Archives of Biochemistry and Biophysics 544 (2014) 128–141

Contents lists available at ScienceDirect

Archives of Biochemistry and Biophysics

journal homepage: www.elsevier.com/locate/yabbi

Review

Structure, mechanism, and dynamics of UDP-galactopyranose mutase

John J. Tanner [a,*], Leonardo Boechi [b], J. Andrew McCammon [b,c], Pablo Sobrado [d,e,*]

[a] Departments of Chemistry and Biochemistry, University of Missouri, Columbia, MO, United States
[b] Departments of Chemistry and Biochemistry, University of California San Diego, La Jolla, CA, United States
[c] Howard Hughes Medical Institute, Department of Pharmacology, University of California San Diego, La Jolla, CA, United States
[d] Department of Biochemistry, Virginia Tech, Blacksburg, VA, United States
[e] Virginia Tech Center for Drug Discovery, Virginia Tech, Blacksburg, VA, United States

ARTICLE INFO

Article history:
Available online 3 October 2013

Keywords:
Flavin-dependent reaction
Galactofuranose
Non-redox reaction
Neglected diseases
Tuberculosis
Redox-switch
Conformational changes
Protein dynamics

ABSTRACT

The flavoenzyme UDP-galactopyranose mutase (UGM) is a key enzyme in galactofuranose biosynthesis. The enzyme catalyzes the 6-to-5 ring contraction of UDP-galactopyranose to UDP-galactofuranose. Galactofuranose is absent in humans yet is an essential component of bacterial and fungal cell walls and a cell surface virulence factor in protozoan parasites. Thus, inhibition of galactofuranose biosynthesis is a valid strategy for developing new antimicrobials. UGM is an excellent target in this effort because the product of the UGM reaction represents the first appearance of galactofuranose in the biosynthetic pathway. The UGM reaction is redox neutral, which is atypical for flavoenzymes, motivating intense examination of the chemical mechanism and structural features that tune the flavin for its unique role in catalysis. These studies show that the flavin functions as nucleophile, forming a flavin–sugar adduct that facilitates galactose-ring opening and contraction. The 3-dimensional fold is novel and conserved among all UGMs, however the larger eukaryotic enzymes have additional secondary structure elements that lead to significant differences in quaternary structure, substrate conformation, and conformational flexibility. Here we present a comprehensive review of UGM three-dimensional structure, provide an update on recent developments in understanding the mechanism of the enzyme, and summarize computational studies of active site flexibility.

© 2013 Elsevier Inc. All rights reserved.

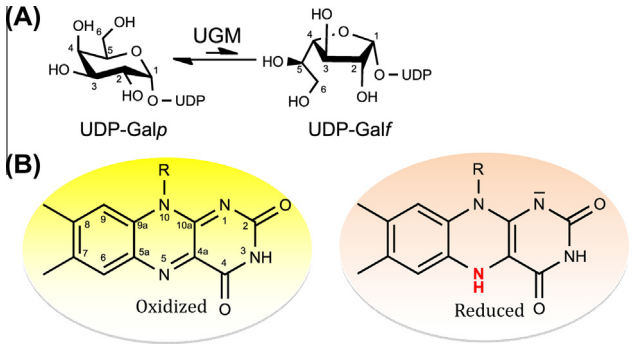
Introduction

Galactofuranose (Gal*f*) is the five membered ring form of the more common 6-membered ring sugar galactopyranose (Gal*p*). Gal*f* is thermodynamically less stable than Gal*p* because of the strain associated with the 5-membered ring. Nevertheless, it has been known for nearly a century that microorganisms produce Gal*f* in several forms. For example, galactocarolose, an extracellular β-D-(1 → 5)-linked polygalactofuranose produced by *Penicillium charlesii*, was the first polysaccharide shown to contain Gal*f* [1], motivating interest in the underlying biosynthetic pathway. Decades later, it was shown that *P. charlesii* could not use exogenous galactose to produce galactocarolose [2], leading to the discovery of a new nucleotide, UDP-Gal*f*, as the precursor in galactocarolose biosynthesis [3]. Similarly, investigations into the origins of Gal*f* in the T1 antigen of *Salmonella typhimurium* identified UDP-Gal*p* in the biosynthetic pathway and suggested the existence of an enzyme that catalyzes the 6-to-5 ring contraction of UDP-Gal*p* to UDP-Gal*f* [4,5]. Interest in Gal*f* biosynthesis stems in part from the observation that this sugar is not present in mammals; however, it is a major component of cell wall and cell surface glycoconjugates in many bacteria and eukaryotic organisms, including the human pathogens *Mycobacteria tuberculosis*, *Klebsiella pneumoniae*, *Trypanosoma cruzi*, *Leishmania major*, and *Aspergillus fumigatus* [6–10]. Targeting cell wall biosynthesis is an effective and well-established method for combating bacterial infections. Since Gal*f* is absent in humans, the enzymes involved in the biosynthesis of Gal*f* are potential drug targets. At the center of Gal*f* biosynthesis is the enzyme UDP-galactopyranose mutase (UGM).[1] The gene encoding for a UGM enzyme was first cloned from

* Corresponding authors. Address: Department of Biochemistry, University of Missouri-Columbia, Columbia, MO 65211, United States (J.J. Tanner). Address: Department of Biochemistry, Virginia Tech, Blacksburg, VA 24061, United States. Fax: +1 540 231 9070 (P. Sobrado).
E-mail addresses: tannerjj@missouri.edu (J.J. Tanner), psobrado@vt.edu (P. Sobrado).

[1] *Abbreviations used*: UGM, UDP-galactopyranose mutase; Gal*p*, galactopyranose; Gal*f*, galactofuranose; UDP-Gal*p*, UDP-galactopyranose; UDP-Gal*f*, UDP-galactofuranose; EcUGM, UDP-galactopyranose mutase from *Escherichia coli*; KpUGM, UDP-galactopyranose mutase from *Klebsiella pneumoniae*; DrUGM, UDP-galactopyranose mutase from *Deinococcus radiodurans*; AfUGM, UDP-galactopyranose mutase from *Aspergillus fumigatus*; TcUGM, UDP-galactopyranose mutase from *Trypanosoma cruzi*; bUGM, bacterial UDP-galactopyranose mutase; eUGM, eukaryotic UDP-galactopyranose mutase; RMSD, root mean square deviation; PIX, positional isotope effects; PDB, Protein Data Bank; MD, molecular dynamics.

0003-9861/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.abb.2013.09.017



Scheme 1. Reaction catalyzed by UGM (A) and structures of the oxidized and reduced flavin cofactor (B).

Escherichia coli in 1996 and given the name *glf* [11], paving the way for detailed structure–function studies that have continued to this day and are the subject of this review.

UGM is flavoenzyme that catalyzes the interconversion of UDP-Gal*p* and UDP-Gal*f* (Scheme 1A). The equilibrium of the UGM-catalyzed reaction favors UDP-Gal*p* by the ratio of 11:1 because of the aforementioned ring strain associated with galactofuranose [11]. Following the cloning of the UGM gene from *E. coli* [11], UGMs from other bacteria, fungi, and parasites have been identified [6,12–14]. Deletion of the gene encoding for UGM in *M. tuberculosis* demonstrated that this enzyme is essential for growth, whereas in *A. fumigatus* and *L. major*, UGM is an important virulence factor [15–17]. Validation of UGM as a drug target has prompted extensive structural and mechanistic studies leading to the elucidation of the chemical mechanism. UGMs have a novel 3-dimensional structure that tunes the flavin cofactor to function as a nucleophile and a scaffold in catalysis. Here, we review the chemical mechanism of UGMs and the structural changes that are required for activity. These unique mechanistic features can be exploited for the development of specific drugs against several bacterial and eukaryotic human pathogens.

Chemical mechanisms of UGM

Catalytic mechanism of UGM, a noncanonical flavoenzyme

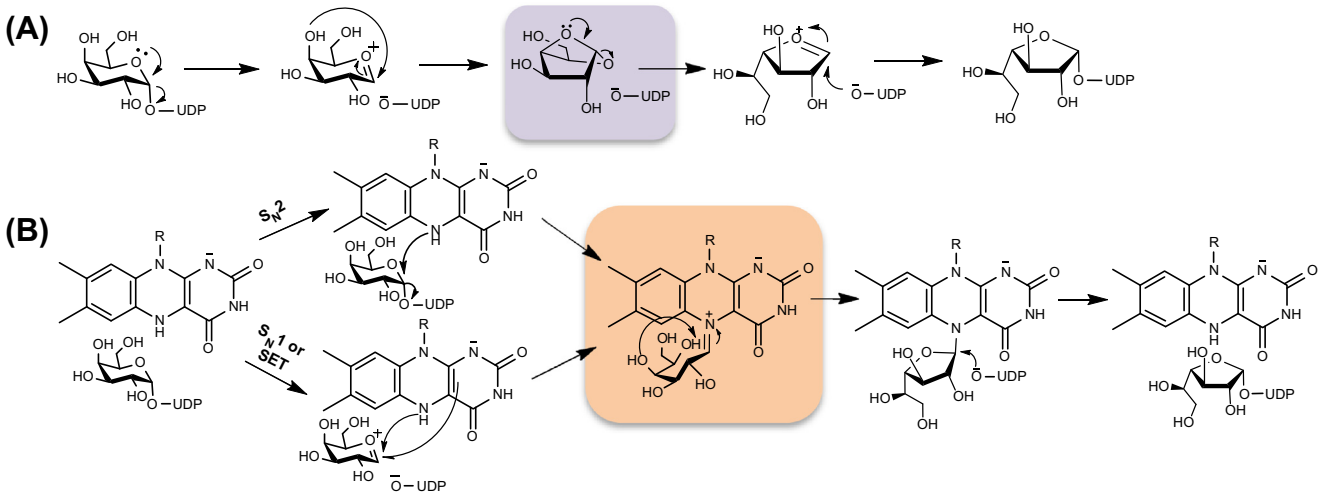
Nassau et al. discovered that *E. coli* UGM is a flavoenzyme [11], and indeed all UGMs characterized to date contain flavin adenine dinucleotide (FAD). Flavoenzymes typically catalyze oxidation–reduction reactions with the flavin serving as the redox center, and thus, the role of FAD in the redox neutral UGM reaction was enigmatic (Scheme 1A). Several mechanistic routes for the interconversion of UDP-Gal*p* to UDP-Gal*f* were initially tested in prokaryotic UGMs using a battery of chemical probes. It was shown that the enzyme was active with UDP-2-F-Gal*p* and UDP-3-F-Gal*f*, eliminating the possibility of oxidation of Gal*p* at the 2-OH or 3-OH moiety [18,19]. Further characterization of the recombinant *E. coli* protein showed that the enzyme was active when the flavin was in the oxidized form, but significantly more active when the enzyme was chemically reduced with dithionite (Scheme 1B)[20]. Subsequent studies showed that only the reduced enzyme exhibits catalytic activity, and the spurious activity attributed to the oxidized enzyme in fact originated from a subpopulation of reduced protein that had persisted in the enzyme preparation [21]. The presence of reduced UGM during purification was clearly established in the UGM from *A. fumigatus*, as the recombinant protein purified under aerobic conditions remarkably stabilizes 50% of the FAD in the reduced form [22].

Blanchard's group demonstrated that the anomeric bond was broken during catalysis using positional isotope effects (PIX) [23], which was later supported by work from Liu's group [18]. Cleavage of the anomeric bond suggested a number of possible mechanisms for UGMs (Scheme 2). One mechanism predicts the formation of 1,4-anhydrogalactopyranose (Scheme 2A) [18,23]. However, activity was not detected when reduced UGM was incubated with 1,4-anhydrogalactopyranose in the presence of UDP, eliminating this species from consideration as an intermediate in the UGM reaction [24]. A mechanism involving a single-electron transfer step during catalysis was supported by potentiometric studies that showed that the flavin semiquinone was stabilized in the presence of substrate [25]. In addition, replacement of the FAD with 5-deaza-FAD, a flavin analog restricted to a net 2-electron process, resulted in inactive UGM. The lack of activity of UGM reconstituted with 5-deaza-FAD was initially interpreted as supporting an electron transfer step in the catalytic cycle (Scheme 2B)[26].

