THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 279, No. 3, Issue of January 16, pp. 2211–2220, 2004

Printed in U.S.A.

High Resolution X-ray Structure of dTDP-Glucose 4,6-Dehydratase from *Streptomyces venezuelae**

Received for publication, September 11, 2003, and in revised form, October 20, 2003
Published, JBC Papers in Press, October 21, 2003, DOI 10.1074/jbc.M310134200

Simon T. M. Allard, W. W. Cleland, and Hazel M. Holden‡

From the Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706

Desosamine is a 3-(dimethylamino)-3,4,6-trideoxyhexose found in some macrolide antibiotics. In *Streptomyces venezuelae*, there are seven genes required for the biosynthesis of this unusual sugar. One of the genes, *desIV*, codes for a dTDP-glucose 4,6-dehydratase, which is referred to as DesIV. The reaction mechanisms for these types of dehydratases are quite complicated with proton abstraction from the sugar 4′-hydroxyl group and hydride transfer to NAD^+, proton abstraction at C-5, and elimination of the hydroxyl group at C-6 of the sugar, and finally return of a proton to C-5 and a hydride from NADH to C-6. Here we describe the cloning, overexpression, and purification, and high resolution x-ray crystallographic analysis to 1.44 Å of wild-type DesIV complexed with dTDP. Additionally, for this study, a double site-directed mutant protein (D128N/E129Q) was prepared, crystallized as a complex with NAD^+ and the substrate dTDP-glucose and its structure determined to 1.35 Å resolution. In DesIV, the phenolate group of Tyr^{151} and O^{γ} of Thr^{127} lie at 2.7 and 2.6 Å, respectively from the 4′-hydroxyl group of the dTDP-glucose substrate. The side chain of Asp^{128} is in the correct position to function as a general acid for proton donation to the 6′-hydroxyl group while the side chain of Glu^{129} is ideally situated to serve as the general base for proton abstraction at C-5. This investigation provides further detailed information for understanding the exquisite chemistry that occurs in these remarkable enzymes.

Macrolide antibiotics belong to a large family of natural products known as polyketides, which are synthesized by bacteria (primarily actinomycetes and myxobacteria) and fungi (1). These drugs constitute an old and well-established class of anti-microbial agents demonstrating excellent clinical activity against Gram-positive bacteria. While the bactericidal properties are still not completely understood, it is clear that these drugs inhibit protein synthesis by interacting with the ribosome or components thereof (2). Some of the structural variability among the polyketides is due to postsynthetic steps that often include glycosylation by unique deoxy sugars (3). The addition of unusual deoxy sugars such as D-desosamine to many polyketides provides or enhances their biological activity. Strikingly, without these appended sugar(s), the pharmacological properties of many of these compounds are either completely lost or dramatically decreased (2). In the case of desosamine, a 3-(dimethylamino)-3,4,6-trideoxyhexose, the sugar is thought to be involved in ionic interactions of the associated antibiotic with the bacterial ribosome and specifically with the peptidyl transferase center (4, 5). Understanding the *in vivo* formation of these unusual sugar residues may lead to the eventual construction of novel antibiotics by genetic manipulation of existing biochemical pathways via "combinatorial biosynthesis" (6).

Among microorganisms, approximately three-fourths of all known bioactive compounds are produced by members of the actinomycete family, mostly belonging to the *Streptomyces* genus (7). *Streptomyces venezuelae* ATCC 15439 produces two distinct groups of macrolide antibiotics with either 12- or 14-membered lactone rings (1). All of these contain the deoxy sugar desosamine. As indicated in Scheme 1, seven genes are thought to be responsible for desosamine biosynthesis, namely *desI*, *desII*, *desIII*, *desIV*, *desV*, *desVI*, and *desVIII*. These genes are located in the *pikB* locus that is part of the pikromycin biosynthetic gene cluster (1, 8). Also present in this cluster is the *desVII* gene that encodes the glycosyltransferase required for the transfer of desosamine to both the 12- and 14-membered lactone rings.

The focus of this structural investigation is dTDP-glucose 4,6-dehydratase, which is encoded by the *desIV* gene and hereafter referred to simply as DesIV.[1] This enzyme catalyzes the second step in the desosamine pathway by oxidizing the 4′-hydroxyl group and removing the 6′-hydroxyl group of the dTDP-glucose substrate (Scheme 1). On the basis of amino acid sequence analyses, DesIV is known to be a member of the short chain dehydrogenase/reductase superfamily (SDR), containing the highly conserved Tyr-*X*-*X*-*X*-Lys catalytic couple (9). Members of this superfamily catalyze a wide range of biochemical reactions with some displaying dehydrogenase activities, while others acting as dehydratases, isomerases, or epimerases, for example (10–12). The x-ray structures of the dTDP 4,6-dehydratases from *Salmonella enterica* and *Streptococcus suis*, alone or in complex with various dTDP sugars, have been determined thus far (13–15). Here we report the high resolution x-ray structure (1.44 Å) of DesIV from *S. venezuelae* complexed with dTDP. In addition, we describe the construction and x-ray structural analysis to 1.35 Å resolution of a double site-directed mutant protein (D128N/E129Q) containing NAD^+ and the substrate, dTDP-glucose. Taken together, these two structures provide a molecular framework within which to elaborate more fully the detailed chemistry that occurs during catalysis.

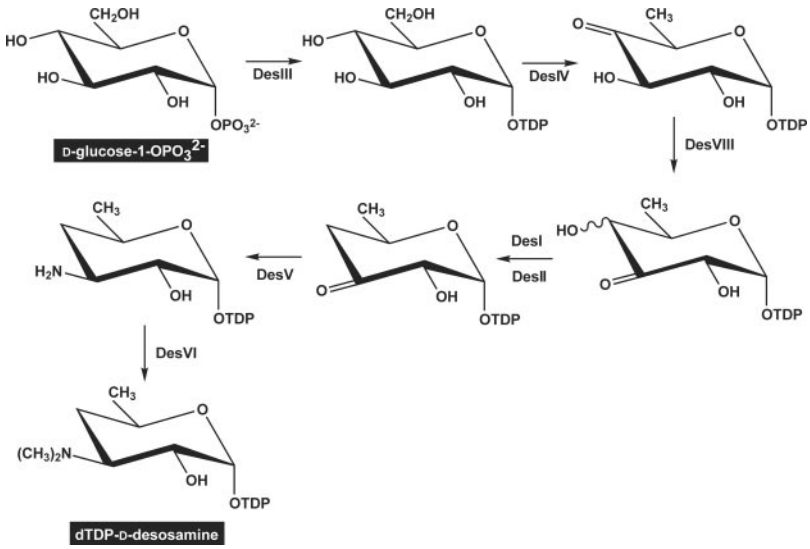
* This research was supported in part by Grants DK47814 (to H. M. H.) and GM18938 (to W. W. C.) from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1R66 and 1R6D) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

‡ To whom correspondence should be addressed. Tel.: 608-262-4988; Fax: 608-262-1319; E-mail: Hazel_Holden@biochem.wisc.edu.

[1] The abbreviations used are: DesIV, dTDP-glucose 4,6-dehydratase; NTA, nitrilotriacetic acid.

SCHEME 1



EXPERIMENTAL PROCEDURES

Cloning of the DesIV Gene—The sequence of the *S. venezuelae desIV* gene has been reported (1) and allowed for the design of primers for gene amplification from genomic DNA. *S. venezuelae* ATCC 15439 genomic DNA was obtained as a freeze-dried pellet from the American Type Culture Collection (ATCC). The oligonucleotide primers were designed to amplify the *desIV* gene for cloning into the pET-28a (Novagen) expression vector to produce a C-terminal His_6-tagged fusion protein. The forward primer, 5′-CATG**CCATGG**CACGGCTTCTGGTGAC-CGGA-3′, contains an NcoI restriction site (in bold). The reverse primer, 5′-CCG**CTCGAG**CGCGGACACCTCCAC-3′ has an XhoI site (in bold). In order to preserve the NcoI recognition sequence of CCATGG, the codon GCC (encoding alanine) was inserted before the second residue of the gene. The gene was PCR-amplified with Platinum *Pfx* DNA polymerase (Invitrogen) according to the manufacturer's instructions and standard cycling conditions. The PCR product was purified with the QIAquick PCR purification kit (Qiagen Inc.), followed by digestion with both NcoI and XhoI at 37 °C overnight. The gene was separated from digestion byproducts on a 1.0% agarose gel, excised from the gel, and purified with QIAquick Gel purification kit. The purified *desIV* gene was then ligated into the NcoI/XhoI sites of pET-28a (Novagen) that had previously been cut with the same restriction enzymes. *Escherichia coli* DH5α cells were transformed with the ligation mixture and subsequently plated onto LB media supplemented with 30 μg/ml kanamycin. Individual colonies were selected, cultured overnight, and plasmid DNA extracted with the QIAprep Spin Miniprep kit (Qiagen Inc.). Positive clones were identified by digestion of the plasmid DNA with appropriate restriction enzymes and visualization of the excised insert by staining the agarose gel with ethidium bromide following electrophoresis. The positive clones were sequenced with the ABI Prism™ Big Dye Primer Cycle Sequencing kit (Applied Biosystems, Inc.) to confirm that no mutations were introduced during PCR amplification.

