Human hydroxymethylbilane synthase: Molecular dynamics of the pyrrole chain elongation identifies step-specific residues that cause AIP

Navneet Bung[a,1], Arijit Roy[a,1], Brenden Chen[b,1], Dibyajyoti Das[a], Meenakshi Pradhan[a], Makiko Yasuda[b], Maria I. New[b,2], Robert J. Desnick[b,2], and Gopalakrishnan Bulusu[a,2]

[a]Life Sciences Division, TCS Innovation Labs–Hyderabad, Tata Consultancy Services Limited, Hyderabad 500081, India; and [b]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029

Contributed by Maria I. New, March 6, 2018 (sent for review November 9, 2017; reviewed by Harry A. Dailey, Pavel Martasek, and Vijay S. Reddy)

Hydroxymethylbilane synthase (HMBS), the third enzyme in the heme biosynthetic pathway, catalyzes the head-to-tail condensation of four molecules of porphobilinogen (PBG) to form the linear tetrapyrrole 1-hydroxymethylbilane (HMB). Mutations in human *HMBS* (*hHMBS*) cause acute intermittent porphyria (AIP), an autosomal-dominant disorder characterized by life-threatening neurovisceral attacks. Although the 3D structure of hHMBS has been reported, the mechanism of the stepwise polymerization of four PBG molecules to form HMB remains unknown. Moreover, the specific roles of each of the critical active-site residues in the stepwise enzymatic mechanism and the dynamic behavior of hHMBS during catalysis have not been investigated. Here, we report atomistic studies of HMB stepwise synthesis by using molecular dynamics (MD) simulations, mutagenesis, and in vitro expression analyses. These studies revealed that the hHMBS active-site loop movement and cofactor turn created space for the elongating pyrrole chain. Twenty-seven residues around the active site and water molecules interacted to stabilize the large, negatively charged, elongating polypyrrole. Mutagenesis of these active-site residues altered the binding site, hindered cofactor binding, decreased catalysis, impaired ligand exit, and/or destabilized the enzyme. Based on intermediate stages of chain elongation, R26 and R167 were the strongest candidates for proton transfer to deaminate the incoming PBG molecules. Unbiased random acceleration MD simulations identified R167 as a gatekeeper and facilitator of HMB egress through the space between the enzyme's domains and the active-site loop. These studies identified the specific active-site residues involved in each step of pyrrole elongation, thereby providing the molecular bases of the active-site mutations causing AIP.

structural biology | enzyme catalysis | molecular dynamics

Tetrapyrrole biosynthesis is essential for virtually all forms of life on Earth. Tetrapyrroles serve as precursors for heme, phycobilins, chlorophyll, and cyanocobalamin (i.e., vitamin B12) and are involved in various biological processes (1). Hydroxymethylbilane synthase (HMBS), also known as porphobilinogen (PBG) deaminase (PBGD; EC 2.5.1.61), the third enzyme in the heme biosynthetic pathway (2), catalyzes the stepwise polymerization of four molecules into the linear tetrapyrrole 1-hydroxymethylbilane (HMB) (3, 4) (Fig. 1). In humans, mutations in *HMBS* (*hHMBS*) cause acute intermittent porphyria (AIP), an autosomal-dominant inborn disorder characterized by life-threatening acute neurovisceral attacks (2, 5).

Previous studies of the purified *Escherichia coli* and human enzymes demonstrated the critical importance of a covalently bound dipyrromethane (DPM) cofactor and the occurrence of stable enzyme substrate complexes in the formation of HMB (4, 6–8). To date, the X-ray crystal structures of *E. coli* (9), *Arabidopsis thaliana* (10), *Bacillus megaterium* (11, 12), and human (13, 14) HMBS (hHMBS) with their DPM cofactors have been reported. The crystal structures of the hHMBS housekeeping isoform are available in the Protein Data Bank (PDB) under ID codes 3EQ1 (13) and 3ECR (14). The hHMBS has three distinct domains: domain 1 (residues 1 to 114, 219 to 236), domain 2 (residues 120 to 212), and domain 3 (residues 241 to 361), connected by interdomain hinge regions (Fig. 2). The active-site region (i.e., catalytic cleft) of hHMBS lies between domains 1 and 2, where PBG binding and polypyrrole elongation reactions occur (9, 13–15). The DPM cofactor, which acts as a primer for HMB formation, is linked by a thioether bond to the conserved C261 residue. Structural studies speculated that domain movements around the hinge region are essential to accommodate the four PBG molecules in polypyrrole elongation (9, 14).

There are several differences between the structure of hHMBS and its homologs from *E. coli* (EcHMBS), *A. thaliana* (AtHMBS), and *B. megaterium*. The major difference is the presence of a 29-residue insert (residues 296 to 324) in domain 3 at the interface between domains 1 and 3 (13, 14) in hHMBS, the role of which is unknown. Also, the active-site loop in the crystal structures of hHMBS (residues 56 to 76) and EcHMBS are disordered and their coordinates are missing, even though this flexible loop has been modeled in the crystal structure of AtHMBS, in which it covers the active-site region (10). Further, the coordinates of N-terminal (residues 1 to 17) and C-terminal

Significance

Human hydroxymethylbilane synthase (hHMBS) is a monomeric enzyme that catalyzes the stepwise head-to-tail condensation of four porphobilinogen (PBG) molecules to form the linear tetrapyrrole 1-hydroxymethylbilane (HMB). Mutations in the *hHMBS* gene cause autosomal-dominant acute intermittent porphyria (AIP). Although crystal structures of hHMBS have been reported, the specific active-site residues and the molecular mechanism of the stepwise PBG chain elongation are unknown. Here, by using molecular-dynamics simulations, the mechanisms and active-site residues for the HMB stepwise synthesis and HMB exit were determined. Mutagenesis of key active-site residues and in vitro expression studies identified the molecular basis of mutations causing AIP.

Author contributions: N.B., A.R., B.C., M.P., R.J.D., and G.B. designed research; N.B., A.R., B.C., and M.P. performed research; N.B., A.R., B.C., D.D., M.P., R.J.D., and G.B. analyzed data; and N.B., A.R., B.C., D.D., M.P., M.Y., M.I.N., R.J.D., and G.B. wrote the paper.

Reviewers: H.A.D., University of Georgia; P.M., Charles University; and V.S.R., The Scripps Research Institute.

The authors declare no conflict of interest.

[1]N.B., A.R., and B.C. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: maria.new@mssm.edu, robert.desnick@mssm.edu, or g.bulusu@tcs.com.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1719267115/-/DCSupplemental.

Published online April 9, 2018.

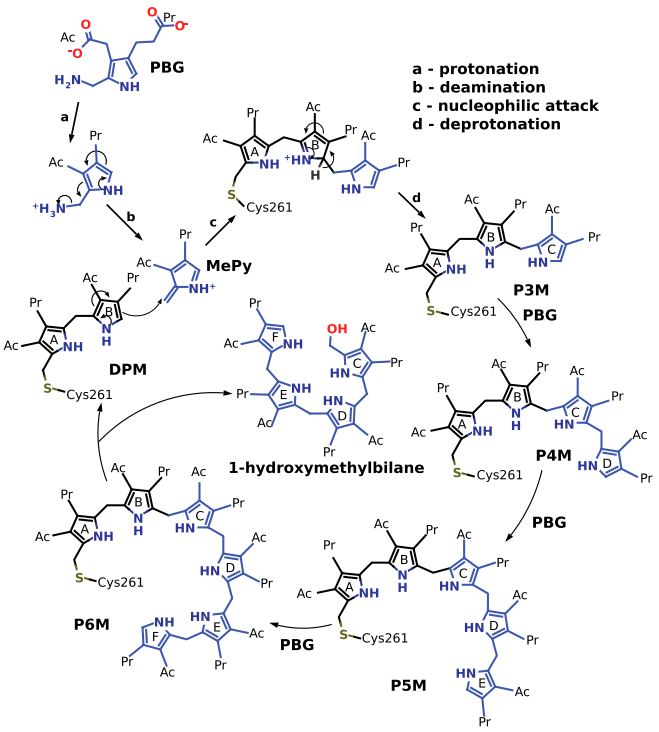


Fig. 1. Schematic representation of mechanism of tetrapyrrole chain elongation catalyzed by hHMBS. The figure shows (*a*) protonation of PBG, (*b*) deamination of PBG to form MePy, and (*c*) nucleophilic attack by ring B of DPM on MePy, forming an intermediate that (*d*) undergoes deprotonation to form a tripyrrole moiety (P3M). Subsequent additions of PBG elongate the chain to form tetrapyrrole (P4M), pentapyrrole (P5M), and hexapyrrole (P6M) moieties. At the end of the last step, the tetrapyrrole product HMB is hydrolyzed, leaving the DPM cofactor attached to protein. The rings of the elongating pyrrole chain are labeled as A, B, C, D, E, and F starting from the pyrrole ring covalently attached to C261. The acetate and propionate side groups of the pyrroles are denoted by "-Ac" and "-Pr," respectively.

