The Three-dimensional Structure of the N-Acetylglucosamine-6-phosphate Deacetylase, NagA, from Bacillus subtilis

Florence Vincent, David Yates, Elspeth Garman, Gideon J Davies, James A Brannigan

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 279, No. 4, Issue of January 23, pp. 2809–2816, 2004

Printed in U.S.A.

The Three-dimensional Structure of the *N*-Acetylglucosamine-6-phosphate Deacetylase, NagA, from *Bacillus subtilis*

A MEMBER OF THE UREASE SUPERFAMILY*

Received for publication, September 12, 2003, and in revised form, October 10, 2003
Published, JBC Papers in Press, October 13, 2003, DOI 10.1074/jbc.M310165200

Florence Vincent‡, David Yates§, Elspeth Garman§, Gideon J. Davies‡¶, and James A. Brannigan‡

From the ‡Structural Biology Laboratory, Department of Chemistry, The University of York, Heslington, York, YO10 5YW, United Kingdom and §Laboratory of Molecular Biophysics, Department of Biochemistry, The University of Oxford, South Parks Road, Oxford, OX1 3QU, United Kingdom

The enzyme *N*-acetylglucosamine-6-phosphate deacetylase, NagA, catalyzes the hydrolysis of the *N*-acetyl group of GlcNAc-6-P to yield glucosamine 6-phosphate and acetate, the first committed step in the biosynthetic pathway to amino-sugar-nucleotides. It is classified into carbohydrate esterase family CE-9 (see afmb.cnrs-mrs.fr/CAZY/). Here we report the cloning, expression, and three-dimensional structure (Protein Data Bank code 1un7) determination by x-ray crystallography of the *Bacillus subtilis* NagA at a resolution of 2.0 Å. The structure presents two domains, a $(\beta/\alpha)_8$ barrel enclosing the active center and a small β barrel domain. The structure is dimeric, and the substrate phosphate coordination at the active center is provided by an Arg/His pair contributed from the second molecule of the dimer. Both the overall structure and the active center bear a striking similarity to the urease superfamily with two metals involved in substrate binding and catalysis. PIXE (Proton-Induced x-ray Emission) data show that iron is the predominant metal in the purified protein. We propose a catalytic mechanism involving proton donation to the leaving group by aspartate, nucleophilic attack by an Fe-bridged hydroxide, and stabilization of the carbonyl oxygen by one of the two Fe atoms of the pair. We believe that this is the first sugar deacetylase to utilize this fold and catalytic mechanism.

The first committed step in the biosynthetic pathway to the amino-sugar precursors required for cell wall peptidoglycan and teichioic acid biosynthesis in *Bacillus subtilis* is the deacetylation of GlcNAc[1]-6-P to yield acetate and GlcN-6-P (Fig. 1). This reaction is catalyzed by the enzyme *N*-acetylglucosamine-6-phosphate deacetylase (EC 3.5.1.25) NagA (Fig. 1*b*). In *Escherichia coli*, where the enzyme is known to be a homotetramer (1–3), the deacetylation of *N*-acetylglucosamine is additionally important in lipopolysaccharide synthesis and cell wall recycling (4). Neither *E. coli* nor *B. subtilis* can convert GlcNAc-6-P to GlcNAc-1-P directly, so the conversion to GlcN-6-P is a prerequisite for conversion to GlcN-1-P by phosphoglucosamine mutase, which can then be acylated and uridylated to UDP-*N*-acetyl-D-glucosamine. Given its position at the crossroads of these bacteria-specific processes, NagA has warranted attention as a potential drug target. Indeed, sugar deacetylation is a validated therapeutic target in other contexts (5, 6).

