Structure and Kinetics of a Monomeric Glucosamine 6-Phosphate Deaminase

Florence Vincent, Gideon J Davies, James A Brannigan

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 280, No. 20, Issue of May 20, pp. 19649–19655, 2005

Printed in U.S.A.

Structure and Kinetics of a Monomeric Glucosamine 6-Phosphate Deaminase

MISSING LINK OF THE NagB SUPERFAMILY?*[S]

Received for publication, February 24, 2005, and in revised form, March 8, 2005
Published, JBC Papers in Press, March 8, 2005, DOI 10.1074/jbc.M502131200

Florence Vincent‡, Gideon J. Davies§, and James A. Brannigan

From the Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York YO10 5YW, United Kingdom

Glucosamine 6-phosphate is converted to fructose 6-phosphate and ammonia by the action of the enzyme glucosamine 6-phosphate deaminase, NagB. This reaction is the final step in the specific GlcNAc utilization pathway and thus decides the metabolic fate of GlcNAc. Sequence analyses suggest that the NagB "superfamily" consists of three main clusters: multimeric and allosterically regulated glucosamine-6-phosphate deaminases (exemplified by *Escherichia coli* NagB), phosphogluconolactonases, and monomeric hexosamine-6-phosphate deaminases. Here we present the three-dimensional structure and kinetics of the first member of this latter group, the glucosamine-6-phosphate deaminase, NagB, from *Bacillus subtilis*. The structures were determined in ligand-complexed forms at resolutions around 1.4 Å. *Bsu*NagB is monomeric in solution and as a consequence is active (k_{cat} 28 s^{-1}, $K_{m(app)}$ 0.13 mM) without the need for allosteric activators. A decrease in activity at high substrate concentrations may reflect substrate inhibition (with K_i of ~4 mM). The structure completes the NagB superfamily structural landscape and thus allows further interrogation of genomic data in terms of the regulation of NagB and the metabolic fate(s) of glucosamine 6-phosphate.

NagB (glucosamine-6-phosphate deaminase; 2-amino-2-deoxy-D-glucose-6-phosphate aminohydrolase (ketol-isomerizing); EC 3.5.99.6) performs the isomerization and deamination reactions that transform glucosamine 6-phosphate (GlcN-6-P)[1] to fructose 6-phosphate (F6P), with the concomitant release of ammonia. This reaction follows the first step in the biosynthetic pathway of amino-sugar nucleotides, which is the formation of glucosamine 6-phosphate, and is the final specific step in the pathway of GlcNAc utilization (1) (Fig. 1). The amino-sugar nucleotides are subsequently used as precursors for the biosynthesis of bacterial cell wall peptidoglycans and teichoic acids.

NagB homologs form a superfamily, whose quaternary and genomic organization is particularly interesting. In both *Escherichia coli* and *Bacillus sphaericus* (2), *nagB* and *nagA* are in the same large operon with the GlcNAc-phosphoenolpyruvate-phosphotransferase system. In most bacilli, including *Bacillus subtilis*, *nagA* and *nagB* exist in a much shorter operon. In *B. subtilis*, the genes encoding NagA and NagB overlap by 2 bp, and in *Bacillus halodurans* they overlap by 13 bp, suggesting that these genes are co-transcribed and translated. The complexity of genomic organization is also reflected, to some extent, in the quaternary structures of the enzymes themselves. In humans (3), mice, and *E. coli* (perhaps the most well-characterized), NagA is a tetramer (4, 5) and NagB a hexamer regulated by an allosteric mechanism (6), whereas in *B. subtilis*, a much simpler organization exists: NagA is dimeric (7) and NagB is a monomer.

Quaternary structure and enzyme specificity thus manifest themselves in a NagB superfamily (Pfam PF01182) of three clusters (Fig. 2). The first cluster contains both human and *E. coli* NagB enzymes. It is amazing that the primary sequence of *Eco*NagB is more similar to its human counterpart than to its Gram-positive bacterial cousins. This branch of the family is defined by the shared quaternary structure and mode of regulation by the allosteric effector GlcNAc-6-P. The second branch of the family comprises the simpler, monomeric NagB proteins exemplified by *Bsu*NagB (described here). Thus, there is a clear split between the Gram-positive and Gram-negative eubacterial divide, exemplified by the first two major classes, which include galactosamine/glucosamine isomerase/deaminase proteins. The third cluster of related sequences are predicted to be phosphogluconolactonases (6-phospho-D-glucono-1,5-lactone lactonohydrolase (6-PGL); EC 3.1.1.31), which catalyze the reaction 6-phospho-D-glucono-1,5-lactone + H_2O to 6-phospho-D-gluconate, the second step in the pentose phosphate pathway (8). The genes for these enzymes are always adjacent to those encoding an NAD-dependent glucose-6-phosphate dehydrogenase. In some organisms, such as *Plasmodium*, 6-PGL and glucose-6-phosphate dehydrogenase are fused to give a single, multifunctional enzyme (9). In bacilli, the annotated glucose-6-phosphate dehydrogenase gene is termed *zwf*, which is associated with a putative 6-P-gluconate dehydrogenase gene (*yqjI*), whereas genes encoding 6-PGL are conspicuous by their absence.

Structural representatives for the oligomeric, allosterically activated NagB are reported, and a Protein Data Bank file for a putative 6-PGL has been deposited. Here we present the expression, purification, kinetic characterization, and the 1.5-Å

* This work was supported by Wellcome Trust Grant 063963. 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 2bkv and 2bkx) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[S] The on-line version of this article (available at http://www.jbc.org) contains three additional figures.

‡ Present address: Lab. AFMB, UMR6098, CNRS, 31 Chemin Joseph Aiguier, 13402 Marseille cedex 02, France.

§ A Royal Society University Research Fellow. To whom correspondence should be addressed. Tel.: 44-1904-328260; Fax: 44-1904-328266; E-mail: davies@ysbl.york.ac.uk.

[1] The abbreviations used are: GlcN-6-P, glucosamine 6-phosphate; F6P, fructose-6-phosphate; 6-PGL, 6-phospho-D-glucono-1,5-lactone lactonohydrolase.

