STRUCTURE NOTE

Crystal Structure of Flavin Binding to FAD Synthetase of *Thermotoga maritima*

Weiru Wang,[1†] Rosalind Kim,[1] Hisao Yokota,[1] and Sung-Hou Kim[1,2*]
[1]*Berkeley Structural Genomics Center, Physical Biosciences Division of the Lawrence Berkeley National Laboratory, Berkeley, California*
[2]*Department of Chemistry, University of California, Berkeley, California*

Introduction. The product of genomic sequence gi:4981319 from *Thermotoga maritima* (TM379) is annotated as flavin adenine dinucleotide (FAD) synthetase,[1] a bifunctional enzyme that catalyzes two reactions: riboflavin kinase (EC 2.7.1.26) and flavin mononucleotide (FMN) adenylyltransferase (EC 2.7.7.2). FAD synthetase is present among all kingdoms of living organisms. Mammals utilize separate enzymes for FMN and FAD formation,[2,3] whereas lower level organisms depend on the bifunctional enzyme. The Berkeley Structural Genomics Center (BSGC) determined that the TM379 gene product possesses a novel fold and determined its crystal structure [Protein Data Bank (PDB)[4] code: 1MRZ].[5] The TM379 structure exhibited a novel combination of two folds: a Rossmann fold at the N-terminus and a flavin binding fold at the C-terminus. While the sequence annotation is still awaiting experimental verification, our recently published structural data[5] provided supportive evidence. Structural analysis suggests that the N-terminal domain of TM379 is involved in FMN adenylyltransferase, and the C-terminal domain contributes to binding of flavin and probably is involved in both steps of the reaction.[5] Despite low sequence identity, the structural similarity between a recently published crystal structure of *Schizosaccharomyces pombe* (Sp) riboflavin kinase (RK)[6] and the C-terminal domain of TM379 further indicated that the C-terminal domain is sufficient for catalyzing the RK reaction. The N-terminal domain also has structural similarity to nucleotidyltransferases, including glycerol-3-phosphate cytidylyltransferase (1COZ),[7] nicotinamide mononucleotide adenylyltransferase (1F9A),[8] and phosphopantetheine adenylyltransferase (1B6T),[9] although sequence identities among them are low. To experimentally characterize the molecular function of TM397, we introduced riboflavin into TM379 crystal and determine the three-dimensional (3D) structure. Here we describe the cocrystal structure of TM379 and flavin, and identify the key residues involved in flavin binding for this family of proteins.

Methods. *Crystallization:* Cloning and purification of TM379 has been described previously.[5] The native protein was concentrated to 21.5 mg/mL in 20 m*M* Tris, pH7.5, 1 m*M* ethylenediaminetetraacetic acid (EDTA), 10% glycerol. Crystallization conditions were tested with a sparse matrix sampling method[10] using hanging drop vapor diffusion at room temperature with commercially prepared reagents (Hampton Research, Laguna Niguel, CA). The optimum conditions were found in 0.2 m*M* sodium citrate, 10% polyethylene glycol 4000 (PEG4K), 5% glycerol, and 10% propanol. The PEG4K concentration was raised to 30% in the reservoir 1 day before data collection. Crystals reached a maximum size of $0.05 \times 0.1 \times 0.2$ mm^3. Solution for soaking experiment was made by adding 5 m*M* of riboflavin to mother liquor. The native crystals were soaked with riboflavin for 24 h before data collection.

TABLE I. X-Ray Data Statistics

Space group	$P2_1$
Unit cell	$a = 66.1$ Å, $b = 82.2$ Å, $c = 66.7$ Å $\alpha = \gamma = 90°$, $\beta = 116.5°$
Wavelength (Å)	0.91963
Resolution (Å)	30.0–2.1
Redundancy	2.2
Unique reflections	40525(2995)[a]
Completeness (%)	94.0(82.7)
I/σ	17.8(1.8)
Rsym[b] (%)	4.9(51.8)

[a]Numbers in parentheses are related to the highest resolution shell, which is 2.15–2.10 Å.
[b]$R_{sym} = \Sigma_{hkl} \Sigma_i |I_{hkl,i} - \langle I \rangle_{hkl}| \Sigma |\langle I \rangle_{hkl}|$

Structure determination and refinement: The X-ray diffraction data were collected at the Advanced Light Source (ALS) beam line 5.0.2 using an Area Detector System Co. (ADSC) Quantum 4 charge-coupled device (CCD) detector. The diffraction data were processed using the programs DENZO and SCALEPACK.[11] X-ray data statistics are shown in Table I.

