon FAD binding. Consequently, $NADP^+$ is not as strongly embedded within FrbG as FAD since both the adenine nucleotide and pyrophosphates of NADPH are exposed to the solvent, whereas FAD's adenine group is deeply embedded within FrbG. Given the volume occupied by the ligands, it is likely that following the reduction of the flavin, $NADP^+$ must be displaced from the active site prior to binding of the CMP-5′-3APn substrate, consistent with biochemical observations on other flavin monooxygenases.

Given our experimental data that only the CMP conjugate of the phosphonate (*i.e.* **4**) is a viable substrate for FrbG, the NADP(H) binding domain likely also serves in substrate selectivity. Specifically, the CMP-5′-3APn (**4**) substrate may be recognized by virtue of the nucleotide group of the conjugate. A docking model of CMP-5′-3APn (**4**) bound to the NADP(H) binding-domain, based on the superposition of the nucleobases of the substrate with cofactor NADPH provides a rationale for this specificity (Fig. S5†). The model suggests that binding of the CMP at a position similar to that of NADPH results in placement of the amine of 3APn directly above the isoalloxazine ring, adjacent to C4α where molecular oxygen is suggested to bind. The orientation of the 3APn amine is further fixed by His-221. FrbG-H221A had significantly higher rates for both NADPH oxidation and substrate hydroxylation compared to the wild-type FrbG. The alanine substitution at this residue is consistent with a role for His-221 in the reaction mechanism (Fig. S4†).

Mechanism of FrbG activity

As a major electron carrier in oxidation–reduction reactions, the isoalloxazine ring of FAD is first reduced to $FADH_2$ *via* a hydride transfer from C4 of NADPH to the N5 atom of the flavin. The reduced flavin subsequently reacts with molecular oxygen to form a C4α-hydroperoxyflavin intermediate, FAD-OOH.[19] This peroxyflavin intermediate has been shown to be stable for minutes to hours at 4 °C and has been spectroscopically observed.[20–23] Once a substrate with a nucleophilic atom is in close proximity to the peroxyflavin intermediate, the nucleophilic attack conducted by the substrate results in the oxygenation or hydroxylation of the substrate by one atom of molecular oxygen, whereas the other atom is released as water during the reaction.[24]

Prior studies of FMOs from *S. pombe* and *Methylophaga* sp. SK1 show that the flavin N5 and the nicotinamide C4 are at an orientation and distance incompatible with hydride transfer (inter-atomic distance of 5.3 Å) (Fig. 4c). This "post hydride transfer" conformation has been suggested to facilitate the