Protein Expression—For protein expression, *E. coli* Rosetta-(DE3)pLysS (Novagen) cells were transformed with the pET28a-*desIV* plasmid and plated onto LB media supplemented with 30 μg/ml kanamycin. After ~16 h of incubation, the plates were scraped, and the cells were suspended in LB media to use for the inoculation of 6 × 2-liter baffled flasks containing 500 ml of Terrific Broth (16) plus 30 μg/ml kanamycin and 30 μg/ml chloramphenicol. The cells were grown at 37 °C with aeration to an OD_{600} of ~0.6, at which time they were induced with 1 mM isopropyl-1-thio-β-D-galactopyranoside. The cells were allowed to grow for an additional 4 h before harvesting by centrifugation at 6000 × *g* for 8 min. The cell paste was frozen in liquid nitrogen and stored at −80 °C.

Protein Purification—All protein purification procedures were carried out on ice or at 4 °C. The cell paste was thawed in 100 ml of cold Ni-NTA lysis buffer containing 50 mM NaH_2PO_4, 300 mM NaCl, and 10 mM imidazole at pH 8.0. Cells were lysed on ice by six cycles of sonication (30 s) separated by 2 min of cooling. Cellular debris was removed by centrifugation at 4 °C for 25 min at 20,000 × *g*. The clarified lysate was loaded onto a 10-ml Ni-NTA agarose (Qiagen Inc.) column pre-equilibrated with Ni-NTA lysis buffer. After loading, the column was washed with about 60 ml of Ni-NTA wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole at pH 8.0), followed by gradient elution of the protein from 20 to 500 mM imidazole in Ni-NTA lysis buffer. Protein-containing fractions were pooled based on SDS-PAGE analysis and dialyzed against 25 mM Tris (pH 7.6) with 200 mM NaCl. The dialyzed protein was concentrated to ~10 mg/ml based on an extinction coefficient of ~0.73 mg/ml as calculated by the program Protean (DNAstar, Inc., Madison, WI).

Enzymatic Assay—The assay is based upon that described in Ref. 17 and follows the formation of a unique, broad UV chromophore centered at 318 nm following incubation of dTDP-4-keto-6-deoxy-D-glucose in 0.1 M NaOH at 310 K for 20 min. The structure of the species responsible for this chromophore has not yet been determined. In addition, several different values for the extinction coefficient for this species have been utilized. The value employed in this investigation was 6500 M^{-1} cm^{-1}, as recently re-determined by Gross *et al.* (18). All readings were done in triplicate. Between 10 and 50 μg of DesIV were transferred into a sterile 1-ml Eppendorf tube containing 86 μl of 50 mM HEPES, pH 7.6 and 2.5 μl of 40 mM NAD^+. A 1-μl sample of 100 mM dTDP-D-glucose was added, and the sample incubated at 37 °C for 20 min. Subsequently, 1 ml of 0.1 M NaOH was then added, and the sample incubated for a further 20 min at 37 °C. The spectrophotometer was zeroed against 0.1 M NaOH, and the absorbance of the sample recorded at 318 nm in a quartz cuvette. Negative controls were set up in the same manner as described above but omitting the substrate. Only active DesIV protein samples were employed for crystallization trials.

Crystallization of DesIV—Initial trials were conducted by the hanging drop method of vapor diffusion with Hampton Research screens I and II (Hampton Research). The protein concentration was typically 10 mg/ml. Prior to crystallization, the protein was incubated with 2 mM dTDP overnight at 4 °C. Single crystals were observed growing after 4 days at 4 °C in 20% poly(ethylene glycol) 8000, 200 mM magnesium acetate tetrahydrate, and 100 mM sodium cacodylate, pH 6.5. Refinement of the crystallization conditions led to large single crystals being grown via the batch method at 4 °C with a precipitant solution containing 7% poly(ethylene glycol) 8000, 100 mM magnesium acetate tetrahydrate, and 100 mM sodium cacodylate, pH 6.3. Crystals achieved typical dimensions of ~0.8 mm × 0.8 mm × 0.8 mm in 3 days. The crystals belonged to space group $P2_12_12$ with unit cell dimensions of a = 75.5 Å, b = 99.8 Å, c = 42.2 Å and one monomer per asymmetric unit.

High Resolution X-ray Data Collection—For high resolution x-ray data collection, the crystals were transferred to solutions of a synthetic mother liquor containing increasing concentrations of glycerol (5–30% in 5% increments). Each crystal was subsequently flash-cooled to −150 °C in a stream of nitrogen gas and subsequently stored under liquid nitrogen until synchrotron beamtime became available. X-ray data were collected from a single crystal on a 3 × 3 tiled "SBC3" CCD detector at the Structural Biology Center 19-BM beamline (Advanced Photon Source, Argonne National Laboratory, Argonne, IL) to a nominal resolution of 1.44 Å. The x-ray data were processed with HKL2000 and scaled with SCALEPACK (19). Relevant x-ray data collection statistics are presented in Table I.

X-ray Structural Analyses—The structure of DesIV was solved by molecular replacement with *AMoRe* (20) as implemented in the CCP4 program suite. The search model was that of the fully refined structure

TABLE I
X-ray data collection statistics

	Wavelength	Resolution	No. independent reflections	Completeness	Redundancy	Avg I/Avg σ(I)	R_{sym}[a]
	Å	Å		%			%
Wild type	0.97930	50.0-1.44	56,518	98.6	5.5	26.4	6.6
		1.49-1.44[b]	5,618	99.9	4.0	6.8	21.2
D128N	1.54178	30.0-1.35	61,664	91.9	3.6	17.3	4.0
E129Q		1.41-1.35[b]	6,573	81.0	1.7	2.0	22.3
Protein							

[a] $R_{sym} = (\Sigma|I - \bar{I}|\Sigma I) \times 100$.
[b] Statistics for the highest resolution bin.

of dTDP-glucose 4,6-dehydratase from *E. coli*.[2] With x-ray data in the resolution range of 10–4.0 Å, one clear solution for the monomer in the asymmetric unit was obtained. Visual inspection of the solution with the graphics program TURBO (21) revealed that the local 2-fold rotational axis of the DesIV dimer was coincident to a crystallographic dyad. Alternate cycles of least-squares refinement with the software package TNT (22) and manual model-building with TURBO reduced the *R*-factor to 17.7% for all measured x-ray data from 30 Å to 1.44 Å resolution. Relevant least-squares refinement statistics are summarized in Table II. DesIV contains 337 amino acid residues per subunit. In the model presented here, the final amino acid observed in the electron density map was Lys^{322}. Other than the C-terminal residues, however, the electron density for the rest of the polypeptide chain was unambiguous. There were no significant outliers in the Ramachandran plot.

Preparation and Structural Analysis of the D128N/E129Q Site-directed Mutant Protein—The double mutant protein, D128N/E129Q, was created with the original plasmid containing the *desIV* gene. The mutations were generated by PCR amplification using the QuikChange™ site-directed mutagenesis kit (Stratagene). Purified primers were purchased from Integrated DNA Technologies, Inc. Plasmids harvested from the mutagenesis reactions were screened and sequenced with the ABI PRISM Big Dye Primer Cycle Sequencing kit (Applied Biosystems, Inc.). Protein expression and purification of the double mutant protein was similar to that described above for the wild-type enzyme. The D128N/E129Q mutant protein demonstrated no activity with the above-described assay.