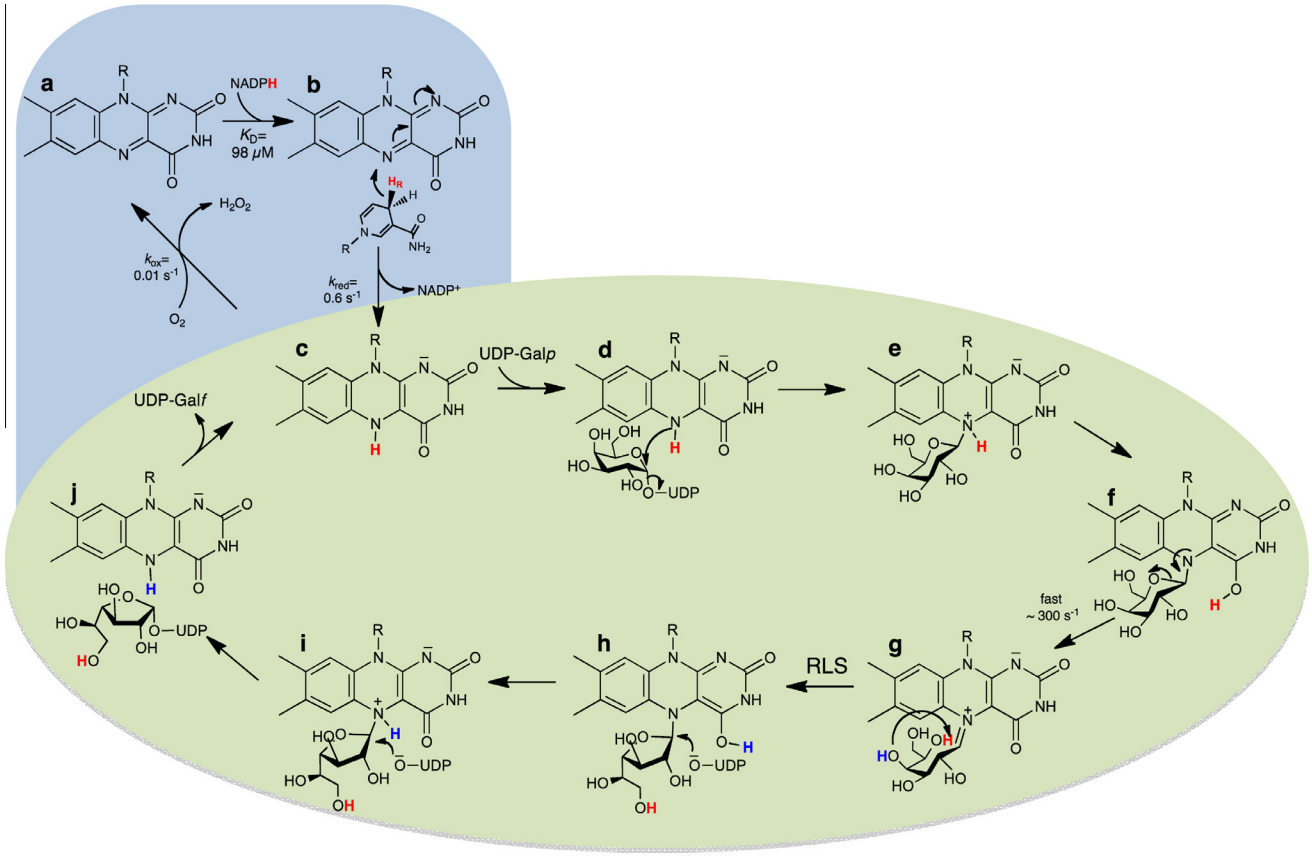
A major breakthrough in our understanding of the mechanism of action of UGM came when Kiessling's group isolated an FAD-galactose covalent intermediate [27]. The covalent intermediate formed between the $N5_{FAD}$ and $C1_{Gal\,p}$ was proposed and later validated by isolation and characterization by mass spectrometry and NMR in bacterial UGM (bUGM) and by UV/vis spectrophotometry and mass spectrometry in eukaryotic UGM (eUGM) [28–30]. The flavin–sugar adduct is important in ring opening and activation of the $C1_{Gal\,p}$ [27]. In this process, the flavin also functions as a molecular scaffold, providing the structural constraints required for ring contraction (Scheme 2B) [31]. Formation of this intermediate made mechanistic sense and was consistent with the breaking of the anomeric bond determined from PIX studies and the lack of activity observed with the enzyme reconstituted with 5-deaza-FAD.

Although identification of the flavin covalent intermediate was a landmark in UGM research, the mechanistic steps leading to formation of this intermediate were not completely understood. Formation of the flavin–sugar adduct was proposed to occur by direct attack of the $N5_{FAD}$ to the $C1_{Gal\,p}$ in an S_N2-type mechanism (Scheme 2B) [27,28]. Alternatively, the intermediate could form by attack of the $N5_{FAD}$ to an oxocarbenium galactose intermediate, in an S_N1-type mechanism [32]. In addition, it was proposed that a single electron transfer step from the flavin to the oxocarbenium intermediate would lead to the formation of a flavin semiquinone and a sugar radical. In this mechanism, the flavin–sugar adduct forms by recombination of the radical pair (Scheme 2B) [26]. Recently, Liu and coworkers probed the mechanism using flavin analogs with different nucleophilicities at the $N5_{FAD}$. The kinetic linear free energy relationship resulted in a slope of $\rho = -2.45$, consistent with a direct attack of the $N5_{FAD}$ in an S_N2-type mechanism [33]. In addition, a flavin-iminium intermediate was observed during time-resolved spectroscopy of the reaction of UDP-Gal*f* with reduced *T. cruzi* UGM (TcUGM) without the formation of a flavin semiquinone intermediate, inconsistent with a single electron transfer mechanism [30]. Furthermore, the structures of the complex of UGM with UDP-Gal*p* clearly show that the $N5_{FAD}$ is the proper distance for direct attack of the $C1_{Gal\,p}$ (described in detailed in the next section).

Recent experimental studies have provided insight into the rate-limiting step. Analysis of the kinetics of eUGMs showed that formation of the flavin iminium ion is very fast ($\sim 300\ s^{-1}$) compared to the k_{cat} ($\sim 12\ s^{-1}$ for TcUGM). In addition, viscosity effects studies demonstrated that product release is not rate limiting. This led to the proposal that ring contraction is the rate-determining step (Scheme 3, g and h) [30]. In order to produce the flavin-iminium ion, the $N5_{FAD}$ must be deprotonated. Although this aspect of the reaction has not been elucidated by biochemical approaches, recent quantum mechanical calculations provided insights into this and other steps in the UGM reaction [34]. Hybrid quantum/classical calculations performed at the density functional theory



Scheme 2. Proposed chemical mechanisms for UGM. (A) Mechanism involving cleavage of the anomeric bond and formation of Gal*f* through a 1,4-anhydrogalactopyranose intermediate. (B) Proposed paths for the formation of UDP-Gal*f* via a FAD-iminium ion intermediate. The covalent flavin intermediate could form by direct attack of the flavin via an S_N2-type reaction. Other routes include an S_N1-type reaction or single electron transfer step to an oxocarbenium ion intermediate. Not shown is the mechanism involving formation of a transient sugar and flavin radical pair upon a single-electron transfer from the reduced flavin, that after recombination form the covalent intermediate.



Scheme 3. Current chemical mechanism for UGM. The activation of eUGM by NADPH is depicted in the blue region (a and b). The rate of flavin reduction (k_{red}) and the K_D value for NADPH and the rate of oxidation (k_{ox}) are those determined for TcUGM [30]. The reduced activated enzyme (c), binds to UDP-Gal*p* and a covalent flavin–galactose adduct is formed via the direct attack of the $N5_{FAD}$ to the $C1_{Galp}$. This step leads to cleavage of the anomeric bond (d). Tautomerization of the flavin permits the transfer of the $N5_{FAD}$-proton (shown in red) to the $C4_{FAD}$ = O. This proton is next transferred to the $C5_{Galp}$-O, facilitating the opening of the sugar ring and formation of the flavin iminium ion (e–g). The $C4_{FAD}$ = O is predicted to accept the proton from the $C4_{Galp}$-OH (shown in blue) during ring contraction. The final step is the direct attack of UDP to the FAD–galactofuranose adduct. The rate-limiting step (RLS) is proposed to be the ring contraction step.

level suggest that the C4 = O of the flavin accepts the proton from the $N5_{FAD}$ (Scheme 3, e and f) and donates it to the $O5_{Galp}$, facilitating ring opening and formation of the iminium ion (Scheme 3, f and g). The simulations also predict that during ring contraction, the proton from the hydroxyl group at $C4_{Galp}$ is transferred to the C4 = O of the flavin (Scheme 3, g and h) [34].

Together, the mechanistic and structural data strongly support a novel role of the flavin cofactor as a nucleophile in the reaction catalyzed by UGMs. Novel roles of flavins in noncanonical redox neutral reactions have been identified in other enzymes [31], however, UGM is the only flavoenzyme that forms a covalent intermediate by acting as a nucleophile [9,35,36].

Mechanism of enzyme activation

Although it was known that reduced UGM is the active enzyme species, detailed characterization of the flavin reduction step was lacking until our studies of eUGMs. Eukaryotic UGMs from the fungus *A. fumigatus* and from the parasites *T. cruzi*, *L. mexicana*, and *L. infantum* have been shown to utilize NADPH as a redox partner (Table 1) [30,37,38]. In addition, these enzymes can stabilize the reduced form of the flavin, even under aerobic conditions. The rate of oxidation is 200–1500-fold slower than the reduction step, thus, the enzymes can turn over several hundred times before becoming oxidized (inactivated) [30,38]. The structures of *A. fumigatus* UGM in complex with NADP(H) have recently been solved and a novel NADPH binding domain was identified [37]. Thus, at least for eUGMs, the mechanism of activation by NAD(P)H has been established.

Bacterial UGMs also can be activated by NAD(P)H; however, activation requires excess coenzyme and extended incubation times, suggesting that NAD(P)H may not be the physiological hydride donor [39]. For example, direct measurement of the rate of flavin reduction of *M. tuberculosis* UGM was shown to be at least 10,000 times slower than for eUGMs [37]. Therefore bUGMs do not effectively react with NAD(P)H for activation. The mechanism of activation of bUGM thus remains to be discovered. One possibility that we favor is that bUGMs might be activated *in vivo* by acquiring reducing equivalents from a variety of redox partners in the cell. This hypothesis is supported by reports that have shown that bacterial nitric oxide synthase functions without a specific redox partner and can acquire reducing equivalents from several "promiscuous" bacterial reductases [40].

Our current understanding of the chemical mechanism of UGMs, including the activation of eUGMs by NAD(P)H, is depicted in Scheme 3. As we will discuss below, this unique chemical mechanism is modulated by extensive protein dynamics regulated by substrate binding and the redox state of the flavin.

Crystal structures of UGMs

The three-dimensional structure of UGM has been characterized extensively using X-ray crystallography. As of this writing, 29 structures of UGM from six organisms have been deposited in the Protein Data Bank (PDB, Table 2). The current portfolio of structures includes UGMs from the prokaryotes *E. coli*, *M. tuberculosis*, *K. pneumoniae*, and *Deinococcus radiodurans* R1, the fungal pathogen *A. fumigatus*, and the protozoan parasite *T. cruzi*. Structures of UGM in both the oxidized and reduced states have been determined, as well as structures of the enzyme complexed with UDP, UDP-Gal*p*, UDP-glucose, UDP-phosphono-galactopyranose (a non-hydrolyzable C-glycosidic phosphonate), NADH, and NADPH. Also, a structure of *K. pneumoniae* UGM complexed with FMN has been deposited, although an interpretation of this structure has not been provided.

Table 1
Kinetic parameters for the reduction of eukaryotic UGMs by NAD(P)H.