(residues 358 to 361) residues are not recorded in the hHMBS crystal structures (13, 14). In addition, the presence of water molecules has been reported in the high-resolution crystal structures of AtHMBS (10), EcHMBS (7), and hHMBS (14). These water molecules were trapped within the active-site cleft and probably stabilized the negatively charged cofactor (7). However, the role of these water molecules in hHMBS enzyme catalysis is largely uncharacterized.

Previous studies of EcHMBS characterized the importance of 12 conserved arginine residues in different steps of pyrrole assembly (15–17) and identified a few key residues that may take part in the HMBS catalytic mechanism. However, in the absence of any structural data for hHMBS with the hexapyrrole and for each intermediate stage of pyrrole polymerization, the exact roles of the residues in hHMBS involved in polypyrrole chain elongation remain largely unknown.

Recently, a stepwise mechanism for EcHMBS was reported with the use of molecular dynamics (MD) simulations (ref. 15; see ref. 18 for an overview of MD simulation for nonstructural biologists). These studies revealed that domain motions and active-site loop movements play a crucial role in accommodating the polypyrrole in the active site. Moreover, these studies revealed that chain elongation takes place through repetition of sequence of steps, including protonation of the PBG substrate, deamination of PBG, nucleophilic attack on the deaminated substrate by the α-carbon atom of the terminal pyrrole ring of the enzyme-bound cofactor, and deprotonation at the same carbon position (9, 15) (Fig. 1).

As the hHMBS crystal structure did not provide information on all of the residues involved in the stepwise pyrrole elongation (13, 14), MD simulations of hHMBS at each stage of polymerization were performed to provide details of this monomeric enzyme's unique catalytic mechanism. These studies revealed the structural changes in the enzyme required for polypyrrole assembly, the effect of the 29-residue insert on the structural dynamics of the enzyme, how the active-site residues accommodate the large, highly negatively charged polypyrrole chain, and which residues are critical for the catalysis at each step of substrate addition. In addition, these studies identified the role of water molecules in the reaction mechanism and the key residues involved in the egress of HMB from hHMBS. Finally, the active-site missense mutations were shown to alter enzyme catalysis and cause AIP.

Results

Structural Fluctuations in hHMBS During Polypyrrole Assembly. MD simulations of hHMBS with the DPM cofactor (DPM stage) and the four subsequent stages of chain elongation were each performed for 50 ns. The intermediate stages of chain elongation were designated as P3M (tripyrrole), P4M (tetrapyrrole), P5M (pentapyrrole), and P6M (hexapyrrole; Fig. 1) (15). These stages have also been referred to as ES, ES2, ES3, and ES4, respectively, in previous reports (4, 7). To study the cumulative structural changes that were responsible for the accommodation of the polypyrrole in hHMBS, a 50-ns MD trajectory of each stage was concatenated to obtain a 250-ns trajectory. The correlated motions of protein dynamics were studied by principal components analysis (PCA) and dynamic cross-correlation matrix (DCCM) (19) analysis using the concatenated trajectory of 250 ns. The principal components showed major movements of the active-site loop region (Fig. S1*A* and Movies S1 and S2). Minor fluctuations were also noticed in domains 1 and 2, where the polypyrrole assembly occurs. Also, the DCCM results were in accordance with the principal component fluctuations (Fig. S1*B*).

The backbone rmsd for each stage of chain elongation was calculated from the previously described concatenated trajectory to determine the stage-specific protein dynamics. After the initial reorganization of the protein during the DPM stage, the rmsd values gradually increased to a magnitude of 1 Å in the P5M stage. In the P6M stage alone, a sharp increase of ∼0.9 Å in the rmsd was observed (Fig. 3*A*). Domain 1, which included the active-site loop, contributed significantly to the rmsd of the protein compared with domains 2 and 3.

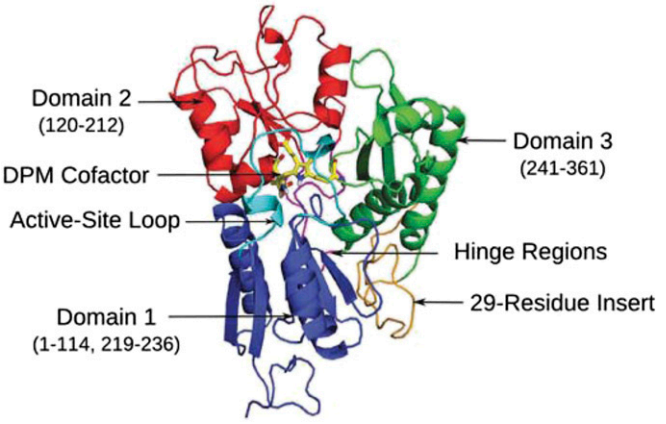


Fig. 2. Structure of hHMBS (PDB ID code 3ECR) with modeled missing residues showing the domains (1, 2, and 3 in blue, red, and green, respectively) with hinge regions (115–119, 213–218, and 237–240 in pink), the additional 29-residue insert (296–324 in orange), and the active-site loop (56–76 in cyan).

The root mean square fluctuations (RMSF) of individual amino acids were analyzed. The active-site loop (residues 56 to 76) showed a significant increase in the RMSF value from the P4M stage onward (Fig. 3*B*). This was mainly because of the active-site loop moving away from the active-site region in response to the elongating polypyrrole. Fluctuations of the 29-residue insert (residues 296 to 324) in domain 3 were also seen, along with the N- and C-terminal regions (Fig. 3*B*). The distance between domains 1 and 2 increased from the P3M stage as a result of the addition of pyrrole ring C (Fig. S1*C*). During the DPM and P3M stages, the active-site loop stayed close to the active-site and moved gradually away during the subsequent stages of chain elongation (Fig. 3*C*).

Cofactor Turn Movement Assists in Polypyrrole Accommodation Within the Catalytic Site. The structures obtained from the trajectory of P3M, P4M, P5M, and P6M stages of chain elongation were aligned to the reference structure, and the average fluctuations of the Cα atom of C261 (linked to the cofactor) were monitored. The average displacements of the Cα atom were 1.8, 2.5, 2.0, and 1.6 Å away from the direction of the active-site loop during P3M, P4M, P5M, and P6M stages of chain elongation, respectively. The effect was greatest in the P4M stage, with an average value of 2.5 Å, and the maximum shift during this stage was 3.5 Å (Fig. 3*D*). These observations, along with the computed rmsd values of the domains, clearly indicated that the cofactor turn reoriented in the P3M, P4M, and P5M stages, along with marginal domain movements to accommodate the elongating polypyrrole.

Active-Site Loop Opens to Accommodate the Polypyrrole. The active-site loop conformations were clustered based on the distance from the active-site residues (R26, Q34, D99, R149, R150), and the rmsd of the active-site loop. Fig. 3*E* shows a probability of cooccurrence plot for different conformations of the active-site loop (computed along the concatenated trajectory). Each bin represented a cluster of active-site loop conformations (Fig. 3*E*). Bin 1 represented the active-site loop conformation in a "closed" state, in which the loop remained close to the active site (Fig. 3*F*). The structures clustered in bin 1 were from DPM and P3M stages. Two active-site loop residues, T58 and D61, formed hydrogen bonds and salt bridges with R26 in the DPM and the P3M stages. These interactions, along with a cation-π interaction between F77 (part of the active-site loop) and R26, kept the active-site loop in a closed conformation (Fig. S2*A*). Bin 2 represented a "semi-open" state of the active-site loop in which the loop had moved away by at least 1 Å (Fig. 3*E*). The conformations clustered in bin 2 were mostly from the P4M and the initial part of the P5M stages, in which the interactions between T58, D61, and R26 were lost as a result of an additional pyrrole ring in the vicinity (Fig. S2*B*). Bin 3 represented the active-site loop conformation in an "open" state, in which the loop moved further away from the active site by ~3 Å (Fig. 3*E*), providing enough space to accommodate the pentapyrrole and hexapyrrole. The conformations clustered in bin 3 were from P5M and P6M stages of chain elongation. Thus, the combined dynamics of the domains, active-site loop, and cofactor