The enzymes involved in the deesterification/de-*N*-acetylation of carbohydrates have been classified into 13 families[2] based upon amino acid sequence similarities. Given the vast diversity of esterases and their frequent lack of specificity, this classification is arguably not as powerful as the related classifications of glycoside hydrolases and glycosyltransferases. However, the enzymes involved in the de-*N*-acetylation of GlcNAc-6-P do all lie in a single sequence-related family termed CE-9. As of September 11, 2003, there are 134 members, the vast majority from bacterial sources. Sequence searches using PSI-BLAST (7) pointed to a three-dimensional structure involving a urease-like bimetallic center different to the known carbohydrate esterases. The majority of carbohydrate esterases/deacetylases whose structures are reported display a classical $\beta/\alpha/\beta$ "serine protease" fold as revealed by three-dimensional structures of the enzymes from families CE-1 (8), CE-5 (9), CE-7 (10), the plethora of enzymes from family CE-10, and the mycolyltransferase "antigen 85C" (5). A small deviation from this canonical fold is displayed in the CE-12 rhamnogalacturonan acetylesterase (11). Thus far, the only two "outliers" to this trend are the CE-8 pectin methylesterase, which instead presents a twin-aspartate catalytic center grafted upon a right-handed parallel β-helix (12), and LpxC zinc-dependent UDP-3-*O*-acetyl-*N*-acetylglucosamine deacetylases from family CE-11, which present an unusual twin-zinc site on a novel α/β framework (6, 13).

Here we present the expression and purification of the *B. subtilis* NagA and an analysis of its metal-ion content through proton-induced x-ray emission data. The three-dimensional structure is reported at 2.0 Å in complex with the reaction product GlcN-6-P (and partial occupancy of unhydrolyzed substrate GlcNAc-6-P). The structure reveals NagA to be a member of the "urease superfamily" with a catalytic center involving a binuclear Fe center. The complexes permit a proposal for the catalytic mechanism for NagA equally applicable to other urease superfamily members.

* This work was funded by the Wellcome Trust. 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 1un7) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

¶ 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: GlcNAc, *N*-acetyl glucosamine; GlcN, glucosamine; CE, carbohydrate esterase (deacetylase); PEG, polyethylene glycol; PIXE, Proton-Induced x-ray Emission; ORF, open reading frame; r.m.s., root mean square.

[2] See afmb.cnrs-mrs.fr/CAZY/.

