Published in final edited form as:

J Mol Biol. 2010 November 19; 404(1): 127–137. doi:10.1016/j.jmb.2010.09.042.

High-Resolution Structure of the Nitrile Reductase QueF Combined with Molecular Simulations Provide Insight into Enzyme Mechanism

Youngchang Kim¹, Min Zhou¹, Shiu Moy¹, Jennifer Morales², Mark A. Cunningham^{2,*}, and Andrzej Joachimiak^{1,*}

¹The Midwest Center for Structural Genomics and Structural Biology Center, Biosciences, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, IL 60439, USA

²Physics and Geology, The University of Texas-Pan American, 1201 West University Drive, Edinburg, TX 78539, USA

Abstract

Here, we report the 1.53-Å crystal structure of the enzyme 7-cyano-7-deazaguanine reductase (QueF) from *Vibrio cholerae*, which is responsible for the complete reduction of a nitrile (C≡N) bond to a primary amine (H₂C–NH₂). At present, this is the only example of a biological pathway that includes reduction of a nitrile bond, establishing QueF as particularly noteworthy. The structure of the QueF monomer resembles two connected ferredoxin-like domains that assemble into dimers. Ligands identified in the crystal structure suggest the likely binding conformation of the native substrates NADPH and 7-cyano-7-deazaguanine. We also report on a series of numerical simulations that have shed light on the mechanism by which this enzyme affects the transfer of four protons (and electrons) to the 7-cyano-7-deazaguanine substrate. In particular, the simulations suggest that the initial step of the catalytic process is the formation of a covalent adduct with the residue Cys194, in agreement with previous studies. The crystal structure also suggests that two conserved residues (His233 and Asp102) play an important role in the delivery of a fourth proton to the substrate.

Keywords

queuosine; oxidoreductase; QueF; nitrile reduction

Introduction

Transfer RNAs (tRNAs) undergo significant post-transcriptional processing, resulting in a variety of structural changes¹ to the canonical nucleosides that provide structural stabilization and improved translational fidelity.^{2,3} One of the heavily modified nucleosides is queuosine,⁴ a guanosine derivative synthesized by bacteria, which is found in the third (wobble) position of the anticodon in tRNA^{Asn}, tRNA^{Asp}, tRNA^{His}, and tRNA^{Tyr}, where it presumably improves the fidelity of translation. Queuosine is utilized but not synthesized in eukaryotic cells, making the queuosine biosynthetic pathway the subject of continued

© 2010 Published by Elsevier Ltd.

*Corresponding authors. cunningham@utpa.edu; andrzej@anl.gov.

Accession numbers

Atomic coordinates and experimental structure factors of this structure have been deposited in the PDB with accession code 3BP1.

Supplementary materials related to this article can be found online at doi:10.1016/j.jmb.2010.09.042

interest: disruption of the pathway in *Shigella flexneri* results in loss of pathogenicity,⁵ for example. Development of a better understanding of the biochemistry involved could well lead to new antibiotic agents.

A key component of the queuosine synthetic pathway shown in Fig. 1 is the enzyme 7-cyano-7-deazaguanine reductase (QueF), which converts the intermediate 7-cyano-7-deazaguanine (preQ₀) to 7-aminomethyl-7-deazaguanine (preQ₁). There are two related sequence subfamilies of QueF enzymes: YqcD (two-domain proteins) and YkvM (single-domain proteins).⁶ The YqcD subfamily, to which the enzyme from *Vibrio cholerae* belongs, is highly conserved across species. BLAST analysis of the protein sequence identified several hundred homologs, of which a representative set of proteins from both subfamilies is aligned in Fig. S1. These proteins (YqcD and YkvM) show sequence similarity to the family of GTP cyclohydrolases I (PF01227) but recent biochemical studies of the orthologous proteins from the YqcD subfamily from *Escherichia coli* (with 63% sequence identity) and YkvM subfamily from *Bacillus subtilis* (33% sequence identity) demonstrated that QueF proteins are NADPH-dependent 7-cyano-7-deazaguanine oxidoreductases.^{7,8} The YqcD subfamily of QueF enzymes is composed of two conserved domains: the N-terminal sequence is categorized as a member of the COG2904 superfamily that is composed of proteins of unknown function and the C-terminal sequence, which encompasses the catalytic residues, is classified as a tunneling fold (T-fold) superfamily. The T-fold is a common structural motif (composed of four β -strands and two α -helices) that oligomerizes to form a $\beta_2\alpha_2$ -type β -barrel (somewhat resembling a TIM barrel unit of $\beta_2\alpha_2$) and is present in a number of different enzymes acting on flat substrates of the purine or pterin families.⁹ The amino acid sequence of the YkvM subfamily maps into the T-fold C-terminal region of the YqcD subfamily. A recent study elucidates details of the enzymatic synthesis of the preQ₀ intermediate; the initial step from GTP to preQ₀, as illustrated in Fig. 1, actually requires four separate enzymatic steps.¹⁰ The enzymes involved, QueC, QueD, and QueE, are part of the same operon as QueF.⁶

The conversion of preQ₀ to preQ₁ catalyzed by QueF requires two molecules of NADPH for each molecule of preQ₁ synthesized and a potential chemical pathway has been proposed (Fig. 2).⁷ As part of the Protein Structure Initiative effort, we have determined a crystal structure of the enzyme QueF from *V. cholerae* via X-ray crystallography, which provides a number of insights into the structure of the enzyme and catalytic mechanism. Additionally, we have conducted a number of numerical simulations of the enzyme, utilizing molecular dynamics and hybrid, quantum/classical (QM/MM) calculations that further shed light on the catalytic mechanism.