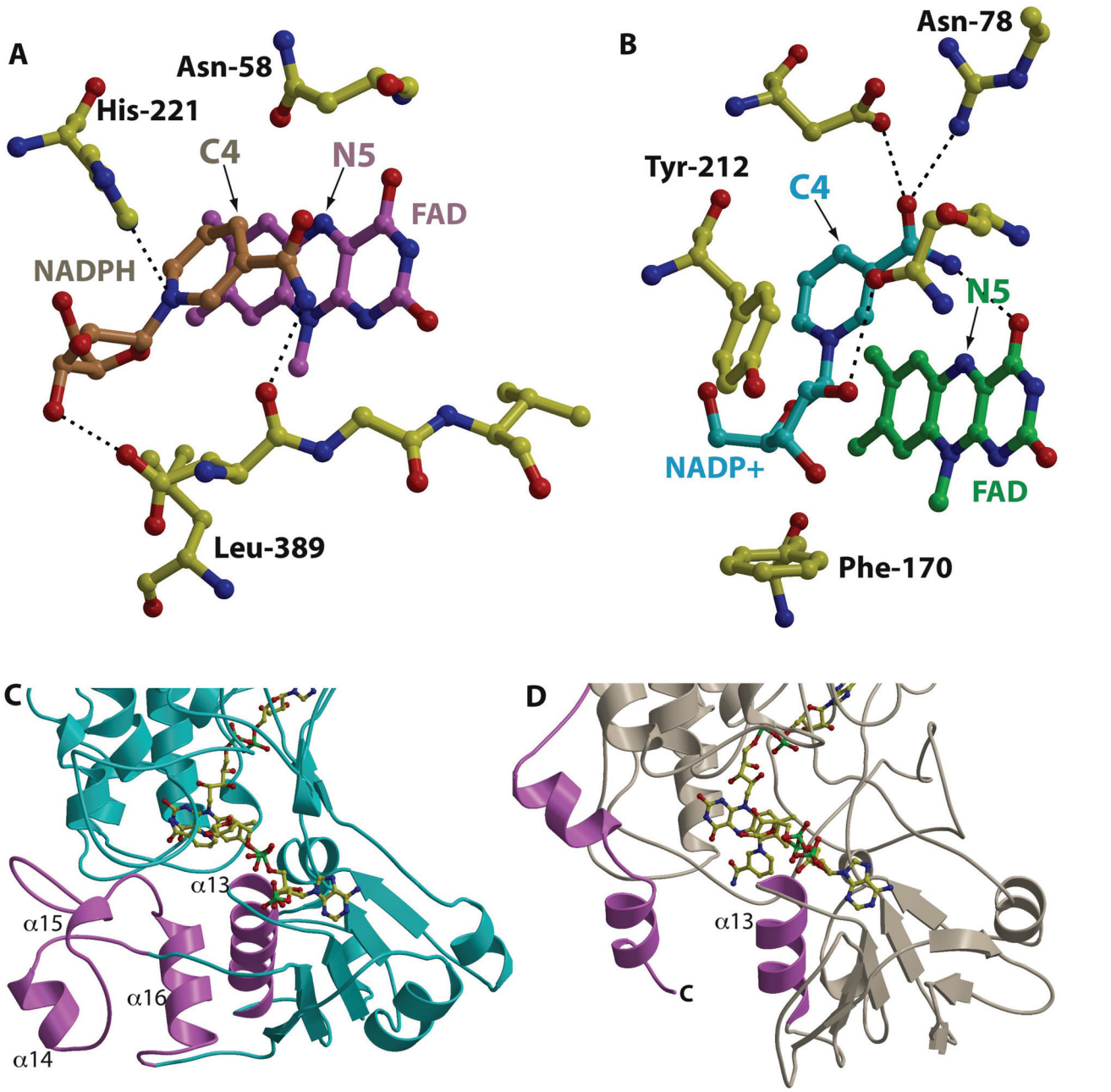


Fig. 4 A comparison of the active sites in FrbG and *S. pombe* FMO near the vicinity of the NADP(H) and FAD. (a, c) The active site of FrbG with protein residues shown in yellow, the FAD shown colored in pink and the NADPH shown colored in tan. The nicotinamide C4 is located near the N5 of the isoalloxazine at a distance and orientation that would facilitate hydride transfer. A number of secondary structure elements (shown in pink) support the position of the NADPH near the flavin. (b, d) In contrast, within the active site of *S. pombe* FMO, the nicotinamide (in blue) is situated too far away from the FAD (in green) for productive hydride transfer. The secondary structural elements (in pink) that could orient the $NADP^+$ are situated away from the active site. A comparison of the two structures may possibly reflect the necessary conformational change that occurs during the catalytic cycle.

stabilization of the C4α-hydroperoxyflavin by shielding the active site and providing a suitable hydrogen-bonding environment for the intermediate. Our FrbG-FAD-NADPH complex structure reveals close stacking of the nicotinamide ring with the isoalloxazine ring, with C4 of the nicotinamide 3.74 Å from N5 of the isoalloxazine ring (Fig. 4a and b). Attempts to soak crystals of FrbG with $NADP^+$ resulted in immediate cracking, consistent with the prevalent hypothesis that similar enzymes undergo a structural rearrangement following the flavin reduction by NADPH, to form the inactive conformation that has previously been observed crystallographically.