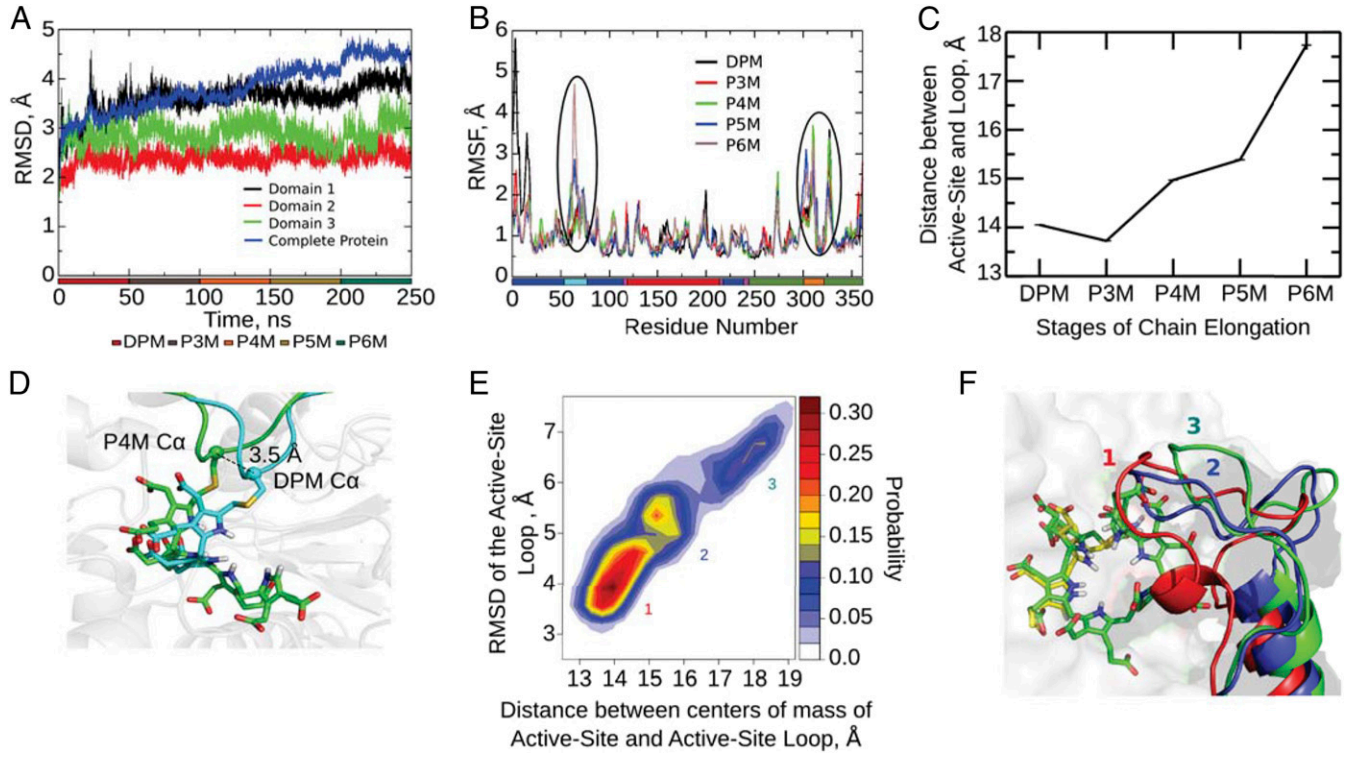


Fig. 3. Structural fluctuations during polypyrrole accommodation. (*A*) rmsd of the protein along with domainwise rmsd, with reference to the DPM stage structure, during stages of chain elongation. (*B*) RMSF of Cα atoms of hHMBS protein at each stage of chain elongation, from DPM to P6M. The encircled regions show high fluctuations in the active-site loop and 29-residue insert regions. (*C*) The distance between the centers of mass of the active-site loop and the active-site residues (R26, Q34, D99, R149, and R150) as a function of time along the stages of chain elongation emphasizes the role of the loop in polypyrrole accommodation. The SE is shown as error bars along *y*-coordinates. (*D*) The cofactor turn shifts in the P4M stage (green) compared with the DPM stage (cyan) to accommodate the growing pyrrole chain during the stages of chain elongation. (*E*) Characterization of the active-site loop conformation as a function of (*i*) distance between the centers of mass of the active-site residues (R26, Q34, D99, R149, and R150) and the active-site loop on the *x*-axis and (*ii*) the rmsd of the active-site loop on the *y*-axis along the concatenated trajectory and (*F*) the conformation of the active-site loop at bins 1, 2, and 3 shown in red, blue, and green, respectively. DPM is shown in yellow and P6M is shown in green sticks.

turn provided enough space to accommodate the four additional PBG molecules.

Charged and Hydrophilic Active-Site Residues Stabilize the Growing Pyrrole Chain. It is important for the incoming PBG to assume a correct conformation so that subsequent chemical reactions can take place. Each of the pyrrole rings added during the catalysis contain two carboxylate groups (an acetate and a propionate). Consequently, the net charge on the polypyrrole increased from −4 (DPM stage) to −12 (P6M stage) during chain elongation. From the crystal structure data of the DPM stage (14) and biochemical studies (9, 13, 14, 16), it was observed that the conserved polar and charged residues in the active site stabilized the DPM cofactor. Atomistic details about the stabilization of the growing pyrrole chain during the various catalytic stages are given in the following sections.

DPM stage. D99, which is part of domain 1, interacted with the nitrogen atoms of each pyrrole ring at all stages of pyrrole elongation (Fig. 4*A*). Side chains of positively charged residues K98, R149, R150, and R173 neutralized the carboxylate groups of the DPM cofactor. Polar residues S146 and S147 also formed hydrogen bonds with the polypyrrole chain (Fig. 4*A*). All of these residues are highly conserved across different species (13). D99 and K98 are part of domain 1, whereas all of the other interacting residues are in domain 2. The hydrogen-bonded stabilizations were also noted in the hHMBS crystal structures (13, 14).

P3M stage. On addition of pyrrole ring C to the DPM cofactor, changes in the active-site dynamics were observed. The distance between domains 1 and 2 marginally increased ∼0.5 Å to accommodate ring C. The carboxylate side chains of ring C were oriented toward domain 1 residues R26, S28, and S96. As a result of the cofactor turn movement, T145 formed a hydrogen bond with the propionate side chain of ring A that persisted through the subsequent stages of chain elongation. Side chains of S28 and S96 interacted with the acetate and propionate side chains of ring C (Fig. 4*B*). R26 also interacted with the propionate side chain of ring C.

P4M stage. When ring D was added, domains 1 and 2 moved another ∼0.5 Å apart and the cofactor turn moved toward domain 2 (Fig. 3*D*) to create space for the new PBG moiety. The active-site loop moved away from the active site at this stage, which allowed T58 to interact with the propionate side chain of ring D. R167 also interacted with the propionate side chain of ring D (Fig. 4*C*). Additionally, R195 interacted with the propionate side chain of ring C.

P5M stage. Ring E of the polypyrrole was accommodated toward the active-site loop, causing the active-site loop to move away from the active site. S69, an active-site loop residue, interacted with the acetate side chain of ring E for the first 27 ns but later shifted the interaction to the propionate side chain of ring D as a result of the loop movement (Fig. 4*D*). S262 and Q356 interacted with the acetate and propionate side chains of ring E, respectively (Fig. 4*D*). Neither S69 or S262 are conserved in HMBS across different species, but play an important role in stabilizing ring E of the polypyrrole in hHMBS.

P6M stage. Because of the addition of ring F (last PBG), the active-site loop moved further away from the active site. Most of the hydrogen bonds that stabilized the polypyrrole up to the P5M stage were retained in the P6M stage (Fig. 4*E*). The pyrrole nitrogen of ring F interacted with the D99 residue through bridging water molecules (Fig. 4*E* and Table S1). R173, previously interacting with the A and B rings of the cofactor, now retained its interaction with the side chains of only ring A (Fig. 4*E*). S262 interacted with both rings E and F, whereas S165 interacted with the propionate side chain of ring F (Fig. 4*E*). Additionally, Q356 interacted with ring F. The side chain of R167 interacted with the acetate side chain of rings E and F instead of only ring D. As a result, no protein residues stabilize the acetate group of ring D (Fig. 4*E*). However, several water molecules that can stabilize the acetate group in the absence of polar amino acid side chains were seen in proximity.