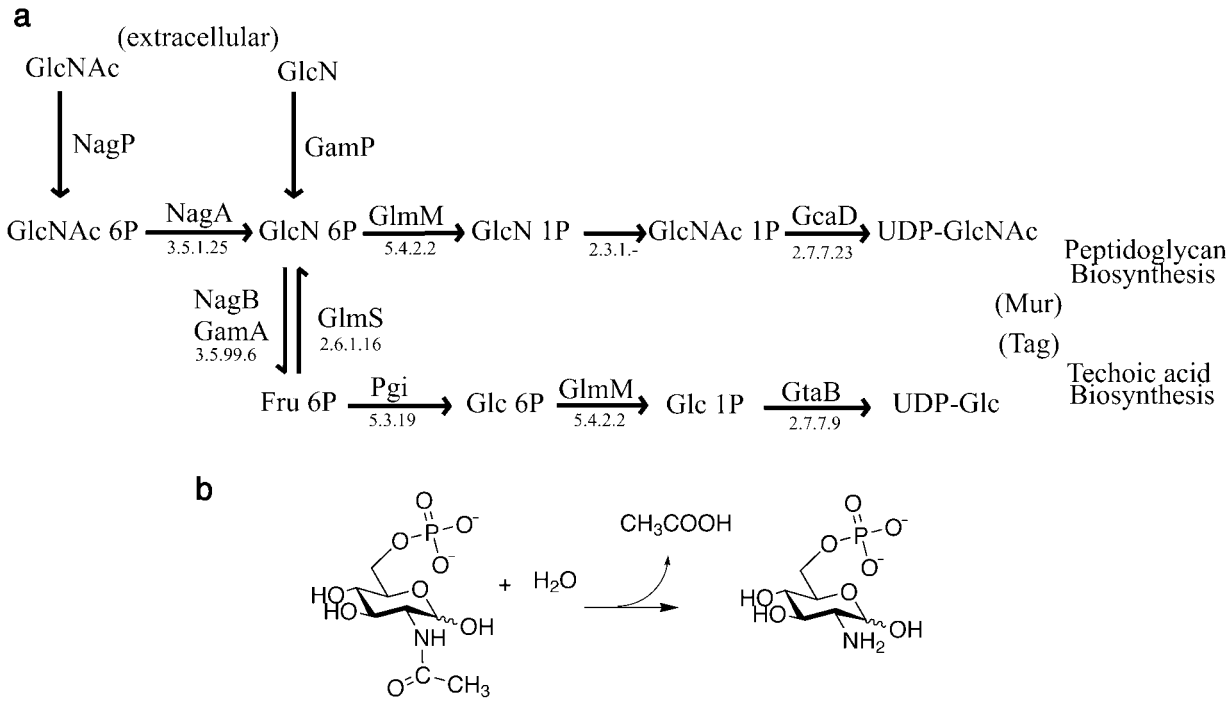

FIG. 1. **The deacetylation of *N*-acetylglucosamine-6-phosphate.** *a*, pathways of amino-sugar metabolism in *B. subtilis*. GlcNAc and GlcN are imported by specific phosphotransferase system components (EC 2.7.1.69) and are converted to precursors of peptidoglycan and anionic polymers essential for cell wall biosynthesis. Proteins catalyzing each step are given with corresponding EC numbers. There seems to be at least three *B. subtilis* gene products able to acetylate GlcN-1-P (PksF, AcuA, and BkdB). Abbreviations: *Fru*, fructose; *Glc*, glucose. Gene products: *NagB*, GlcN-6-P isomerase; *GlmS*, L-glutamate-D-fructose-6-P amidotransferase; *GlmM*, phosphoglucomutase; *GcaD*, UDP-GlcNAc pyrophosphorylase; *Pgi*, Glc-6-P isomerase; *GtaB*, UTP-Glc-1-P uridyltransferase. *b*, the chemical reaction catalyzed by NagA, the *N*-acetylglucosamine-6-phosphate de-*N*-acetylase.

EXPERIMENTAL PROCEDURES

Cloning and Protein Production—The *nagA* gene was cloned into the Gateway Entry vector pDONR201 (Invitrogen) by incorporating *attB* sequences into the PCR oligonucleotides used to amplify the coding region from chromosomal DNA of *B. subtilis* strain IG20 (168 Trp^-). The cloned fragment was transferred into the T7 promoter-based expression vector pDEST14 by recombination and expressed in *E. coli* BL21(DE3), a strain with an inducible T7 RNA polymerase gene (Novagen).

For protein production, cell cultures were grown in LB medium containing 100 μg/ml ampicillin to an optical density of 0.7 at 600 nm before induction of protein expression by addition of isopropyl-1-thio-β-D-galactopyranoside to a final concentration of 1 mM. After 3 h of incubation, the cells were harvested by centrifugation and disrupted by sonication. The lysate was clarified by centrifugation, applied to a Q-Sepharose column, and resolved using a linear gradient of increasing sodium chloride. Fractions containing NagA were concentrated and further purified by gel filtration (Superdex200 HR10/30, Amersham Biosciences) in a 50 mM Tris-HCl (pH 8), 200 mM NaCl buffer with the apparent molecular mass estimated by comparison to standard protein markers. Pure fractions of NagA were pooled, washed into 50 mM Tris-HCl (pH 8) and concentrated to 50 mg/ml using a 10-kDa centrifugation membrane (Vivaspin).

Crystallization, Data Collection, and Processing—In the absence of substrate, NagA crystals could be grown using 0.2 M lithium acetate dihydrate, 20% polyethylene glycol (PEG) 3350 (pH 7.8) as a precipitant but they diffracted poorly. NagA (0.23 mM) was incubated for 20 min at 4 °C in a 1 mM GlcNAc-6-P solution. Crystals were grown by vapor-phase diffusion using the hanging drop method with an equal volume (1 μl) of protein/substrate and reservoir solution composed of 8% (w/v) PEG 20,000, 8% (v/v) PEG 550 monomethyl ether, 0.1 M HEPES (pH 7.5), and 0.3 M sodium acetate. A single NagA/substrate co-crystal was transferred to a solution of mother liquor with 25% PEG 550 monomethyl ether as a cryoprotectant and flash-cooled to 120 K in a rayon loop. Diffraction quality was assessed using a home source. A single wavelength experiment was conducted on beamline ID14-1 at the European Synchrotron Radiation Facility at Grenoble at a temperature of 100 K using an ADSC CCD detector. Data were integrated, scaled, and reduced using DENZO and SCALEPACK (14). The crystals are orthorhombic with unit cell dimensions of a = 51.68, b = 107.73, and c = 188.25 Å, corresponding to a solvent content of 60% with two molecules (hereafter designated Amol and Bmol) in the asymmetric unit. All further crystallographic computations were carried out using the CCP4 suite of programs (15).