UGM	k_{red}, s^{-1}	K_D (μM)	k_{red}/K_D (10^4 M^{-1} s^{-1})	Ref.
L. infantum[a]	1.30 ± 0.03	78 ± 7.8	1.6 ± 0.1	[38]
L. mexicana[a]	0.27 ± 0.003	102 ± 4	0.30 ± 0.01	[38]
T. cruzi[a]	0.60 ± 0.01	98 ± 3	0.60 ± 0.01	[30]
A. fumigatus[a]	3.0 ± 0.1	25 ± 2	12 ± 1	[37]
A. fumigatus[b]	0.172 ± 0.003	260 ± 20	0.067 ± 0.005	[37]
T. cruzi[b]	0.0850 ± 0.0006	550 ± 10	0.015 ± 0.0002	[30]

Conditions: 50 mM sodium phosphate pH 7.0, under anaerobic conditions at 15 °C.
[a] Kinetic parameters determined with NADPH.
[b] Kinetic parameters determined with NADH.

The basic UGM fold

The first UGM crystal structure was reported five years after the *E. coli glf* gene was cloned [21]. The structure of *E. coli* UGM (EcUGM) revealed a three-domain architecture that would prove to be emblematic of UGMs (Fig. 1A). The topology of UGM is rather complex because domains 1 and 3 are formed by noncontiguous residues, with domain 1 being tripartite and domain 3 bipartite.

Domain 1 binds FAD and features an abbreviated Rossmann dinucleotide-binding fold at its core. The Rossmann fold of UGM (and some other FAD-binding enzymes) lacks the final strand of the parallel 6-stranded β-sheet found in the classic Rossmann fold (see lactate dehydrogenase, for example). Nevertheless, as in other Rossmann fold enzymes, the dinucleotide adopts an extended conformation and binds at the C-terminal edge of the β-sheet with the pyrophosphate interacting with the Gly-rich loop and a conserved water molecule [41]. Domain 1 also includes two smaller subdomains, denoted 1A and 1B. Subdomain 1A interrupts the Rossmann fold at the C-terminus of β2 and consists of a two-stranded antiparallel β-sheet. Subdomain 1A is followed by subdomain 1B, which is a 3-stranded antiparallel β-sheet and associated α-helix. The number of strands in the sheet varies from two to three in different UGMs. The connector between 1A and 1B is very important for function, as it provides the conserved histidine loop as well as an aromatic residue that serves as the backstop for the substrate Gal*p* moiety. As described below, the histidine loop of eukaryotic UGMs exhibits large conformational changes when the oxidized enzyme is activated by reducing agents such as dithionite and NAD(P)H. Domain 1 also includes a C-terminal helix that is outside of the Rossmann fold. The connector between Rossmann β5 and the C-terminal helix is functionally important as it provides active site residues that cluster near the substrate Gal*p*.

Domain 2 is the only contiguous domain of UGM and is essentially a bundle of five α-helices. This domain functions in binding the uridine group of the substrate and in dimerization. Domain 2 also contains a mobile active site flap that responds to substrate binding. The flap is located on the loop between the last helix of domain 2 and the final strand of subdomain 1B (Fig. 1) and contains a key Arg residue that interacts with the substrate (described below).

Domain 3 features a twisted, 6-stranded anti-parallel β-sheet. This domain is noncontiguous because strand 2 connects to domain 2 rather than strand 3 of its own sheet. As a consequence, strands 2 and 3 of the 6-stranded sheet are separated by ~150 residues. These intervening 150 residues form several structural elements distributed throughout the fold, including all of domain 2, the final strand of subdomain 1B, and about one-half of the Rossmann fold before returning to domain 3 at β-strand 3. Strand 6 of domain 3 exits to domain 1 to complete the UGM fold.

All bUGMs exhibit essentially the same structure as EcUGM. For example, analysis with SSM [42] shows that 73–96% of the secondary structure elements of EcUGM are preserved in other bUGMs, implying conservation of secondary and tertiary structure despite substantial variation in primary structure (37–44% pairwise identity with EcUGM). Also, the root mean square deviation (RMSD) of the other bUGMs to EcUGM spans the range 1.1–1.5 Å. This range is quite low considering that the two EcUGM protomers in the crystallographic asymmetric unit align with an RMSD of 0.97 Å.

Table 2
UGM crystal structures deposited in the Protein Data Bank (PDB).[a]

UGM[b]	PDB code	Resolution	Redox state(s)[c]	Active site ligand	Release Date	Ref.
Ec	1I8T	2.40	O		2001	[21]
Mt	1V0 J	2.25	O		2005	[64]
Kp	2BI7	2.00	O		2005	[64]
Kp	2BI8	2.35	R		2005	[64]
Kp	1WAM	2.35	R		2006	[64]
Kp	3GF4	2.45	O	UMP, UDP-glucose	2009	[29]
Kp	3INR	2.30	O	UDP-Gal*p*	2009	[28]
Kp	3INT	2.51	R	UDP, UDP-Gal*p*	2009	[28]
Dr	3HDQ	2.36	O	UDP-Gal*p*	2009	[65]
Dr	3HDY	2.40	O/R	UDP-Gal*p*	2009	[65]
Dr	3HE3	2.40	O	UDP	2009	[65]
Dr	3MJ4	2.65	O	UDP, UDP-CH_2-Gal*p*	2010	[66]
Kp	3KYB	2.30	O	FMN	2010	-
Af	3UTE	2.35	O	Sulfate	2012	[43]
Af	3UTF	2.25	R		2012	[43]
Af	3UTG	2.25	R	UDP	2012	[43]
Af	3UTH	2.25	R	UDP-Gal*p*	2012	[43]
Af	3UKA	2.64	O		2012	[45]
Af	3UKF	2.50	R	UDP-Gal*p*	2012	[45]
Af	3UKH	2.30	O/R	UDP-Gal*p*	2012	[45]
Af R182K	3UKK	2.75	O/R	UDP	2012	[45]
Af	3UKL	2.63	O	UDP	2012	[45]
Af R327A	3UKP	3.10	O	UDP-Gal*p*	2012	[45]
Af R327K	3UKQ	3.15	O	UDP-Gal*p*	2012	[45]
Tc	4DSG	2.25	O	UDP	2012	[54]
Tc	4DSH	2.25	R	UDP	2012	[54]
Af	4GDC	2.75	O	NADPH	2012	[37]
Af	4GDD	2.75	O	NADH	2012	[37]
Af	4GDE	2.20	R		2012	[37]

[a] Listed in chronological order of release date.
[b] Enzyme sources abbreviated as follows: Ec, *Escherichia coli*; Kp, *Klebsiella pneumoniae*; Dr, *Deinococcus radiodurans*; Af, *Aspergillus fumigatus*; Tc, *Trypanosoma cruzi*.
[c] *Abbreivations*: O, oxidized; R, reduced; O/R, oxidized and reduced forms present in the same crystal.

Eukaryotic UGMs – elaboration of the basic UGM fold

Structures of eUGMs first appeared in the literature more than a decade after the EcUGM structure (Table 2). Part of this time lapse was likely due to the many technical challenges encountered during structure determination, including translational pseudosymmetry and/or twinning observed in crystals of *A. fumigatus* UGM (AfUGM) [43–45] and *L. major* UGM [46]. The surface mutagenesis strategy employed by Dhatwalia et al. proved to be the most effective solution to these problems [43]. The method involves modifying the protein surface using site-directed mutagenesis focused on long, charged residues, especially Lys and Glu [47]. Since the target structure is unknown, one typically uses bioinformatics methods such as the surface entropy reduction server at UCLA to guide mutagenesis [48]. Using this approach, Dhatwalia and coworkers identified a double mutant of AfUGM (K344A/K345A) that reproducibly forms crystals that diffract to high resolution and are free of crystallographic pathologies [43].

AfUGM was the first eUGM to be structurally characterized and is the prototype for its class. Two groups reported the structure in early 2012 using different crystal forms [43,45]; the two forms revealed essentially identical structures.

AfUGM exhibits a variation of the three-domain architecture of bUGMs (Fig. 1B). Familiar features include the abbreviated Rossmann fold in domain 1, ancillary subdomains 1A and 2A, a bundle of helices forming domain 2, a large, twisted anti-parallel β-sheet reminiscent of domain 3, and the complex topology. However, the polypeptide of AfUGM is about 100 residues longer than those of bUGMs, which results in additional secondary and tertiary structural elements (Fig. 1B, red).

Domain 1 has an additional 4-stranded anti-parallel β-sheet that packs against the Rossmann fold sheet. Domain 1 also has ~30 extra residues at the C-terminus, which form a U-shaped substructure that traverses 64 Å. As described below, the U-shaped C-terminus is involved in the unique tetrameric assembly of AfUGM.

Domain 2 of AfUGM shows two variations from the bUGM blueprint. First, it has an extra helix at the end of the domain (residues 188–197). This variation is important because it occurs in the region corresponding to the mobile active site flap of bUGMs. Insertion of the new helix into the basic UGM fold generates a second mobile flap that is absent in bUGMs. The two flaps flank the new helix and correspond to residues 179–187 (flap 1) and 203–209 (flap 2) in AfUGM. Flap 1 of AfUGM corresponds to the mobile flap of bUGMs, while flap 2 is unique to eUGMs. The second variation of domain 2 involves the second helix (residues 115–134), which is about seven residues longer in AfUGM and rotated by about 90°. Both the extended helix and the extra one are involved in tetramerization, as described below.

Domain 3 has two long inserts compared to bUGMs. These sections of the polypeptide form a β-strand, an α-helix, and a long coiled loop (residues 335–366). As a result, the sheet of domain 3 has 7 strands rather than 6 as in bUGMs.

The crystal structure of TcUGM (45% identical to AfUGM) has also been determined (Table 2). The protomer structure is very similar to that of AfUGM, as evidenced by an RMSD of 1.1 Å (Fig. 2A). Despite the obvious global structural similarity, the two structures have meaningful differences. First, TcUGM lacks the helix at the end of the U-shaped C-terminal extension of AfUGM (Fig. 2A). Also, the second helix of domain 2 is one turn shorter in TcUGM (Fig. 2B). As described below, these differences result in TcUGM being monomeric rather than tetrameric.

Oligomeric state and quaternary structure

Despite substantial conservation of protein fold, UGMs display a variety of oligomeric states. Bacterial UGMs form dimers and possibly decamers, whereas eUGMs form monomers and tetramers.