Grant sponsor: National Institutes of Health; Grant number: GM 62412.

†Current address: Plexxikon, Inc., 91 Bolivar Drive, Berkeley, CA 94710

*Correspondence to: Sung-Hou Kim, Berkeley Structural Genomics Center, Physical Biosciences Division, Lawrence Berkeley National Laboratory, Department of Chemistry, University of California, Berkeley, CA 94720. E-mail: shkim@cchem.berkeley.edu

Received 11 February 2004; Accepted 20 February 2004

Published online 1 October 2004 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.20207

TABLE II. Refinement Statistics

Space group	$P2_1$
Total number residues	549
Total non-H atoms	4717
Number of water molecules	228
Average temperature factors	43.3 Å^2
R factor	21.2%
Free *R* factor	25.9%
Stereochemical ideality:	
Bond	0.013 Å
Angle	1.67°
Improper	1.04°

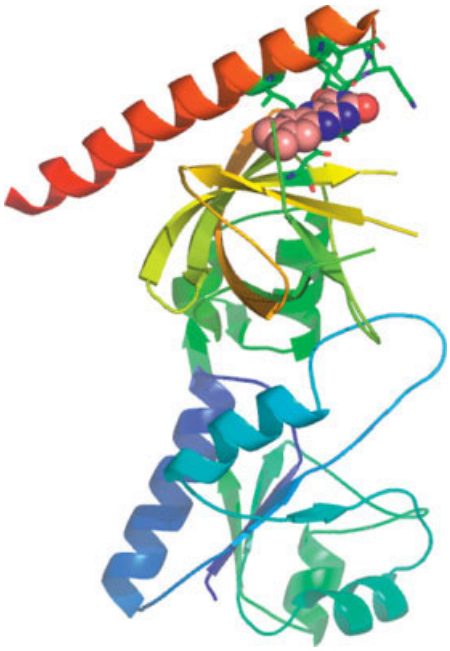

Fig. 1. A ribbon diagram of the TM379 structure. The rainbow colors are progressively applied from N- to C- terminus. The ball model indicates the flavin molecule bound to the protein structure.

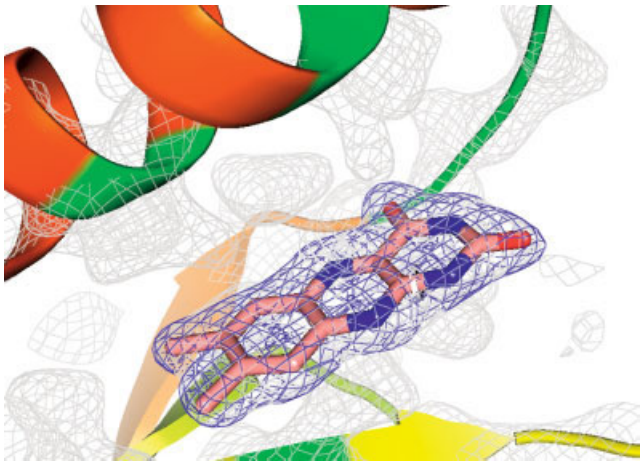

Fig. 2. Electron density map of the region of the protein (white) bound with flavin (blue).