Water-Mediated Interactions Stabilize the Polypyrrole. The elongating polypyrrole has a number of negatively charged acetate and propionate side chains. Based on the simulations of different

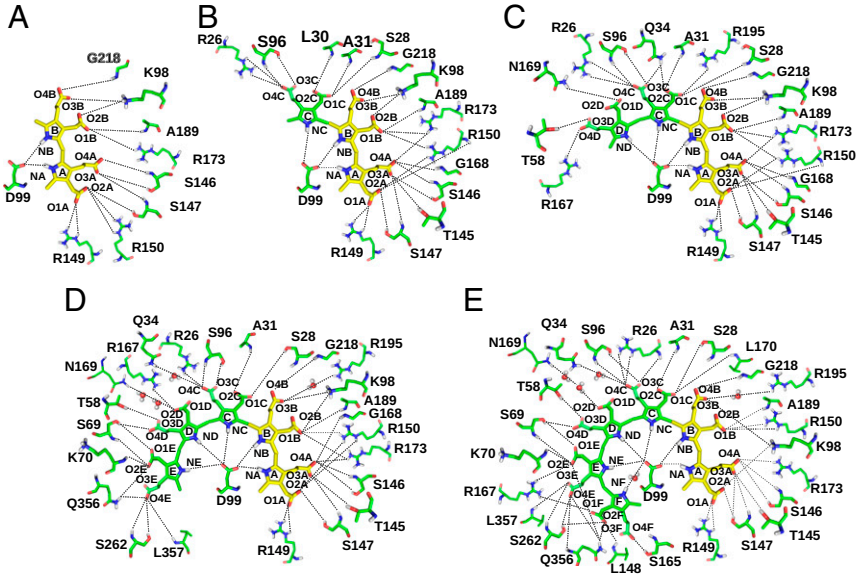


Fig. 4. Residues in active site stabilize growing polypyrrole chain. Interactions of the residues lining the active site with the growing pyrrole chain at (*A*) DPM, (*B*) P3M, (*C*) P4M, (*D*) P5M, and (*E*) P6M stages. The interactions are shown by dotted lines. The cofactor is shown in yellow, and water molecules are represented as spheres. The carboxylate oxygen (marked as "O") of acetate/propionate side chains and the pyrrole nitrogen (marked as "N") for each of the pyrrole rings are labeled. The numeric suffixes (labeled 1–4) indicate the positions of the oxygen on the pyrrole, and the letter suffixes ("A" through "F") indicate the respective pyrrole rings from which the carboxylate oxygen is attached.

stages of chain elongation, it was observed that the polar and charged amino acid residues in the HMBS active site might not be sufficient to stabilize the negatively charged polypyrrole. For example, the acetate side chain of ring D in the P6M stage had no direct interaction with any polar or positively charged amino acid of the enzyme. To investigate possible water-mediated interactions, water molecules that could form hydrogen bonds with the carboxylate groups of the polypyrrole were identified. The number of hydrogen-bonded water molecules during subsequent stages of pyrrole chain elongation increased with the increasing number of carboxylate groups (Fig. 5). For P5M and P6M elongation, the number of water molecules increased significantly as a result of the opening of the active-site loop. These water molecules interacted with the negatively charged acetate and propionate side chains (Fig. 5). Table S1 summarizes the water-mediated stabilization for all charged groups of the polypyrrole. Any loss of charge stabilization from the residues around the active site could be compensated by water-mediated interactions (*Water-Mediated Interactions Stabilize the Negative Charges of the Polypyrrole* provides additional details).

R26 and R167 Are Critical Residues for Enzymatic Catalysis. As the catalytic mechanism is carried out in four successive steps (Fig. 1), the residues involved in the protonation of each incoming PBG molecule were investigated. The PBG moieties were docked in the active site at each stage of chain elongation. Based on the distance between the polar side chains of residues lining the active site near the terminal ring of the cofactor and the site of docked PBG, R26 was the most probable proton donor for the formation of P3M and P4M moieties (Fig. 6 *A* and *B*). However, the residues R26, Q34, and R195 cannot act as the proton donor for the formation of P5M and P6M stages because they are far from the docked position of the respective PBG moieties (Fig. S4). Therefore, R167, which is at a favorable position near the terminal rings of the docked PBG moieties (Fig. 6 *C* and *D*), is an ideal proton donor for the formation of the P5M and P6M moieties.

R167 Dynamics Is Important for HMB Exit. To investigate the exit mechanism of HMB from the catalytic cleft, the hexapyrrole was cleaved to form HMB, leaving the DPM cofactor attached to the protein ("HMB stage"). The HMB stage was simulated for 50 ns

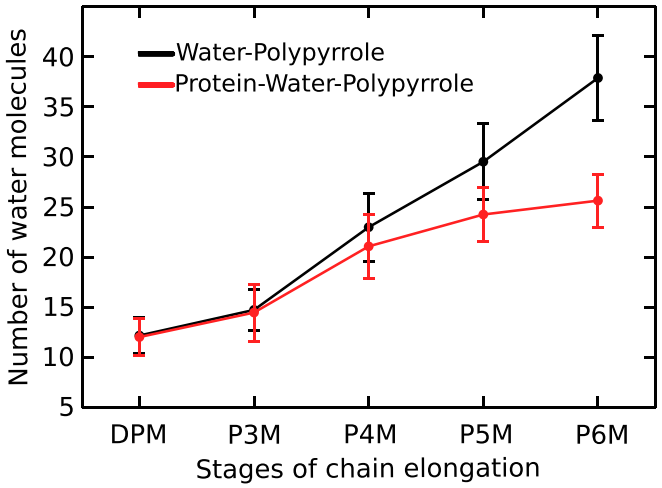


Fig. 5. Water-mediated interactions stabilize the negative charge on polypyrrole. Graph showing the total number of water molecules interacting with the polypyrrole during the stages of chain elongation (black); total number of water-mediated interactions between the growing polypyrrole and protein (red).

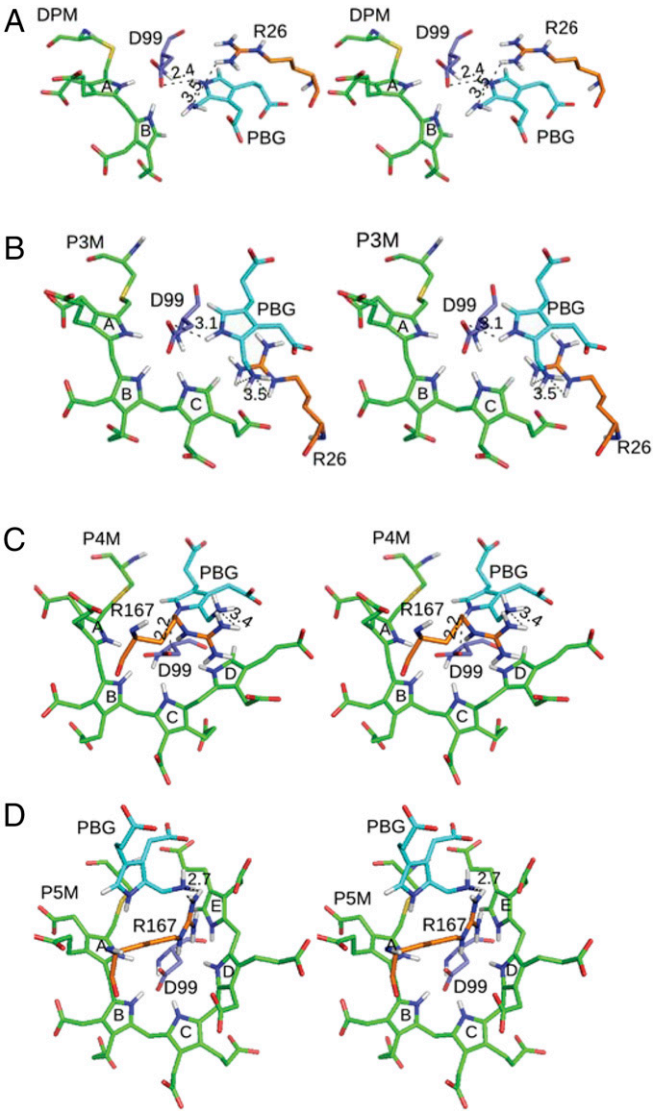


Fig. 6. R26, in DPM and P3M stages, and R167, in P4M and P5M stages, mediate the catalytic mechanism during substrate addition. Stereograms of PBG docked in the active site of hHMBS at (*A*) DPM, (*B*) P3M, (*C*) P4M, and (*D*) P5M stages showcase the role of R26 and R167 in proton donation to the incoming PBG. All measurements are in angstroms. The PBG, polypyrrole, arginine (R26 and R167), and aspartate (D99) residues are shown in cyan, green, orange, and violet colors, respectively.