Potential crystallization conditions were examined with Hampton Research screens I and II at 4 °C by the hanging drop method of vapor diffusion. The D128N/E129Q mutant protein, at 10 mg/ml, was incubated with 20 mM dTDP-glucose for 4 h at 4 °C. Crystals were observed growing at 4 °C in 20% poly(ethylene glycol) 8000, 200 mM magnesium acetate tetrahydrate, and 100 mM sodium cacodylate, pH. 6.5. Refinement of the crystallization conditions led to large single crystals grown by macroseeding into batch experiments with precipitant concentrations of 5% poly(ethylene glycol) 8000, 50 mM $MgCl_2$, and 100 mM cacodylate (pH 6.5). Crystals achieved typical dimensions of $\sim 0.6 \times 0.2 \times 0.2$ mm^3 in 1–2 weeks and belonged to the orthorhombic space group $P2_12_12$ with unit cell dimensions of a = 71.7 Å, b = 99.4 Å, and c = 42.2 Å and one molecule in the asymmetric unit.

Crystals were harvested from the batch experiments and soaked for several hours in a synthetic mother liquor composed of 15% poly(ethylene glycol) 8000, 50 mM $MgCl_2$, 100 mM NaCl, 10.0 mM dTDP-glucose, and 100 mM cacodylate (pH 6.5). These crystals were then serially transferred to a cryoprotectant solution composed of 25% poly(ethylene glycol) 8000, 50 mM $MgCl_2$, 200 mM NaCl, 10.0 mM dTDP-glucose, 15% ethylene glycol, and 100 mM cacodylate (pH 6.5). The crystals were flash-cooled to −150 °C in a stream of nitrogen gas. X-ray data were collected with a HiStar (Bruker AXS) area detector system using CuKα radiation generated from a Rigaku RU200 rotating anode generator operated at 50 kV and 90 mA and equipped with Göbel focusing optics. The x-ray data were processed with XDS (23, 24) and scaled with XSCALIBRE.[3] Relevant x-ray data collection statistics are presented in Table I.

The structure of the D128N/E129Q mutant protein was solved via molecular replacement with the wild-type structure as the search model and refined in a similar manner to that described above. Refinement statistics are presented in Table II and a representative portion of the electron density map near the ligands is shown in Fig. 1. Again, there were no significant outliers in the Ramachandran plot.

Determination of the Redox State of the Dinucleotide in D128N/E129Q Protein Crystals—All procedures were conducted at 4 °C. Thirty crystals from the same batch experiment employed for x-ray data collection were carefully removed and swiftly washed in 0.1 micron-filtered water (Anatop 10, Whatman). They were then transferred to 100 μl of fresh 0.1-micron filtered water and gently dissolved. Aliquots of this crystal mixture (20 μl) were subsequently added to 80 μl of the following solutions: 8 M urea, 100% ethanol, and 200 mM phosphate buffer, pH 7.0. The ethanol solution was mixed and centrifuged at 14,000 × *g* for 5 min to precipitate the protein. The supernatant containing the free dinucleotide was carefully removed for analysis. The 8 M urea solution was used to denature the protein to enable the measurement of free dinucleotide in the presence of denatured enzyme, while the phosphate buffer was utilized to allow observation of the undenatured complex. All solutions were flash frozen in liquid nitrogen and stored at −20 °C until required.

Absorption spectra were measured from 220 to 600 nm for all three samples with a Cary 300 Scan UV-visible spectrophotometer (Varian). If present, NADH would be detectable at these protein concentrations (1.03 mg/ml as estimated from the phosphate buffer spectrum A_{280}) by its absorbance at 340 nm (ϵ_{340} = 6.22 mM^{-1} cm^{-1}). No significant absorbance features, however, were observed at this wavelength. The protein free sample (80% ethanol supernatant) showed a strong 262 nm band (A_{262} = 0.814) and none at 340 nm, consistent with the assignment of NAD^+ to the electron density observed in the crystalline lattice. In summary, these data indicate, within the threshold of detection, that the bound nucleotide is in the oxidized state in the D128N/E129Q mutant protein crystals.

TABLE II
Least-squares refinement statistics

	Wild-type enzyme	D128N/E129Q mutant protein
Resolution limits (Å)	30.0-1.44	30.0-1.35
R-factor (overall)[a] %/no. reflections	17.7/57,828	17.6/61,664
R-factor (working) %/no. reflections	17.6/52,099	17.4/55,477
R-factor (free) %/no. reflections	23.1/5,729	21.6/6,187
No. protein atoms	2,492	2,511
No. hetero-atoms	348	452
Average *B* values (Å^2)		
Protein atoms	25.1	19.2
NAD^+	17.1	13.3
dTDP	18.5	16.4
Solvents	37.8	34.7
Weighted root-mean square deviations from ideality		
Bond lengths (Å)	0.013	0.014
Bond angles (deg)	2.34	2.35
Trigonal planes (Å)	0.008	0.007
General planes (Å)	0.015	0.013
Torsional angles[b] (deg)	15.4	15.3

[a] *R*-factor = $(\Sigma|F_o - F_c|/\Sigma|F_o|) \times 100$ where F_o is the observed structure-factor amplitude and F_c is the calculated structure-factor amplitude.
[b] The torsional angles were not restrained during the refinement.

RESULTS

Structure of the Wild-type DesIV Complexed with dTDP—A ribbon representation of the DesIV monomer is displayed in Fig. 2*a*. The molecule has overall dimensions of $\sim 48 \times 60 \times 45$ Å^3 and consists of two domains: the N-terminal motif formed by Met^1 to Tyr^{182} and Thr^{220} to Gly^{245} and the C-terminal region delineated by Gly^{183} to His^{219} and Leu^{246} to Lys^{322}. The N-terminal domain, responsible for cradling the NAD^+, contains seven strands of parallel β-sheet (Met^1 to Val^5, Glu^{32} to Asp^{37},

[2] J. B. Thoden and H. M. Holden, unpublished results.
[3] I. Rayment and G. Wesenberg, unpublished results.

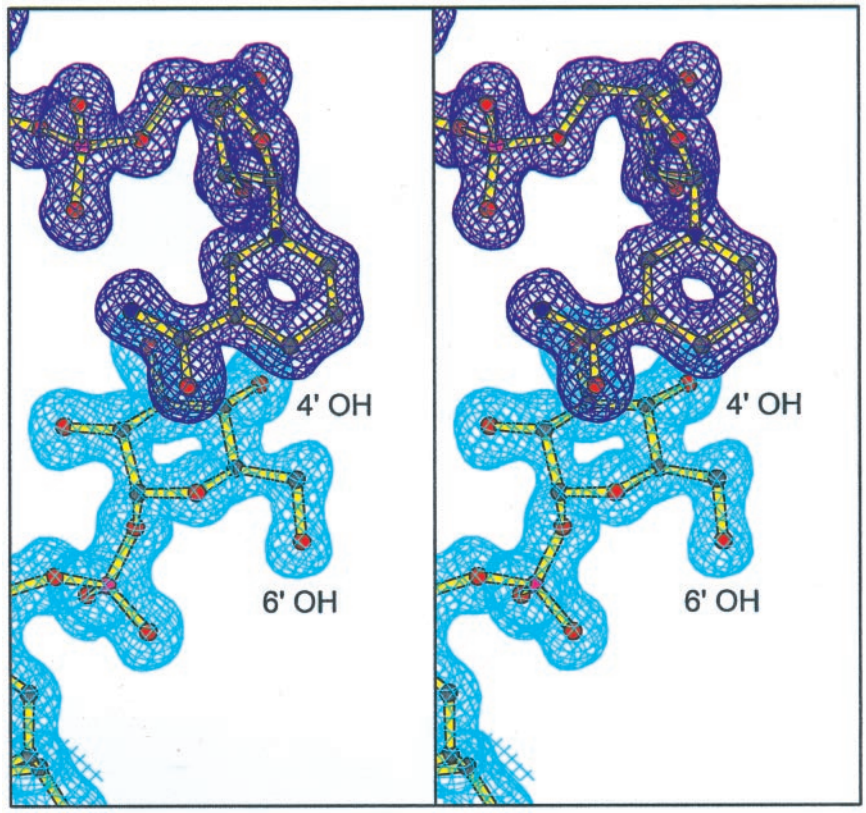


FIG. 1. **Quality of the x-ray model.** Electron density corresponding to the bound dTDP-glucose and NAD^+ in the D128N/E129Q mutant protein structure is displayed. The map was calculated with coefficients of the form $(F_o - F_c)$, where F_o was the native structure factor amplitude and F_c was the calculated structure factor amplitude from the model lacking the coordinates for the ligand. The map was contoured at 3 σ. All figures were prepared with the software package MOLSCRIPT (38).