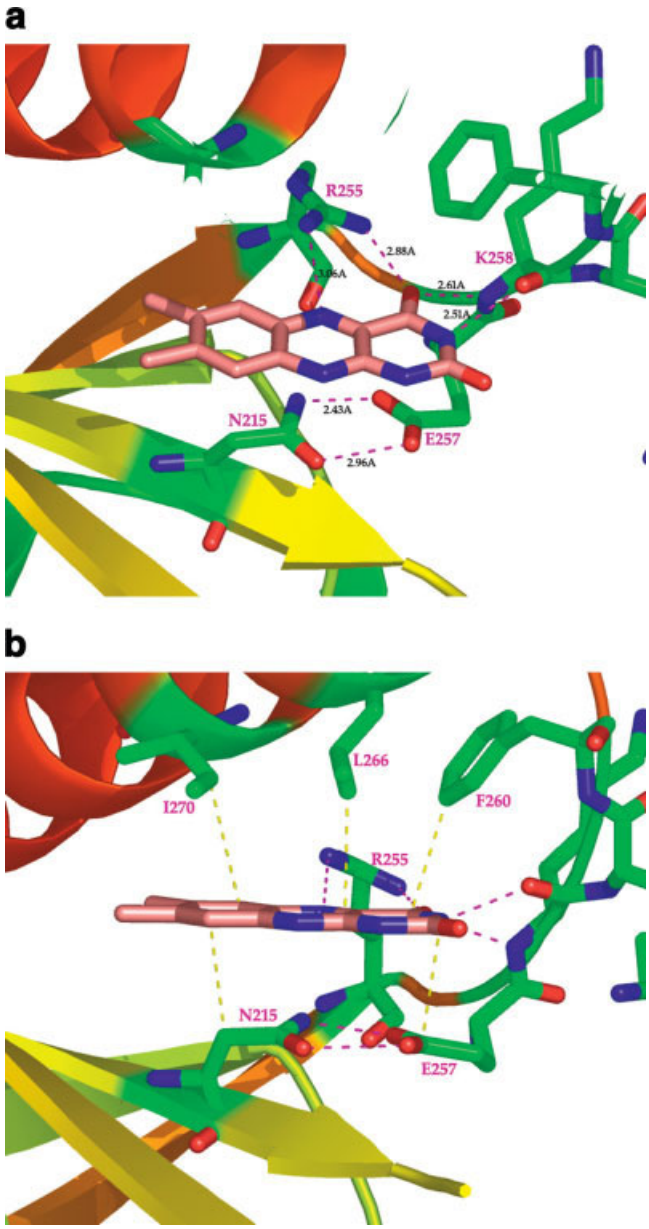

Fig. 3. Interactions between flavin and active site. Each dotted line represents an interaction. Magenta indicates H-bonds. Yellow indicates VDW contacts. (**a**) Shows hydrogen bond interactions only. (**b**) Shows all interactions.

The first set of phases was obtained by molecular replacement using the TM379 apo structure (PDB code: 1MRZ) as a search model. The model was refined against 2.1Å data using the Crystallography & NMR System (CNS) program[12] and REFAM5.[13] The model building was performed using O.[14] A model containing 549 residues was derived from progressive improvement of the electron density map using rounds of refinement and manual building. The noncrystallographic symmetry (NCS) constraints and restraints were completely released throughout the refinement. Ten percent of the data was randomly picked out for free *R* factor cross-validation. The refinement statistics are shown in Table II. Atomic

coordinates have been deposited at the PDB with access code 1S4M.

Results. The overall structure of TM379 flavin complex is the same as that without flavin as previously described.[5] TM379 is present as a dimer in the crystallographic asymmetric unit. Each monomer structure is composed of two domains connected by a single-loop linkage. The N-terminal domain (residues 2–134) adopts a typical nucleotide binding fold (Rossmann fold). The C-terminal domain (residues 135–288) contains a 6-stranded antiparallel β-barrel architecture (Fig. 1).