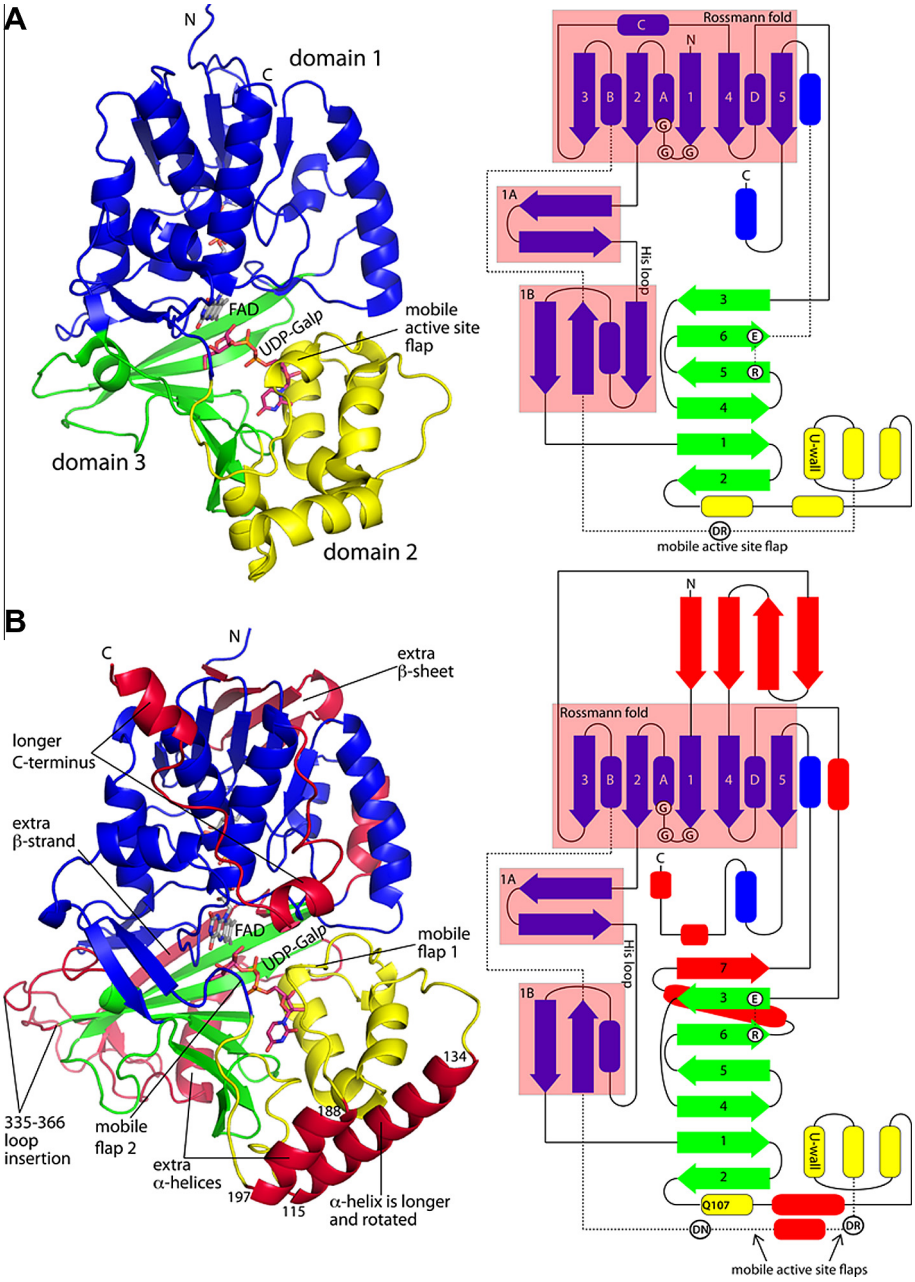


Fig. 1. The UGM fold. (A) The protomer structure (left) and topology diagram (right) of bacterial UGM, as exemplified by DrUGM. Domains 1, 2, and 3 are colored blue, yellow, and green, respectively. (B) The protomer structure (left) and topology diagram (right) of eukaryotic UGM, as demonstrated by AfUGM. Domains 1, 2, and 3 are colored blue, yellow, and green, respectively. Structural elements that are unique to eukaryotic UGMs are colored red. In the topology diagrams, the shaded boxes indicate the three subdomains of domain 1. Also, selected residues are indicated by circled letters as follows: G, Gly-rich loop of the Rossmann fold; R–E, Arg–Glu ion pair that interacts with the pyrophosphate; DR, dynamic Arg involved in substrate recognition; DN, dynamic Asn of eUGMs. This figure and others were created with PyMOL [63].

Bacterial UGMs

Dimer seems to be the prominent form of bUGMs. Dynamic light scattering analysis of eUGM suggests a dimer in solution [49]. Analysis of protein–protein interfaces in crystal lattices reveals a common dimeric assembly for UGMs from *E. coli*, *M. tuberculosis*, and *K. pneumoniae*. The observation of the same 2-body assembly in different crystal lattices of different proteins strongly suggests that this unit is formed in solution. The dimer is a semi-circular particle with the interface formed by domain 2 of one protomer packing against the β-sheet of domain 3 of another protomer (Fig. 3A).

The crystal structure of *D. radiodurans* (DrUGM) suggests a different oligomeric state. Inspection of the crystal lattice shows, surprisingly, that the semi-circular dimer is absent. However, analysis of protein–protein interfaces with PDBePISA [42] suggests that a different dimer may be formed in solution. In contrast to the semi-circular dimer, the DrUGM dimer is formed primarily via domain 1 (Fig. 3B). A notable interaction in the dimer interface is a short stretch of intermolecular β-sheet formed by β3 of the Rossmann fold (Fig. 3B, right panel). In the crystal, five such dimers assemble into a pentamer-of-dimers decamer (Fig. 3C). Analysis with PDBePISA suggests that this large assembly is also potentially stable in solution. Biophysical studies are needed to determine the solution oligomeric state of DrUGM.

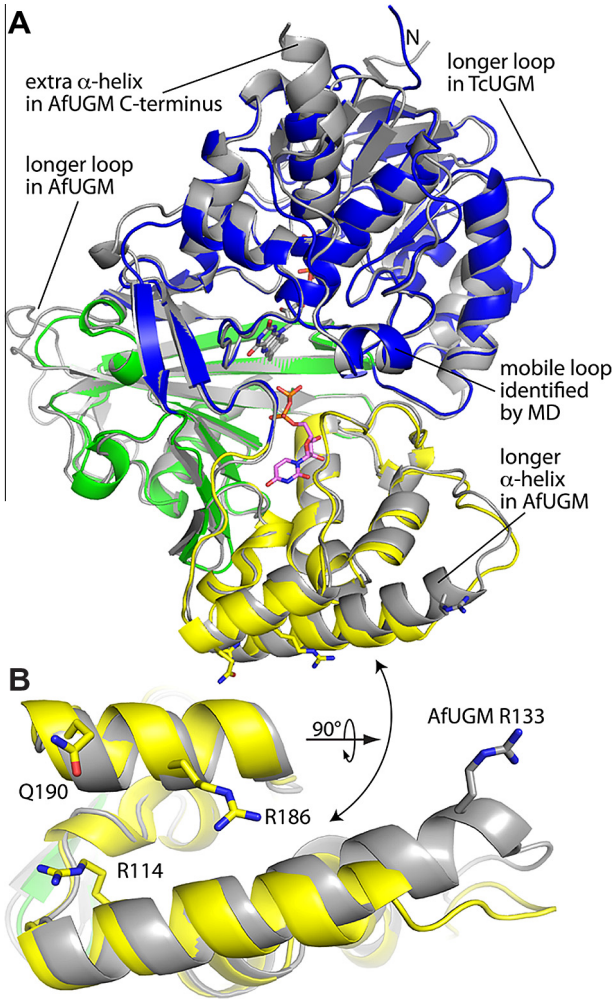


Fig. 2. Structure of TcUGM. (A) Superposition of TcUGM (blue, yellow, green) and AfUGM (gray). (B) Close-up view of two helices of domain 2 that differ in TcUGM and AfUGM. The view is rotated from panel A by ~90° around the horizontal axis.

Eukaryotic UGMs

The oligomeric state and quaternary structure of AfUGM were determined using a combination of small-angle X-ray scattering and X-ray crystallography [43]. These studies showed that AfUGM forms a dimer-of-dimers tetramer in solution (Fig. 3D). The tetramer has 222 point group symmetry and comprises three different 2-body assemblies. Interestingly, none of these assemblies resemble the bUGM dimers, and thus AfUGM exhibits a unique quaternary structure. As described previously [43], the additional structural features of AfUGM, such as the extra helix in domain 2 and the inserts in domain 3, preclude the assembly of bUGM-like dimers.

The interfaces of AfUGM have been analyzed in detail to identify structural features that stabilize the tetramer [43]. This analysis suggests that the unique tetramer results from secondary structure elements of AfUGM that are absent in bUGMs. These include the U-shaped C-terminal extension, the extra helix of domain 2, and the elongated helix of domain 2. All three substructures figure prominently in the tetramer interfaces. In particular, the extra and elongated helices of domain 2 form an intersubunit 4-helix bundle at one of the 2-fold axes (Fig. 3D, right panel). In summary, the unique variation of the UGM fold displayed by AfUGM precludes formation of the classic semicircular dimer and promotes formation of a new tetrameric assembly.

Although the solution structural properties of TcUGM have not been studied as rigorously as AfUGM, the available data suggest a monomeric protein in solution. Size exclusion chromatography data are consistent with a monomer [22]. Also, neither the AfUGM-type tetramer nor any of the previously observed UGM dimers are present in the TcUGM crystal lattice. Curiously, a disulfide-linked dimer is present in the lattice, but this species is likely an artifact of crystal packing.

The monomeric state of TcUGM can be understood in terms of the tertiary structure. As with AfUGM, the extra helix of domain 2 and the inserts in domain 3 likely prevent formation of bUGM-like dimers. Why TcUGM does not form the AfUGM-like tetramer can also be rationalized. As described above, structural differences between TcUGM and AfUGM are evident in domains 1 and 2 (Fig. 2). Because of these differences, TcUGM lacks some of the components of the tetramer interfaces, such as the C-terminal helix of domain 1 and an Arg residue (Arg 133 in AfUGM, Fig. 2B) that forms intersubunit hydrogen bonds in the AfUGM tetramer. Also, long side chains protruding from a helix of domain 2 (Fig. 2B) likely prevent formation of the intersubunit 4-helix bundle that stabilizes the AfUGM tetramer. In summary, the unique fold-level variations exhibited by TcUGM are responsible for the monomeric state. Thus, as with AfUGM, tertiary structure dictates the oligomeric state.

Substrate recognition

The structural basis of substrate recognition has been characterized. Crystal structures of bUGMs and eUGMs complexed with UDP-Gal*p* or UDP have been determined (Table 2), allowing conserved and unique features of substrate recognition to be identified.

Conserved themes of substrate recognition

Some aspects of substrate recognition appear to be shared by all UGMs. For example, the substrate consistently binds with the Gal*p* moiety next to the flavin isoalloxazine such that the $N5_{FAD}$ and $C1_{Galp}$ atoms are in close (3.3–3.5 Å) contact (Figs. 4 and 5). This aspect of the structures is consistent with the chemical mechanism showing that the flavin functions as a nucleophile that attacks the $C1_{Galp}$ atom (Scheme 3, d). Also, the general position of the substrate in relation to the protein domains is conserved, and thus the division of labor in substrate binding is consistent: domain 1 is important for positioning Gal*p* for nucleophilic attack; domain 2 provides most of the interactions with the uridine group; and domain 3 figures prominently in binding the pyrophosphate.

Common themes of substrate recognition are also evident in detailed enzyme–substrate noncovalent interactions (Table 3, Fig. 5A and B). One wall of the uridine pocket is highly conserved. The uridine wall consists of four residues belonging to the fourth helix of domain 2. These residues are conserved at the sequence level and include Trp and Asn/Thr side chains that hydrogen bond to the ribose hydroxyls and two aromatic residues (Tyr and Phe) that stack against the uracil. Although the precise orientations of the Phe and Tyr residues vary between bUGMs and eUGMs, their role in shaping the uracil pocket remains a constant (compare Fig. 5A and B).

All UGMs appear to share a common theme of pyrophosphate recognition. Two conserved aspects are evident (Fig. 5A and B). First, a pair of Tyr residues forms hydrogen bonds with the pyrophosphate. These residues are located on the C-terminus of the final strand of domain 3 and the loop following Rossmann fold β5. A conserved ion pair also stabilizes the pyrophosphate. The ion pair consists of Arg and Glu residues located on strands 5 and 6 of do-

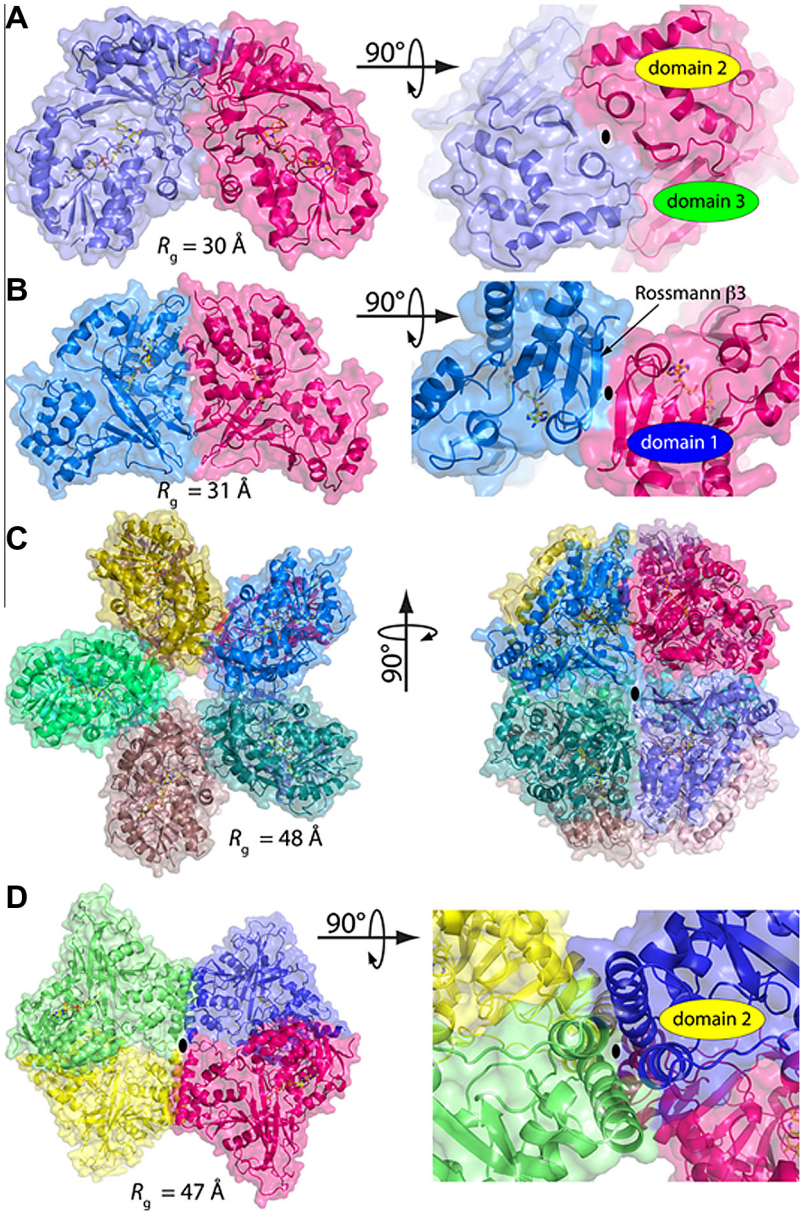


Fig. 3. UGM oligomers. (A) The typical dimer of bacterial UGM, as demonstrated by EcUGM. (B) The atypical dimer of DrUGM. (C) The DrUGM decamer. (D) The AfUGM tetramer. In all panels, the black ovals denote 2-fold axes. The radius of gyration (R_g) of each oligomer is indicated.

main 3, respectively (Fig. 1). The Arg side chain forms ionic interactions with the pyrophosphate (Fig. 5A and B).