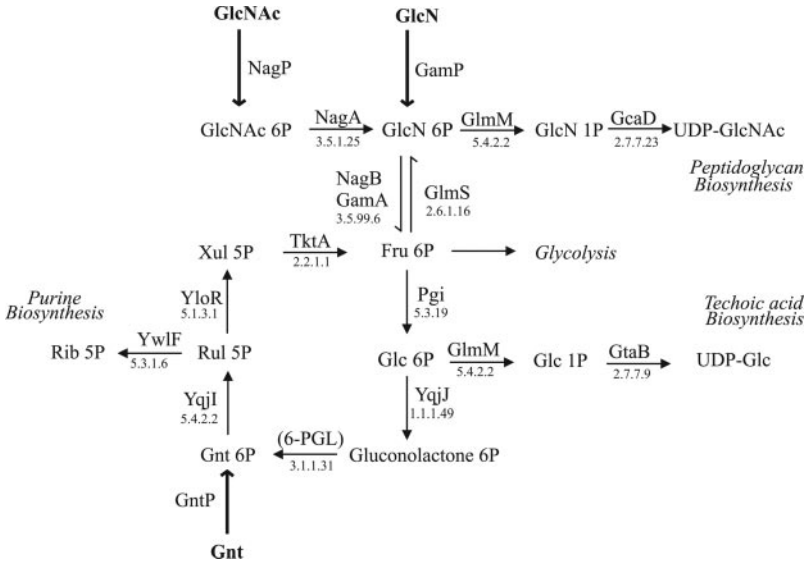

FIG. 1. **Pathway of amino-sugar metabolism in *B. subtilis*.** GlcNAc, GlcN, and gluconate (*Gnt*) are imported by specific phosphotransferase system components (EC 2.7.1.69) and are converted to form the precursors of peptidoglycan and cell wall biosynthesis. NagB, along with GlmS, has been termed a "gate" with a pivotal role in the utilization of GlcN-6P, since it lies on the fulcrum of peptidoglycan formation and glycolysis (30). A *Bacillus* gene encoding 6-PGL has yet to be identified.

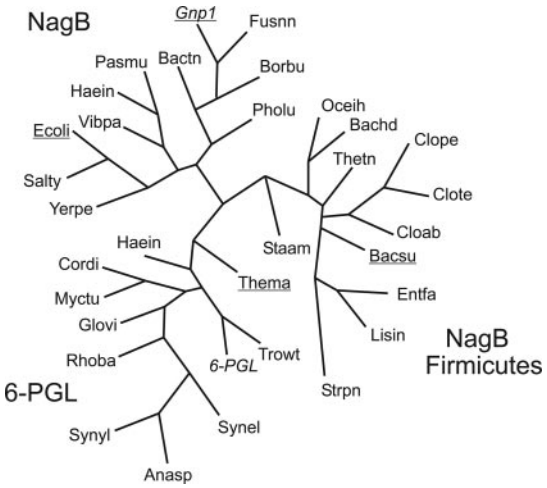

FIG. 2. **Phylogenetic tree of NagB-related protein sequences.** Selected bacterial protein sequences (see Supplementary Material) were aligned with those of human (in *italics*) 6-PGL and glucosamine-6-phosphate deaminase (*GNP1*). Proteins of known three-dimensional structure are *underlined*. The tree was generated using the protpars program (parsimony method) in the Phylip package, using default parameters and bootstrap. The sequences partition into three clusters, with NagB from *B. subtilis* in a monophyletic clade from Firmicutes bacteria. The bacterial 6-PGL proteins are mainly from cyanobacteria and actinobacteria, whereas proteobacterial sources predominate in the cluster with Gnp1. Codes are based on SwissProt species designations as follows. *Anasp*, *Anabaena* sp.; *Bachd*, *B. halodurans*; *Bacsu*, *B. subtilis*; *Bactn*, *Bacteroides thetaiotaomicron*; *Borbu*, *Borrelia burgdorferi*; *Cloab*, *Clostridium acetobutylicum*; *Clope*, *C. perfringens*; *Clote*, *C. tetani*; *Cordi*, *Corynebacterium diptheriae*; *Ecoli*, *Escherichia coli*; *Entfa*, *Enterococcus faecalis*; *Fusnn*, *Fusobacterium nucleatum*; *Glovi*, *Gleobacter violaceus*; *Haein*, *Haemophilus influenzae*; *Lisin*, *Listeria innocua*; *Myctu*, *Mycobacterium tuberculosis*; *Oceih*, *Oceanobacillus iheyensis*; *Pasmu*, *Pasteurella multocida*; *Pholu*, *Photorhabdus luminescens*; *Rhoba*, *Rhodopseudomonas palustris*; *Salty*, *Salmonella typhimurium*; *Staam*, *Staphylococcus aureus*; *Strpn*, *Streptococcus pneumoniae*; *Synel*, *Thermosynechococcus elongatus*; *Synyl*, Synechocystis sp.; *Thema*, *Thermatoga maritima*; *Thetn*, *Thermoanaerobacter tengcongensis*; *Trowt*, *Tropheryma whipplei*; *Vibpa*, *Vibrio parahemeolyticus*; *Yerpe*, *Yersinia pestis*.

resolution crystal structure of the "missing link" of the superfamily, the *B. subtilis* NagB. The three-dimensional structure of this monomeric enzyme has also been solved in complex with the products of the reaction with GlcN-6-P and glucose 6-phosphate. *Bsu*NagB is active without the need for allosteric regulation, but its kinetic activity displays substrate inhibition at high concentrations of GlcN-6-P. *Bsu*NagB completes the structural portfolio and thus allows further interrogation of genomic data in terms of the regulation of NagB and the metabolic fate(s) of glucosamine 6-phosphate.

MATERIALS AND METHODS

Cloning and Protein Production—The *nagB* gene was amplified from chromosomal DNA of *B. subtilis* strain IG20 (168 *trp*$^-$) by the PCR method. Oligonucleotides used to prime the PCR were designed to incorporate suitable restriction endonuclease sites (NdeI and XhoI) for convenient cloning into the T7-promoter-based expression vectors pET28a and pET26b, to yield three separate constructs. These encoded a native protein plus hexahistidine-tagged versions at the N or C terminus. These constructs were designed to test the hypothesis that NagA and NagB could interact; we were unable to demonstrate any such interaction either *in vitro* or through co-expression. The genes were expressed in *Escherichia coli* BL21(DE3), a strain with an inducible T7 RNA polymerase gene (Novagen).