Arg^{58} to His^{61}, Ala^{79} to His^{82}, Arg^{121} to Thr^{127}, Val^{173} to Tyr^{182}, and Ile^{239} to Ile^{242}). These β-strands are connected by seven α-helices (ranging in length from six to twenty amino acids) and nine reverse turns (six Type I, two Type II, and one Type III). The two largest of these α-helices, Ala^{96} to Ala^{114} and Pro^{150} to Tyr^{169}, are involved in subunit:subunit interactions as discussed below.

The C-terminal domain, responsible for dTDP positioning, is dominated by a three stranded mixed β-sheet formed by Val^{214} to His^{219}, Leu^{246} to Thr^{249}, and Arg^{280} to Leu^{283} and a two stranded parallel β-sheet (Leu^{205} to Leu^{207} and Val^{268} to Val^{271}). Additionally, there are five α-helices and six classical reverse turns (three Type I, two Type II, and one Type II′) in the C-terminal domain. The β-phosphoryl group of dTDP lies within ~6 Å of the nicotinamide ring of the dinucleotide within the cleft formed by the N- and C-terminal domains.

DesIV is known to be a dimer. The enzyme crystallized with its local 2-fold rotational axis coincident to a crystallographic dyad, thereby reducing the contents of the asymmetric unit to one monomer. A ribbon representation of the dimer, generated by crystallographic symmetry, is displayed in Fig. 2*b*. The dimer has overall dimensions of $\sim 74 \times 88 \times 45$ $Å^3$ and a total buried surface area of ~2440 $Å^2$ as calculated according the algorithm of Lee and Richards (25). Three regions from the N-terminal domains of each subunit contribute to the dimeric interface: Ile^{93} to Gln^{108}, Pro^{144} to Ser^{149}, and Leu^{160} to Tyr^{169}. This interface is characterized by numerous hydrogen bonds formed between side chain functional groups from one subunit and backbone atoms from another or between side chain groups contributed by each subunit. In addition, there are numerous water molecules located between the two subunits and several hydrophobic side chains project into the dimeric interface including Ile^{93}, Pro^{144}, Leu^{145}, Pro^{147}, Leu^{160}, and Val^{161} (from each subunit). The major α-helices forming the subunit:subunit boundary run parallel to one another within the individual monomers but anti-parallel between subunits.

Active Site Geometry—A close-up view of the DesIV active site with bound NAD^+ and dTDP is presented in Fig. 3*a*. As can be seen, the thymine ring of the dTDP ligand is wedged into a fairly hydrophobic pocket composed of Leu^{191}, Leu^{194}, Phe^{195}, Pro^{206}, Leu^{207}, and Tyr^{208}. The aromatic side chain of Tyr^{208} participates in a parallel stacking interaction with the thymine ring of the ligand. A schematic of the hydrogen bonding pattern around the dTDP is presented in Fig. 3*b*. The carbonyl oxygens of the thymine ring, at positions 2 and 4, lie within hydrogen bonding distance to the peptidic NH group of Tyr^{208} and a water molecule, respectively. The ring nitrogen, at position 3, forms a hydrogen bond with the carbonyl oxygen of Pro^{206}. As expected, the ribose moiety of the dTDP ligand adopts the $C_{2'}$-*endo* form. The 3-hydroxyl group of the ribose lies within 3.0 Å of $N^{\delta 2}$ of Asn^{250} and 2.6 Å of $N^{\epsilon 2}$ of His^{277}. One of the α-phosphoryl oxygens forms electrostatic interactions with a water molecule and the backbone peptidic NH group of Leu^{191} while the second α-phosphoryl oxygen is positioned within hydrogen bonding distance to the side chains of His^{88} and Arg^{274} and a water molecule. This particular water molecule also lies within ~3 Å of the bridging oxygen between the α- and β-phosphorus atoms, and one of the β-phosphoryl oxygens. Two of the β-phosphoryl oxygens form salt bridges with the guanidinium groups of Arg^{215} and Arg^{274}, respectively. Interestingly, the side chain of Glu^{129}, a residue thought to be important in the reaction mechanism, is located at 2.5 Å from the third β-phosphoryl oxygen thereby suggesting that in this complex the carboxylate group of Glu^{129} is protonated (Fig. 3*b*).

The NAD^+ lies across the C-terminal edge of the seven-stranded parallel β-sheet and is surrounded by ten water molecules within 3.2 Å. Both riboses of the NAD^+ adopt the $C_{2'}$-*endo* conformation and the nicotinamide ring is in the *syn*-conformation as would be expected for a B-side specific enzyme. There is an intramolecular hydrogen bond (2.8 Å) between the carboxamide group of the nicotinamide ring and one of its phosphoryl oxygens. Numerous backbone peptidic groups point

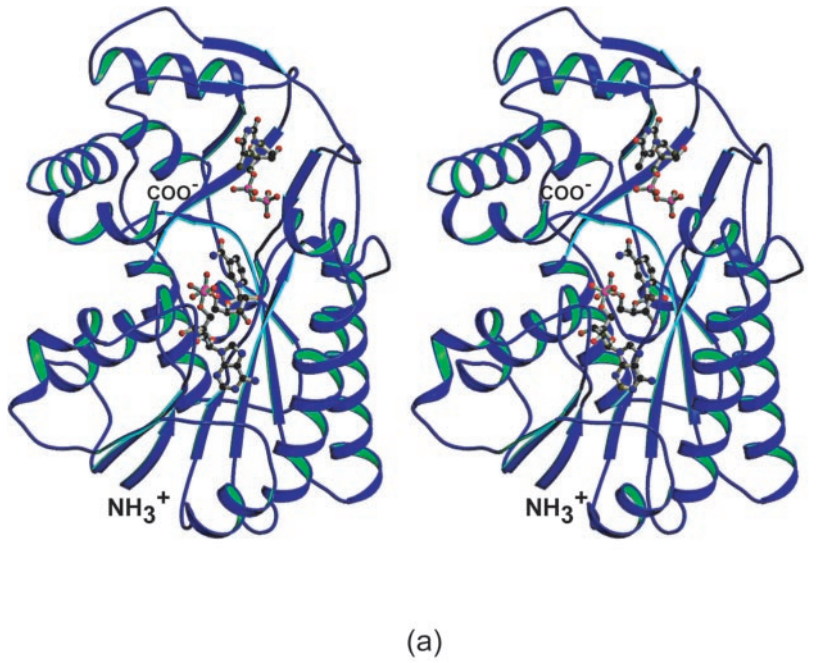


FIG. 2. **Ribbon representation of DesIV.** Shown in *a* is a *ribbon representation* of one subunit of DesIV with the bound ligands, NAD^+ and dTDP, shown in a *ball-and-stick* representation. The dimeric form of the enzyme is presented in *b*.

toward the dinucleotide including those donated by Phe^{11}, Ile^{12}, Ser^{38}, Thr^{40}, Ile^{64}, and Asn^{181}. Five side chains serve to anchor the NAD^+ to the protein: Asp^{37}, Asp^{63}, Ser^{87}, Tyr^{151}, and Lys^{155}. Specifically, the carboxylate of Asp^{63} hydrogen bonds to the C-6 amino group of the adenine ring, the carboxylate of Asp^{37} bridges the 2- and 3-hydroxyl groups of the adenine ribose, the hydroxyl group of Ser^{87} interacts with a phosphoryl oxygen, the amino group of Lys^{155} lies at 2.8 Å from the 3-hydroxyl group of the nicotinamide ribose, and O^{η} of Tyr^{151} is located within 2.7 Å of the nicotinamide ribose 2-hydroxyl. It has been speculated that the positive charge associated with the ϵ-amino group of Lys^{155} (conserved in the SDR family) aids in deprotonation of the conserved tyrosine residue, which for DesIV, is Tyr^{151}. In DesIV the amino group of Lys^{155} and O^{η} of Tyr^{151} are separated by 5.3 Å. Interestingly, the distance between the nitrogen of the nicotinamide ring, which carries the positive charge in the oxidized form of the dinucleotide, and O^{η} of Tyr^{151} is shorter at 4.1 Å. It is possible that the conserved lysine in the SDR superfamily functions primarily for proper dinucleotide positioning and that the positive charge on the nicotinamide ring lowers the pK_a of the conserved tyrosine hydroxyl group.