Results and Discussion

The structure of QueF from *V. cholerae* was determined at 1.53 Å, and its monomer and dimer architecture and active sites are shown in Fig. 3. The protein shares 63% sequence identity with the experimentally characterized QueF protein from *E. coli*, and therefore, it is likely to possess the same enzymatic activity and serves as a good model for structural studies of the family. The protein crystallizes as a tetramer, a dimer of dimers (Fig. S2). Solution data (Fig. S3) suggest that the protein is a dimer and it may exist in equilibrium with monomer. Size-exclusion chromatography showed two peaks and a lower molecular weight shoulder. The two peaks correspond to molecular masses of 67.9 kDa (dimer) and 43.7 kDa (1.33×monomer) and a shoulder of ~33 kDa (monomer), suggesting dimer/monomer equilibrium in solution.

Each monomer is composed of a seven-stranded, antiparallel β -sheet, eight α -helices, and several loops. Seven helices are on the concave side of the β -sheet and one 1-turn helix is on

the convex side. The monomer seems to be formed by duplication of a simpler ferredoxin-like structural module. One unit is composed of a three-stranded β -sheet and two α -helices (residues 57–136, part of the COG2904 superfamily sequence) and the other is composed of a four-stranded β -sheet with two α -helices (residues 176–271, similar to a T-fold) (Fig. 3).

Therefore, the QueF monomer is made up of two ferredoxin-like domains aligned together with their β -sheets that have additional embellishments. An example of such embellishment is a 30-residue loop (residues 138–168) that connects two ferredoxin-like sub-domains and crosses over to the convex side of the β -sheet. These loops contribute to the dimer interface as well as to the formation of the tetramer in the crystal (Fig. S2). Each loop makes contact with the loops from the other three monomers. QueF was co-crystallized with GTP, but only guanine, phosphate, and pyrophosphate molecules were observed in the crystal structure. These are depicted as space-filling representations in Fig. 3. One guanine molecule is bound to each of the four monomers in the crystal asymmetric unit (for clarity, only two are shown in Fig. 3), in the cavity formed near the dimer interface on the concave side of each β -sheet between two sub-domains. The two pyrophosphates bind on the interface between two monomers in the dimer, at equal distance from two guanines and are presumably positioned close to where the diphosphate moiety of NADPH binds. This arrangement suggests that all four sites in a tetramer can be occupied by a substrate, but only one molecule of NADPH cofactor can bind to a dimer (two orientations are possible, see below) and can serve only one of the two sites occupied by substrate. Therefore, as is consistent with QueF solution studies and previous reports by Van Lanen *et al.*,⁶ the dimer is a minimal biological unit and QueF should exhibit half-site reactivity.

Figure 4 displays one of the guanine molecules in the active site where it mimics the binding of a substrate: preQ₀. The guanine is bound into a groove between two ferredoxin-like modules and is sandwiched between the two N-termini helices α_2 and α_5 . It is precisely oriented by a series of hydrogen bonds and a ring-stacking interaction with a phenylalanine residue (Phe232). The guanine N¹ and N² nitrogen atoms are coordinated by the carboxyl moiety of Glu234, which is strictly conserved (Fig. S1). From the chain A monomer, we observe the distances to be $d(N^1-O^{e1})=2.81$ Å and $d(N^2-O^{e2})=2.80$ Å, with similar distances (± 0.03 Å) for other monomers. The backbone carbonyl oxygen atom of Ile93, which is sometimes substituted by valine, also coordinates the guanine N² nitrogen atom, with $d(N^2-O)=2.74$ Å. The guanine N³ nitrogen atom is coordinated by the amide nitrogen atom of the Ser95 residue $d(N^3-N)=2.85$ Å and the guanine N⁹ nitrogen atom is coordinated by the carboxylate moiety of Glu94 $d(N^9-O^{e2})=2.74$ Å; both residues are strictly conserved (Fig. S1). The hydroxyl moiety of Ser95, found in two conformations in the crystal structure, also interacts with the N⁹ nitrogen atom of guanine: $d(N^9-O\gamma)=3.03$ Å. Finally, the guanine O⁶ oxygen atom is coordinated by the backbone amide nitrogen of the strictly conserved His233 residue with $d(O^6-N)=2.77$ Å. We observe the His233 residue in close proximity to the strictly conserved Asp201. The hairpin loop near the active site containing the catalytic⁶ Cys194 is disordered in all four monomers.