For protein production, cell cultures were grown in LB medium containing 35 μg/ml kanamycin to an optical density of 0.7 at 600 nm before induction of protein expression by the addition of isopropyl 1-thio-β-D-galactopyranoside to a final concentration of 1 mM. After a further 3 h of growth, the cells were harvested by centrifugation and disrupted by sonication. The lysate was clarified by centrifugation and applied to a Q-Sepharose column resolved using a linear gradient of increasing sodium chloride. NagB-containing fractions were concentrated and further purified by gel filtration (Superdex75 16/60; Amersham Biosciences) in a 20 mM Tris-HCl (pH 8), 200 mM NaCl buffer, with apparent molecular mass estimated by comparison with standard protein markers. Pure fractions of NagB were pooled, washed into 20 mM Tris-HCl (pH 8), and concentrated to 43 mg/ml using a 10-kDa cut-off ultracentrifugation membrane (Vivaspin).

Crystallization, Data Collection, and Processing—Crystals of NagB were grown by vapor phase diffusion using the hanging drop method with an equal volume (1 μl) of protein (10 mg/ml) and reservoir solution composed of 20% polyethylene glycol 8000, 0.1 M Tris-HCl (pH 8), and 0.2 M calcium acetate. Substrate soaks were performed by incubating crystals in reservoir solution containing 1 mM glucosamine 6-phosphate or glucose 6-phosphate. A single crystal of NagB was transferred to a solution containing the mother liquor components with 25% polyethylene glycol 550 MME as a cryoprotectant. The crystal was suspended in a film of solution using a rayon loop and flash-cooled to 120 K. The diffraction quality of the crystals was assessed using a home source, after which the crystals were taken to a synchrotron. A single wavelength experiment was conducted on beamline ID29 at the European Synchrotron Radiation Facility (ESRF), Grenoble, at a temperature of 100 K using an ADSC CCD detector. Data were integrated, scaled, and reduced using HKL2000 and SCALEPACK (23). The crystals are monoclinic, with unit cell dimensions $a = 63$, $b = 48$, $c = 72$ Å, and $\beta = 91°$. This corresponds to a solvent content of 35.9%, assuming 2 molecules per asymmetric unit. All further crystallographic computations were carried out using the CCP4 suite of programs (24).

Phasing, Model Building, and Refinement—The structure was solved by molecular replacement, using AMoRe (25), with the NagB from

TABLE I
Crystal, data, and refinement statistics

Crystal Parameters	Native	+F6P from GlcN-6-P
Space group	$P2_1$	$P2_1$
Cell dimensions		
a (Å)	63.0	63.0
b (Å)	48.0	48.0
c (Å)	71.8	71.8
β (degrees)	91.0	91.0
Molecules/asymmetric unit	2	2
Data quality		
Wavelength (Å)	0.9168 (ID29)	0.934 (ID14–1)
Resolution of data (Å)	29–1.5	26–1.4
Resolution of data (outer shell) (Å)	1.55–1.50	1.45–1.40
Unique reflections	67,231	78,148
R_{merge} (outer shell)[a]	0.107 (0.498)	0.048 (0.299)
Mean $I/\sigma I$ (outer shell)	14.7 (4.6)	15.8 (4.3)
Completeness (outer shell) %	98.8 (97.7)	92.9 (61.6)[b]
Multiplicity (outer shell)	3.8 (3.5)	3.6 (3.2)
Refinement		
Protein atoms	3811	3818
Solvent waters	402	855
Ligand	Disordered sugar-phosphate	F6P
R_{cryst}	0.155	0.123
R_{free}	0.210	0.165
Root mean square deviation 1–2 bonds (Å)	0.019	0.020
Root mean square deviation 1–3 angles (degrees)	1.769	1.737

[a] $R_{merge} = (\Sigma_{hkl} \Sigma_i |I_{hkl} - (I_{hkl})| / \Sigma_{hkl} \Sigma_i [I_{hkl}])$.
[b] This incompleteness reflects data from the corners of a square detector. The 1.5-Å shell data are 97% complete.

Escherichia coli (38% sequence identity with *B. subtilis* NagB) as a search model (Protein Data Bank code 1dea) (6). The rotation and translation functions gave two solutions, resulting in two molecules in the asymmetric unit. After a rigid body refinement performed in AMoRe, we obtained a correlation coefficient of 48.5% and an *R*-factor of 45.4%. The electron density map calculated from the model was of sufficient quality to allow tracing of the C^α chain of the molecule using the REFMAC (26)/ARP-wARP (27) programs. The side chains were built with the program QUANTA (Accelrys, San Diego, CA) using Xautofit (28), and the model was refined using the CCP4 program REFMAC and QUANTA. The final model contains 3811 nonhydrogen protein atoms with 402 water molecules and two phosphate ions. The crystallographic R_{cryst} and R_{free} values are 15.5 and 21%, respectively (see Table I).

The structure of the F6P complex was solved using the native structure. A rigid body refinement was performed in REFMAC, maintaining an identical set of cross-validation reflections. The structure was subsequently refined using REFMAC, as above, with ligand target geometry defined using QUANTA. The final model has 3818 protein atoms, 855 waters, and a Fru-6-P molecule in each molecule of the asymmetric unit. The crystallographic *R*-factor and *R*-free values are 12.3 and 16.5% respectively.