Structure of the D128N/E129Q Mutant Protein Complexed with dTDP-glucose—In previous studies, the structural equivalents of Asp^{128} and Glu^{129} in DesIV have been implicated to play key roles in the reaction mechanism of the *E. coli* dTDP-

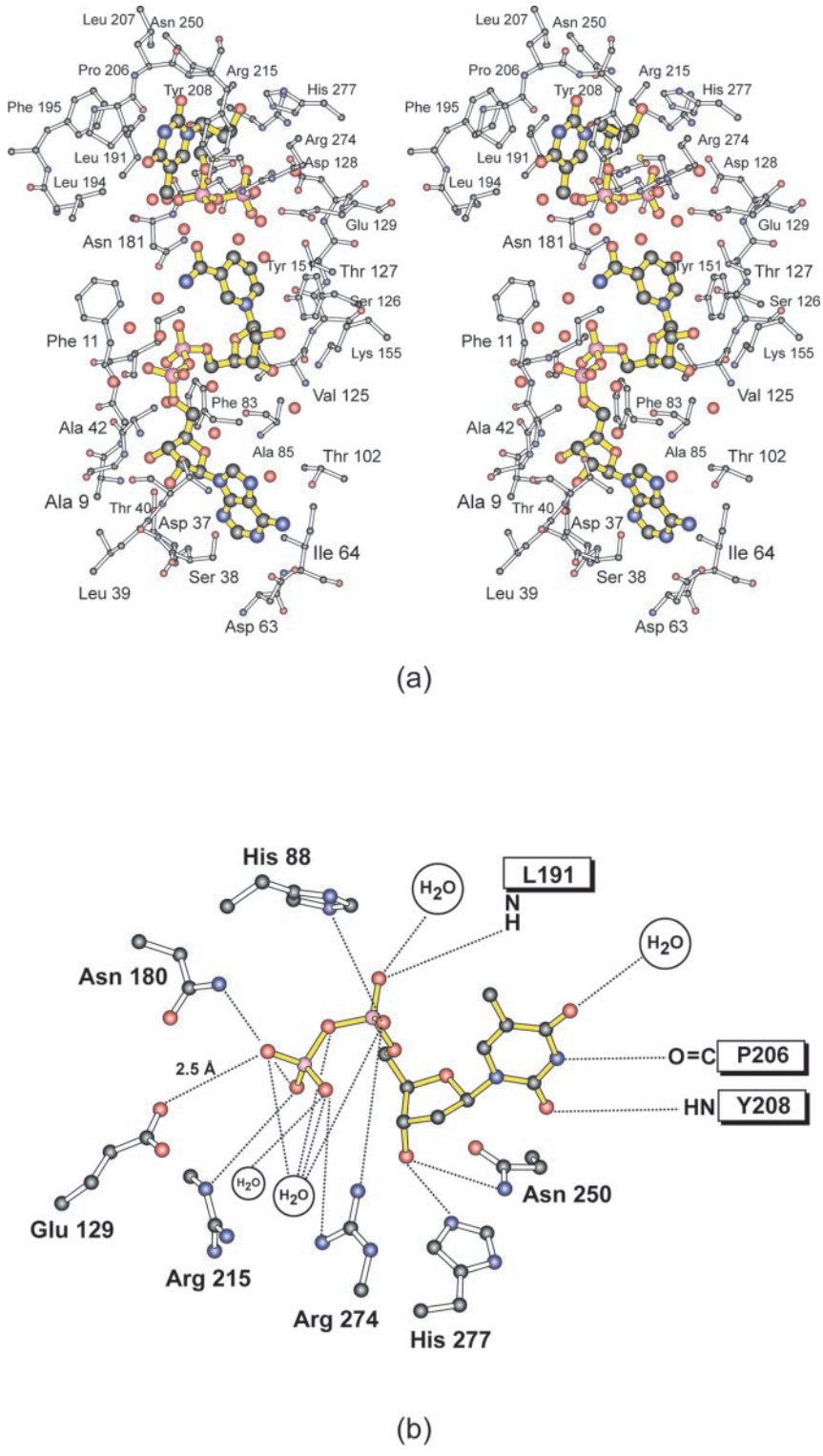
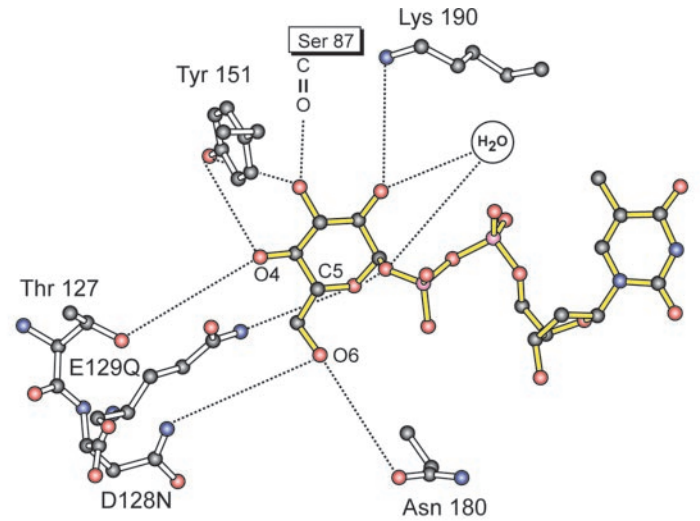


FIG. 3. **Close up view of the DesIV active site.** A close up view of the active site, within ~3.2 Å of the NAD^+ and dTDP ligands, is shown in *a*. For the sake of clarity, Ser^{87} and His^{88} were omitted from the figure. The side chain hydroxyl group of Ser^{87} interacts with a phosphoryl oxygen of the NAD^+ while $N^{\epsilon 2}$ of His^{88} forms a hydrogen bond with a phosphoryl oxygen of the dTDP. A schematic of the hydrogen-bonding pattern around dTDP is displayed in *b*. Possible hydrogen-bonding interactions within ~3.2 Å are indicated by *dashed lines*.

glucose 4,6-dehydratase (26–28). Consequently, in an attempt to solve the structure of DesIV in the presence of its substrate, dTDP-glucose, Asp^{128} and Glu^{129} were replaced with an asparagine and a glutamine, respectively. As can be seen in Fig. 1, the electron density map for the double mutant protein clearly shows the presence of dTDP-glucose in the active site. The α-carbon backbones for the wild-type enzyme and the D128N/E129Q mutant protein superimpose with a root-mean-square deviation of 0.19 Å.

The positions of the dTDP moieties in both the wild-type and mutant protein structures are virtually indistinguishable. Indeed, there are few structural differences in the active sites between these two protein models. Specifically, there are four water molecules in the active site of the wild-type/dTDP complex that are excluded upon binding dTDP-glucose in the D128N/E129Q mutant protein. Additionally, in the wild-type enzyme, the ϵ-amino group of Lys^{190} lies within hydrogen bonding distance to two waters and the C=O group of Ser^{87}. Upon binding dTDP-glucose in the mutant protein, this lysine side chain moves out of the way to accommodate the 2′-hydroxyl group of the sugar. The only other slight change is the position of the side chain of Glu^{129} (Gln^{129} in the mutant protein). As