N-Glycosylase activity

We note that, as depicted in Fig. 4, the ligand observed in our X-ray structure is guanine and not GTP: there is no electron density corresponding to the ribophosphate moiety of GTP, suggesting that QueF retains some residual nucleosidase activity. Indeed, the QueF protein shows sequence similarity to GTP cyclohydrolases and was initially categorized as such.⁶ Van Lanen *et al.* reported that the QueF proteins from *B. subtilis* and *E. coli* do not have cyclohydrolase activity, which includes hydrolysis of the N⁹-C^{1'} bond, but are oxidoreductases.⁶ A Dali search¹¹ of the Protein Data Bank (PDB) for proteins structurally similar to QueF finds several GTP cyclohydrolase I proteins with *Z* scores above 3. The closest is GTP cyclohydrolase I from *E. coli* (PDB ID: 1A9C; *Z*=6.7, RMSD of 11.2 Å

between C^α atoms by SSM). The structure of the dimer of the cyclohydrolase with bound GTP is depicted in Fig. S3. This enzyme monomer also has one ferredoxin-like domain composed of a four-stranded β-sheet with two α-helices and assembles to a pentamer. The GTP binds in a groove between two ferredoxin-like modules coming from two adjacent subunits (Fig. S4). The guanine moiety of the GTP rests in a pocket between two α-helices in a mode reminiscent of guanine binding to QueF, but, unlike what we observed in the QueF structure, each of the α-helices is provided by a separate monomer. Moreover, the cyclohydrolase active-site pocket has a very different amino acid composition; for example, two of the three cyclohydrolase His residues are not present in QueF. The catalytic mechanism of *N*-glycosyl bond cleavage by the GTP cyclohydrolase I enzymes utilizes a Zn ion to facilitate the reaction.¹² We find no such evidence for ordered metal ion binding in our structure.

Nevertheless, the crystal structure of QueF suggests potential activity of the *N*-glycosylase reaction. From Fig. 4, the conserved residue His233 is in an appropriate position to donate a proton to the N⁷ nitrogen of the guanine moiety of GTP. Protonation of the nucleotide will result in cleavage of the N⁹–C¹ bond.¹³ This reaction can potentially be facilitated by the conserved residue Arg262; this residue is not oriented appropriately in the crystal structure but molecular dynamics simulations indicate that motion of the Arg262 side chain can lead to catalytically competent conformations of these key residues. Both His233 and Arg262 are contributed by the C-terminal catalytic domain. This work is still ongoing and will be reported in a subsequent communication. Recently, we have solved the structure of QueF with ATP (data not shown), and in the structure, we observe adenine and pyrophosphate molecules, similar to the structure reported here. Therefore, it is possible that the apparent QueF nucleosidase activity may be the residual reaction carried out in this unique QueF catalytic site. This activity is not essential for the primary function of preQ₀ reduction, and because of the unique nature of the product and its benefits to bacteria, it is being tolerated. Nevertheless, we cannot exclude, at present, other possible explanations for the *N*-glycosylase activity, such as (i) the presence of a contaminating nucleosidase activity in the QueF preparation or (ii) the presence of guanine and pyrophosphate in the commercial GTP preparation. We note that the substrate NADPH also has an *N*-glycosyl bond, which potentially could impact its binding and integrity during reaction (see discussion below).

Covalent adduct

It has been proposed that the initial step of the reduction of PreQ₀ involves formation of a covalent adduct through the nucleophilic attack of a conserved cysteine residue (Cys194 in *V. cholerae*) on the cyano group of the substrate (Fig. 2).⁶ This covalent bond can hold the substrate in place while two NADPH molecules are recruited to the active site to donate hydrides to the substrate. The catalytic Cys194 residue resides on the hairpin loop between strands β₇ and β₈ of the protein (Fig. S1), which is disordered in all four monomers in the crystal structure. To investigate the formation of the covalent adduct, we have conducted a number of numerical simulations.

Simulation models were constructed from the atomic coordinates of the crystal structure and solvated in a large box of water. The box dimensions were chosen to provide a diffuse limit, in which protein atoms do not interact directly with protein atoms in the periodic images. Sodium and chlorine ions were added to the simulation model to provide overall charge neutrality and an ionic strength of 0.15 mM. Molecular dynamics simulations were conducted with the CHARMM 27 force field, using the program NAMD. The tetramer, as observed in the crystal structure, was used in the simulations, providing four nominally equivalent sites and two biological units. Simulations included four preQ₀ molecules in addition to the protein to investigate formation of the covalent adduct. Missing residues were built into the structure with the PSFGEN module of VMD. Not surprisingly, and

consistent with the crystal structure, the simulations demonstrated significant mobility of the hairpin loop. This is illustrated in Fig. S5, in which a single time step from a 2-ns simulation is depicted. The resulting four models of monomers have been stacked by aligning their protein backbones away from the loop segment.