Domain movements have been invoked as necessary for the catalytic cycles of FMOs and structurally related proteins, such as phenylacetone monooxygenase.[8] Similar domain move-

ments have been observed in crystallographic studies of the topologically related thioredoxin reductase, in which the NADPH and FAD binding domains rotate by 67° in order to achieve a conformation that is compatible with hydride transfer.[25] A comparison of the structures of FrbG ("pre hydride transfer") with those of FMOs from *S. pombe* and *Methylophaga* sp. SK1 ("post hydride transfer") suggests some of the structural changes that may accompany the flavin reduction (Fig. 4d and e). The topology of FrbG is characterized by three helical insertions (α14–α16) between the fourth and fifth β-strand of the smaller nucleotide-binding domain. These helices are situated at similar locations as seen for the two carboxy-terminal helices in the structures of other FMOs, and are suggested to undergo a conformational movement in response to hydride transfer. The high thermal (B) factors observed for the *Methylophaga* sp. SK1 FMO in this region are thought to indicate conformational flexibility that may facilitate structural reorganization.[6] Based on our structural comparison, the relative orientations of these helices are representative of the reorganization that accompanies hydride transfer. In the FrbG structure, helix α13 is situated near the NADPH in order to orient His-221 for stabilization of the nicotinamide. The analogous helix in both from *Schizosaccharomyces pombe* and *Methylophaga* sp. SK1 FMOs orient away from the nucleotide,[6,7] and this shift may reflect another reorganization that occurs following hydride transfer.

The biochemical analysis and crystal structure described here led us to propose a reaction mechanism for the *N*-hydroxylation catalyzed by FrbG (Fig. 5). Similar to known *N*-hydroxlases, the cofactor NADPH first enters the enzyme and reduces FAD to form $FADH_2$. In the absence of substrate, the C4α-hydroperoxyflavin intermediate can typically now be formed by the addition of molecular oxygen. Alternatively, the UV-vis spectroscopic analysis indicates slow formation of the intermediate and susceptibility towards NADPH oxidase activity, which may be consistent with formation of the C4α-hydroperoxyflavin intermediate. Our crystallographic structure of FrbG-NADPH-FAD complex demonstrates the existence of this stable "pre hydride transfer" intermediate, which is also consistent with our inability to obtain a fully reduced FrbG.

The lack of a substrate-binding pocket within the active site in the FrbG-NADPH-FAD complex suggests that substrate and cofactor occupy overlapping binding sites, consistent with observations in the *S. pombe* FMO, where the substrate methimazole occupies the same binding site as $NADP^+$.[7] Inhibition of the initial velocity was monitored at increasing concentrations of the substrate (Fig. S6†). These data suggest that the CMP-5′-3APn (**4**) substrate and NADPH cofactor compete for the same binding pocket. Prior structural studies on *S. pombe* FMO also show that the enzyme remains in the "post hydride transfer" conformation as the $NADP^+$ is displaced from the binding site, suggesting that cofactor displacement must be