Enzyme Assays—NagB activity was measured in a coupled assay relying on the increase in absorbance at 340 nM due to the conversion of NADP to NADPH. The assays were performed on a GBC Cintra10 spectrophotometer equipped with a thermoequilibrated cellblock. Unless otherwise indicated, the reactions were performed at 22 °C, pH 8, in acrylic cuvettes. The assay mixture contained 20 mM Tris-HCl, pH 8, 5 mM $MgCl_2$, 1 mM NADP, phosphoglucoisomerase (2 μg/ml), and glucose-6-phosphate dehydrogenase (3 μg/ml) in a total volume of 980 μl. Varying concentrations of substrate diluted in 20 mM Tris-HCl, pH 8 (from 0.0312 to 32 mM) were used in experiments for the studies of substrate inhibition kinetics. The reactions were initiated by the addition of 20 μl of NagB to give a final concentration of 0.3 μg/ml in the reaction mixture. Experimental data describing the dependence of NagB activity on substrate concentration were fitted to the substrate inhibition equation (Equation 1, where V_i is the initial rate of the reaction) by a nonlinear regression method using the program GraFit (29).

$$V_i = \frac{V_{max}[S]}{K_m + [S] + [S]^2/K_i} \quad \text{(Eq. 1)}$$

RESULTS AND DISCUSSION

*Overall Structure—Bsu*NagB is a monomer of 244 residues in solution, as evidenced by analytical ultracentrifugation, gel filtration, and dynamic light scattering results (data not shown). Crystals of NagB (see "Materials and Methods") contain two molecules in the asymmetric unit. The structure was solved by molecular replacement using a single monomer (from the hexamer) of the R-state *E. coli* NagB structure (Protein Data Bank code 1dea) (6). The structure of *Bsu*NagB is indeed similar to that of a monomer from *Eco*NagB, with a root mean square deviation of 0.97 Å over 239 equivalent C^α atoms (calculated with LSQMAN (10)). The secondary structure elements superimpose, almost identically, apart from the C terminus of *Eco*NagB, which is 18 residues longer than *Bsu*NagB in the form of an additional α-helix that is involved in allosteric regulation and hexamer formation (Fig. 3).

The *Bsu*NagB protein fold is a three-layer (α/β/α) structure, with a main β-sheet of seven parallel β-strands flanked on both sides with α- and 3_{10}-helices, Fig. 4*A*. The central body (residues 1–66, 95–153, and 192–239) of the enzyme is essentially a "Rossmann-like" fold and encloses the active site. Such an arrangement is similar to that of aldose-ketose hydrolases, such as ribose-5-phosphate isomerase (11), and their mode of phosphate binding also appears to be conserved. Residues 158–178 form a helix-loop motif that has been suggested to represent an active site "lid" of NagB (6) and are part of a domain that constitutes a mixed β-sheet of three β-strands surrounded by three α-helices (Fig. 4).

Substrate Binding and Catalysis—The architecture of the catalytic site is formed by the loops β_2-α_2, α_5-β_8, β_6-β_7, and β_3-β_4 and helix α_5 (Fig. 4*A*). After refinement, the "native" structure of *Bsu*NagB exhibited significant additional electron density near the catalytic center. This was modeled as a phosphate group plus two carbons and an oxygen atom, which probably represent the carbon atoms C-6 and C-5 and the oxygen O-5 of the extremity of a sugar phosphate compound. We were not able to build the rest of the molecule, and waters have been placed to match the remaining density. This ligand is probably a disordered sugar-phosphate that has been sequestered from *E. coli* during protein overexpression.

The structure of NagB with GlcN-6-P was subsequently studied through crystal soaking experiments, and the structure was solved at 1.4-Å resolution. The "unbiased" electron density unambiguously shows a molecule of fructose 6-phosphate