We identified one conformation from these simulations in which the Cys194 was oriented in an appropriate attack conformation as the starting point for a series of QM/MM calculations using the program NWChem. Molecular mechanics atoms utilized the AMBER force field, and quantum atoms were treated via the density function method B3LYP, using an Ahlrichs pVDZ basis set. A reactant state was defined by optimizing the structure from the molecular dynamics trajectory, keeping atoms more than 12 Å from the C¹⁰ carbon atom of the substrate fixed. A product state was defined by optimizing the structure with constraints on the sulfur–carbon (S–C¹⁰) distance to form the bond and then relaxing the constraints and re-optimizing. The product-state conformation is shown in Fig. 5. A minimum-energy pathway between reactant and product states was defined by the nudged elastic band (NEB) method, using 10 replicas (beads) of the system, resulting in the pathway shown in Fig. 6. We observe a large barrier between reactant and product states that we believe is due to a relatively poor choice for the reactant state. In this conformation, there are no other residues in proximity to the catalytic cysteine to aid the reaction. It is possible that additional dynamics simulations will identify a better conformation of the mobile loop and reduce the observed barrier. We observe, however, that the relative barrier to the serine-catalyzed reaction is significantly higher than the cysteine reaction (~80 kJ/mol). This is consistent with the observation in *B. subtilis* that the equivalent C55S mutation resulted in loss of catalytic function.⁷

NADPH binding

In the reduction of preQ₀, two hydrides are transferred to the substrate sequentially from two molecules of NADPH. We have investigated binding of the NADPH molecules via molecular dynamics simulations. From Fig. 3, there are two equivalent sites; each is occupied by a guanine molecule and potentially each could form a covalent adduct. We have conducted a series of simulations with preQ₀ and adenine molecules in the active sites, initially stacking the adenine molecules on the side opposite the Phe232 residues. The results of these simulations indicate that there is insufficient space in the active sites to accommodate the adenine moiety of NADPH when preQ₀ also occupies the site.

It is possible that the NADPH can displace one of the preQ₀ molecules from its binding site and that the adenine moiety will occupy the same pocket as preQ₀. This is illustrated in Fig. 7 where we plot the relative orientation of adenine in the binding pocket. This conformation is taken from one frame of a dynamics simulation and is representative of the conformations observed. Note that the strictly conserved Lys96 residue is now coordinating the phosphate group from the NADPH and the Glu94 residue has adopted an alternative conformation. The N⁶ nitrogen atom of the adenine is coordinated by the carboxylate of Glu234, in the same fashion as the N² nitrogen atom of preQ₀. Indeed, the recently solved structure of QueF with ATP confirms that adenine can bind in the active site and binding of adenine is consistent with numerical simulations (data not shown).

A model of the overall binding of NADPH to the enzyme is depicted in Fig. 8, where we use a surface representation of the protein to emphasize the binding cleft. The nicotinamide moiety of NADPH is positioned to deliver a hydride to the C¹⁰ carbon of the substrate. In this simulation, the preQ₀ was not modeled as covalently bound to the Cys194 side chain. The diphosphate moiety of the NADPH occupies a position that is comparable to the position of the pyrophosphate shown in Fig. 3. The orientation of the nicotinamide reaffirms our earlier observation that the adenine moiety will not fit into the active site in a

conformation in which it is stacked against preQ₀. The nicotinamide fits into the narrow cleft in a conformation where the nicotinamide ring is transverse to the preQ₀ ring plane.

Final proton transfer

We propose that the fourth proton delivered to the substrate is provided by the strictly conserved His233 residue. This residue is found in close proximity to the strictly conserved Asp201 residue in the crystal structure and throughout our simulations, as is indicated in Figs. 3–5. Formation of the covalent adduct, as shown in Fig. 5, modifies the geometry of the C¹⁰ carbon atom of the substrate from linear to planar, pushing the N¹¹ nitrogen atom towards the His233. The transfer of the final proton to the N¹¹ nitrogen atom will be facilitated by the carboxyl moiety of Asp201.

This final step completes the process of the reaction mechanism at work in QueF (illustrated in Fig. 2) and accounts for the four protons and electrons delivered to the substrate. The initial proton is provided by Cys194 as a consequence of the formation of the covalent adduct. Two hydride transfers from NADPH account for the four electrons and two additional protons. The final proton is obtained from His233. Use of the crystal structure and numerical simulations to fill in missing details has enabled us to provide a more complete view of the catalytic process at work in this enzyme system.

The proposed mechanism suggests that the preQ₀ binds first and forms a covalent adduct with the enzyme (Fig. 2). The Cys194 attacks the C¹⁰ atom of the preQ₀ cyano group, donating one proton to the N¹¹ nitrogen atom of the substrate (Fig. 5). NADPH binds subsequently and, if necessary, displaces one of the preQ₀ molecules and positions itself for hydride transfer. This commits the reaction to one substrate-binding site. In the next step, a hydride is transferred from NADPH to the C¹⁰ carbon atom of the substrate and a third proton is provided by His233. During all of these steps, the substrate remains covalently attached to the enzyme, assuring productive completion of the reaction. Finally, a second molecule of NADPH provides a second hydride to the C¹⁰ carbon atom, breaking the adduct and allowing the product (preQ₁) to leave the active site. Additional simulations are planned to study the detailed energetics along the proposed reaction pathway.