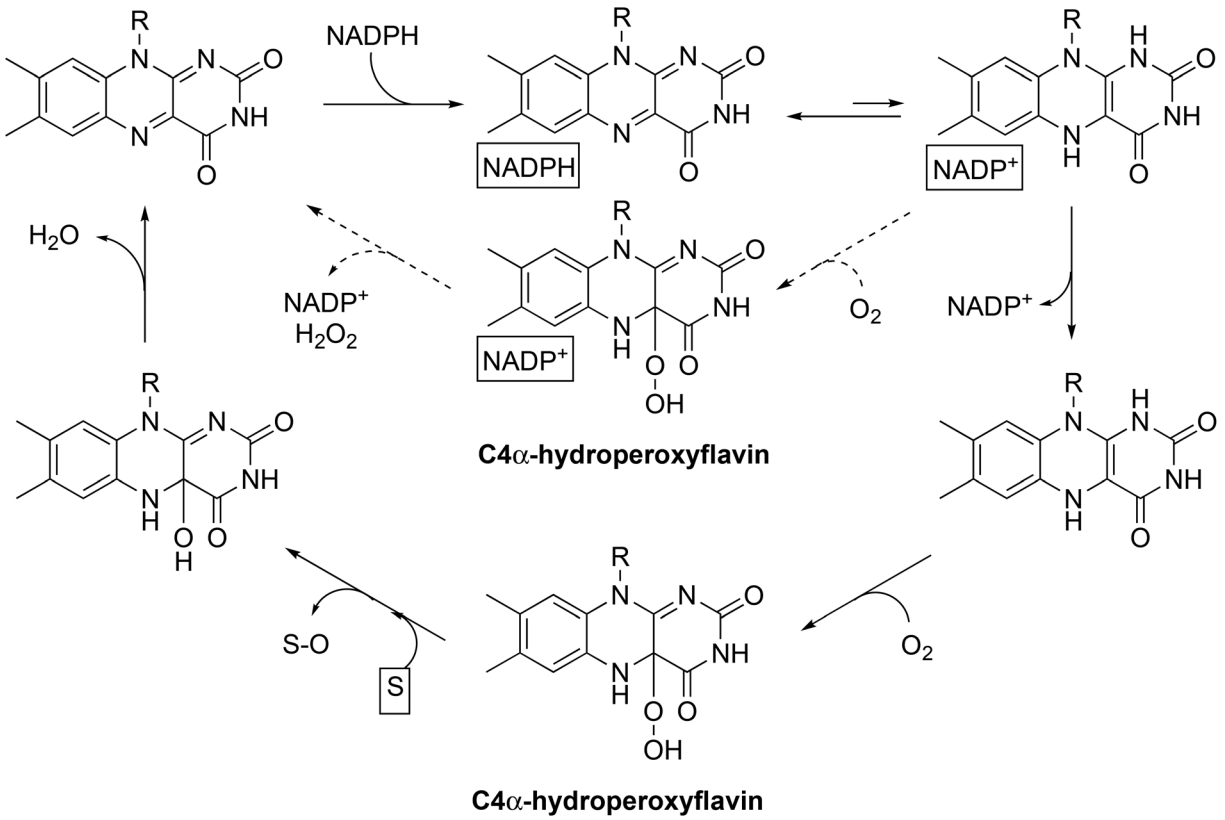


Fig. 5 Proposed catalytic mechanism of FrbG. As drawn, acceptance of molecular oxygen to form the C4α-hydroperoxyflavin intermediate occurs prior to binding of the substrate. However, it is also formally possible that C4α-hydroperoxyflavin formation occurs after substrate is bound.

fast enough that the C4α-hydroperoxyflavin does not decay. Likewise, in FrbG, following hydride ion transfer to the flavin, $NADP^+$ is displaced from the active site in order to accommodate binding of the CMP conjugated 3-aminopropylphosphonate substrate, with the C4α-hydroperoxyflavin then rapidly transferring bound molecular oxygen onto the substrate. Following the *N*-hydroxylation of the substrate, a water molecule is released and $FADH_2$ is oxidized to FAD in preparation for another catalytic cycle. CMP recognition by the NADPH/substrate binding site of FrbG contributes to tight substrate specificity in that the substrate would be of the appropriate length to access the C4α-hydroperoxyflavin intermediate within the channel for catalysis to occur.

Following FrbG-catalyzed hydroxylation, a compound is observed by LC-MS with the correct *m*/*z* to be the nitroso derivative of the hydroxylamine product. This compound could arise *via* a second FrbG-catalyzed hydroxylation cycle to the dihydroxy-intermediate, followed by a spontaneous dehydration to the nitroso. However, rapid turnover of the hydroxylamine product to the nitroso derivative ensued even following removal of FrbG, and was further observed for the chemically synthesized hydroxylamine following exposure to air (Fig. S7†). As a result, this second oxidation step is most likely non-enzymatic, as has been observed for other hydroxylamine-containing compounds.[26]