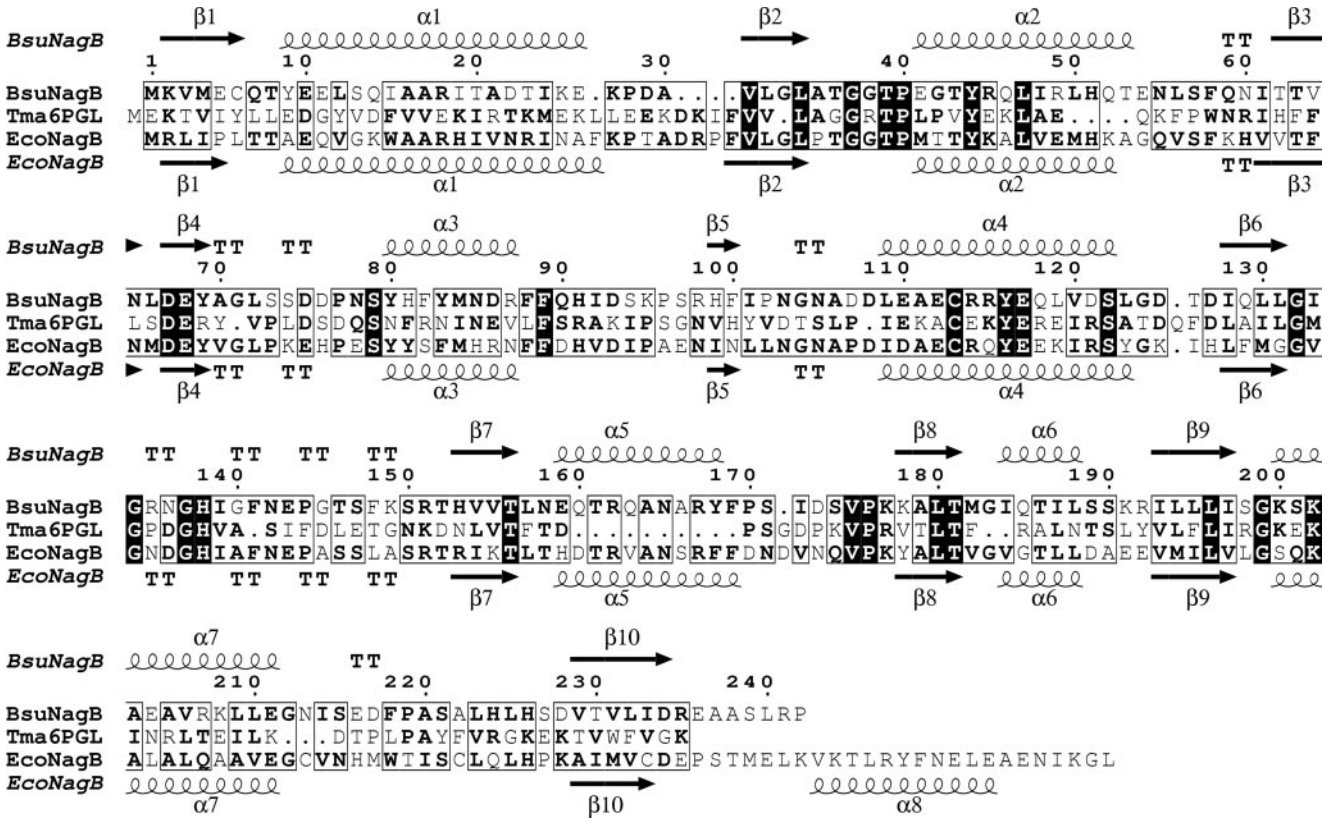

FIG. 3. **Structure-based sequence alignment.** The sequences of *B. subtilis* NagB, *T. maritima* 6-PGL, and *E. coli* NagB are aligned, with the secondary structures of the top and bottom sequences presented. Note the additional α-helix at the C terminus of *E. coli* NagB and the absence of an α-helix in 6-PGL between strands β_7 and β_8. More comprehensive sequence alignments are supplied as Supplementary Material. This figure was drawn using ESPript (31).

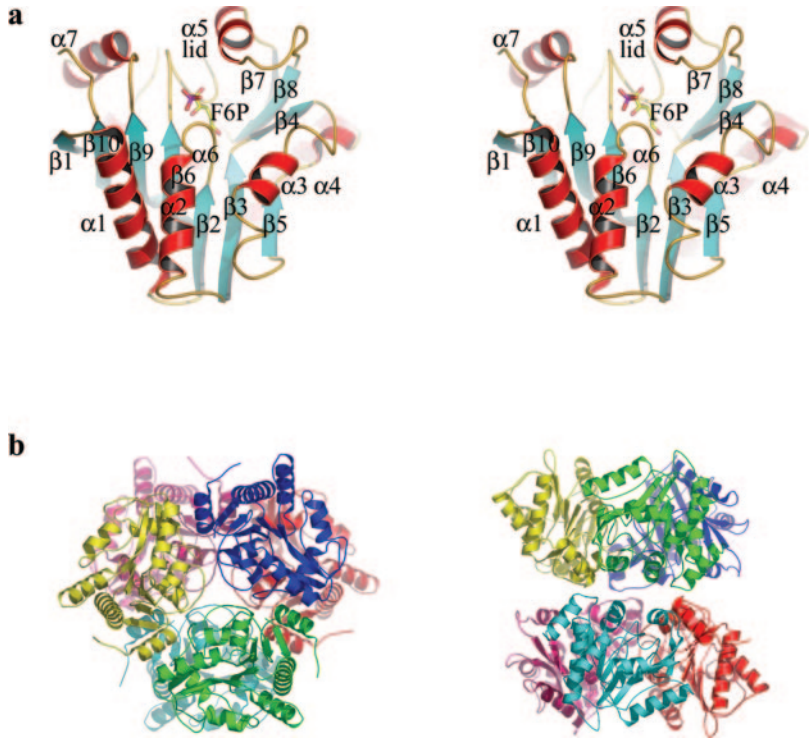

FIG. 4. **The three-dimensional structures of NagB from *B. subtilis* and *E. coli*.** *a*, protein schematic diagram of the three-dimensional structure of the fructose-6-phosphate complex of NagB showing the $\alpha/\beta/\alpha$ sandwich fold. The helices (*red*) are numbered from α_1 to α_7, and the β-strands (*cyan*) are numbered from β_1 to β_{10}. This figure is in divergent ("wall-eyed") stereo. *b*, hexamer of NagB *E. coli* shown in schematic diagram representation view along the 3-fold axis and at 90° along one of the three 2-fold axes. The structure is *rainbow-colored*. Both *pictures* were drawn using PyMOL (32).

(F6P), in an extended open chain form, in the catalytic site of each monomer of the asymmetric unit, Fig. 5. Since F6P is the product of the reaction, *Bsu*NagB has been able to perform the isomerization/deamination reaction both in the absence of added co-factors and "in-crystal." This is in marked contrast to *Eco*NagB, which requires both the allosteric activator GlcNAc-6-P and the subsequent conformational change from the T- to the R-state of the enzyme for activity (6, 12, 13). Superposition of the "native" *Bsu*NagB structure with the *Bsu*NagB·F6P complex reveals very similar structures with a root mean square

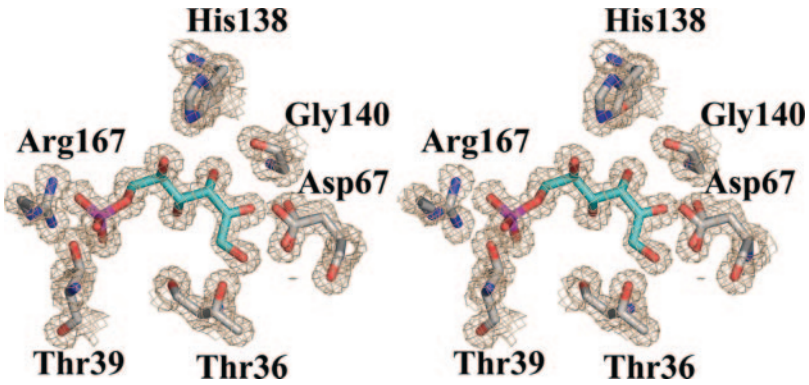

FIG. 5. **Electron density at the active center of *B. subtilis* NagB.** The reaction product F6P is represented surrounded by the interacting residues of the active center. This figure in divergent ("wall-eyed") stereo was drawn using PyMOL (32) and shows a REFMAC maximum likelihood/σ_A-weighted $2F_o - F_c$ electron density at ~0.53 electrons/Å^2. Asp^{67} is visible in two similar conformations in the electron density maps.

deviation of 0.32 Å for 241 equivalent C$^\alpha$ atoms. This is in contrast with the drastic structural changes that occur in *Eco*NagB upon binding of allosteric activator. The *Bsu*NagB structure is thus very close to that of the active R conformer of *Eco*NagB rather than the ligand-free T conformer, which undergoes large structural differences in the lid region (and also some quaternary rotation movement of subunits in the hexameric structure) upon ligand binding. The equivalent "lid" region in *Bsu*NagB does not appear to move upon binding of substrate, although the native structure does contain a partial occupancy of disordered residual ligand and thus may not be a true native structure. In an attempt to obtain a complex of intact substrate, *Bsu*NagB crystals were incubated with glucose 6-phosphate, whose lack of a 2-amino group should render it unreactive or, at least, a poor substrate. The observed electron density corresponded, however, not to glucose 6-phosphate, but to the product F6P, showing that in-crystal glucose 6-phosphate is indeed able to act as a substrate (data not shown).