Conclusions

The high-resolution structure of the QueF protein reported here provides a structural framework for mechanistic studies and supports previous biochemical characterizations of the enzyme.^{7,8} Consistent with previous reports and this work, the QueF dimer is a minimal catalytic unit and the enzyme should exhibit half-site reactivity. QueF is unique and facilitates the transfer of four electrons and four protons to its PreQ₀ substrate using NADPH as a cofactor, reducing the nitrile bond to a primary amine. Utilizing the structure as a starting state, we conducted a series of numerical simulations to provide structural information about the catalytic cysteine residue that is disordered and is not observed in the electron density and propose the catalytic role of several residues in the active site. Combining these calculations with the experimental X-ray structure that contains bound guanine and pyrophosphate has enabled us to begin to characterize each of the four steps of the reaction (two proton and two hydride transfers) and the roles of the active-site residues in binding and catalysis. We also found unexpected nucleosidase activity for the QueF enzyme and propose a potential mechanism, although further experimental work will be required to validate that model.

Materials and Methods

Protein cloning, expression, and purification

The open reading frame of the *queF* gene from *V. cholerae* O1 biovar El Tor str. N16961 (GenBank accession no. AAF94064.1) was amplified from genomic DNA with *KOD* DNA polymerase using conditions and reagents provided by Novagen (Madison, WI). The gene was cloned into pMCSG7 vector using a modified LIC protocol. This process generated an expression clone producing a fusion protein with an N-terminal His₆ tag and a tobacco etch virus protease recognition site. A selenomethionine derivative of the expressed protein was prepared as described by Walsh *et al.*¹⁴ and purified using standard procedures on an AKTExpress automated purification system (GE/Amersham) as described by Kim *et al.*¹⁵ The concentration of the purified protein was determined utilizing an ND-1000 Spectrophotometer System (NanoDrop Technologies). The fusion tag was then removed by adding recombinant tobacco etch virus protease at a ratio of approximately 1:75 (mg) and incubated for 48 h at 4 °C. The cleaved protein was then separated on a nickel-nitrilotriacetic acid agarose nickel charged resin column (Qiagen Inc.). The purified protein solution was dialyzed in a crystallization buffer (20 mM Hepes, pH 8.0, 250 mM NaCl, and 2 mM DTT) for 24 h and concentrated using a Centricon Plus-20 concentrator with a nominal molecular weight limit of 5000 (Millipore Corp.).

Size-exclusion chromatography

Size-exclusion chromatography was performed on a Superdex-200 10/300GL column using AKTAExplorer (GE Biosciences). The column was pre-equilibrated with crystallization buffer (20 mM Hepes, pH 8.0, 250 mM NaCl, and 2 mM DTT) and calibrated with premixed protein standards, including ovalbumin (43 kDa), con-albumin (75 kDa), aldolase (158 kDa), and Blue Dextran (2000 kDa). A 100-μl protein sample at 29 mg/ml was injected into the column. The chromatography was carried out at room temperature at a flow rate of 0.3 ml/min. The calibration curve of K_{av} versus log molecular weight was prepared using the equation $K_{av} = (V_e - V_0) / (V_t - V_0)$, where V_e is the elution volume for the protein, V_0 is the column void volume, and V_t is the total bed volume. Size-exclusion chromatography indicates a split peak of protein monomer and dimer (Fig. S3).

Protein crystallization

The protein was crystallized using hanging drop vapor diffusion at 289 K in a CrystalQuick® standard profile—LBR round bottom plate (Greiner Bio-One North America, Inc.). A 400-nl droplet of protein (199 mg/ml) was mixed with a 400-nl droplet of crystallization reagent and allowed to equilibrate over 135 μl of crystallization reagent. Nanopipetting was performed using the Mosquito® nanoliter liquid handling system (TTP LabTech). The finished plate was then incubated at 16 °C within a RoboIncubator® automated plate storage system (Rigaku). Automated crystal visualization was utilized in locating several crystals [Minstrel III® (Rigaku)]. These crystals were cryoprotected and flash frozen in liquid nitrogen. The protein crystallized in space group *P*1 with cell dimensions of $a = 71.52 \text{ \AA}$, $b = 72.58 \text{ \AA}$, $c = 71.51 \text{ \AA}$, $\alpha = 119.25^\circ$, $\beta = 110.18^\circ$, and $\gamma = 99.58^\circ$. The crystallization buffer included 0.04 M sodium dihydrogen phosphate, 0.96 M di-potassium hydrogen phosphate, and 10 mM GTP.

Data collection

Diffraction data were collected at 100 K at the 19BM beamline of the Structural Biology Center at the Advanced Photon Source, Argonne National Laboratory. The single-wavelength anomalous dispersion (SAD) data at 0.9793 Å (near selenium absorption white line, 12.6605 keV) up to 1.53 Å were collected from a single (0.1 mm×0.04 mm×0.05 mm)

selenomethionine-labeled protein crystal. The crystal was exposed for 3 s per 1.0° rotation of ω with a crystal-to-detector distance of 280 mm. The data were recorded on a CCD detector scanning a full 360° on ω .

Structure determination and refinement

The structure was determined by SAD phasing using HKL-3000,¹⁶ SHELX,¹⁷ MLPHARE,¹⁸ and SOLVE/RESOLVE¹⁹ and refined to 1.53 Å using REFMAC²⁰ in CCP4.¹⁹ The initial model was completed by using ARP/wARP²¹ and manual tailoring using Coot.²² The final R was 14.3% with an R_{free} of 18.3% with all data (Table 1).