to study the exit mechanism of HMB from the enzyme. During this simulation, the HMB moiety was relaxed and moved marginally away from the D99 residue (Fig. S5). As it would require a longer simulation time for the product to exit from hHMBS by natural diffusion, an enhanced sampling method, random acceleration MD (RAMD), was performed (16). RAMD simulation can predict the ligand exit path in an unbiased manner provided the force applied on the ligand is optimal. As noted earlier, a large number of hydrophilic and charged residues stabilized the acetate and propionate side chains of HMB. The external force required to break such hydrogen bonds/salt bridges would be large. From all 72 trajectories (*Methods*), it was seen that the applied force was a critical factor determining the successful exit of HMB. None of the trajectories with 10 kcal/mol/Å of force or less were successful. However, 50% of trajectories with a force magnitude of 20 or 25 kcal/mol·Å were successful. The simulation time for protein exit also depended on the applied

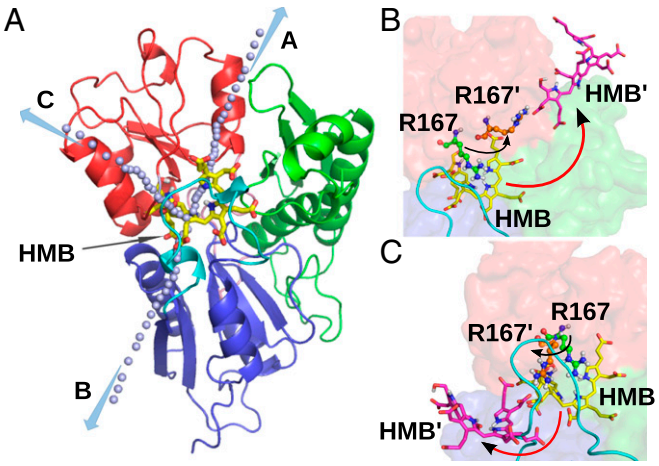


Fig. 7. Exit of HMB from the catalytic cleft of hHMBS. (*A*) Three possible exit paths for HMB are depicted using gray spheres. The position of the HMB is shown in sticks (yellow); the arrows are indicative of the probable exit paths A, B, and C. Exit path A is shown in *B* and exit path B is shown in *C*. Domains 1, 2, and 3 are represented in blue, red, and green, respectively, and the active-site loop is shown in cyan. R167 and HMB, in green and yellow, respectively, are shown in the stage at which HMB is beginning to exit from the protein. R167′ and HMB′, in orange and pink, respectively, represent the stage at which HMB is almost outside the protein. R167 acts as a gatekeeper for the HMB exit.

force. With an external force of 90 kcal/mol·Å, the ligand exit time was as low as 33 ps, whereas, with a force of 25 kcal/mol·Å, it took 1,068 ps. Analysis of all of the trajectories showed that there were minor modifications in each path studied as a result of randomly chosen force vector direction. However, all successful paths were clustered into three paths, labeled as A, B, and C (Fig. 7*A*). Among all of the successful trajectories, HMB exited through path A in 46% of cases, whereas the exits through paths B and C were 33% and 25%, respectively.

R167 acted as a gatekeeper during the exit of HMB in all of the three exit paths (Movie S3). Although a major conformational change was observed for R167 in all three paths, a few other charged residues also showed conformational changes, depending on the exit path studied. Until HMB was formed (P6M stage), the R167 side chain was oriented toward the active site. In all three exit paths, R167 interacted with the acetate and propionate side chains of HMB with large conformational fluctuations. This facilitated the exit of HMB from the enzyme (Fig. 7*A*). During the P6M stage, the active-site loop moved further away from the active site (Fig. 3*B*), thereby creating a channel for the exit of HMB from the enzyme. In path A, HMB used this channel for the exit (Fig. 7*B*). Apart from R167, R173 also interacted with HMB until it exited from the protein through path A. Path B was through the region between domains 1/2 and the active-site loop. The exit of HMB through path B required further active-site loop dynamics (Fig. 7*C*). In path B, K62 also interacted with HMB during its exit. HMB exit via path C was through the region between domain 2 and the active-site loop (Fig. 7*A*). The salt bridge between R167 and the acidic side chains of HMB caused a major distortion in domain 2 (rmsd change of 4 Å observed). This major structural distortion made path C less likely than paths A and B.

Mutagenesis and in Vitro Expression of Active-Site Residues Predicted by MD Simulations. To further support the results of the MD simulations, identified active-site residues were mutagenized and expressed in vitro to determine the effects of the amino acid substitutions of these residues on hHMBS enzymatic activity (Table 1). The five amino acids, T145, S146, S147, R149, and R173, are highly conserved in evolution and interacted with ring A (Fig. 4*A*). Amino acid substitutions T145I, T145N, S147P, R149L, R149Q, R173W, and R173Q essentially abolished hHMBS enzymatic activity (0.1–2.3% of expressed WT activity; Table 1). For S146, mutagenesis to cysteine (S146C) or glycine (S146G) only slightly impaired the enzyme activity (70% and 81% of expressed WT, respectively). However, mutagenesis of S146 to isoleucine (S146I) dramatically decreased the enzyme activity to ∼2.3% of expressed WT activity (Table 1). For the three highly conserved amino acids, K98, R150, R195, which directly interacted with ring B, amino acid substitutions significantly impaired hHMBS enzymatic activity (0.8–3% of expressed WT).

For the five amino acids, R26, S28, A31, Q34, and S96, that interacted with ring C (Fig. 4*B*), mutagenesis and in vitro expression of each resulted in severe impairment of hHMBS activity (0.2–1.2% of expressed WT; Table 1). For L170, which interacted with the polypyrrole through the amino acid backbone, mutagenesis of leucine to valine (L170V) caused a minimal reduction of expressed enzymatic activity (72.6%), whereas substitution of leucine to arginine (L170R) drastically reduced

Table 1. Mutagenesis and in vitro expression of hHMBS mutants with residues interacting with the polypyrrole during chain elongation

Polypyrrole ring	Amino acid*	Mutation	Expressed enzymatic activity†
Ring A	R149	R149L‡	0.1 ± 0.1
		R149Q‡	0.5 ± 0.9
	S147	S147P‡	0.2 ± 0.0
	T145	T145I	0.4 ± 0.1
		T145N	0.8 ± 0.6
	S146	S146C	69.9 ± 1.7
		S146G	81 ± 9
		S146I	2.3 ± 1.1
	R173	R173Q‡	0.4 ± 0.1
		R173W‡	0.7 ± 0.2
Ring B	K98	K98R‡	0.70 ± 1.1
	R150	R150I	0.02 ± 0.03
		R150Q	0.8 ± 1.2
	R195	R195C‡	3 ± 1
	G218	G218R‡	0.1 ± 0.1
Ring C	L170	L170V	72.6 ± 5.1
		L170R	0.6 ± 0.6
	S28	S28N‡	0.8 ± 0.4
		S28C	1.2 ± 0.2
	A31	A31T‡	0.50 ± 0.4
		A31P‡	0.6 ± 0.5
	R26	R26C‡	0.3 ± 0.3
		R26H‡	0.3 ± 0.3
	S96	S96F‡	1.0 ± 1.7
	Q34	Q34R‡	0.7 ± 0.3
		Q34K‡	0.2 ± 0.3
Ring D	N169	N169I	6.2 ± 0.5
	T58	T58I	0.70 ± 0.1
Ring E	R167	R167Q‡	1.0 ± 1.0
		R167W‡	2.3 ± 1.0
Ring F	S262	S262C	46.3 ± 8.6
	Q356	Q356E	95.7 ± 11.2
	S165	S165C	69.9 ± 1.7
Ring A-B-C-D-E-F	D99	D99G‡	3.2 ± 5.4
		D99H‡	3.3 ± 5.4

Mutation of the residues that were interacting with the polypyrrole during stages of chain elongation shows a significant reduction in enzymatic activity in the in vitro expression studies.
*A missense mutation of this amino acid is likely to cause AIP.
†Mean ± SD percentage of expressed WT activity.
‡HGMD reported mutations that cause AIP.

the enzymatic activity (0.6%; Table 1). This was unexpected, but it was possible that the positively charged arginine residue destabilized the local hydrophobic environment more drastically than the hydrophobic valine.

Of the six amino acids (T58, S165, R167, N169, S262, Q356) that were directly involved in the addition of rings D, E, and F, mutagenesis and in vitro expression of T58I, R167Q, R167W, and N169I significantly reduced hHMBS activity (0.7%, 1%, 2.3%, and 6.2% of expressed WT activity, respectively), indicating their importance in pyrrole elongation. In contrast, substitution of S165C, S262C, or Q356E resulted in somewhat reduced activity (70%, 46%, and 95%, respectively), suggesting that these residues were not critical for elongation (Table 1).