Conclusion

FrbG is a novel flavin dependent monooxygenase that catalyzes the synthesis of the CMP-*N*-hydroxy-3-aminopropylphosphonate intermediate during FR900098 biosynthesis. The crystal structure of FrbG-NADPH-FAD complex reveals a two domain architecture with each domain harboring nucleotide binding motifs to anchor the prosthetic group (FAD) and stabilize the cofactor (NADPH) required for catalysis. Prior studies of other FMOs in the "post hydride transfer" conformation suggest a "moonlighting" role for the cofactor, where NADPH promotes FAD reduction and then stabilizes the C4α-hydroperoxyflavin intermediate. A comparison of these structures with that of FrbG reveals structural reorganizations that likely occur following flavin reduction. Our current studies add to what little is known about the biosynthesis of phosphonates and assist in our efforts of increasing the efficiency, improving the efficacy, and enhancing the production of FR-900098, an antibiotic that has been clinically proven to be effective against malaria. This work has the potential of expanding the therapeutic applications of FR-900098 and its derivatives beyond parasitic and bacterial infections.

Materials and methods

Cloning, expression and purification

Full-length FrbG was cloned into the pET28a(+) vector and transformed into *E. coli* BL21(DE3) competent cells. Transformed cells were cultured overnight at 37 °C in Luria broth medium (BD Biosciences, San Jose, CA) with the appropriate antibiotic (50 μg ml^{-1} of kanamycin). The cultures were diluted 1 : 100 with the same medium and grown to an OD_{600} ~ 0.8. Protein overexpression was induced with a final concentration of 1 mM isopropyl-β-D-thiogalactoside and cultures were incubated at 18 °C with agitation for 18 hours. After centrifugation at 4000 rpm for 30 minutes the supernatant was discarded and the pellet was resuspended in 10 mM PBS, pH 8.0. Samples were lysed *via* the freeze/thaw and French Press (3000–4000 psi) methods and centrifuged at 15000 rpm for 1 hour. The soluble fraction was collected and subjected to affinity chromatography with buffer containing 20 mM Tris, pH 8.0 and 500 mM NaCl in the absence and presence of 250 mM imidazole, pH 8.0. SDS-PAGE gel electrophoresis was used to assess the presence of FrbG in the collected fractions. Cleavage and removal of the 6xHis-tag from FrbG was performed by the addition of thrombin (1 unit per ml) and $CaCl_2$ (1 mM) as the collected fractions dialyzed into 20 mM Tris, pH 8.0 and 100 mM KCl. The collected fractions were further purified by cation exchange and size exclusion chromatography with buffer exchange into 20 mM HEPES, pH 7.5 and 100 mM KCl. Concentration using a 10 kDa molecular mass limit centrifugal filter device (Millipore, Billerica, MA) resulted in a protein yield of ~45 mg ml^{-1}. Selenomethionine-labeled protein was produced and purified in a similar manner with the exception of growth in minimal media and addition of amino acids and selenomethionine prior to induction.

LC-MS analysis of the FrbG reaction

After incubating purified FrbG with the substrate CMP-5′-3-aminopropylphosphonate (CMP-5′-3APn) and NADPH, conversion to CMP-5′-*N*-hydroxy-3-aminopropylphosphonate (CMP-5′-H3APn) was monitored by HPLC. Product formation was further confirmed by LC-MS, where the new peak displayed the expected fragmentation (*m*/*z* 459 → 322) of the hydroxylamine.

UV/Vis absorbance spectrum analysis

Using a Cary 300 Bio UV-vis spectrophotometer (Varian, Cary, NC), yellow, purified FrbG (20 μM) in solution was analyzed and compared with the standard spectra of the buffer, 20 mM HEPES, pH 7.5, and 100 mM KCl, within which FrbG was stored.

Crystallization

Purified and concentrated (5–10 mg ml^{-1}) full-length FrbG was crystallized using the hanging drop vapor diffusion method. Crystals grew at 8 °C in 3 μl drops in 0.2 M KCl, 0.1 M magnesium acetate tetrahydrate, 0.05 M sodium cacodylate trihydrate, pH 6.5 and 10% (w/v) polyethylene glycol 8000 against a reservoir of the same composition. FrbG crystals took ~14–30 days to form. Selenomethionine incorporated FrbG crystals were more reproducible than native FrbG crystals in the same crystallization condition.