All UGMs have a dynamic Arg that participates in substrate recognition. This key residue is located in the middle of a mobile active site flap that moves substantially when the active site transitions from the open, ligand-free conformation to the substrate-bound, closed state (Fig. 4). As described above, bUGMs have one active site flap (Fig. 4B), whereas eUGMs have two mobile flaps. The dynamic Arg of eUGMs is part of flap 1 (Fig. 4A). In both UGM classes, the dynamic Arg is solvent exposed in the open state, and interacts with the substrate in the closed state. These interactions vary somewhat in different structures. The dynamic Arg forms a hydrogen bond with the Gal*p* 2-OH in AfUGM (Fig. 5A) and DrUGM (Fig. 5B). This interaction is missing in *K. pneumoniae* UGM (KpUGM) (Fig. 4B). In bacterial UGMs, the dynamic Arg additionally interacts with the pyrophosphate (Fig. 4B), and/or the Gal*p* 3-OH (Fig. 5B). Despite these variations, the dynamic Arg is always stabilized by a conserved Asn/Asp contributed by the loop following Rossmann β5 and positioned near the pyrimidine ring of the isoalloxazine. The location of the Arg–Asn/Asp link near the Gal*p* moiety and isoalloxazine suggests that this interaction is important for locking down the closed active site conformation and positioning the sugar for catalysis.

Conserved steric interactions are also important for substrate recognition. All UGMs have an aromatic residue that contacts the O3–O4 locus of Gal*p* and appears to serve as a backstop that helps position the sugar for catalysis (Fig. 5A and B). The backstop residue is part of the conserved histidine loop and corresponds to Phe in eUGMs (Phe66 of AfUGM) and His in bUGMs (His88 of DrUGM) (Fig. 6). The backstopping His of bUGM is well outside of hydrogen bonding distance to the Gal*p* hydroxyls (3.9–4.7 Å), consistent with a steric role (Fig. 5B).

The final conserved aspect of substrate recognition involves a "third Arg" residue in the active site (Arg447 of AfUGM, Arg364 of DrUGM). This residue is located on the loop following Rossmann β5 and is positioned near the dimethylbenzene ring of the isoallox-

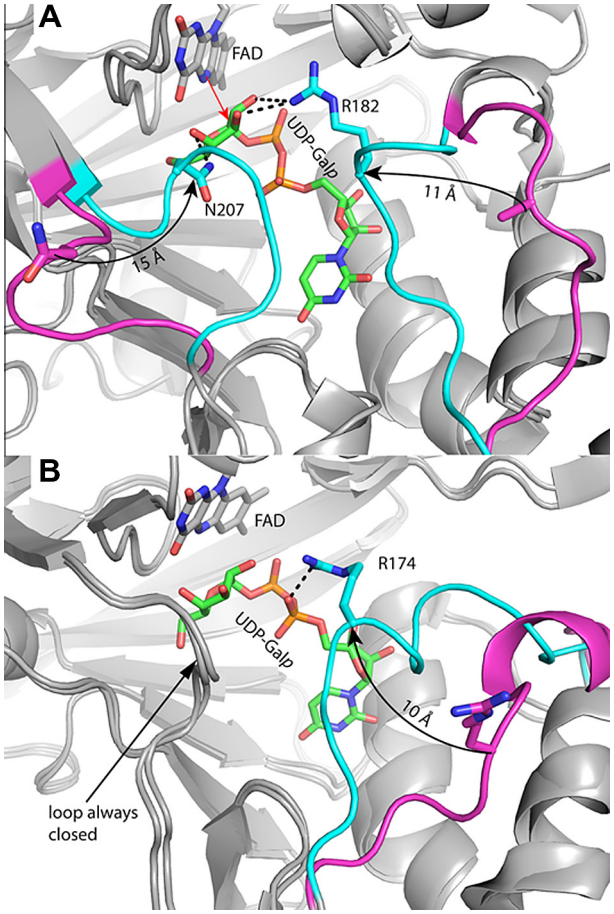


Fig. 4. The open and closed conformations of UGM. (A) Superposition of reduced AfUGM without and with bound UDP-Gal*p* (PDB codes 3UTF, 3UTH). The active site flaps are colored magenta in the open conformation and cyan in the closed conformation. Flaps 1 and 2 are on the right and left sides of the active site, respectively. (B) Superposition of the structures of reduced KpUGM without and with bound UDP-Gal*p* (PDB codes 2BI8 and 3INT). The mobile flap is colored magenta in the open conformation and cyan in the closed conformation.

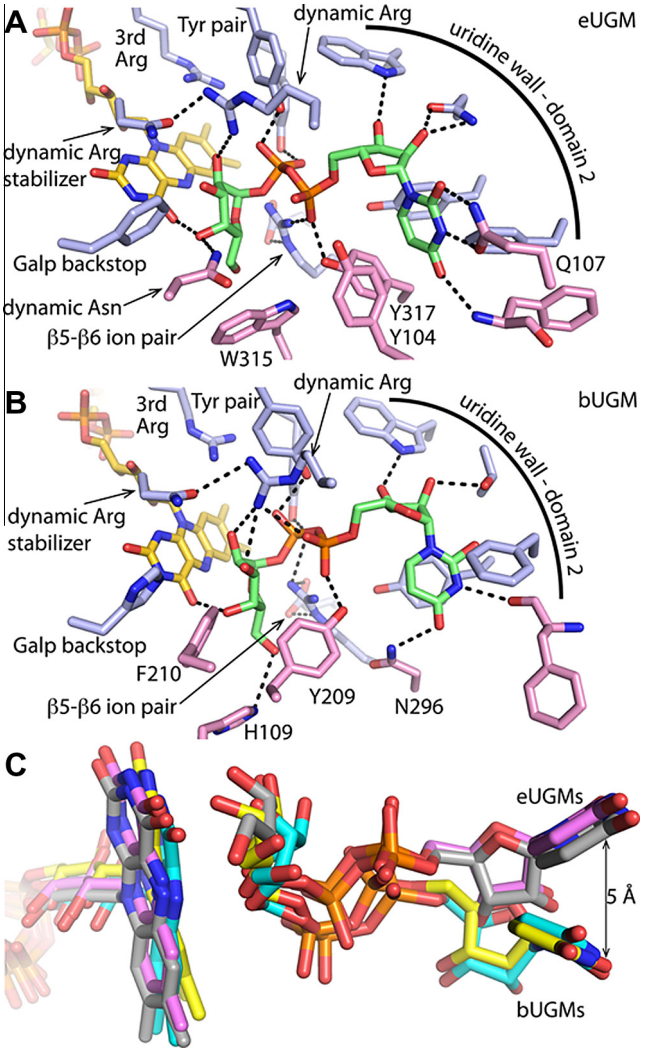


Fig. 5. Substrate recognition by UGMs. (A) Active site reduced AfUGM complexed with UDP-Gal*p*. Blue residues denote conserved interactions present in all UGMs, and pink residues indicate interactions unique to eUGMs. (B) Active site of reduced DrUGM complexed with UDP-Gal*p*. Blue residues denote conserved interactions present in all UGMs, and pink residues indicate interactions unique to bUGMs. (C) Superposition of the UDP-Gal*p* complexes of AfUGM (gray), TcUGM (pink), DrUGM (cyan), and KpUGM (yellow) showing the difference in the position of the UMP in eUGMs and bUGMs.

azine (Fig. 5A and B). The third Arg does not directly contact the substrate, but forms water-mediated hydrogen bonds to the pyrophosphate and Gal*p* in AfUGM.

Unique aspects of substrate recognition

Despite the numerous conserved features described above, eUGMs and bUGMs differ in some aspects of substrate binding. The largest of these involve the UMP moiety. The UMP bound to AfUGM or TcUGM is displaced by 3–5 Å compared to the bacterial enzymes, with the uridine exhibiting the largest shift (Fig. 5C). This variation reflects both sequence and structural differences in domain 2. In AfUGM, the uridine packs against Tyr104 and Tyr317, while forming hydrogen bonds with Gln107 (Fig. 5A). Identical interactions are observed in TcUGM complexed with UDP. These three residues are different in bUGMs. For example, Tyr317 is replaced by Asn in bUGMs (Fig. 5A, Asn296 in DrUGM), and bUGMs lack a residue analogous to Gln107 because of local protein conformational differences. Furthermore, bUGMs have nonpolar residues occupying the space corresponding to the uridine pocket of eUGMs.

The UMP phosphate interactions are also different in the two UGM classes. A Tyr residue from β4 of domain 3 interacts with the UMP phosphate eUGM (Tyr317 in eUGM, Fig. 5A), whereas a different Tyr from the linking peptide between domain 2 and subdomain 1B performs this function in bUGMs (Tyr209 in DrUGM, Fig. 5B). This difference likely contributes to the observed difference in the pyrophosphate dihedral angle.

The two classes of UGM also differ in the region near the O4–O5 locus of Gal*p*. Eukaryotic UGMs have a dynamic Asn residue and a Trp side chain in addition to the Gal*p* backstop (Fig. 5A). The dynamic Asn (Asn207 in AfUGM) is located on mobile flap 2, a region that is stationary in bUGMs. This key residue is solvent exposed in the open state and moves 15 Å to form a hydrogen bond to the O4 hydroxyl in the closed state (Fig. 4A). The Trp residue (Trp315 of AfUGM) forms a steric interaction with C6 of Gal*p* (Fig. 5A). In contrast, bUGMs have a second backstopping residue (Phe210 in DrUGM) and no hydrogen-bonding residue (Fig. 5B). These protein structural differences likely account for the different orientations of the O6 hydroxyl in the two UGM classes.

In summary, bUGMs and eUGMs share many common elements of substrate recognition. However, there are also aspects of substrate binding that are unique to each class of UGM, and the two classes differ most substantially in the region near the UMP. This

Table 3
Conserved elements of substrate recognition.