For residue D99, which was involved in all elongation steps, mutagenesis and in vitro expression confirmed that known mutations causing AIP (D99G and D99H) (20) markedly reduced hHMBS enzymatic activity (~3.2% and ~3.35% of expressed WT activity, respectively; Table 1).

MD Simulations of hHMBS Mutations That Impaired Chain Elongation. To further investigate the structural effect of D99G, R26C, and R167W active-site mutations, we carried out 50-ns simulations for each of the three catalytically important active-site residue mutations at DPM, P3M, and P5M stages of chain elongation, respectively (*Validation of Polypyrrole Conformation* and *MD Simulation of Three Catalytically Important Active-Site Residue Mutations*). The D99G mutation caused the cofactor to be displaced by 3.5 Å and the pyrrole nitrogens moved closer to each other, presumably as a result of loss of hydrogen bonds between pyrrole nitrogens and the D99 side chain (Fig. S6 and *MD Simulation of Three Catalytically Important Active-Site Residue Mutations*). In the P3M stage, in the presence of the R26C mutation, the polypyrrole was accommodated in a different conformation compared with that in WT hHMBS. The R26C-induced orientation of ring C disrupted further substrate addition (Fig. S7). Even though the R167W mutation did not show any major structural changes, its effect in causing AIP can be attributed to the low proton-donating propensity of tryptophan at physiologic pH.

Active-site mutations A31T, A31P, A31V, and S96F that have been reported to cause AIP (20) were deleterious based on the structural analysis of the growing pyrrole. The mutations were predicted to hinder the polypyrrole chain elongation as a result of severe steric clash with the terminal pyrrole ring at the P3M stage. Although G260 was not an active-site residue, G260D, which causes AIP (20), hindered the polypyrrole elongation in the P5M stage and was therefore deleterious. Similarly, Q217 did not interact with the polypyrrole ring at any stage, but was close to the active-site residues K98 and R150. Hence, mutations encoding Q217R and Q217H, which cause AIP (20), affected the side chain orientation of active-site residues K98 and R150, thereby disrupting hHMBS function.

Discussion

The specific residues and mechanisms by which hHMBS performed each of the stepwise additions of four molecules of PBG to the DPM cofactor were unknown as a result of the lack of structural data for the intermediate elongation stages. From the crystal structures of hHMBS, it was evident that the active site was not large enough to accommodate four PBG molecules (13, 14).

MD simulations revealed that the cofactor turn movement along with minor domain motions provided space for the addition of first two PBG moieties to the DPM cofactor in hHMBS. The movement of the active-site loop away from the active-site region was the necessary next step to facilitate the accommodation of the next two PBG moieties. Our previous MD study of EcHMBS (15) indicated the importance of the domain movements and the active-site loop dynamics in the formation of HMB. In EcHMBS, domains 1 and 2 moved toward domain 3 to create space for pyrrole elongation (Fig. 8 *A* and *C*) along with loop movement. However, hHMBS differed from EcHMBS by the presence of an additional 29-residue insert in domain 3. The 29-residue insert was wedged between domains 1 and 3, pressing domain 3 against domain 2, precluding the large movements of domains 1 and 2 observed for EcHMBS (Fig. 8 *B* and *D*). In hHMBS, the domain movement was compensated by the cofactor turn movement and the large active-site loop movement. It is important to note that hHMBS and EcHMBS have 53% sequence similarity and most of the active-site residues are conserved between these two species. Therefore, the environments for catalysis by EcHMBS and hHMBS were similar with notable exceptions: Q34 in hHMBS (Q19 in EcHMBS) interacted only with ring C at the P4M stage, whereas, in subsequent stages, the interactions were mediated by water molecules. In contrast, Q19 in EcHMBS interacted with ring C from the P3M stage onward. The role played by Q19 in EcHMBS was played mostly by S29 in hHMBS, as MD simulations suggested that S28 interacted with ring C from the P3M stage onward. In vitro studies of the S28N mutation further confirmed its importance. The limitation of steered MD, which was used to study EcHMBS and to specify the initial force direction vector, was overcome by using RAMD simulation to study HMB exit from hHMBS. Among the three most likely paths that were evaluated for HMB exit, the path between the interface of domains and the active-site loop was the most probable for hHMBS.

It was noted that there were eight and six bound water molecules around the charged side chains of the DPM cofactor in the EcHMBS and AtHMBS crystal structures (7, 10), respectively, whereas only one water molecule was reported in the hHMBS crystal structure (14). Although water molecules are present on each side of the DPM (9, 10) and have been hypothesized to stabilize the DPM-hHMBS structure, these studies demonstrate the important role of water molecules in the subsequent steps of

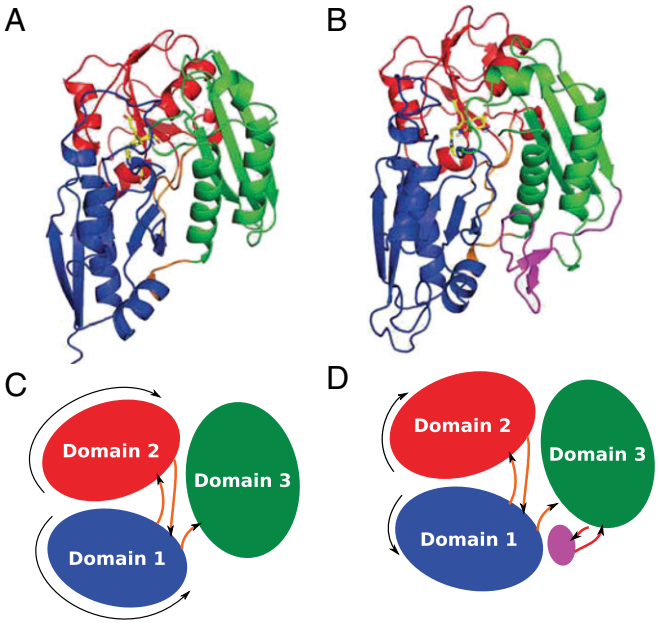


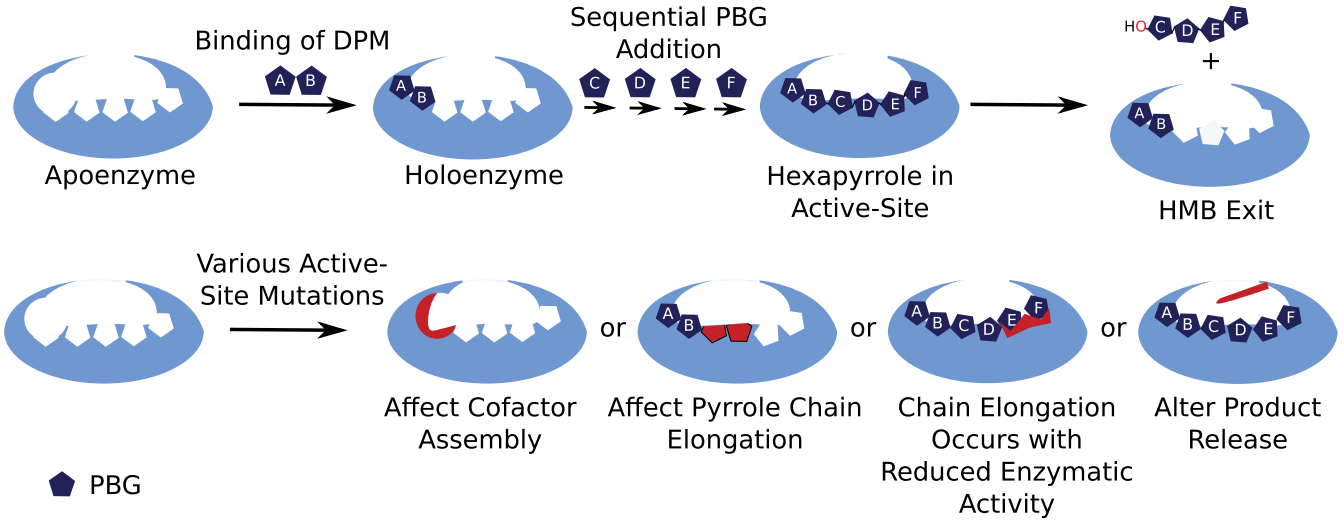
Fig. 8. Effect of the 29-residue insert on domain dynamics. Representative structure of the DPM stage in (*A*) EcHMBS and (*B*) hHMBS. Domains 1, 2, and 3 are colored in blue, red, and green, respectively. The 29-residue insert is shown in magenta color. Schematic representation of the domain dynamics in (*C*) EcHMBS and (*D*) hHMBS. The 29-residue insert prevents the motion of domains 1 and 2. The length of the black arrows in *C* and *D* is proportional to the extent of the domain motions observed in EcHMBS and hHMBS, respectively.

polypyrrole elongation. The MD simulations also indicated that any loss of charge stabilization caused by amino acid substitutions in the P4M to P6M steps could be compensated by water-mediated interactions.