TABLE I
Crystal, data, and refinement statistics

Crystal parameters	
Space group	$P2_12_12$
Cell dimensions (Å)	
a	51.68
b	107.73
c	188.25
No. of molecules/AU	2
Data quality	
Wavelength (Å)	0.9791
Resolution of data (Å)	22–2.00
(Outer shell)	2.07–2.0
Unique reflections	69,518
R_{merge} (outer shell)[a]	0.109 (0.460)
Mean $I/\sigma I$ (outer shell)	7.5 (2.4)
Completeness (outer shell) (%)	97.4 (94.3)
Multiplicity (outer shell)	2.7 (2.4)
Refinement	
Protein atoms	5902
Solvent waters	626
Ions	2Fe
R_{cryst}	0.20
R_{free}	0.25
R.m.s. deviation 1–2 bonds (Å)	0.016
R.m.s. deviation 1–3 angles (°)	1.636
Mean Protein B (Å^2) Amol/Bmol	29/31
Mean GlcN B (Å^2) Amol/Bmol	26/28
Mean Solvent B (Å^2)	38

[a] $R_{merge} = (\Sigma_{hkl}\Sigma_i I_{hkl} - (I_{hkl})/\Sigma_{hkl}\ \Sigma_i[I_{hkl}])$.

Phasing, Model Building, and Refinement—The structure was solved by molecular replacement using AMoRe (16) using the structure of the protein from the *Thermotoga maritima* ORF TM0184 (Protein Data Bank code 1O12 derived from a Structural Genomics program, www.jcsg.org/), which has 33% identity with *B. subtilis* NagA, as a search model. The data exhibit pseudosymmetry characterized by a peak at