Element	Role	DrUGM	AfUGM
Uridine wall	Four residues of domain 2 that contact the uridine group	F176, Y179, T180, W184	F158, Y162, N163, W167
Tyr pair	Two Tyr resides that hydrogen bond with the pyrophosphate	Y335, Y370	Y419, Y453
Ion pair	The Arg of an Arg–Glu ion pair interacts with the pyrophosphate	R305–E325	R327–E373
Dynamic Arg	Located on a mobile active site flap; interacts with Gal*p*; stabilized by Asn or Asp	R198–N372	R182–N457
Gal*p* backstop	Contacts the O3–O4 locus of Gal*p*; last residue of the histidine loop	H88	F66
3rd Arg	One of three Arg residues in the active site; near dimethylbenzene ring of isoalloxazine; does not directly contact the substrate	R364	R447

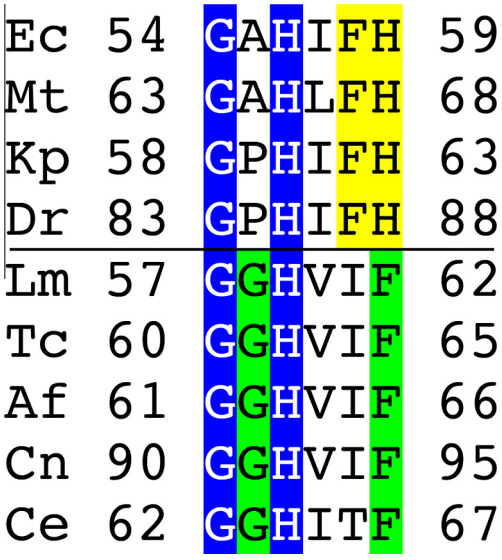


Fig. 6. Alignment of the histidine loops of UGMs. The line separates bacterial (top) and eukaryotic (bottom) UGM sequences. *Abbreviations*: Ec, *E. coli*; Mt, *M. tuberculosis*; Kp, *K. pneumoniae*; Dr, *D. radiodurans*; Lm, *L. major*; Tc, *T. cruzi*; Af, *A. fumigatus*; Cn, *Cryptococcus neoformans*; Ce, *Caenorhabditis elegans*.

analysis suggests that inhibitors mimicking UMP could potentially exhibit broad specificity against bacterial or eukaryotic UGMs, but not both.

The NAD(P)H site of eukaryotic AfUGM

The crystal structures of oxidized AfUGM complexed with NADPH or NADH were recently reported [37]. NAD(P)H interacts exclusively with domains 1 and 3. One face of NAD(P)H packs into a groove at the junction of these domains, while the other face is exposed to the vacant substrate-binding cavity (Fig. 7A). NAD(P)H adopts a compact conformation in which the adenine forms a hydrogen bond with a ribose hydroxyl (Fig. 7B).

The nicotinamide is near the FAD isoalloxazine, and the ADP moiety extends toward mobile active site flap 2 (Fig. 7A). Mobile flap 1 is in the open conformation, while flap 2 is disordered, indicating that NAD(P)H binding does not induce closure of the active site. Binding of the nicotinamide near the isoalloxazine is consistent with the known chemistry of hydride transfer from NAD(P)H to FAD. The nicotinamide, however, is not optimally aligned for hydride transfer (Fig. 7B), which could result from an adventitious sulfate ion in the active site or suggest that the trapped complex represents a transient species that precedes the active hydride transfer complex [37]. The latter interpretation is consistent with the fact that long soaks (~30 min) of oxidized crystals in NAD(P)H result in complete reduction of the enzyme and release of $NAD(P)^+$. The nicotinamide riboside binding site overlaps that of the substrate Gal*p* moiety (Fig. 7C), indicating that NAD(P)H and substrate binding are mutually exclusive. This aspect of the structure is consistent with the mechanism of UGM (Scheme 3).

The ADP group of NADPH and the UDP group of the substrate occupy distinct pockets (Fig. 7C), consistent with the observation that NADPH makes no interactions with domain 2. (Recall that domain 2 functions prominently in binding the substrate uridine.) Thus, the NADPH complex reveals a new ligand-binding pocket and provides new opportunities for inhibitor discovery. The new pocket is defined by several residues that form nonpolar and electrostatic interactions with the AMP (Fig. 7B). The adenosine of NADPH binds in a hydrophobic pocket formed by Ile65, Phe66, Tyr104, and the non-polar chain of Arg91. The adenine base forms hydrogen bonds with Ser93 and Tyr317. The pyrophosphate is stabilized by His68 and Asn457. Tyr104 makes a hydrogen bond with the 2′-phosphoryl of NADPH. This interaction is responsible, at least in part, for the 179-fold higher k_{red}/K_D value for NADPH compared to NADH observed with AfUGM [37].

All of the residues of the NADPH site are identically conserved in other eukaryotic UGMs, such as TcUGM and *L. major* UGM, two enzymes that are of interest for inhibitor design (Fig. S3 of [37]). This high degree of sequence conservation suggests that the identified binding site is present in other eUGMs. In contrast, none of the residues that contact the ADP half of NADPH are present in the sequences of bUGMs, suggesting that the identified NAD(P)H site is not present in bUGMs. We note that the binding site for NADPH in AfUGM is entirely different from the Rossmann dinucleotide-binding fold, which is the most common structural motif for binding nicotinamide adenine dinucleotide cofactors [41]. Furthermore, the compact NADPH conformation observed in AfUGM differs substantially from the extended conformations typically seen in other enzymes. These observations suggest that compounds targeted at the AfUGM NADPH site should have low affinity for other enzymes that use nicotinamide dinucleotide cofactors. Thus, the NADPH site of AfUGM provides a potential platform for discovering eUGM-specific inhibitors.

Conformational changes associated with enzyme activation – a new flavin switch protein?

Some flavoenzymes are activated by a change in the redox state of the flavin [50]. These flavin switch proteins exhibit large conformational changes upon flavin oxidation or reduction that promotes functions such as light-induced signaling, transcription regulation, and membrane binding. Since UGM is activated by flavin reduction, it is potentially another example of a flavin switch protein. Studies of bUGMs reduced *in crystallo* showed that reduction induces bending of the isoalloxazine but no perceptible change in the protein structure [28,51]. However, recent studies of eUGMs have re-

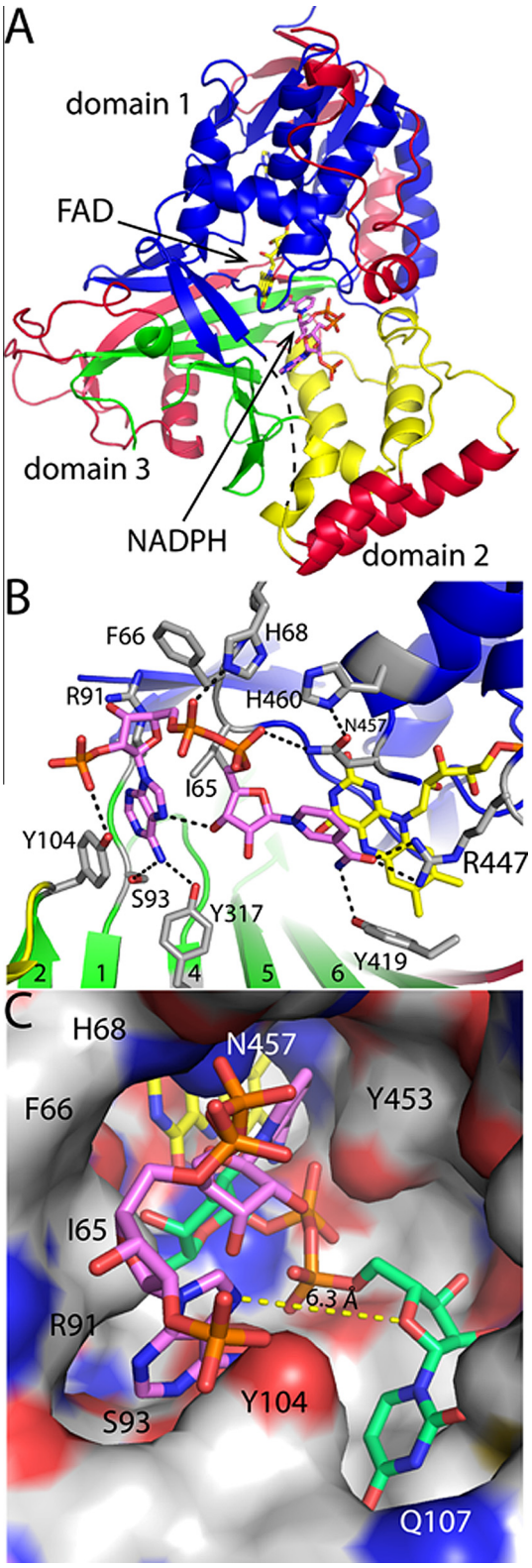


Fig. 7. The NAD(P)H site of AfUGM. (A) Ribbon drawing of oxidized AfUGM complexed with NADPH. The domains are colored as in Fig. 1B. The dashes indicate disordered mobile flap 2. (B) Interactions between AfUGM and NADPH. NADPH and FAD are colored pink and yellow, respectively. Residues that interact with NADPH are shown in gray. (C) Spatial proximity of the NADPH and UDP-Gal*p* binding sites. The surface of oxidized AfUGM–NADPH is shown with NADPH in pink and FAD in yellow. For reference, the AfUGM–NADPH structure has been overlaid with the AfUGM–UDP-Gal*p* complex and the UDP-Gal*p* (green) included in this image. The dashed yellow line denotes the distance between the adenine and uracil sites.

vealed that the protein conformation near the flavin changes dramatically in response to flavin reduction, likely heralding a new chapter in UGM biochemistry.

Redox-linked conformational changes observed in AfUGM

Structural studies of AfUGM provide the best evidence for large redox-linked conformational changes. The AfUGM structure has been determined from two different crystal forms, hexagonal [43] and triclinic [45]. Both forms yielded structures of the oxidized and *in crystallo* reduced enzymes. In addition, the hexagonal form was used to trap a complex of oxidized AfUGM with NADPH (see previous section) [37]. The hexagonal and triclinic reduced enzyme structures are essentially identical. The fact that the two crystal forms give similar structures is reassuring and strongly suggests that the crystalline enzyme conformation is an excellent model for the solution conformation. Likewise, the structures of the NADPH complex and triclinic oxidized enzyme agree well. Again, different crystal forms yielding the same conformation implies a faithful representation of the protein in solution. In contrast, the hexagonal structure of ligand-free oxidized AfUGM displays an unusual conformation of the histidine loop and is likely and artifact of the high sulfate ion concentration and low pH used in crystallization, as we have previously suggested [37,43]. Thus, the hexagonal NADPH complex and triclinic oxidized enzyme structures provide the correct conformation of oxidized AfUGM for comparison to the reduced enzyme.

The AfUGM structures imply that enzyme activation induces profound protein conformational changes (Fig. 8C). In the oxidized enzyme, His63 is near the pyrimidine portion of the isoalloxazine and oriented parallel to Trp315, while Arg327 donates a hydrogen bond to the $N5_{FAD}$. This hydrogen bond is diagnostic of the oxidized enzyme because Arg is an obligate hydrogen bond donor at physiological pH, and $N5_{FAD}$ is an obligate acceptor only in the oxidized state (Scheme 1B). Note also that the imidazole of His63 is not in contact with the isoalloxazine. Flavin reduction by either dithionite or NAD(P)H induces a dramatic reconfiguration of the conserved histidine loop in which His63 and Gly62 move by 6 and 5 Å, respectively (Fig. 8C). His63 moves to the *si* face of the isoalloxazine where it forms hydrogen bonds with the carbonyl oxygen of Gly61 and the 2′-OH of the ribityl chain (Fig. 8E). Furthermore, Gly62 changes conformation in order to accept a hydrogen bond from the $N5_{FAD}$ atom. This interaction is diagnostic of the reduced flavin because the main chain carbonyl is an obligate hydrogen bond acceptor, and N5 is a donor only in the reduced state (Scheme 1B). The conformational change of the histidine loop is coordinated with an 8-Å movement of Trp315 to evade Gly62 and rupture of the Arg327-N5 hydrogen bond (Fig. 8C).

Reduction of AfUGM also changes the conformation of the flavin itself. Upon reduction, the isoalloxazine tilts by 2 Å (Fig. 8C). Although smaller in magnitude, this movement is reminiscent of the mobile flavin in p-hydroxybenzoate hydroxylase [52,53]. Also, at the crystallographic resolutions of these structures (~2.2 Å) it is possible to discern the planarity of the isoalloxazine. The maps show that the isoalloxazine is planar in the oxidized state (see Figure S1 of [43] and Fig. 4 of [37]), whereas the reduced flavin exhibits a butterfly-like conformation in which the pyrimidine ring bends 7° out of the plane such that the *si* face is concave (Fig. 8A).