Notably, the MD simulations identified the critical active-site residues involved in DPM binding and subsequent steps of polypyrrole elongation at atomistic levels. Amino acid residues from all three domains interacted with the polypyrrole. Totals of 11, 14, and 3 residues from domains 1, 2 and 3, respectively, took part in stabilizing the highly negatively charged polypyrrole. These studies highlighted previously unknown roles of R167 in hHMBS catalysis and HMB exit. Not only was R167 a critical proton donor for PBG, in all three major exit paths, R167 also remained salt-bridged to one of the carboxylates of the polypyrrole chain and facilitated HMB exit. Thus, the mechanistic findings predicted the pathogenic basis of the common R167Q and R167W mutations that severely impair hHMBS activity (Table 1) and cause AIP (20) by failing to elongate the pyrrole chain or to facilitate the HMB exit. These mutations are among the most frequently identified in patients with AIP (20).

These studies also provided the molecular basis for reported missense AIP-causing mutations that impaired chain elongation (Fig. 9). Of the 136 reported missense mutations causing AIP (20), 27 mutations (19.8%) were found at the active-site residues, 11 of which altered eight residues that were critical to the covalent binding of the DPM cofactor to apo-hHMBS (Table 1, rings A and B); 11 other mutations altered 5 residues involved in the addition of the third pyrrole (ring C, P3M stage; Fig. 9); and R167Q and R167W were involved in the addition of rings E and F. There was no reported missense hHMBS mutation involved in the addition of ring D, with the exception of three reported D99 mutations (D99N, D99G, and D99H) (20), which impaired the addition of all six PBG molecules. Interestingly, the three reported mutations of D99 did not totally prevent HMB assembly, as D99G and D99H had approximately 3% of expressed WT activity. The MD simulations showed that the hexapyrrole was able to form a compact helicoidal conformation within the tight active site as a result of the interactions between D99 and all of the nitrogens on the six pyrrole rings. It is notable that an analogous aspartate (D86) in the heme biosynthetic enzyme, uroporphyrinogen decarboxylase, was shown to coordinate with the cyclic uroporphyrinogen III substrate and presumably stabilized the intermediates formed during the enzyme-catalyzed decarboxylation of its substrate (21).

In summary, these studies describe the atomistic details of the hHMBS catalytic mechanism and provide insights into the molecular basis of active-site mutations that cause AIP (Fig. 9). The active-site loop dynamics and cofactor turn movement in hHMBS reorganized the active-site region such that an arginine was present near the terminal ring for protonation of the PBG substrate. The presence of the 29-residue insert in hHMBS drastically reduced the domain motions in comparison with EcHMBS enzyme. These studies also delineate the role played by the active-site residues and water molecules in binding and



Type of Alteration	Critical Active-Site Residue	Consequence
Altered Cofactor Binding Site	K98, D99, T145, S147, R149, R150, R173, R195	No Cofactor Assembly
Altered PBG Binding Site	R26, S28, Q34, T58, S96, R167	No Pyrrole Chain Elongation
Affecting Polypyrrole Charge stabilization	S165, N169, S262	Reduced Enzymatic Activity
Affecting Product Release	R167	Product Release Blocked

Fig. 9. Effects of mutations on the active-site residues. Classification of amino acid mutations that are responsible for altering (*A*) cofactor binding, (*B*) incoming PBG binding sites for pyrrole chain elongation, (*C*) polypyrrole charge stabilization, and (*D*) HMB release.

stabilizing the negative charges of the elongating polypyrrole at each stage. In addition to the active-site residues, these water molecules may also act as proton donors. To date, 27 mutations of 14 of the 27 highly conserved active-site residues have been reported to cause AIP (20). Notably, R167 played a crucial role in enzyme catalysis and facilitated HMB exit. Finally, these studies provide a starting point for quantum chemical calculations. The intermediate snapshots at different pyrrole elongation stages should be extremely useful to determine the reaction coordinates and to investigate the enzyme's reaction mechanism at an electronic level.

Methods

Preparation of Starting Structures. The structure of WT hHMBS [PDB ID code 3ECR (14)] was used as the basis for molecular modeling. The missing coordinates of the residues 60–78 (part of the active-site loop and adjoining region) were modeled based on the structure of AtHMBS [PDB ID code 4HTG (10)] using Modeler 9v8 (22). The missing residues at the N-terminal (1–20) and the C-terminal (residues 358–361) were also modeled.

The starting structures for each of the intermediate stages were generated by docking a moiety of methylene pyrrolenine (MePy) in the active-site cavity by using Autodock Vina (23) and covalently attaching it to the polypyrrole of the protein. The resultant structure was energy-minimized to ensure that there were no bad contacts. For each stage, a new pyrrole ring was oriented in such a manner that its nitrogen atom could interact with the side chain of D99 positioned between domains 1 and 2.

Pyrrole Chain Elongation. The holoenzyme and the intermediate stages of chain elongation were simulated by using Gromacs 4.5.5 (24) with G53a6 united-atoms force field (25). The ATB server (26) was used to obtain parameters for the covalently attached DPM cofactor and subsequent chain extensions, P3M, P4M, P5M, and P6M moieties. The systems were soaked in an octahedral solvent box with 9 Å padding of SPC/E (extended simple point charge model) water molecules (27). The net charge on the system was neutralized by adding sodium ions. The steepest descent algorithm was used for energy minimization until convergence was reached. The procedure for NVT, NPT, and the production run simulations were similar to that described by Bung et al. (15). Each of the intermediate stages of pyrrole elongation was simulated for 50 ns.

RAMD. RAMD (28, 29) was used to study the exit path of HMB from hHMBS. RAMD produces unbiased pathways on a nanosecond timescale by enhancing the dynamics of the ligand (29–31). During RAMD simulations, an external force (F_{ext}) was applied to the center of mass of the ligand in a random direction for a predetermined number of steps (n).

$$F_{ext} = F_0 \hat{r}$$

Here, F_0 is the constant force acting on the ligand with unit vector $\hat{r}$. In practice, random acceleration is provided as input based on the mass of the ligand. The ligand keeps a certain velocity during the chosen step, n. After completion of n steps, the distance between the centers of mass of ligand and protein was computed. If this distance was less than the user-defined minimum distance (d_{min}) that the ligand is supposed to travel, a new random direction was chosen. All RAMD simulations were performed by using the CHARMM22 force field (CgenFF) (32) and the NAMD program (33). The protein was solvated in a cubic box, and ions were added to ensure the charge neutrality of the system. The force field parameters of HMB and the DPM cofactor were prepared by using CgenFF. The TIP3P water model (34) was used to solvate the system. The RAMD simulations were preceded by a short energy minimization step and 2 ns of equilibration. Several trial runs were performed with an external force ranging from 90 kcal/mol/Å to 5 kcal/mol/Å with an interval of 5 kcal/mol/Å. The step values (n) of 40 or 80 and the minimum distance (d_{min}) of 0.008 or 0.016 Å were tested. For each force value, four simulations were performed by varying the n and d_{min}.

Trajectory and Structural Analyses. VMD 1.9.1 (35) and Gromacs 4.5.5 (24) were used to analyze the trajectories. RMSFs of the Cα atoms of the entire protein were calculated by using the g_rmsf module of Gromacs. The energy-minimized starting structure of DPM stage was considered as the reference structure. The concatenated trajectory of all of the stages of pyrrole chain elongation was used to analyze the cumulative effect of domain motions. PCA was performed on backbone atoms by using the Normal Mode Wizard plugin of VMD (36). Dynamic cross-correlation maps (19) were calculated for Cα atoms from the concatenated trajectory with respect to the reference structure.

The distance between domains 1 and 2 was computed by using the centers of mass of interface residues from domain 1 (residues 25–38) and domain 2 (residues 165–175 and 193–198). The volume of the active-site cavity along each trajectory was calculated by using Epock (37). Hydrogen bond interactions of the growing pyrrole chain with the protein and solvent were calculated by using the hbond tool in VMD. For hydrogen bond calculations, a donor (D)–acceptor (A) cutoff distance of 3.5 Å and D-H···A angle cutoff of 30° was considered. To study the water-mediated interactions, water molecules that lie within 3.5 Å of polar groups of the growing pyrrole chain and the protein molecule were selected. The cooccurrence probability to characterize the loop conformation was represented as a 2D histogram and plotted by using R (38).