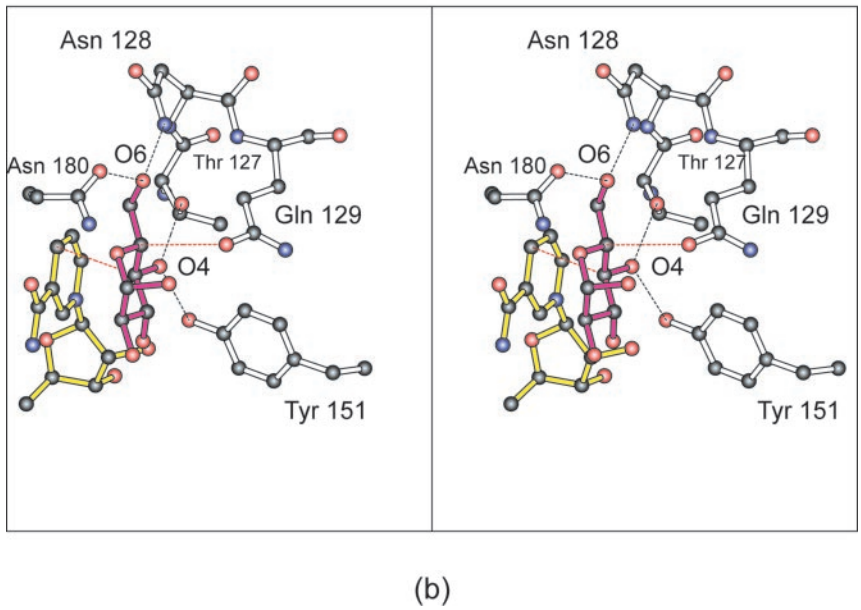


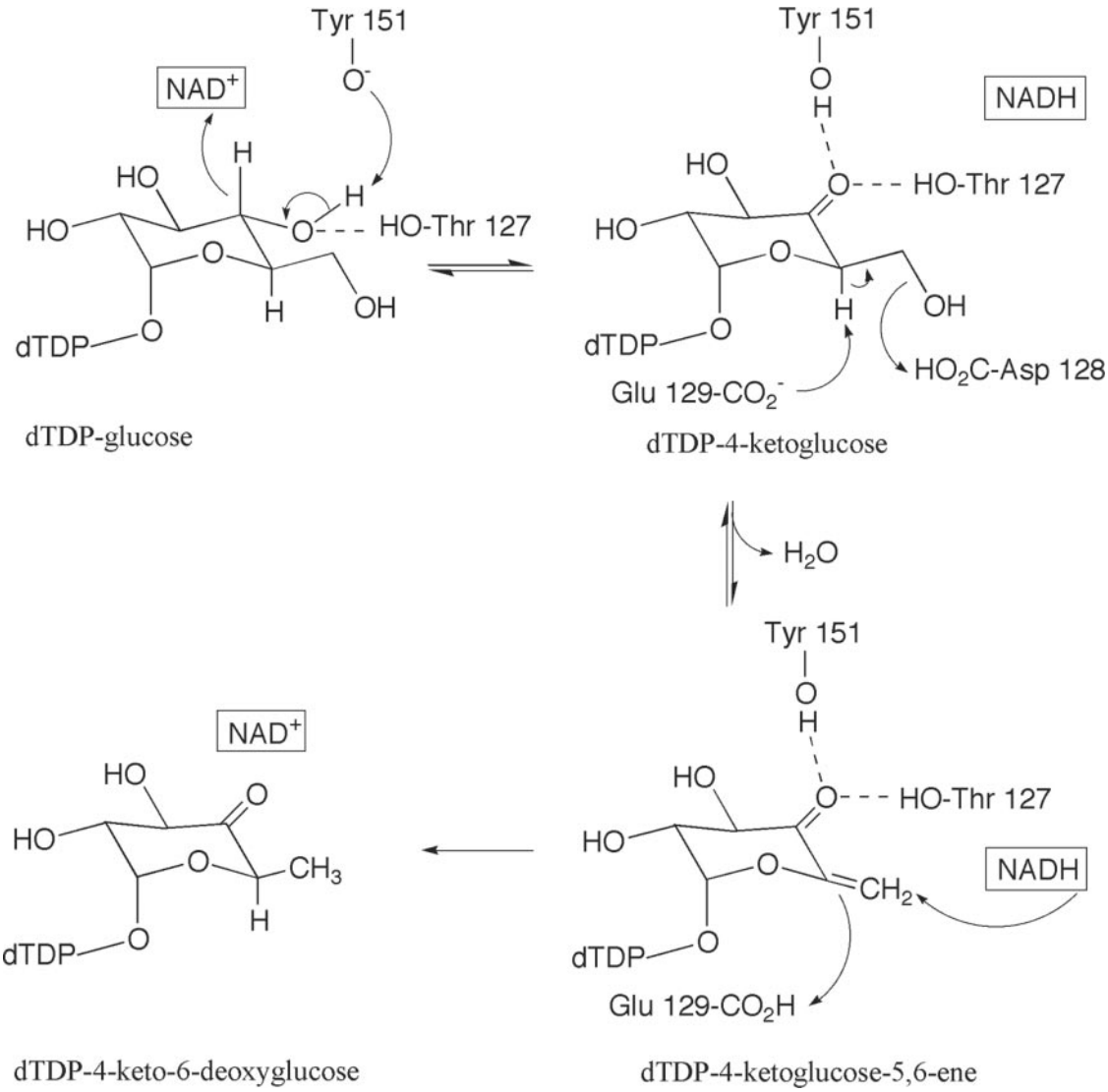
FIG. 4. **Schematic of hydrogen-bonding interactions between the protein and dTDP-glucose.** Possible electrostatic interactions within ~3.2 Å between the D128N/E129Q mutant protein and glucosyl moiety are indicated by the *dashed lines* in *a*. A close up stereo view of the key interactions between the protein and the substrate is presented in *b*. Portions of the NAD^+ and the dTDP-glucose are highlighted in *yellow* and *pink*-filled bonds, respectively. Hydrogen bonds are indicated by *black dashed lines*. *Red dashed lines* show the direction of either hydride transfer from the sugar C-4 to the NAD^+ or proton abstraction from the sugar C-5 by Glu^{129} in the wild-type enzyme (represented here as Gln^{129}).

indicated above, in the wild-type/dTDP complex, the side chain of Glu^{129} is within 2.5 Å of a β-phosphoryl oxygen. In the D128N/E129Q protein complex, the side chain shifts slightly away so that it is now located at ~2.8 Å from the equivalent oxygen in the dTDP-glucose.

A schematic of the hydrogen bonding pattern between the protein and the glycosyl moiety of the dTDP-glucose is presented in Fig. 4*a*. Most of the interactions between the sugar group and the protein occur via side chain functional groups. Specifically $O^{\delta 1}$ of Asn^{128} and $O^{\delta 1}$ of Asn^{180} lie at 2.7 Å and 2.8 Å of the sugar O-6, respectively. Likewise, the hydroxyl groups of Thr^{127} and Tyr^{151} are positioned within 2.6 Å and 2.7 Å of the sugar O-4, respectively. The ϵ-amino group of Lys^{190} sits at 2.7 Å from the sugar O-2 while the carbonyl oxygen of Ser^{87} is situated within 2.7 Å of the sugar O-3. There is only one solvent involved in binding the sugar to the protein and it is located within hydrogen bonding distance to OP-1 and O-2 of the ligand. A stereo view highlighting the geometry between the sugar moiety and those residues thought to be important for catalysis is depicted in Fig. 4*b*. Strikingly, C-4 of the sugar and C-4 of the dinucleotide are separated by 3.3 Å. Importantly, both the side chain hydroxyl groups of Thr^{127} and Tyr^{151} are in the ideal orientation to allow for overlap of their protons with the lone pairs of electrons on O-4 when this group is oxidized to a keto moiety during catalysis.

DISCUSSION

One of the first enzymes to be extensively characterized in the short chain dehydrogenase/reductase superfamily was UDP-galactose 4-epimerase from *E. coli* (29, and references



SCHEME 2

therein). This enzyme catalyzes the interconversion of UDP-galactose and UDP-glucose during normal galactose metabolism. The reaction mechanism is thought to occur through the following steps: (1) abstraction of 4′-hydroxyl proton by a conserved tyrosine and transfer of a hydride from C-4 of the sugar to C-4 of the nicotinamide ring of NAD^+ to yield NADH and a 4′-ketopyranose intermediate, (2) rotation of this intermediate in the active site to present the opposite face of the sugar to the reduced dinucleotide, and (3) finally, transfer of the hydride back to C-4 of the sugar and donation of the proton from the conserved tyrosine back to the 4′-keto oxygen. A key feature of this reaction mechanism is an active site cleft large enough to accommodate the presumed movement of the 4′-ketopyranose intermediate.

The first half of the 4,6-dehydratase reaction mechanism, outlined in Scheme 2, again involves the abstraction of the 4′-hydroxyl proton from the sugar by the conserved tyrosine residue and transfer of a hydride from the sugar C-4 to C-4 of the nicotinamide ring. The catalytic mechanisms for the epimerase and the dehydratase, however, diverge at this point. In the 4,6-dehydratases, a proton is removed from C-5 of the sugar and the C-6 hydroxyl group is protonated, leading to dehydration and a 4-keto-5,6-ene intermediate (18). The final product is generated when the hydride from NADH is transferred to C-6, and a proton is donated back to C-5. Evidence for the hydride being transferred from C-4 of the sugar to the dinucleotide and then back to C-6 of the sugar comes from previous labeling studies (30, 31). As would be expected in light of the differences in their reaction mechanisms, the volume for the active site of the *E. coli* UDP-galactose 4-epimerase is 25% larger than that for DesIV as calculated with the software VOIDOO (32, 33).