Validation and deposition

The stereochemistry of the structure was checked with PROCHECK²³ and a Ramachandran plot. The main-chain torsion angles for all residues are in allowed regions.

Molecular dynamics

Molecular dynamics simulations were performed with the code NAMD,²⁴ developed by the Theoretical and Computational Biophysics Group in the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign. The CHARMM 27 force field²⁵ parameters were used. Analysis of dynamics trajectories was conducted with the program VMD.²⁶ Initial coordinates for protein atoms were obtained from the crystal structures. Missing atom coordinates (hydrogens) were defined with the PSFGEN module of VMD, which was also used to solvate the protein and neutralize the total charge of the simulation model. The solvation box was extended for 10 Å beyond the protein. With a 12-Å cutoff for electrostatic interactions, the box dimensions ensured a “diffuse” simulation model, in which no protein atoms interacted directly with protein atoms in the periodic images. Electrostatic interactions were computed using a smooth particle-mesh Ewald method,²⁷ with a grid size of approximately 1 Å spacing. Simulations began with a small amount of minimization (~1000 steps) and then utilized typically 100,000 steps of NVT dynamics and 100,000 steps of NPT dynamics run with 1-fs time steps to equilibrate the system. Constant temperature was maintained by a Langevin method,²⁸ and constant pressure conditions were enforced through a modified version of the Langevin piston²⁹ and Hoover^{30,31} methods. Typical production runs were of 1-ns duration and were conducted using NPT dynamics with 2-fs time steps, recording coordinate information at 1-ps intervals.

Simulations were conducted for the tetrameric conformation of the protein, as observed in the crystal structure. There were approximately 100,000 atoms in the simulations, in a box of dimension 84 Å×97 Å×132 Å.

QM/MM studies

To study energetics along the proposed reaction pathway, we performed QM/MM calculations with the program NWChem.³² We utilized a NEB method³³ implemented recently. The NEB method requires the definition of reactant and product states. Initial coordinates for reactant states were taken from snapshots of the dynamics trajectories in which the attacking water molecule was positioned in what we would describe as a near attack conformation.³⁴ The quantum partitions for all simulations included the side chain of residue 194 and the ¹⁰preQ₀ substrate. Atoms beyond 15 Å from the target C carbon atom of the substrate were frozen in place and only atoms within that spherical region were allowed to move. The reactant state was defined by a process in which the original model was optimized using the density functional method B3LYP³⁵ and a 3–21 g* basis set to speed convergence to a coarse estimate of the state geometry. Subsequent optimizations and dynamics used the Ahlrichs pVDZ basis set.³⁶ An initial estimate of the product-state

geometry was obtained by constraining the S^g–C¹⁰ distance between the attacking cysteine and the substrate to 1.8 Å and the H^g–N¹¹ distance to 1 Å and by repeating the optimizations at the B3LYP/Ahlrichs pVDZ level of theory. All atom constraints were then relaxed and the model geometry was re-optimized to define the product state. The NEB simulations were conducted using 10 replicas (beads) to define the reaction pathway.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank the members of the Midwest Center for Structural Genomics and Structural Biology Center for their support and Marat Valiev at Pacific Northwest National Laboratory for his help with NWChem. This research has been funded in part by a grant from the National Institutes of Health (GM074942) and by the U.S. Department of Energy, Office of Biological and Environmental Research, under Contract DE-AC02-06CH11357. One of us (M.A.C.) has also received support through the National Science Foundation's FaST program (HRD-0703584), administered by the Department of Educational Programs at Argonne National Laboratory. The authors acknowledge the Texas Advanced Computing Center at The University of Texas at Austin for providing HPC resources that have contributed to the research results reported within this article[†].

Abbreviations used

QueF	7-cyano-7-deazaguanine reductase
preQ₀	7-cyano-7-deazaguanine
preQ₁	7-aminomethyl-7-deazaguanine
T-fold	tunneling fold
QM/MM	hybrid quantum/classical
PDB	Protein Data Bank
NEB	nudged elastic band
SAD	single-wavelength anomalous dispersion

References

1. Iwata-Reuyl D. An embarrassment of riches: the enzymology of RNA modification. *Curr Opin Chem Biol.* 2008; 12:126–133. [PubMed: 18294973]
2. Dunin-Horkawicz S, Czerwoniec A, Gajda MJ, Feder M, Grosjean H, Bujnicki JM. MODOMICS: a database of RNA modification pathways. *Nucleic Acids Res.* 2006; 34:D145–D149. [PubMed: 16381833]
3. Björk GR, Durand JMB, Hagervall TG, Leipuvien R, Lundgren HK, Nilsson K, et al. Transfer RNA modification: influence on translational frameshifting and metabolism. *FEBS Lett.* 1999; 452:47–51. [PubMed: 10376676]
4. Iwata-Reuyl D. Biosynthesis of the 7-deaza-guanosine hypermodified nucleosides of transfer RNA. *Bioorg Chem.* 2002; 31:24–43. [PubMed: 12697167]
5. Durand JM, Dagberg B, Uhlin BE, Björk GR. Transfer RNA modification, temperature and DNA superhelicity have a common target in the regulatory network of the virulence of *Shigella flexneri*: the expression of the *virF* gene. *Mol Microbiol.* 2000; 35:924–935. [PubMed: 10692168]