A catalytic mechanism has been proposed for *Eco*NagB based upon the seminal kinetic observations by Rose and Midelfort (14) and the three-dimensional structure of a competitive inhibitor in the active site (6). The same general reaction scheme may be applied to *Bsu*NagB, since F6P interacts with the equivalent conserved residues as observed for the R-state *Eco*NagB (Figs. 6 and 7). Most authors believe that the reaction occurs on the open-chain form of the sugars. Midelfort and Rose (14) proposed that the α-anomer of GlcN-6-P is the favored ring form substrate, despite the reaction occurring on the open form of the sugar, since the ring-closed substrate in $^4C_1$ (chair) conformation is then consistent with a *cis*-enolamine intermediate, and kinetics indicated a preference for the α-anomeric form. Given that the conformational change from a ring-closed to extended ring-open species, as observed in-crystal, is enormous compared with the rotation around a single bond, such an interpretation seems unlikely. One can certainly model a ring form in the active center (although we and others (see Fig. 5 of Ref. 6)) find that a β-anomeric form docks with less steric clashes. His^{138} is perfectly poised to give general acid assistance to ring opening through protonation of O-5, as also observed on the recent ribose-5-phosphate isomerase structures (15); indeed, it also hydrogen-bonds to the open chain O-5 in the F6P structure (Fig. 6).

Whatever the mechanism or requirement for ring opening, proton abstraction from C-2 is most likely performed by Asp^{67} with the C–H σ-bond in plane with the *p*-orbital of the sugar carbonyl. Formation of the enamine intermediate would then be followed by attack of water at C-2. It is likely that this passes through the more electrophilic iminium ion with subsequent departure of the ammonia (or the ammonium ion) both driving the reaction and thus yielding fructose 6-phosphate (Fig. 7).

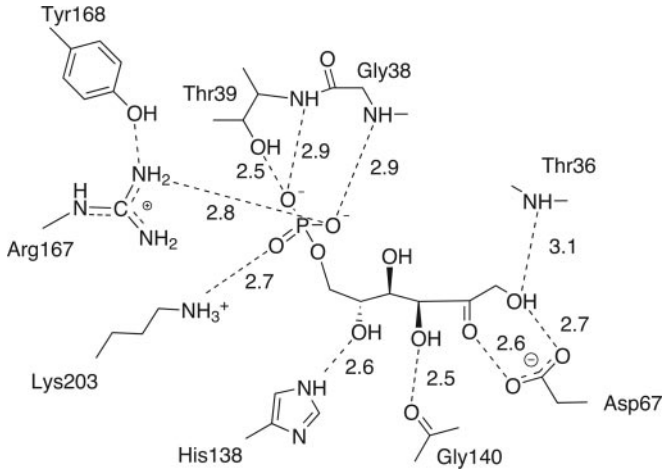

FIG. 6. **Schematic diagram of the active center interactions of NagB.** Distances of direct hydrogen bonds, in Å, are given.

Enzyme Activity—In the R-state of *Eco*NagB, two phosphates ions can be observed per monomer, in the catalytic site and at the subunit interface, the latter representing the allosteric site. To reach the ligand-free T-state, the enzyme needs to bind the allosteric activator GlcNAc-6-P. Our attempts to soak the crystals of *Bsu*NagB in GlcNAc-6-P failed to indicate any binding of this compound. The lack of multimers and the absence of an activator-binding site suggest that *Bsu*NagB is not regulated by an allosteric mechanism. Modified Michaelis-Menten kinetic analysis of NagB, including a model for substrate inhibition, generates $k_{cat} = 28 \pm 6\ s^{-1}$ and $K_{m(GlcN\text{-}6\text{-}P)} = 0.13 \pm 0.02$ mM. At high substrate concentrations, *Bsu*NagB displays a decrease in activity in what is classically interpreted as (excess of) substrate inhibition, with an apparent K_i value of 4 ± 0.6 mM (Fig. 8). The exact cause of substrate inhibition is not known; most likely, a second molecule of substrate obscures the binding site in some way, although we have no structural evidence to support this. A second molecule could bind adjacent to the (catalytically relevant) substrate-binding site perpendicular to the ligand and directly above the C-5 carbon of the sugar. The residues covering the wall of this entrance (Tyr^{80}, Asn^{165}, Gly^{37}, Ser^{79}, Tyr^{69}, and Tyr^{168}) are conserved in Gram-positive bacterial NagB proteins. Such a cavity also exists on the surface of *Eco*NagB, but its shape is probably too distorted to support a similar hypothesis. Perhaps more likely is that the inhibiting substrate simply binds with its phosphate group in the phosphate site, but with the sugar pointing outward instead of inward.