Structure determination of enzyme-FAD-NADPH complex

FrbG native and selenomethionine crystals were cryoprotected with 1-minute soaks at 4 °C in mother liquor containing 30% glycerol and immediately flash frozen with liquid nitrogen. Selenomethionine-incorporated amplitudes were collected, under standard cryogenic conditions, at a high-flux insertion device synchrotron beam line (Argonne National Laboratories, sector 21D, F, G), to a limiting resolution of 1.6 Å and the corresponding data processed using HKL2000.[27] As the anomalous signal from this data set was small owing to the low symmetry of the crystals, another 1.8 Å resolution data set was collected using the inverse-beam method for crystallographic phasing. Heavy-atom sites were located using HySS[28] and imported into SHARP[29] for maximum likelihood refinement. An electron density map calculated from the resultant phases was of excellent quality, permitting automated and manual fitting of the entire polypeptide chain using XtalView.[30] Iterative cycles of model building followed by round of refinement using REFMAC5 [31] further improved the model. Water molecules were added when the free *R* factors fell below 30% and cofactors and prosthetic groups were only added during the very late stages of refinement. The stereochemistry of the models was routinely monitored throughout the course of refinement using PROCHECK.[32] Crystal parameters, data collection parameters and refinement statistics for each of the structures are summarized in Table 2.

Sequence and structural comparisons

A preliminary search for sequence and structural homology was performed using Basic Local Alignment Search Tool (BLAST) on the National Center for Biotechnology Information (NCBI) Web server[33] and the Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank.[34] The DALI server[35] was used to observe protein fold similarity between FrbG and other known structures.

Kinetic analysis

Kinetic assays were performed using a Varian Cary 100 Bio UV-visible spectrophotometer (Varian) in triplicate at 30 °C. A 100 μL mixture containing 1 μM FrbG and substrate (concentrations varied from 0–2 mM) in 50 mM HEPES buffer, pH 7.5 was incubated for 5 minutes. The reaction was initiated by addition of 200 μM NADPH and then monitored by the decrease in absorbance at 340 nm or by product formation on LC-MS. The C4α-hydroperoxyintermediate was generated by addition of 7 μM NADPH to FrbG in aerated buffer solution (7 μM in 50 mM HEPES, pH 7.5) at 4 °C.[6]

Mutational analysis

A megaprimer PCR method was used to generate the N58A and H221A site specific mutants of FrbG. Briefly, primers containing the designed mutations were used in the first PCR to create the megaprimers. The megaprimers were then used in a second PCR to obtain the full-length mutants. Expression and purification of the two mutants followed the protocol listed above. A modified protocol to the kinetic analysis was used. A 100 μL mixture containing 5 μM wild type or mutant FrbG and 500 μM CMP-5′-3APn (not included for NADPH oxidase activity) in 50 mM HEPES buffer was incubated for 5 minutes. The reaction was initiated by addition of 200 μM NADPH and then monitored by the decrease in absorbance at 340 nm.

Analysis of non-enzymatic oxidation of the hydroxylamine product

Loss of the hydroxylamine product *via* non-enzymatic oxidation to the nitroso derivative was observed by two methods. First, the FrbG enzyme was removed from reaction mixtures (as described above) *via* diafiltration in a 10 kDa cutoff filter. The resulting mixtures were then incubated at 30 °C, with samples collected and analyzed by LC-MS for the presence of the hydroxylamine product. Second, the nitroso derivative was purified by HPLC fractionation and anaerobically reduced back to the hydroxylamine using SmI_2 following the method of Kende and Mendoza.[36] The hydroxylamine product of this reaction was analyzed by LC-MS both before and after 30 minutes of aerobic exposure.