[†]<http://www.tacc.utexas.edu>

6. Van Lanen SG, Reader JS, Swairjo MA, Crécy-Lagard V, Lee B, Iwata-Reuyl D. From cyclohydrolase to oxidoreductase: discovery of nitrile reductase activity in a common fold. *Proc Natl Acad Sci USA*. 2005; 102:4264–4269. [PubMed: 15767583]
7. Lee BWK, Van Lanen SG, Iwata-Reuyl D. Mechanistic studies of *Bacillus subtilis* QueF, the nitrile oxidoreductase involved in queuosine biosynthesis. *Biochemistry*. 2007; 46:12844–12854. [PubMed: 17929836]
8. Swairjo MA, Reddy RR, Lee B, Van Lanen SG, Brown S, de Crécy-Lagard V, et al. Crystallization and preliminary X-ray characterization of the nitrile reductase QueF: a queuosine-biosynthesis enzyme. *Acta Crystallogr Sect F*. 2005; 61:945–948.
9. Colloc'h N, Poupon A, Mornon JP. Sequence and structural features of the T-fold, an original tunneling building unit. *Proteins: Struct Funct Genet*. 2000; 39:142–154. [PubMed: 10737935]
10. McCarty RM, Somogyi A, Lin G, Jacobsen NE, Bandarian V. The deazapurine biosynthetic pathway revealed: in vitro enzymatic synthesis of preQ₀ from guanosine 5'-triphosphate in four steps. *Biochemistry*. 2009; 48:3847–3852. [PubMed: 19354300]
11. Holm L, Kaariainen S, Rosenstrom P, Schenkel A. Searching protein structure databases with DaliLite v.3. *Bioinformatics*. 2008; 24:2780–2781. [PubMed: 18818215]
12. Tanaka Y, Nakagawa N, Kuramitsu S, Yokoyama S, Masui R. Novel reaction mechanism of GTP cyclohydrolase I. High-resolution X-ray crystallography of *Thermus thermophilus* HB8 enzyme complexed with a transition state analogue, the 8-oxoguanine derivative. *J Biochem*. 2005; 138:263–275. [PubMed: 16169877]
13. Jordan F, Niv H. Glycosyl conformational and inductive effects on the acid catalyzed hydrolysis of purine nucleosides. *Nucleic Acids Res*. 1977; 4:697–709. [PubMed: 17100]
14. Walsh M, Dementieva I, Evans G, Sanishvili R, Joachimiak A. Taking MAD to the extreme: ultra fast protein crystal structure determination. *Acta Crystallogr Sect D: Biol Crystallogr*. 1999; 55:1168–1173. [PubMed: 10329779]
15. Kim Y, Dementieva I, Zhou M, Wu R, Lezondra L, Quartey P, et al. Automation of protein purification for structural genomics. *J Funct Struct Genomics*. 2004; 5:111–118.
16. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol*. 1997; 276:307–326.
17. Sheldrick GM, Schneider TR. SHELXL: high resolution refinement. *Methods Enzymol*. 1997; 277:319–343. [PubMed: 18488315]
18. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr Sect D: Biol Crystallogr*. 1994; 50:760–763. [PubMed: 15299374]
19. Murshudov GN, Lebedev A, Vagin AA, Wilson KS, Dodson EJ. Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr Sect D: Biol Crystallogr*. 1999; 55:247–255. [PubMed: 10089417]
20. Minor W, Cymborowski M, Otwinowski Z, Chruszcz M. HKL-3000: the integration of data reduction and structure solution—from diffraction images to an initial model in minutes. *Acta Crystallogr Sect D: Biol Crystallogr*. 2006; 62:859–866. [PubMed: 16855301]
21. Perrakis A, Sixma TK, Wilson KS, Lamzin VS. wARP: improvement and extension of crystallographic phases by weighted averaging of multiple-refined dummy atomic models. *Acta Crystallogr Sect D: Biol Crystallogr*. 1997; 53:448–455. [PubMed: 15299911]
22. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr Sect D: Biol Crystallogr*. 2004; 60:2126–2132. [PubMed: 15572765]
23. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr*. 1993; 26:283–291.
24. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. *J Comput Chem*. 2005; 26:1781–1802. [PubMed: 16222654]
25. MacKerell AD Jr, Bashford D, Bellott M, Dunbrack RL Jr, Evanseck J, Field MJ, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B*. 1998; 102:3586–3616.
26. Humphrey W, Dalke A, Schulten K. VMD—visual molecular dynamics. *J Mol Graphics*. 1996; 14:33–38.

27. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pederson LG. A smooth particle mesh Ewald method. *J Chem Phys.* 1995; 103:8577–8593.
28. Brünger A, Brooks CB, Karplus M. Stochastic boundary conditions for molecular dynamics simulations of ST2 water. *Chem Phys Lett.* 1984; 105:495–500.
29. Feller SE, Zhang Y, Pastor RW, Brooks BR. Constant pressure molecular dynamics: the Langevin piston method. *J Chem Phys.* 1995; 103:4613–4621.
30. Hoover WG. Canonical dynamics: equilibrium phase–space distributions. *Phys Rev A.* 1985; 31:1695–1697. [PubMed: 9895674]
31. Hoover WG. Constant pressure equations of motion. *Phys Rev A.* 1986; 34:2499–2500. [PubMed: 9897546]
32. Kendall RA, Apra E, Bernholdt DE, Bylaska EJ, Dupuis M, Fann GI, et al. High performance computational chemistry: an overview of NWChem, a distributed parallel application. *Comput Phys Commun.* 2000; 128:260–283.
33. Valiev M, Yang J, Adams JA, Taylor SS, Weare JH. Phosphorylation reaction in cAPK protein kinase-free energy quantum mechanical/molecular mechanics simulations. *J Phys Chem B.* 2007; 111:13455–13464. [PubMed: 17983217]
34. Bruice TC, Benkovic SJ. Chemical basis for enzyme catalysis. *Biochemistry.* 2000; 39:6267–6274. [PubMed: 10828939]
35. Becke AD. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys Rev A.* 1988; 38:3098–3100. [PubMed: 9900728]
36. Schäfer A, Horn H, Ahlrichs R. Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *J Chem Phys.* 1992; 97:2571–2577.

Fig. 1.

Queuosine biosynthetic pathway. PreQ₀ is reduced to preQ₁ by an NADPH-dependent reaction catalyzed by the enzyme QueF. Two NADPH molecules are utilized for each molecule of preQ₁ formed.

Fig. 2. Reaction mechanism. The reduction of preQ₀ (1) to preQ₁ (2) proceeds through the creation of a covalent adduct (3). Residues identified as important in the catalytic function in the text are also displayed.

Fig. 3. Crystal structure Of QueF. The protein crystallized as a tetramer (dimer of dimers). On the left, the secondary structures are colored to emphasize the ferredoxin-like structural modules and are labeled according to the annotation in Fig. S1. A guanine (G) molecule and phosphate (PO₄) and pyrophosphate (PP_i) molecules are shown with a space-filling representation (right): carbon, green; nitrogen, blue; oxygen, red; and phosphorus, gold. On the right, the dimer structure is illustrated.

Fig. 4. Stereo view of the QueF active site (chain A). Residues interacting with the guanine substrate are depicted as thin tubes and colored according to atom type: nitrogen, blue; oxygen, red; and carbon, tan. The guanine substrate (with carbons colored green) is oriented by a number of hydrogen bonds, indicated by the broken lines. Water molecules are rendered as red spheres. Residues are labeled with their single-letter codes: D, aspartic acid; E, glutamic acid; F, phenylalanine; I, isoleucine; K, lysine; S, serine; and W, tryptophan.

Fig. 5. Simulation model of the covalent adduct. This image represents the product state of the QM/MM simulations, where the preQ₀ substrate is covalently bound to the enzyme via a thioester bond. Atom colors are the same as in the previous figure, with the addition of atom types sulfur, yellow, and hydrogen, white.

Fig. 6. Minimum energy pathway. The minimum energies for the native protein are plotted as circles for each replica (bead) from the reactant (1) to product (10) states. Energies are normalized to the reactant state. The C194S mutant results are plotted as squares.

Fig. 7. Simulation model of NADPH binding. The pocket that binds preQ₀ adapts to bind the adenine moiety of NADPH. Note that the Lys96 residue has moved to bind the 2'-phosphate group of NADPH.

Fig. 8. Simulation model of the NADPH binding pocket. The protein is represented as a surface plot with each dimer colored separately, indicating the binding cleft. NADPH and preQ₀ are represented as tubes, with the atoms colored as follows: carbon, tan; nitrogen, blue; oxygen, red; phosphorus, gold; and sulfur, yellow. Carbon atoms in preQ₀ are colored green. The adenine moiety of NADPH is bound in a pocket below the lower phosphate group.

Table 1

Summary of the QueF protein crystallographic data

<i>Data collection</i>	
Experiment	SAD
Space group	P1
Unit cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	71.52, 71.58, 71.51
α , β , γ (°)	119.25, 110.18, 99.58
Protein molecule/asymmetric unit	4
Energy (keV)	0.9793
<i>Data processing</i>	
Resolution range (Å)	61–1.53 (1.57–1.53)
Unique reflections	149,957 (12,259)
Redundancy	4.5 (3.9)
Completeness (%)	94.1 (76.5)
<i>I</i> / σ <i>I</i>	7.5 (2.4)
<i>R</i> _{merge} (%)	7.0 (64)
<i>Refinement</i>	
Resolution range (Å)	26.16–1.53 (1.53)
<i>R</i> _{work} (%)	14.3 (18.3)
<i>R</i> _{free} (%)	17.8 (25.8)
RMSD bond length (Å)	0.015
RMSD bond angle (°)	1.58
Protein residues	1015 of 1160 ^a
Water molecules	1177
Mg ion	1
Ramachandran plot statistics (%)	
Residues in most favored regions	97.3
Residues in additional allowed regions	2.7
Residues in disallowed regions	0.0

Values in parentheses are those for the highest-resolution shell.

^aIncludes three N-terminal residues per protein chain that are cloning artifacts.