While the role of the conserved tyrosine in the first half of the dehydratase reaction mechanism has been well-established (34), the identities of the general acid and general base required for the second half of catalysis, as outlined in Scheme 2, are still open to question. It was originally suggested that in the *E. coli* dTDP-glucose 4,6-dehydratase, a conserved glutamate, namely Glu^{129} in DesIV, might function in both capacities by acting as a proton shuttle from C-5 to the 6′-hydroxyl group (26). As noted in Ref. 26, this was an especially attractive proposal in that the glutamate would be in the proper ionization state for the next round of catalysis. When the structures of the *S. enterica* and *S. suis* dTDP 4,6-dehydratases were subsequently solved, it was suggested that an aspartic acid/glutamate pairing was responsible for the dehydration step (13, 14). This idea was put forth in light of the fact that in the *S. suis* dehydratase/NAD^+/dTDP-glucose model, Asp^{126} was situated at 2.6 Å from the C-6 hydroxyl oxygen while Glu^{127} was located at 3.1 Å from C-5 of the glucosyl moiety.

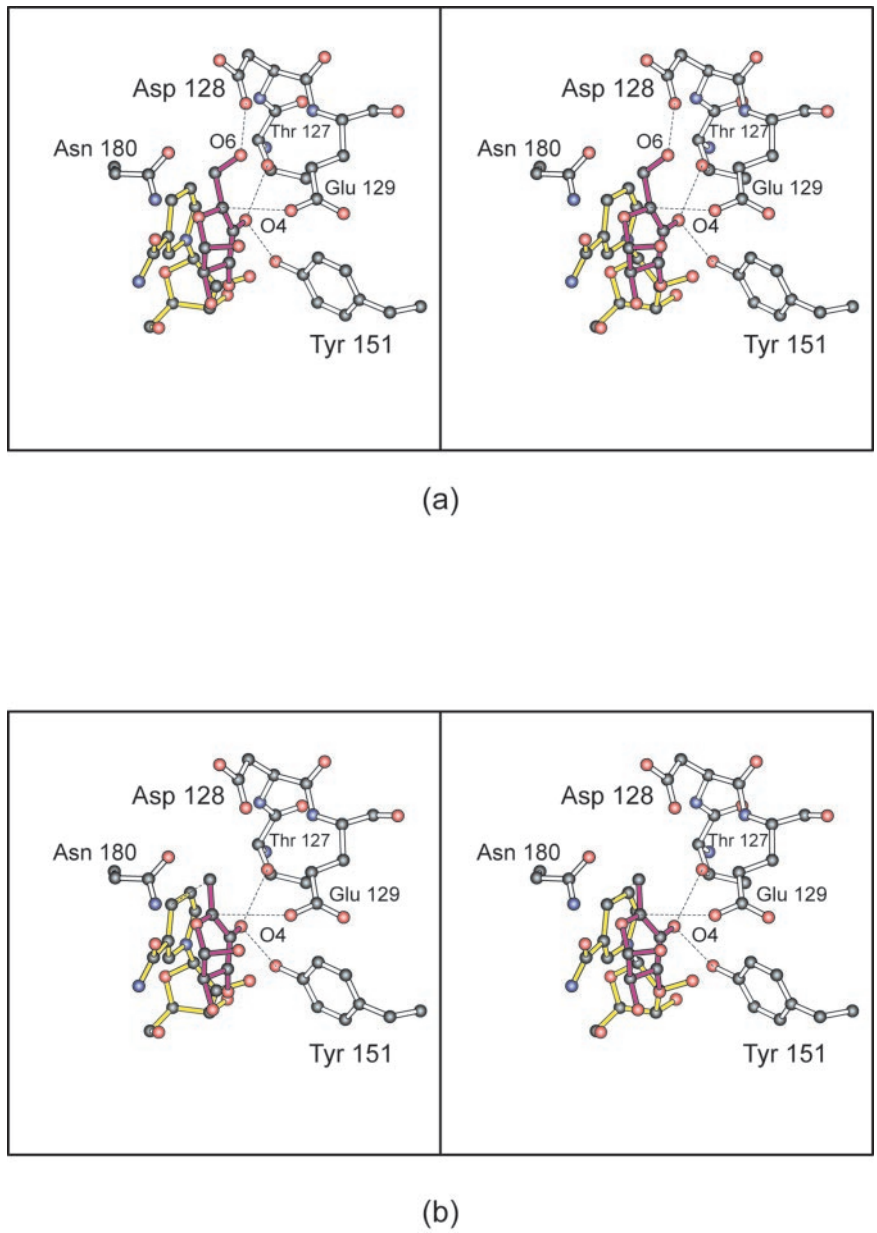


FIG. 5. **Models for the possible intermediates occurring during the reaction mechanism of DesIV.** Shown in *a* is a model for the 4′-keto intermediate. The *black dashed lines* indicate hydrogen-bonding interactions between the C=O of the sugar and the side chains of Thr^{127} and Tyr^{151}. In the 4′-keto intermediate it is presumed that there is an additional hydrogen bond between the sugar C-6 hydroxyl and the carboxylate of Asp^{128}, again as indicated by the *black dashed line*. The direction of attack of the Glu^{129} carboxylate as it abstracts the C-5 hydrogen in the next step of the reaction pathway is also indicated by a *dashed black line*. Note that the hybridization about the sugar C-4 is sp^2 in this figure and the nicotinamide ring of the dinucleotide is slightly puckered according to that observed in previous studies with UDP-galactose 4-epimerase (39). A model for the 4-keto-glucose-5,6-ene intermediate is presented in *b*. In this case, C-4, C-5, and C-6 are sp^2 hybridized. Potential hydrogen bonds are indicated by *black dashed lines*. The direction of hydride transfer from the reduced cofactor back to C-6 and proton donation by Glu^{129} back to C-5 of the sugar are indicated by *black dashed lines*.

On the basis of the active site geometries observed in this high resolution x-ray analysis, it is now possible to elaborate more fully on the reaction mechanism of DesIV in particular and on the dTDP-glucose 4,6-dehydratases in general. In the first step of the reaction mechanism, the phenolate group of Tyr^{151} acts as the general base to abstract a proton from the 4′-hydroxyl group of the sugar while the hydride on C-4 is transferred to NAD^+. A number of structural events must occur as the hydride is transferred to C-4 of the nicotinamide ring. First, the sugar C-4 changes its hybridization state from sp^3 to sp^2 and as a consequence moves within the active site by ~0.4 Å. Secondly, the proton on the sugar 4′-hydroxyl group is shifted toward the phenolate of Tyr^{151}. Finally, the nicotinamide ring of the dinucleotide adopts a puckered conformation. The net result of these changes is oxidation of the sugar 4′-hydroxyl group to a keto moiety. With these considerations in mind, a three-dimensional model of this 4-ketose intermediate was built into the active site of DesIV and is presented in Fig. 5*a*. Note that the two lone pairs of electrons on the oxygen of the keto functionality are in the proper orientation for ideal hydrogen bonding to both the side chains of Tyr^{151} and Thr^{127}. This 4-dehydrogenation of the substrate serves as the activating step for the subsequent β-elimination of water between C-5 and C-6 by lowering the pK_a of the C-5 proton into the range of 18–19 (35). Furthermore, Thr^{127} is absolutely conserved among the dTDP-glucose 4,6-dehydratase amino acid sequences determined to date thus suggesting that the formation of a hydrogen bond with the sugar 4-keto group plays a role in promoting catalysis.

In the second step of the reaction mechanism as outlined in Scheme 2, Glu^{129} removes the proton from C-5 of the sugar. As proton removal occurs at C-5, the C-6 hydroxyl will eliminate as water by donation of a proton from the side chain of Asp^{128}, thereby forming a dTDP-4-ketoglucose-5,6-ene intermediate. A model for this intermediate was built into the DesIV active site and is shown in Fig. 5*b*. The elimination of water has been shown to proceed via *syn* stereochemistry with respect to the C-5(H) and the C-6(OH) (36) and as a concerted mechanism (28). Although this pathway reflects concerted proton abstraction and leaving group elimination, it will involve asynchronous bond cleavage, with abstraction of the C-5 proton leading

the elimination of the C-6 hydroxyl, an idea also suggested for polysaccharide lyases (37).