Mutagenesis and in Vitro Expression Studies. The hHMBS cDNA constructs for each of the 35 reported and predicted missense variants were individually generated by using the QuikChange Lightning Single-Site Mutagenesis kit (Agilent Genomics) to alter the h*HMBS* WT cDNA in the pKK223 vector (39). Each construct was resequenced to confirm its authenticity. WT and mutant constructs were each expressed in *E. coli* strain BL21(DE3)pLysS (Promega), which had low endogenous HMBS activity. The enzymatic activity of each mutant enzyme in the lysate was calculated as the percentage of the expressed WT activity, which was separately expressed in the same experiment as described previously (39). All results are presented as the mean activities and SDs of at least three independent experiments.

ACKNOWLEDGMENTS. We thank Dr. Alpesh Malde (University of Queensland) for help with the force field parameters of the pyrrole moieties, Dr. Monika Sharma (Indian Institute of Science Education and Research, Mohali) for help with DCCM calculations, and Dr. Siladitya Padhi (TCS Innovation Labs–Hyderabad) for helpful comments and suggestions.

1. Hamza I, Dailey HA (2012) One ring to rule them all: Trafficking of heme and heme synthesis intermediates in the metazoans. *Biochim Biophys Acta* 1823: 1617–1632.
2. Anderson KE, Sassa S, Bishop DF, Desnick RJ (2001) Disorders of heme biosynthesis: X-linked sideroblastic anemia and the porphyrias. *The Metabolic and Molecular Bases of Inherited Disease*, eds Valle D, et al. (McGraw-Hill, New York), Chap 124. Available at ommbid.mhmedical.com/content.aspx?bookid=971&Sectionid=62638866.
3. Bogorad L (1958) The enzymatic synthesis of porphyrins from porphobilinogen. I. Uroporphyrin I. *J Biol Chem* 233:501–509.
4. Anderson PM, Desnick RJ (1980) Purification and properties of uroporphyrinogen I synthase from human erythrocytes. Identification of stable enzyme-substrate intermediates. *J Biol Chem* 255:1993–1999.
5. Anderson PM, Reddy RM, Anderson KE, Desnick RJ (1981) Characterization of the porphobilinogen deaminase deficiency in acute intermittent porphyria. Immunologic evidence for heterogeneity of the genetic defect. *J Clin Invest* 68:1–12.
6. Jordan PM, Warren MJ (1987) Evidence for a dipyrromethane cofactor at the catalytic site of E. coli porphobilinogen deaminase. *FEBS Lett* 225:87–92.
7. Louie GV, et al. (1996) The three-dimensional structure of Escherichia coli porphobilinogen deaminase at 1.76-A resolution. *Proteins* 25:48–78.
8. Shoolingin-Jordan PM, Al-Dbass A, McNeill LA, Sarwar M, Butler D (2003) Human porphobilinogen deaminase mutations in the investigation of the mechanism of dipyrromethane cofactor assembly and tetrapyrrole formation. *Biochem Soc Trans* 31:731–735.
9. Louie GV, et al. (1992) Structure of porphobilinogen deaminase reveals a flexible multidomain polymerase with a single catalytic site. *Nature* 359:33–39.
10. Roberts A, et al. (2013) Insights into the mechanism of pyrrole polymerization catalysed by porphobilinogen deaminase: High-resolution X-ray studies of the Arabidopsis thaliana enzyme. *Acta Crystallogr D Biol Crystallogr* 69:471–485.
11. Azim N, et al. (2014) Structural evidence for the partially oxidized dipyrromethene and dipyrromethanone forms of the cofactor of porphobilinogen deaminase: Structures of the Bacillus megaterium enzyme at near-atomic resolution. *Acta Crystallogr D Biol Crystallogr* 70:744–751.
12. Guo J, Erskine P, Coker AR, Wood SP, Cooper JB (2017) Structural studies of domain movement in active-site mutants of porphobilinogen deaminase from Bacillus megaterium. *Acta Crystallogr F Struct Biol Commun* 73:612–620.
13. Gill R, et al. (2009) Structure of human porphobilinogen deaminase at 2.8 A: The molecular basis of acute intermittent porphyria. *Biochem J* 420:17–25.
14. Song G, et al. (2009) Structural insight into acute intermittent porphyria. *FASEB J* 23: 396–404.
15. Bung N, Pradhan M, Srinivasan H, Bulusu G (2014) Structural insights into E. coli porphobilinogen deaminase during synthesis and exit of 1-hydroxymethylbilane. *PLoS Comput Biol* 10:e1003484.
16. Jordan PM, Woodcock SC (1991) Mutagenesis of arginine residues in the catalytic cleft of Escherichia coli porphobilinogen deaminase that affects dipyrromethane cofactor assembly and tetrapyrrole chain initiation and elongation. *Biochem J* 280:445–449.

17. Lander M, et al. (1991) Studies on the mechanism of hydroxymethylbilane synthase concerning the role of arginine residues in substrate binding. *Biochem J* 275:447–452.
18. Schlick T (2010) *Molecular Modeling and Simulation: An Interdisciplinary Guide*, ed Schlick T (Springer, New York).
19. Harte WE, Jr, et al. (1990) Domain communication in the dynamical structure of human immunodeficiency virus 1 protease. *Proc Natl Acad Sci USA* 87:8864–8868.
20. Stenson PD, et al. (2017) The human gene mutation database: Towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. *Hum Genet* 136:665–677.
21. Phillips JD, Whitby FG, Kushner JP, Hill CP (2003) Structural basis for tetrapyrrole coordination by uroporphyrinogen decarboxylase. *EMBO J* 22:6225–6233.
22. Fiser A, Sali A (2003) Modeller: Generation and refinement of homology-based protein structure models. *Methods Enzymol* 374:461–491.
23. Trott O, Olson AJ (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 31:455–461.
24. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4:Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* 4:435–447.
25. Oostenbrink C, Villa A, Mark AE, van Gunsteren WF (2004) A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J Comput Chem* 25:1656–1676.
26. Malde AK, et al. (2011) An automated force field topology builder (ATB) and repository: Version 1.0. *J Chem Theory Comput* 7:4026–4037.
27. Berendsen HJC, Grigera JR, Straatsma TP (1987) The missing term in effective pair potentials. *J Phys Chem-Us* 91:6269–6271.
28. Lüdemann SK, Lounnas V, Wade RC (2000) How do substrates enter and products exit the buried active site of cytochrome P450cam? 1. Random expulsion molecular dynamics investigation of ligand access channels and mechanisms. *J Mol Biol* 303: 797–811.
29. Vashisth H, Abrams CF (2008) Ligand escape pathways and (un)binding free energy calculations for the hexameric insulin-phenol complex. *Biophys J* 95:4193–4204.
30. Winn PJ, Lüdemann SK, Gauges R, Lounnas V, Wade RC (2002) Comparison of the dynamics of substrate access channels in three cytochrome P450s reveals different opening mechanisms and a novel functional role for a buried arginine. *Proc Natl Acad Sci USA* 99:5361–5366.
31. Long D, Mu Y, Yang D (2009) Molecular dynamics simulation of ligand dissociation from liver fatty acid binding protein. *PLoS One* 4:e6081.
32. Vanommeslaeghe K, Raman EP, MacKerell AD, Jr (2012) Automation of the CHARMM general force field (CGenFF) II: Assignment of bonded parameters and partial atomic charges. *J Chem Inf Model* 52:3155–3168.
33. Phillips JC, et al. (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26: 1781–1802.
34. Jorgensen WL, Chandrasekhar J, Buckner JK, Madura JD (1986) Computer simulations of organic reactions in solution. *Ann N Y Acad Sci* 482:198–209.
35. Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. *J Mol Graph* 14:33–38.
36. Bakan A, Meireles LM, Bahar I (2011) ProDy: Protein dynamics inferred from theory and experiments. *Bioinformatics* 27:1575–1577.
37. Laurent B, et al. (2015) Epock: Rapid analysis of protein pocket dynamics. *Bioinformatics* 31:1478–1480.
38. Team R (2014) R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna).
39. Chen CH, Astrin KH, Lee G, Anderson KE, Desnick RJ (1994) Acute intermittent porphyria: Identification and expression of exonic mutations in the hydroxymethylbilane synthase gene. An initiation codon missense mutation in the housekeeping transcript causes "variant acute intermittent porphyria" with normal expression of the erythroid-specific enzyme. *J Clin Invest* 94:1927–1937.