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

We thank Joseph Brunzelle and Keith Brister for facilitating data collection at the Life Sciences Collaborative Access Team (Sector 21) at Argonne National Labs. This work was supported by a program grant from the National Institute of General Medical Sciences GM077596 (H. Z. and S. K. N.). M. A. D. acknowledges support from the US National Institutes of Health under Ruth L. Kirschstein National Research Award 5 T32 GM070421 from the National Institute of General Medical Sciences. The funding sources had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1 W. W. Metcalf and W. A. van der Donk, Biosynthesis of phosphonic and phosphinic acid natural products, *Annu. Rev. Biochem.*, 2009, **78**, 65–94.
2 A. C. Eliot, B. M. Griffin, P. M. Thomas, T. W. Johannes, N. L. Kelleher, *et al.*, Cloning, expression, and biochemical characterization of Streptomyces rubellomurinus genes required for biosynthesis of antimalarial compound FR900098, *Chem. Biol.*, 2008, **15**, 765–770.
3 T. W. Johannes, M. A. DeSieno, B. M. Griffin, P. M. Thomas, N. L. Kelleher, *et al.*, Deciphering the late biosynthetic steps of antimalarial compound FR-900098, *Chem. Biol.*, 2010, **17**, 57–64.

4 J. R. Cashman, Some distinctions between flavin-containing and cytochrome P450 monooxygenases, *Biochem. Biophys. Res. Commun.*, 2005, **338**, 599–604.
5 D. M. Ziegler, Recent studies on the structure and function of multisubstrate flavin-containing monooxygenases, *Annu. Rev. Pharmacol. Toxicol.*, 1993, **33**, 179–199.
6 A. Alfieri, E. Malito, R. Orru, M. W. Fraaije and A. Mattevi, Revealing the moonlighting role of NADP in the structure of a flavin-containing monooxygenase, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 6572–6577.
7 S. Eswaramoorthy, J. B. Bonanno, S. K. Burley and S. Swaminathan, Mechanism of action of a flavin-containing monooxygenase, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 9832–9837.
8 E. Malito, A. Alfieri, M. W. Fraaije and A. Mattevi, Crystal structure of a Baeyer-Villiger monooxygenase, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 13157–13162.
9 N. B. Beaty and D. P. Ballou, The oxidative half-reaction of liver microsomal FAD-containing monooxygenase, *J. Biol. Chem.*, 1981, **256**, 4619–4625.
10 L. L. Poulsen and D. M. Ziegler, The liver microsomal FAD-containing monooxygenase. Spectral characterization and kinetic studies, *J. Biol. Chem.*, 1979, **254**, 6449–6455.
11 N. B. Beaty and D. P. Ballou, Transient kinetic study of liver microsomal FAD-containing monooxygenase, *J. Biol. Chem.*, 1980, **255**, 3817–3819.
12 C. Kemal and T. C. Bruice, Simple synthesis of a 4a-hydroperoxy adduct of a 1,5-dihydroflavine: preliminary studies of a model for bacterial luciferase, *Proc. Natl. Acad. Sci. U. S. A.*, 1976, **73**, 995–999.
13 L. L. Poulsen and D. M. Ziegler, Multisubstrate flavin-containing monooxygenases: applications of mechanism to specificity, *Chem.-Biol. Interact.*, 1995, **96**, 57–73.
14 J. K. Suh, L. L. Poulsen, D. M. Ziegler and J. D. Robertus, Molecular cloning and kinetic characterization of a flavin-containing monooxygenase from Saccharomyces cerevisiae, *Arch. Biochem. Biophys.*, 1996, **336**, 268–274.
15 J. R. Cashman, Structural and catalytic properties of the mammalian flavin-containing monooxygenase, *Chem. Res. Toxicol.*, 1995, **8**, 166–181.
16 T. N. Gustafsson, T. Sandalova, J. Lu, A. Holmgren and G. Schneider, High-resolution structures of oxidized and reduced thioredoxin reductase from Helicobacter pylori, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2007, **63**, 833–843.