When the C-6 hydroxyl is eliminated, the distance between C-6 and the water that is formed increases to van der Waals distance, a scenario pointed out in earlier work (14). If the water is not free to move, the sugar moiety might adjust in the active site by moving C-6 toward NADH and thus lining it up for hydride transfer from the reduced cofactor. Such a movement is not unreasonable given that in the D128N/E129Q mutant protein structure, C-4 of the NAD^+ is located within 3.6 Å of the sugar C-6.

It can be postulated that the hydrogen bonds between the sugar 4′-keto group and the side chains of Tyr^{151} and Thr^{127} serve to enhance the electrophilicity of C-6, thereby inducing hydride transfer from NADH (Fig. 5*b*). But rather than form an enol intermediate, which is a high energy compound, there will likely be proton transfer from Glu^{129} to C-5 to give the final product. This is thus a concerted reaction, but asynchronous, with C-H bond formation at C-6 leading C-H formation at C-5. The ionization states of both Tyr^{151} and Asp^{128} are reset upon product release from the active site. It has recently been suggested, on the basis of *ab initio* electronic structure calculations, that this final aspect of the dehydratase reaction might be stepwise, with an enol intermediate (15). However, the consideration of a model in which there is both hydrogen bonding to the 4′-keto group and concerted proton transfer to C-5 was not taken into account.

The present structure of the double mutant protein contains bound NAD^+ and unchanged substrate. Although the mutations at positions 128 and 129 prevent the second step of the reaction, there is no reason why the first step should not have occurred. The electron density for the glucosyl moiety clearly shows a hydroxyl rather than keto functionality (Fig. 1). The equilibrium constant for the oxidation of a secondary alcohol such as the one at C-4 (based on the value for oxidation of glycerol-3-phosphate to dihydroxyacetone-phosphate) is $\sim 5 \times 10^{-5}$ at pH 7, increasing a factor of 10 per pH unit at higher pH. While the equilibrium constant can certainly be altered in the enzyme active site, it is very unlikely to exceed 10^{-2}, and thus the keto intermediate would not be detected in this x-ray study. In agreement with this, Gross *et al.* (18) failed to detect the keto intermediate during steady state turnover, while the glucosene intermediate was observed. The level of keto intermediate formed from the substrate, while small, is sufficient to permit the thermodynamically favorable elimination of water in the second step of the reaction.

In summary, the structures of wild-type DesIV complexed with dTDP, and a double site-directed mutant protein with bound NAD^+ and dTDP-glucose have been determined to high resolution. From this analysis, the roles of the conserved tyrosine, threonine, aspartate, and glutamate in the reaction mechanism have been further defined for both this enzyme and for other 4,6-dehydratases.

Acknowledgments—We thank Dr. James B. Thoden for aid in x-ray data collection and for helpful discussions. We also acknowledge Dr. Kit-Yin Ling and Nicole Koropatkin for help with cloning and protein purifications. The aid of Dr. Adrian D. Hegeman in determining the NAD(H) content of crystals is gratefully acknowledged.

REFERENCES

1. Xue, Y., Zhao, L., Liu, H. W., and Sherman, D. H. (1998) *Proc. Natl. Acad. Sci. U. S. A* **95,** 12111–12116
2. Weymouth-Wilson, A. C. (1997) *Nat. Prod. Rep.* **14,** 99–110
3. Tang, L., and McDaniel, R. (2001) *Chem. Biol.* **8,** 547–555
4. Schlünzen, F., Zarivach, R., Harms, J., Bashan, A., Tocilj, A., Albrecht, R., Yonath, A., and Franceschi, F. (2001) *Nature* **415,** 814–821
5. Hansen, J. L., Ippolito, J. A., Ban, N., Nissen, P., Moore, P. B., and Steitz, T. A. (2002) *Mol. Cell.* **10,** 117–128
6. Khosla, C., and Zawada, R. J. (1996) *Trends Biotechnol.* **14,** 335–341
7. Mendez, C., and Salas, J. A. (2001) *Trends Biotechnol.* **19,** 449–456
8. Chen, S., Roberts, J. B., Xue, Y., Sherman, D. H., and Reynolds, K. A. (2001) *Gene* (*Amst.*) **263,** 255–264
9. Kallberg, Y., Oppermann, U., Jörnvall, H., and Persson, B. (2002) *Eur. J. Biochem.* **269,** 4409–4417
10. Jörnvall, H., Persson, B., Krook M., Atrian, S., González-Duarte, R., Jeffery, J., and Ghosh, D. (1995) *Biochemistry* **34,** 6003–6013
11. Duax, W. L., Griffin, J. F., and Ghosh, D. (1996) *Curr. Opin. Struct. Biol.* **6,** 813–823
12. Duax, W. L., Ghosh, D., and Pletnev, V. (2000) *Vitam. Horm.* **58,** 121–148
13. Allard, S. T. M., Giraud, M. F., Whitfield, C., Graninger, M., Messner, P., and Naismith, J. H. (2001) *J. Mol. Biol.* **307,** 283–295
14. Allard, S. T. M., Beis, K., Giraud, M. F., Hegeman, A. D., Gross, J. W., Wilmouth, R. C., Whitfield, C., Graininger, M., Messner, P., Allen, A. G., Maskell, D. J., and Naismith, J. H. (2002) *Structure* **10,** 81–92
15. Beis, K., Allard, S. T. M., Hegeman, A. D., Murshudov, G., Philp, D., and Naismith, J. H. (2003) *J. Am. Chem. Soc.* **125,** 11872–11878
16. Sambrook, J., and Russell, D. W. (2001) *Molecular Cloning: a Laboratory Manual*, 3rd Ed., Cold Spring Harbor Laboratory Press, New York
17. Okazaki, R., Okazaki, T., Strominger, J. L., and Michelson, A. M. (1962) *J. Biol. Chem.,* **237,** 3014–3026
18. Gross, J. W., Hegeman, A. D., Vestling, M. M., and Frey, P. A. (2000) *Biochemistry* **39,** 13633–13640
19. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
20. Navaza, J. (1994) *Acta Crystallogr.* **A50,** 157–163
21. Roussel, A., Fontecilla-Camps, J. C., and Cambillau, C. (1990) *Acta Cryst.* **A4,** C66-C67
22. Tronrud, D. E., Ten Eyck, L. F., and Matthews, B. W. (1987) *Acta Crystallogr. Sect. A* **43,** 489–501
23. Kabsch, W. (1988) *J. Appl. Crystallogr.* **21,** 67–71
24. Kabsch, W. (1988) *J. Appl. Crystallogr.* **21,** 916–924
25. Lee, B., and Richards, F. M. (1971) *J. Mol. Biol.* **55,** 379–400
26. Hegeman, A. D., Gross, J. W., and Frey, P. A. (2001) *Biochemistry* **40,** 6598–6610
27. Gross, J. W., Hegeman, A. D., Gerratana, B., and Frey, P. A. (2001) *Biochemistry* **40,** 12497–12504
28. Hegeman, A. D., Gross, J. W., and Frey, P. A. (2002) *Biochemistry* **41,** 2797–2804
29. Thoden, J. B., and Holden, H. M. (1998) *Biochemistry* **37,** 11469–11477
30. Melo, A., Elliott, W. H., and Glaser, L. (1968) *J. Biol. Chem.* **243,** 1467–1474
31. Gabriel, O., and Lindquist, L. C. (1968) *J. Biol. Chem.* **243,** 1479–1484
32. Kleywegt, G. J., and Jones, T. A. (1993) *Newsl. Prot. Crystallogr.* **29,** 26–28
33. Kleywegt, G. J., and Jones, T. A. (1994) *Acta Crystallogr. Sect. D.* **50,** 178–185
34. Gerratana, B., Cleland, W. W., and Frey, P. A. (2001) *Biochemistry* **40,** 9187–9195
35. March, J. (1985) *Advanced Organic Chemistry*, 3rd Ed., pp. 218–236, John Wiley and Sons, New York
36. Snipes, C. E., Brillinger, G. U., Sellers, L., Mascaro, L., and Floss, H. G. (1977) *J. Biol. Chem.* **252,** 8113–8117
37. Charnock, S. J., Brown, I. E., Turkenburg, J. P., Black, G. W., and Davies, G. J. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 12067–12072
38. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950
39. Thoden, J. B., Frey, P. A., and Holden, H. M. (1996) *Biochemistry* **35,** 5137–5144