The NagB Superfamily—As described previously (see Fig. 2), the NagB "superfamily" contains three major clusters: allosteric NagBs, monomeric NagBs, and 6-PGLs (and all of their respective homologs). Genomic sequence analysis is hindered

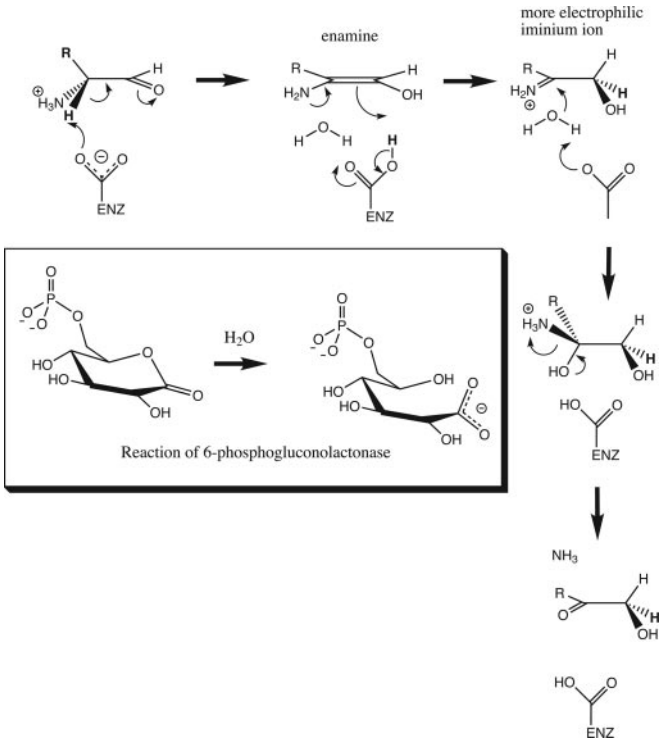

FIG. 7. **Putative reaction mechanism of NagB.** The reaction catalyzed by 6-PGL is shown in the *inset*, and a small section of the fructose 6-phosphate product of NagB (drawn as in Fig. 5) is also included for reference. Formation of the enamine intermediate is followed by attack of water at C-2. It is likely that this passes through the more electrophilic iminium ion with subsequent departure of the ammonia (or the ammonium ion), yielding fructose 6-phosphate given appropriate proton transfer (not shown).

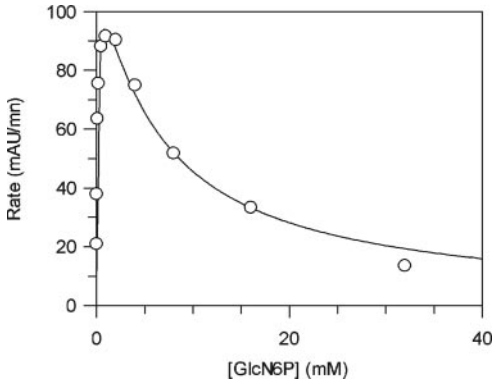

FIG. 8. **Modified Michaelis-Menten kinetics of *Bsu*NagB, which may be interpreted in terms of substrate inhibition at high [GlcN-6-P].**

by the presence of paralogs in some genomes such as *yieK* and *agaI* in *E. coli* and *gamA* in *B. subtilis*. Clues of the physiological function of each open reading frame may be gleaned from their encoding gene context on the chromosome; for instance, AgaI is a GalNAc epimerase as inferred from the phosphotransferase system components with which it is co-transcribed. A further complication is that related carbohydrates are transported and metabolized by similar metabolic enzymes through pathways that have common intermediates (16). Despite this imperspicuity, the Gram-positive NagB structure presented here means that we now have a structural representative from each major class within the superfamily, since the 6-PGL from *Thermotoga maritima* (*Tma*6PGL) has recently been solved as part of a "structural genomics" program (Protein Data Bank code 1pbt).

We can now assign differences, based upon the three-dimensional structures of the three representatives that allow functional dissection of the genomic data. The major difference between the NagB enzymes thus pertain to the oligomeric status: the allosteric hexamer from *E. coli* and human sources (3) and the monomeric form exemplified by the NagB from *B. subtilis*. The *E. coli* hexamers are arranged in a dimer of trimers with a 3-fold axis perpendicular to three 2-fold axes and crossing each other in the center of the particle (6) (Fig. 4*B*). The intersubunit contacts form the allosteric activator binding site and are not conserved in *Bsu*NagB. In particular, residues of β_7 and its preceding loop (*e.g.* Ser^{151}, Arg^{158}, and Lys^{160}), and the C-terminal α-helix extension (*e.g.* Tyr^{254}) are characteristic of this hexameric group.

The 6-PGL structure (root mean square deviation between *Tma*6PGL and *Bsu*NagB is 1.67 Å over 185 equivalent C^{α} atoms) (Fig. 3) also allows us to dissect out the conserved structural and catalytic features. Across the three groups, equivalent catalytic side chains are conserved (*i.e.* Lys^{202} at the beginning of α_7 (*Eco*NagB Lys^{209}, *Tma*6PGL Lys^{190}), His^{138} in the turn after β_6 (*Eco*NagB His^{143}, *Tma*6PGL His^{138}) and Asp^{67} at the start of β_4 (*Eco*NagB Asp^{72}, *Tma*6PGL Asp^{68}). These active site residues must be correctly positioned, and, consequently, a network of interacting residues appear to be somewhat conserved between the NagB superfamily sequences. *Bsu*NagB His^{138} is oriented by Glu^{143}, Thr^{161}, and Asn^{136}. In *Eco*NagB, Glu^{148} has the same role as *Bsu*NagB Glu^{143}, but the Asn^{136} is replaced by Asp^{141} (Asp^{127} in *Tma*6PGL). Interestingly, a D141N mutant of *Eco*NagB converts the allosteric effector GlcNAc-6P into an inhibitor (17). 6-PGL sequences appear to lack a structural equivalent of *Bsu*NagB Glu^{143} and so this position provides a signature to distinguish between the different NagB family members.