Redox-linked conformational changes observed in TcUGM

Crystal structures of TcUGM show a different set of conformational changes induced by flavin reduction (Fig. 8D). The structures of oxidized and reduced TcUGM complexed with UDP have been determined [54]. The active site structure of the reduced TcUGM–UDP complex is essentially identical to that of reduced

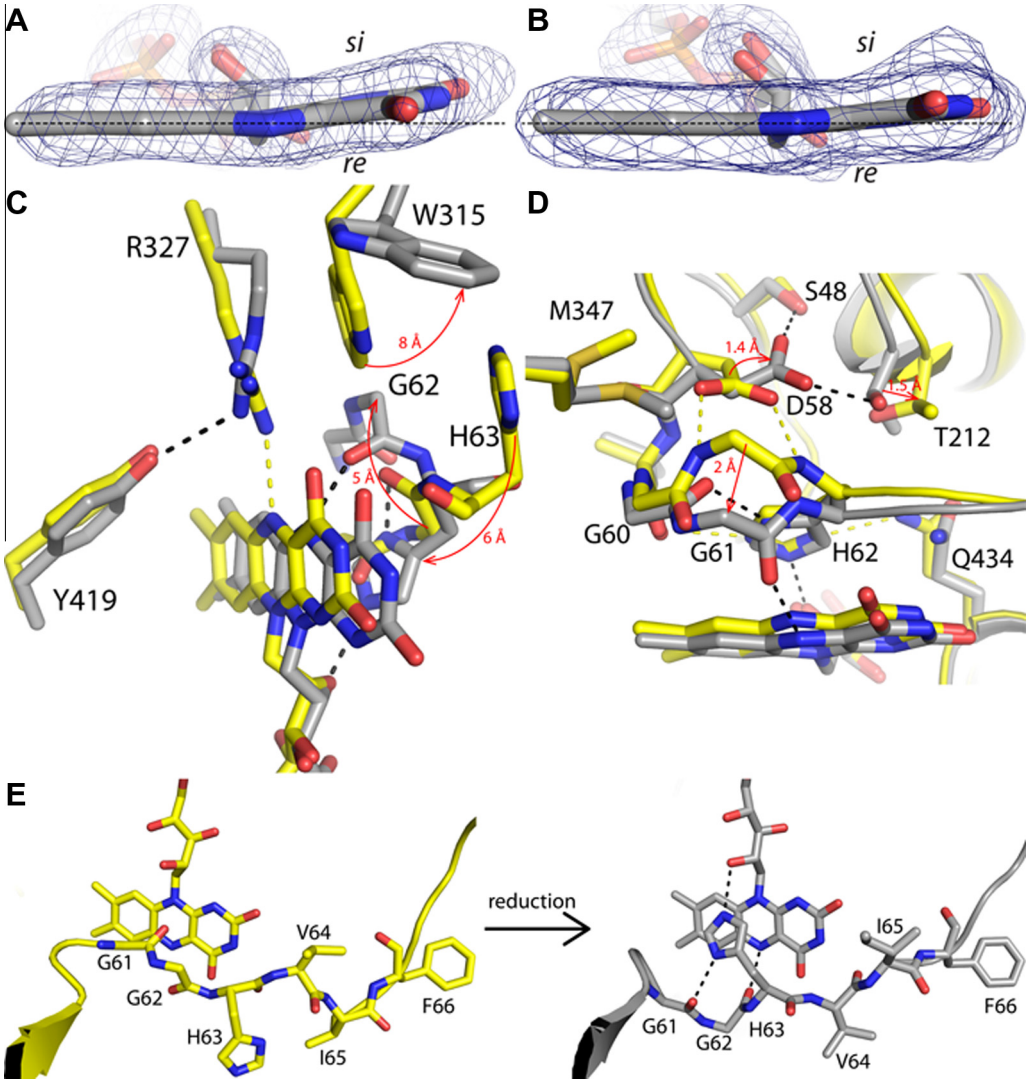


Fig. 8. Structural changes induced in eUGMs by flavin reduction. The bent isoalloxazines of (A) reduced AfUGM (PBD code 3UTF) and (B) reduced TcUGM (PDB code 4DSH). The cages represent F_o–F_c omit electron density maps (3σ). The dashed lines guide the eye to help see the deviation from planarity. (C) Superposition of oxidized (yellow) and reduced (gray) AfUGM. Yellow and black dashes denote hydrogen bonds in the oxidized and reduced enzymes, respectively. The red arrows denote the direction of conformational changes induced by flavin reduction. (D) Superposition of oxidized (yellow) and reduced (gray) TcUGM. Yellow and black dashes denote hydrogen bonds in the oxidized and reduced enzymes, respectively. The red arrows denote the direction of conformational changes induced by flavin reduction. (E) Conformational changes in the histidine loop of AfUGM.

AfUGM–UDP; even water molecules that bridge UDP to the enzyme are located similarly in the two enzymes. Also, reduction induces bending of the isoalloxazine ring as in AfUGM (Fig. 8B and Fig. 4 of [54]). However, the oxidized TcUGM structure revealed a new conformation of the histidine loop, implying a different set of protein conformational changes associated with enzyme activation. In oxidized TcUGM, the conserved histidine loop is retracted from the FAD isoalloxazine (Fig. 8D). This conformation is stabilized by hydrogen bonds between Asp58 and the backbone of the histidine loop. Upon reduction, the histidine loop is released and shifts 2.3 Å toward the isoalloxazine, allowing Gly61 to accept a hydrogen bond from the flavin N5 (Fig. 8D). As noted above for AfUGM, this hydrogen bond is observed in all reduced UGM structures and is considered essential for stabilizing the reduced flavin. Movement of the histidine loop also triggers rotations of Asp58 and Thr212 so that they engage in a hydrogen bond in the reduced state. Finally, analysis of hydrogen bond interactions suggests that flavin reduction induces protonation of His62 and rotation of its imidazole ring by 180° [54].

Molecular dynamics studies of active site flexibility

The dramatically different conformations of the active site flaps observed in various UGM crystal structures indicate that these regions are highly flexible. This flexibility is presumably important for function, since the flaps contain residues that directly contact the substrate (Fig. 4). Furthermore, movement of the flaps is essential for creating the closed active site. As suggested previously [43], flap closure not only assembles the constellation of residues needed for substrate recognition but also creates a protected environment for catalysis and prevents the severed UDP from migrating out of the active site during the catalytic cycle (Scheme 3, e–h). Thus, protein flexibility is a fundamental aspect of UGM catalysis, particularly for eUGMs, which have an additional mobile active site flap and a highly flexible histidine loop.

Molecular dynamics (MD) simulation is used widely to study enzyme flexibility [55] and has been applied to UGM. The first model of a prokaryotic substrate-bound UGM structure was based on short (5 ns) MD simulations in which UDP-Gal*p* and analogs

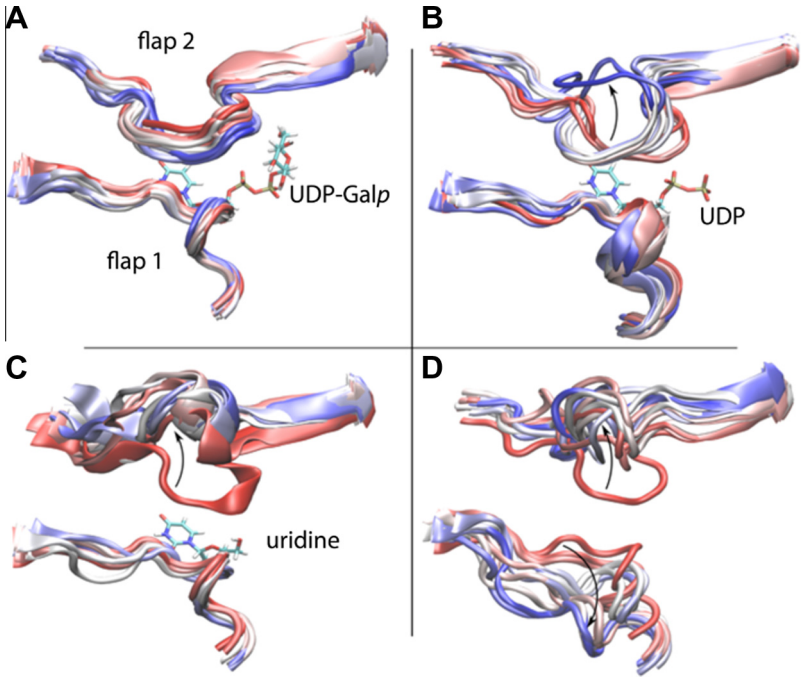


Fig. 9. Conformations of the mobile flaps from MD simulations of TcUGM complexed with (A) UDP-Gal*p*, (B) UDP, and (C) uridine. (D) Simulation of the apo enzyme. Snapshots are colored according to timestep, on a red–white–blue color scale. Arrows show the direction of the flaps displacement.

were docked into the substrate-free active site [56,57]. The simulations predicted that the flexible flap changed from an open conformation to a closed conformation in the presence of the substrate. This prediction would be confirmed by crystal structures of UGMs complexed with UDP and UDP-Gal*p*.

MD simulations have also been used to study substrate-associated active site motion in TcUGM [58]. Classical and accelerated MD simulations agree with the proposed hypothesis that ligand binding influences the open-closed equilibrium of the mobile flaps. After extensive MD simulations of both substrate-free and substrate-bound TcUGM, it was observed that both flaps open in the absence of ligand, thus creating a channel for ligand uptake [58]. When UDP-Gal*p* is bound, intramolecular interactions help maintain both flaps closed. Interestingly, MD simulations show that another flap occasionally moves increasing the channel size. This new mobile loop (residues 462–469 of TcUGM, Fig. 2A) is located in the U-shaped C-terminal region of domain 1, which is absent in bUGMs.

The simulations also suggest a modular mechanism in which each moiety of the substrate controls the flexibility of a distinct flap. When either UDP or UDP-Gal*p* is in the active site, both flaps remain closed, although some distortions appear in the flap 2 in the UDP-bound structure, suggesting that Gal*p* is important in the dynamics of flap 2 (Fig. 9). When the diphosphate moiety is removed from the ligand, and only the uridine moiety was left in the active site, flap 2 displays large distortions similar to those observed in the apo form, although no distortion was observed in flap 1. The dynamic Arg in TcUGM (Arg176), which lies on flap 1, was observed to interact with the bound UDP in the TcUGM structures, however this flap only opens as widely as observed in the apo structure when the ligand is completely removed (Fig. 9). The conformation of flap 1, although substrate-dependent, does not seem to be strictly a result of the observed interactions between the dynamic Arg and the bound ligand.

Visual inspection of the TcUGM conformations suggests that the uridine moiety controls flap 1 through a hydrogen bond network that involves Ala178, Gln103, and Phe102. Similar stabilization of the uridine moiety is observed in KpUGM with Phe151, Thr156, and Trp160 (the uridine wall-described previously). On the other hand, the diphosphate and Gal*p* moieties seem to control flap 2 through Asn201 in TcUGM, although no homologue for Asn201 seems to exist in bUGMs, further highlighting the differences between the two UGM classes.

Summary and Outlook

The structural and biochemical studies of UGMs have revealed a unique 3-dimensional structure and novel role for flavins in biological processes. Catalysis by UGMs (in particular in eUGMs) is coupled to extensive conformational changes that are modulated by substrate binding and the redox state of the flavin. The aggregate research on UGMs sets the stage for new discoveries.

Exploring the relationship between dynamics and catalytic activity should be a productive area of inquiry. In particular, the MD simulations raise new questions about the mechanism of product release. MD simulations show that UMP is essential for closure of the active site. Since UMP is present in both the substrate and product, the molecular events that promote opening of the active site for product release are unclear. One possibility is that the formation of Gal*f* shifts the equilibrium toward the open state. Determining structures of UGM complexed with UDP-Gal*f* should shed light on this aspect of the mechanism.