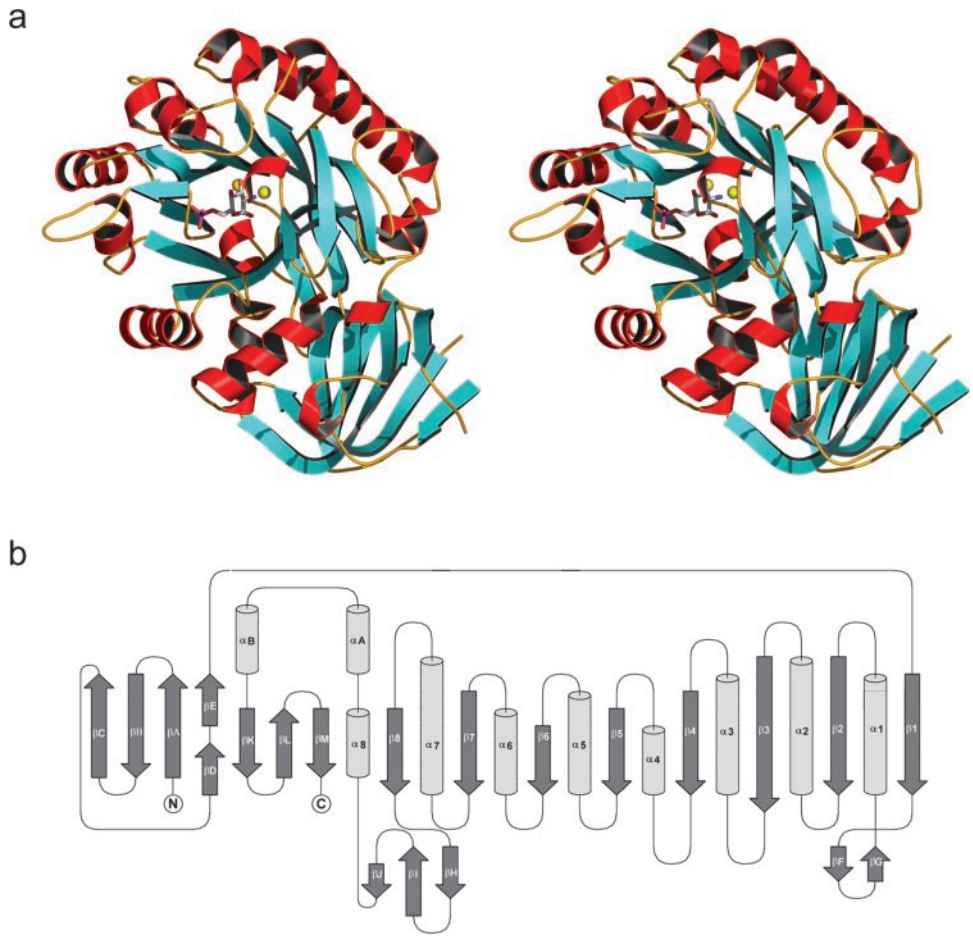

FIG. 2. **The three-dimensional structure of *B. subtilis* NagA.** *a*, protein "cartoon" of the three-dimensional structure of NagA showing the two-domain fold. The product, GlcN-6-P, is drawn in *liquorice*, and the twin Fe sites are shown as *yellow spheres*. This figure in divergent ("wall-eyed") stereo was drawn using BOBSCRIPT (49). *b*, topology diagram for NagA based upon the program DSSP (23).

$x = 0.5$, $y = 0.5$, and $z = 0.5$ in the native Patterson map with a peak height of 80% of the origin peak (when calculated with data between 10 and 2.0Å). Therefore, the data can nearly be described as belonging to space group I222, making it difficult to assign the exact primitive space group because the systematic absences, which could be attributed to screw axes, are also caused by the pseudosymmetry. The space group was determined by running AMoRe (including rigid-body optimization of the best solutions) in every subgroup of I222. The clearest solution was obtained in space group $P2_12_12$ with a correlation coefficient of 55.4% and an *R*-factor of 51.7%, some 1.5% "better" on each statistic in this space group than in any other.

The electron density map calculated from the model was difficult to interpret. Several cycles of the REFMAC/ARP-WARP program in warpNtrace mode (17, 18) resulted in the automatic building of the main chain of most of the α-helices and β-sheets, but attempts at automatic assignment of the sequence failed. Manual rebuilding was therefore performed with the X-Autofit module in the program Quanta (Accelrys Inc., San Diego, CA). The model was refined using REFMAC (17) interspersed with manual rebuilding using X-Autofit in Quanta. The final model contains 5902 non-hydrogen protein atoms, 626 water molecules, 4 metal-ions, 2 polyethylene glycol molecules, and 2 GlcN-6-P molecules. The crystallographic *R*-factor and *R*-free values (19) are 0.20 and 0.25, respectively (Table I). The stereochemistry of the model was assessed with the program PROCHECK (20) prior to deposition.

MicroPIXE Determination of Bound Metals—MicroPIXE (Proton-Induced x-ray Emission) measurements were carried out on the National Ion Beam Centre at the University of Surrey on a beamline arranged as described previously (21). A 2–3-MeV proton beam of 1-μm diameter was used to induce characteristic x-ray emission from crystals dried onto a 2-μm thick mylar film pre-tensioned over a 1-cm hole in an aluminum target holder and held in vacuum. X-rays were detected in a solid-state lithium-drifted silicon detector with high energy resolution. The proton beam was scanned spatially in X and Y and by placing software windows round the x-rays in the spectrum associated with particular elements, and by sorting these events into an X-Y grid, maps could be obtained of all elements heavier than neon that were present in the sample. The Rutherford backscattered proton spectrum measured in the chamber using a silicon surface barrier detector was used to get an accurate thickness of the sample to correct for self-absorption of the x-rays. Quantitative information was obtained by collecting spectra at 3 or 4 points on the crystal and also on the backing foil. These spectra are analyzed using GUPIX (22) to extract the areal density of each element of interest in the sample. The number of atoms of each element per protein molecule can then be computed from the ratio of that element to the sulfur measurement and from knowledge of the number of Cys and Met residues in the protein. This internal sulfur calibration allows much more accurate quantitation (normally approximately ± 10%) than would be possible if absolute measurements were necessary.