The TM379 crystals were soaked in mother liquor containing 5 m*M* of riboflavin. However, only the flavin tricyclic base was clearly defined in the electron density map, missing the ribityl side-chain, possibly due to structural flexibility (Fig. 2). Flavin binds to the outside surface of the β-barrel within a cavity surrounded by a long α-helix and the β-barrel (Fig. 1). The hydrophobic isoalloxazine side of the tricyclic ring is embedded inside the cavity, while the hydrophilic edge is partially exposed to solvent. Arg 255 interacts with O4 and N5 of flavin by donating an H-bond to each atom. The conformation of Arg 255 side-chain was stabilized by a salt bridge with Glu 273. The N3 of the flavin base donates an H-bond to the main-chain carbonyl of Lys 258 [Fig. 3(a)]. The side-chains of Asn 215 and Glu 257 couple together to form a plenary geometry that is parallel to the flavin tricyclic plane. In the TM379 apo structure, side-chains of Asn 215 and Glu 257 adopted completely different configurations.[5] Therefore, coupling of these side-chains is dependent on flavin binding. The interactions of the coupled sides with flavin appear to be van der Waals (VDW) contacts and cover an entire face of the flavin base. Hydrophobic interactions between flavin and its binding sites involve Phe 260, Val 197, Val 213, Leu 266, and Ile 270 [Fig. 3(b)].

Binding of flavin to the C-terminal domain of TM379 provides positive evidence supporting the sequence based prediction of the enzymatic function. Moreover, comparison of TM379 C-terminal domain structure with the Sp RK structure (PDB code: 1N07[6]) revealed significant similarity in overall as well as active site geometry. FMN in the Sp RK structure is located at the same site as that of flavin in TM379. They shared some key features in substrate-active site interactions (e.g., the interactions involving the side-chain of Arg 255 and main-chain of Lys 258), as well as extensive hydrophobic contacts. Based on the above data, we predict that the C-terminal domain is sufficient to catalyze RK activity, the first half of FAD synthetase activity.

Acknowledgments. We thank Dr. David King for mass spectrometric analysis of the protein and Drs. Thomas Ernest and Gerry McDermott for help during the data collection at the ALS beam line 5.0.2. We are grateful to Barbara Gold and Andrew Waight for cloning of the gene of TM379.

REFERENCES

1. Manstein DJ, Pai EF. Purification and characterization of FAD synthetase from *Brevibacterium ammoniagenes*. J Biol Chem 1986;261:16169–16173.
2. Merrill AH, McCormick DB. Affinity chromatographic purification and properties of flavokinase (ATP:riboflavin 5′-phosphotransferase) from rat liver. J Biol Chem 1980;255:1335–1338.
3. Oka M, McCormick DB. Complete purification and general characterization of FAD synthetase from rat liver. J Biol Chem 1987;262: 7418–7422.
4. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. Protein Data Bank. Nucleic Acids Res 2000;28:235–242.
5. Wang W, Kim R, Jancarik J, Yokota H, Kim S-H. Crystal structure of a flavin-binding protein from *Thermotoga maritima.* Proteins 2003;52:633–635.
6. Bauer S, Kemter K, Bacher A, Huber R, Fischer M, Steinbacher S. Crystal structure of *Schizosaccharomyces pombe* riboflavin kinase reveals a novel ATP and riboflavin-binding fold. J Mol Biol 2003;326:1463–1473.
7. Weber CH, Park YS, Sanker S, Kent C, Ludwig ML. A prototypical cytidylyltransferase: CTP:glycerol-3-phosphate cytidylyltransferase from *Bacillus subtilis*. Struct Fold Des 1999;7:1113–1124.
8. D'Angelo I, Raffaelli N, Dabusti V, Lorenzi T, Magni G, Rizzi M. Structure of nicotinamide mononucleotide adenylyltransferase: a key enzyme in NAD(+) biosynthesis. Struct Fold Des 2000;8:993–1004.
9. Izard T, Geerlof A. The crystal structure of a novel bacterial adenylyltransferase reveals half of sites reactivity. EMBO J 1999;18:2021–2030.
10. Jancarik J, Kim S-H. Sparse matrix sampling: a screening method for crystallization of proteins. J Appl Crystallogr 1991;24:409–411.
11. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 1996;276:307–326.
12. Brunger AT, Adams PD, Clore GM, DeLano Wl, Gros P, Grosse-Kuntsleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR System: a new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 1998; 54:905–921.
13. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 1997;53:240–255.
14. Jones TA, Zou J-Y, Cowan SW, Kjeldgaard M. Improved methods for binding protein models in electron density maps and the location of errors in these models. Acta Crystallogr D Biol Crystallogr 1991;47: 110–119.