The study of redox-linked conformational changes should also be an exciting future area of UGM research. It is clear that activation of eUGMs triggers profound conformational changes. The initial data suggest that the inactive conformations of AfUGM and TcUGM differ, and thus different types of conformational changes bring each enzyme into the active form. However, it should be noted that the structure of ligand-free TcUGM has not yet been determined, and the presence of UDP in the current structures could certainly bias the conformation of the histidine loop. Therefore, a critical goal should be to determine the structure of TcUGM

and other eUGMs in the oxidized, ligand-free state in order to assess conservation of the oxidized enzyme conformation and redox-linked conformational changes. Another goal should be to identify the sequence and structural elements that enable redox-linked conformational changes in eUGMs. Sequence analysis suggests one working hypothesis that an additional Gly residue in the histidine loop of eUGMs, which is Ala or Pro in bUGMs, imparts the flexibility to the histidine loop (Fig. 6).

Finally, UGM presents several challenges for inhibitor design that need to be addressed. For the eUGMs, a simple, robust, high-throughput screening assay has not been developed. Application of other methods, such as ThermoFAD and ThermoFluor with focused libraries might be worth pursuing [59,60]. The wealth of structural information on UGM-substrate complexes would seem to bode well for computer-aided drug design. However, the large conformational changes associated with substrate binding must be taken into account during docking, which is challenging. UGM is perhaps a good test case for developing and testing *in silico* docking methods that incorporate receptor flexibility. Finally, covalent inactivators represent a new potential area of UGM research. The reduced flavin is a reactive nucleophile and could be targeted by an electrophilic warhead fused to substrate-like scaffold. Although the drug design community has trended away from covalent inactivators because of immune-mediated toxicity, this opinion appears to be changing [61,62]. The design of such mechanism-based inactivators will rely heavily upon the accumulated knowledge of the structure and mechanism of UGM summarized here.

Acknowledgments

This work was supported by a National Institutes of Health Grant GM094469 (P.S. and J.J.T). Work at UCSD is supported in part by NIH, NSF, HHMI, NBCR, CTBP, and the NSF Supercomputer Centers. L.B. is the recipient of a Pew Latin American Fellowship.

References

[1] W.N. Haworth, H. Raistrick, M. Stacey, Biochem. J. 31 (1937) 640–644.
[2] J.E. Gander, Arch. Biochem. Biophys. 91 (1960) 307–309.
[3] A.G. Trejo, J.W. Haddock, G.J. Chittenden, J. Baddiley, Biochem. J. 122 (1971) 49–57.
[4] H. Nikaido, M. Sarvas, J. Bacteriol. 105 (1971) 1073–1082.
[5] M. Sarvas, H. Nikaido, J. Bacteriol. 105 (1971) 1063–1072.
[6] S.M. Beverley, K.L. Owens, M. Showalter, C.L. Griffith, T.L. Doering, V.C. Jones, M.R. McNeil, Eukaryot. Cell 4 (2005) 1147–1154.
[7] R.M. de Lederkremer, R. Agusti, Adv. Carbohydr. Chem. Biochem. 62 (2009) 311–366.
[8] B. Tefsen, A.F. Ram, I. van Die, F.H. Routier, Glycobiology 22 (2012) 456–469.
[9] M.R. Richards, T.L. Lowary, Chembiochem 10 (2009) 1920–1938.
[10] M. Oppenheimer, A.L. Valenciano, P. Sobrado, Enzyme Res. 2011 (2011) 415976.
[11] P.M. Nassau, S.L. Martin, R.E. Brown, A. Weston, D. Monsey, M.R. McNeil, K. Duncan, J. Bacteriol. 178 (1996) 1047–1052.
[12] R. Koplin, J.R. Brisson, C. Whitfield, J. Biol. Chem. 272 (1997) 4121–4128.
[13] A. Weston, R.J. Stern, R.E. Lee, P.M. Nassau, D. Monsey, S.L. Martin, M.S. Scherman, G.S. Besra, K. Duncan, M.R. McNeil, Tuber. Lung. Dis. 78 (1997) 123–131.
[14] J.F. Novelli, K. Chaudhary, J. Canovas, J.S. Benner, C.L. Madinger, P. Kelly, J. Hodgkin, C.K. Carlow, Dev. Biol. 335 (2009) 340–355.
[15] F. Pan, M. Jackson, Y. Ma, M. McNeil, J. Bacteriol. 183 (2001) 3991–3998.
[16] P.S. Schmalhorst, S. Krappmann, W. Vervecken, M. Rohde, M. Muller, G.H. Braus, R. Contreras, A. Braun, H. Bakker, F.H. Routier, Eukaryot. Cell 7 (2008) 1268–1277.
[17] B. Kleczka, A.C. Lamerz, G. van Zandbergen, A. Wenzel, R. Gerardy-Schahn, M. Wiese, F.H. Routier, J. Biol. Chem. 282 (2007) 10498–10505.
[18] Q. Zhang, H.W. Liu, J. Am. Chem. Soc. 123 (2001) 6756–6766.
[19] J.N. Barlow, J.S. Blanchard, Carbohydr. Res. 328 (2000) 473–480.
[20] Q. Zhang, H.W. Liu, J. Am. Chem. Soc. 122 (2000) 9065–9070.
[21] D.A. Sanders, A.G. Staines, S.A. McMahon, M.R. McNeil, C. Whitfield, J.H. Naismith, Nat. Struct. Biol. 8 (2001) 858–863.
[22] M. Oppenheimer, M.B. Poulin, T.L. Lowary, R.F. Helm, P. Sobrado, Arch. Biochem. Biophys. 502 (2010) 31–38.
[23] J.N. Barlow, E.M. Girvin, J.S. Blanchard, J. Am. Chem. Soc. 121 (1999) 6968–6969.
[24] A. Caravano, P. Sinay, S.P. Vincent, Bioorg. Med. Chem. Lett. 16 (2006) 1123–1125.
[25] S.W. Fullerton, S. Daff, D.A. Sanders, W.J. Ingledew, C. Whitfield, S.K. Chapman, J.H. Naismith, Biochemistry 42 (2003) 2104–2109.
[26] Z. Huang, Q. Zhang, H.W. Liu, Bioorg. Chem. 31 (2003) 494–502.
[27] M. Soltero-Higgin, E.E. Carlson, T.D. Gruber, L.L. Kiessling, Nat. Struct. Mol. Biol. 11 (2004) 539–543.
[28] T.D. Gruber, W.M. Westler, L.L. Kiessling, K.T. Forest, Biochemistry (2009).
[29] T.D. Gruber, M.J. Borrok, W.M. Westler, K.T. Forest, L.L. Kiessling, J. Mol. Biol. 391 (2009) 327–340.
[30] M. Oppenheimer, A.L. Valenciano, K. Kizjakina, J. Qi, P. Sobrado, PLoS One 7 (2012) e32918.
[31] P. Sobrado, Int. J. Mol. Sci. 13 (2012) 14219–14242.
[32] K. Itoh, Z. Huang, H.W. Liu, Org. Lett. 9 (2007) 879–882.
[33] H.G. Sun, M.W. Ruszczycky, W.C. Chang, C.J. Thibodeaux, H.W. Liu, J. Biol. Chem. 287 (2012) 4602–4608.
[34] W. Huang, J.W. Gauld, J. Phys. Chem. B 116 (2012) 14040–14050.
[35] S. Bornemann, Nat. Prod. Rep. 19 (2002) 761–772.
[36] C.T. Walsh, T.A. Wencewicz, Nat. Prod. Rep. 30 (2013) 175–200.
[37] R. Dhatwalia, H. Singh, L.M. Solano, M. Oppenheimer, R.M. Robinson, J.F. Ellerbrock, P. Sobrado, J.J. Tanner, J. Am. Chem. Soc. 134 (2012) 18132–18138.
[38] I. Da Fonseca, K. Kizjakina, P. Sobrado, Arch. Biochem. Biophys (2013) 103–110.
[39] J.N. Barlow, J. Marcinkeviciene, J.S. Blanchard, in: P.A. Frey, D.B. Northrop (Eds.), Enzymaic Mechanisms, IOS Press, Burke, VA, 1999, p. 98106.
[40] I. Gusarov, M. Starodubtseva, Z.Q. Wang, L. McQuade, S.J. Lippard, D.J. Stuehr, E. Nudler, J. Biol. Chem. 283 (2008) 13140–13147.
[41] C.A. Bottoms, P.E. Smith, J.J. Tanner, Protein Sci. 11 (2002) 2125–2137.
[42] E. Krissinel, K. Henrick, Acta Crystallogr. D Biol. Crystallogr. 60 (2004) 2256–2268.
[43] R. Dhatwalia, H. Singh, M. Oppenheimer, D.B. Karr, J.C. Nix, P. Sobrado, J.J. Tanner, J. Biol. Chem. 287 (2012) 9041–9051.
[44] G.A. Penman, D.E.A. Lockhart, A. Ferenbach, D.M.F. van Aalten, Sect. F Struct. Biol. Cryst. Commun. 68 (2012) 705–708.
[45] K.E. van Straaten, F.H. Routier, D.A. Sanders, J. Biol. Chem. 287 (2012) 10780–10790.
[46] K.E. van Straaten, F.H. Routier, D.A. Sanders, Sect. F Struct. Biol. Cryst. Commun. 68 (2012) 455–459.
[47] Z.S. Derewenda, Structure (London, England: 1993) 12 (2004) 529–535.
[48] L. Goldschmidt, D.R. Cooper, Z.S. Derewenda, D. Eisenberg, Protein Sci. 16 (2007) 1569–1576.
[49] S.A. McMahon, G.A. Leonard, L.V. Buchanan, M.F. Giraud, J.H. Naismith, Acta Crystallogr. D Biol. Crystallogr. 55 (1999) 399–402.
[50] D.F. Becker, W. Zhu, M.A. Moxley, Antioxid. Redox Signal 14 (2011) 1079–1091.
[51] S.K. Partha, K.E. van Straaten, D.A. Sanders, J. Mol. Biol. 394 (2009) 864–877.
[52] D.L. Gatti, B.A. Palfey, M.S. Lah, B. Entsch, V. Massey, D.P. Ballou, M.L. Ludwig, Science 266 (1994) 110–114.
[53] B. Entsch, L.J. Cole, D.P. Ballou, Arch. Biochem. Biophys. 433 (2005) 297–311.
[54] R. Dhatwalia, H. Singh, M. Oppenheimer, P. Sobrado, J.J. Tanner, Biochemistry 51 (2012) 4968–4979.
[55] S.A. Adcock, J.A. McCammon, Chem. Rev. 106 (2006) 1589–1615.
[56] X. Yao, D.W. Bleile, Y. Yuan, J. Chao, K.P. Sarathy, D.A. Sanders, B.M. Pinto, M.A. O'Neill, Proteins 74 (2009) 972–979.
[57] Y. Yuan, D.W. Bleile, X. Wen, D.A. Sanders, K. Itoh, H.W. Liu, B.M. Pinto, J. Am. Chem. Soc. 130 (2008) 3157–3168.
[58] L. Boechi, C.A. de Oliveira, I. Da Fonseca, K. Kizjakina, P. Sobrado, J.J. Tanner, J.A. McCammon, Protein Sci. (2013) in press (Epub ahead of print).
[59] F. Forneris, R. Orru, D. Bonivento, L.R. Chiarelli, A. Mattevi, FEBS J. 276 (2009) 2833–2840.
[60] M.D. Cummings, M.A. Farnum, M.I. Nelen, J. Biomol. Screen. 11 (2006) 854–863.
[61] A.S. Kalgutkar, D.K. Dalvie, Exp. Opin. Drug Discov. 7 (2012) 561–581.
[62] D.S. Johnson, E. Weerapana, B.F. Cravatt, Med. Chem. 2 (2010) 949–964.
[63] W.L. DeLano, The PyMOL User's Manual, DeLano Scientific, San Carlos, CA, USA, 2002.
[64] K. Beis, V. Srikannathasan, H. Liu, S.W. Fullerton, V.A. Bamford, D.A. Sanders, C. Whitfield, M.R. McNeil, J.H. Naismith, J. Mol. Biol. 348 (2005) 971–982.
[65] S. Karunan Partha, S.A. Bonderoff, K.E. van Straaten, D.A. Sanders, Sect. F Struct. Biol. Cryst. Commun. 65 (2009) 843–845.
[66] S.K. Partha, A. Sadeghi-Khomami, K. Slowski, T. Kotake, N.R. Thomas, D.L. Jakeman, D.A. Sanders, J. Mol. Biol. 403 (2010) 578–590.