RESULTS

Overall Structure—The *B. subtilis* NagA exists as a dimer in solution as determined by gel filtration and dynamic light scattering data (data not shown). The crystals of NagA contain two molecules in the asymmetric unit. The two molecules involved in what we believe is the active dimer (see below) are related by the crystallographic 2-fold symmetry around the *c* axis. The polypeptide chain is visible from residues 3 to 394 in each molecule. The monomer folds into two structural domains (Fig. 2*a*), an $(\alpha/\beta)_8$ barrel (residues 59–332) enclosing the catalytic site of the enzyme and a small β-strand barrel made up from secondary structure elements contributed by the N and C termini (residues 3–58 and 368–394). The secondary structure of NagA as defined by the program DSSP (23) is shown in Fig. 2*b*. The $(\alpha/\beta)_8$ barrel is a somewhat distorted "TIM-

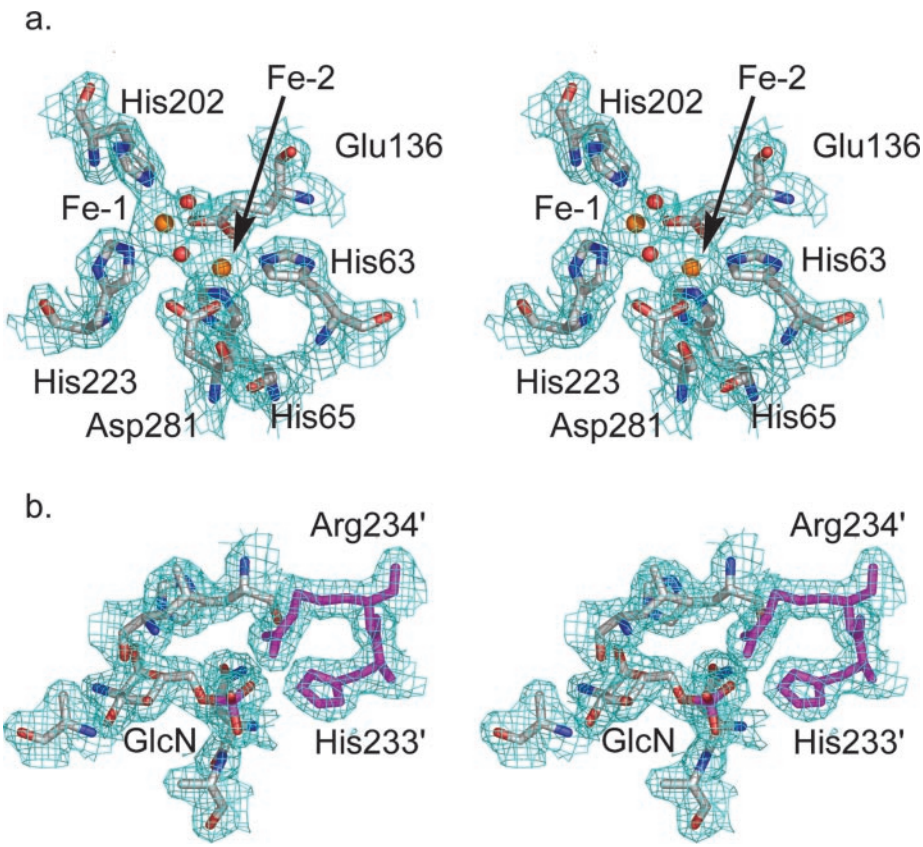

FIG. 3. **Electron density at the active center of *B. subtilis* NagA.** *a*, the two Fe sites with their coordinating residues. *b*, the reaction product GlcN-6-P. Arg-324 and His-233 shown in *purple* are derived from the second molecule of the dimer. This figure in divergent ("wall-eyed") stereo was drawn using PyMOL (pymol.sourceforge.net/) and shows maximum likelihood/σ_A (50) weighted $2F_{obs} - F_{calc}$ electron density at 0.25 electrons/Å^2.