The catalytic site of *Tma*6PGL is similar to that of NagB, but no enzymatic data are available for this enzyme; all of the enzymatic assays have been performed with eukaryotic 6-PGL enzymes (18–21). Differences in activity are expected between 6-PGL and NagB proteins, since recombinant human 6-PGL is indeed unable to deaminate/isomerize GlcN-6-P (22). The majority of catalytic residues are conserved in 6-PGL sequences, although the equivalent of Arg^{167} that coordinates the (open-chain) 6-phosphate group in *Bsu*NagB is absent, due to the deletion of the "lid" region between strands β_7 and β_8. In the *Tma*6PGL structure, Arg^{167} is functionally substituted by Arg^{43} from a quite remote piece of three-dimensional structure; however, this residue is not conserved across the 6-PGL family. A signature, however, does characterize the 6-PGL family and is thus not present in *Bsu*NagB. By definition, 6-PGL acts on the substrate 6-phosphogluconolactone. The enzyme must therefore catalyze a direct attack of water on the C-1 (as opposed to C-2) atom of the substrate, yet 6-phosphogluconolactone exists only as a ring-form sugar (Fig. 7, *inset*), and an overlay (not shown) of the structure of 6-PGL with *Bsu*NagB reveals that a 6-PGL "signature" motif of Arg^{70} (at the end of β_4) and Arg^{167} (at the beginning of β_8), which, in the deposited 6-PGL structure, interact with a second phosphate site (observed as sulfate in-crystal), is not present in NagB enzymes. This sulfate site is exactly where one would predict the 6-phospho group of a ring-form sugar would lie, as one expects for an enzyme necessarily active on the ring-form sugar. On the basis of sequence work alone, it had been proposed that the equivalent of Arg^{70} (Arg^{81}) of the human enzyme (22) might instead stabilize an intermediate anion species. From comparison of the three-dimensional structures, we would conclude that it more likely interacts with the ring-form substrate 6-phospho group directly. The *Bsu*NagB structure, together with the

three-dimensional structures of the hexameric NagB and the 6-PGL, thus "completes" the structural tree and allows one to mine the genomic data. Sequence fingerprints, viewed in light of structure, reveal the key functional elements that allow for both specificity toward glucosamine 6-phosphate or phosphogluconolactone and the basis for allosteric multimeric assembly or, as illustrated here, unregulated monomeric state.

Acknowledgments—We thank David Vocaldo (Simon Fraser University) for advice on enzyme kinetic analysis and the ESRF (Grenoble) for excellent beamline facilities.

REFERENCES

1. Warren, L. (1972) *Biosynthesis and Metabolism of Amino Sugars and Amino Sugar-containing Heterosaccharide: Glycoproteins* (Gottschalk, A., ed) Elsevier, Amsterdam
2. Alice, A. F., Perez-Martinez, G., and Sanchez-Rivas, C. (2003) *Microbiology* **149,** 1687–1698
3. Arreola, R., Valderrama, B., Morante, M. L., and Horjales, E. (2003) *FEBS Lett.* **551,** 63–70
4. Ferreira, F. M., Mendoza-Hernandez, G., Calcagno, M. L., Minauro, F., Delboni, L. F., and Oliva, G. (2000) *Acta. Crystallogr. Sect. D* **56,** 670–672
5. Souza, J. M., Plumbridge, J. A., and Calcagno, M. L. (1997) *Arch. Biochem. Biophys.* **340,** 338–346
6. Oliva, G., Fontes, M. R., Garratt, R. C., Altamirano, M. M., Calcagno, M. L., and Horjales, E. (1995) *Structure* **3,** 1323–1332
7. Vincent, F., Yates, D., Garman, E., Davies, G. J., and Brannigan, J. A. (2004) *J. Biol. Chem.* **279,** 2809–2816
8. Berg, J. M., Tymoczko, J. L., and Stryer, L. (2002) *Biochemistry*, W. H. Freeman and Co., New York
9. Clarke, J. L., Scopes, D. A., Sodeinde, O., and Mason, P. J. (2001) *Eur. J. Biochem.* **268,** 2013–2019
10. Kleywegt, G. J., and Jones, T. A. (1994) *ESF/CCP4 Newsletter* **31,** 9–14
11. Hamada, K., Ago, H., Sugahara, M., Nodake, Y., Kuramitsu, S., and Miyano, M. (2003) *J. Biol. Chem.* **278,** 49183–49190
12. Horjales, E., Altamirano, M. M., Calcagno, M. L., Garratt, R. C., and Oliva, G. (1999) *Struct. Fold Des.* **7,** 527–537
13. Rudino-Pinera, E., Morales-Arrieta, S., Rojas-Trejo, S. P., and Horjales, E. (2002) *Acta Crystallogr. Sect. D* **58,** 10–20
14. Midelfort, C. F., and Rose, I. A. (1977) *Biochemistry* **16,** 1590–1596
15. Roos, A. K., Burgos, E., Ericsson, D. J., Salmon, L., and Mowbray, S. L. (2005) *J. Biol. Chem.* **280,** 6416–6422
16. Shakeri-Garakani, A., Brinkkotter, A., Schmid, K., Turgut, S., and Lengeler, J. W. (2004) *Mol. Genet. Genomics* **271,** 717–728
17. Cisneros, D. A., Montero-Moran, G. M., Lara-Gonzalez, S., and Calcagno, M. L. (2004) *Arch. Biochem. Biophys.* **421,** 77–84
18. Rakitzis, E. T., and Papandreou, P. (1995) *Biochem. Mol. Biol. Int.* **37,** 747–755
19. Kawada, M., Kagawa, Y., Takiguchi, H., and Shimazono, N. (1962) *Biochim. Biophys. Acta* **57,** 404–407
20. Bauer, H. P., Srihari, T., Jochims, J. C., and Hofer, H. W. (1983) *Eur. J. Biochem.* **133,** 163–168
21. Hofer, H. W., and Bauer, H. P. (1987) *Cell Biochem. Funct.* **5,** 97–99
22. Collard, F., Collet, J. F., Gerin, I., Veiga-da-Cunha, M., and Van Schaftingen, E. (1999) *FEBS Lett.* **459,** 223–226
23. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
24. Collaborative Computational Project 4 (1994) *Acta Crystallogr. Sect. D* **50,** 760–763
25. Navaza, J., and Saludijan, P. (1997) *Methods Enzymol.* **276,** 581–594
26. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. Sect. D* **53,** 240–255
27. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6,** 458–463
28. Oldfield, T. J. (2001) *Acta Crystallogr. Sect. D* **57,** 82–94
29. Leatherbarrow, R. J. (2001) *GraFit*, 5th Ed., Erithacus Software, Ltd., London
30. Komatsuzawa, H., Fujiwara, T., Nishi, H., Yamada, S., Ohara, M., McCallum, N., Berger-Bachi, B., and Sugai, M. (2004) *Mol. Microbiol.* **53,** 1221–1231
31. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) *Bioinformatics* **15,** 305–308
32. DeLano, W. L. (2002) *PyMOL*, DeLano Scientific, San Carlos, CA