17 A. Karlsson, J. V. Parales, R. E. Parales, D. T. Gibson, H. Eklund, *et al.*, Crystal structure of naphthalene dioxygenase: side-on binding of dioxygen to iron, *Science*, 2003, **299**, 1039–1042.
18 I. Schlichting, J. Berendzen, K. Chu, A. M. Stock, S. A. Maves, *et al.*, The catalytic pathway of cytochrome p450cam at atomic resolution, *Science*, 2000, **287**, 1615–1622.
19 V. Massey, Activation of molecular oxygen by flavins and flavoproteins, *J. Biol. Chem.*, 1994, **269**, 22459–22462.
20 N. B. Beaty and D. Ballou, *Microsomes, Drug oxidation, and chemical carcinogenesis*, Academic Press, New York, 1980, pp. 299–302.
21 D. M. Ziegler, L. L. Poulsen and M. W. Duffel, *Microsomes, Drug oxidation, and chemical carcinogenesis*, Academic Press, New York, 1980, pp. 637–645.
22 D. E. Williams, D. M. Ziegler, D. J. Nordin, S. E. Hale and B. S. Masters, Rabbit lung flavin-containing monooxygenase is immunochemically and catalytically distinct from the liver enzyme, *Biochem. Biophys. Res. Commun.*, 1984, **125**, 116–122.
23 K. C. Jones and D. P. Ballou, Reactions of the 4a-hydroperoxide of liver microsomal flavin-containing monooxygenase with nucleophilic and electrophilic substrates, *J. Biol. Chem.*, 1986, **261**, 2553–2559.
24 S. K. Krueger and D. E. Williams, Mammalian flavin-containing monooxygenases: structure/function, genetic polymorphisms and role in drug metabolism, *Pharmacol. Ther.*, 2005, **106**, 357–387.
25 B. W. Lennon, C. H. Williams Jr. and M. L. Ludwig, Twists in catalysis: alternating conformations of Escherichia coli thioredoxin reductase, *Science*, 2000, **289**, 1190–1194.
26 N. Li, V. K. Korboukh, C. Krebs and J. M. Bollinger Jr., Four-electron oxidation of p-hydroxylaminobenzoate to p-nitrobenzoate by a peroxodiferric complex in AurF from Streptomyces thioluteus, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 15722–15727.
27 Z. Otwinowski, D. Borek, W. Majewski and W. Minor, Multiparametric scaling of diffraction intensities, *Acta Crystallogr., Sect. A: Found. Crystallogr.*, 2003, **59**, 228–234.
28 R. W. Grosse-Kunstleve and P. D. Adams, Substructure search procedures for macromolecular structures, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2003, **59**, 1966–1973.
29 G. Bricogne, C. Vonrhein, C. Flensburg, M. Schiltz and W. Paciorek, Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2003, **59**, 2023–2030.
30 D. E. McRee, XtalView/Xfit–A versatile program for manipulating atomic coordinates and electron density, *J. Struct. Biol.*, 1999, **125**, 156–165.
31 G. N. Murshudov, A. A. Vagin and E. J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 1997, **53**, 240–255.
32 R. A. Laskowski, J. A. Rullmannn, M. W. MacArthur, R. Kaptein and J. M. Thornton, AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, *J. Biomol. NMR*, 1996, **8**, 477–486.
33 S. F. Altschul, T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, *et al.*, Gapped BLAST and PSI-BLAST: a new gene-

ration of protein database search programs, *Nucleic Acids Res.*, 1997, **25**, 3389–3402.
34 H. M. Berman, T. Battistuz, T. N. Bhat, W. F. Bluhm, P. E. Bourne, *et al.*, The Protein Data Bank, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2002, **58**, 899–907.
35 L. Holm and C. Sander, Protein structure comparison by alignment of distance matrices, *J. Mol. Biol.*, 1993, **233**, 123–138.
36 A. S. Kende and J. S. Mendoza, Controlled reduction of nitroalkanes to alkyl hydroxylamines or amines by samarium diiodide, *Tetrahedron Lett.*, 1991, **32**, 1699–1702.