barrel" fold, the connections between the β-strands and the helices are equivalent apart from an excursion of a small three-stranded β-sheet between β_8 and α_8 that closes one end of the $(\alpha/\beta)_8$ barrel and contribute residues involved in ligand interaction (see below). The second domain is an eight-stranded sheet that wraps into an incomplete β barrel. Five of the β-strands come from the N-terminal region of the enzyme (βA to E) and form the main part of the β barrel with two curved β-strands (βA and B). Finally, three other β-strands lie "on top" of the others and belong to the C-terminal region of the enzyme (βK to M). A DALI search (24) reveals a structural relationship between the catalytic domain and a large family of metal-dependent hydrolases with a similar barrel as described by SCOP (25), which includes urease, phosphotriesterase, and hydantoinase (26, 27). A subset of this family, now including NagA, is characterized by the presence of an all-β-strand domain made up of composite elements from the N and C termini (25).

The interface-accessible surface area is 1054 Å^2, which represents 7% of the total (28). The two molecules are related by 180°-rotational symmetry, and a loop between β_6 and α_6 forms an arm, which contacts its neighbor subunit across the dimer interface. There are two salt bridges between Asp-246 and Lys-265 and a number of hydrogen bond contacts between α_6 and α_7 to their symmetric equivalents with water excluded from the interface.

A comparison with the original search model derived from the *T. maritima* ORF TM0184, which crystallized in space group $P3_121$, shows a very similar fold with an r.m.s. deviation of 1.96 Å for 354 C_α atoms of the monomer (calculated with the program QUANTA (29)). The relative orientation of the two molecules of the *T. maritima* dimer is slightly different and reflected in an r.m.s. deviation of 6.18 Å for 706 C_α atoms of the dimer. The highest deviations occur in the β barrel domain (~6.36 Å for βC, βD, and βE) and in the loops connecting α_2 to β_3 and α_4 to β_5 (~5 Å). Only one iron ion is observed in the structure of ORF TM0184, and it superposes with Fe-2 of NagA and interacts with the same residue types (His-176, His-197, and Glu-115). The implications of this finding are described below.

The Catalytic Site and Substrate Binding—Electron density, corresponding to the reaction product GlcN-6-P, was clearly visible in the cavity of the (α/β) barrel (Fig. 3*b*) (30). In one molecule, there is residual difference density suggesting a partial (<0.3) occupancy of unhydrolyzed substrate, GlcNAc-6-P, and although partial occupancy of substrate sheds light on catalysis (see "Discussion"), it was not refined due to the constraint of limited resolution. GlcN-6-P binds in the cavity through side chain interactions with Asn-226 to the phosphate group, His-258 to the 1′-sugar hydroxyl, and main chain hydrogen bonds (Fig. 4). A putative catalytic acid (discussed below), Asp-281, is positioned close to the scissile bond. Strikingly, the phosphate also makes contact with a pair of side chains, His-233 and Arg-234, which form a pincer-like extension from the β_6-α_6 loop contributed from the partner subunit (Fig. 5). In the *T. maritima* structure, the equivalent residues (His-207 and Arg-208) point toward the solvent, suggesting that they are solicited only in the presence of the ligand in the cavity. The recruitment of residues from an adjacent partner subunit in a homo-oligomer toward substrate binding is unusual but not unique. The 180°-rotational symmetry between two monomers of granzymeA is used in the functional dimer to form an extended substrate-binding cleft (31, 32). Examples of an Arg residue cooperating (from adjacent subunits in a hexamer) to ligate a substrate phosphate ion include uridine phosphorylase (33) and purine nucleoside phosphorylase (34).

Additional electron density at the active center was assumed to correspond to bound metal-ions. The microPIXE technique applied to elemental analysis of proteins (35) was used to determine the metal content. To estimate metal content, a

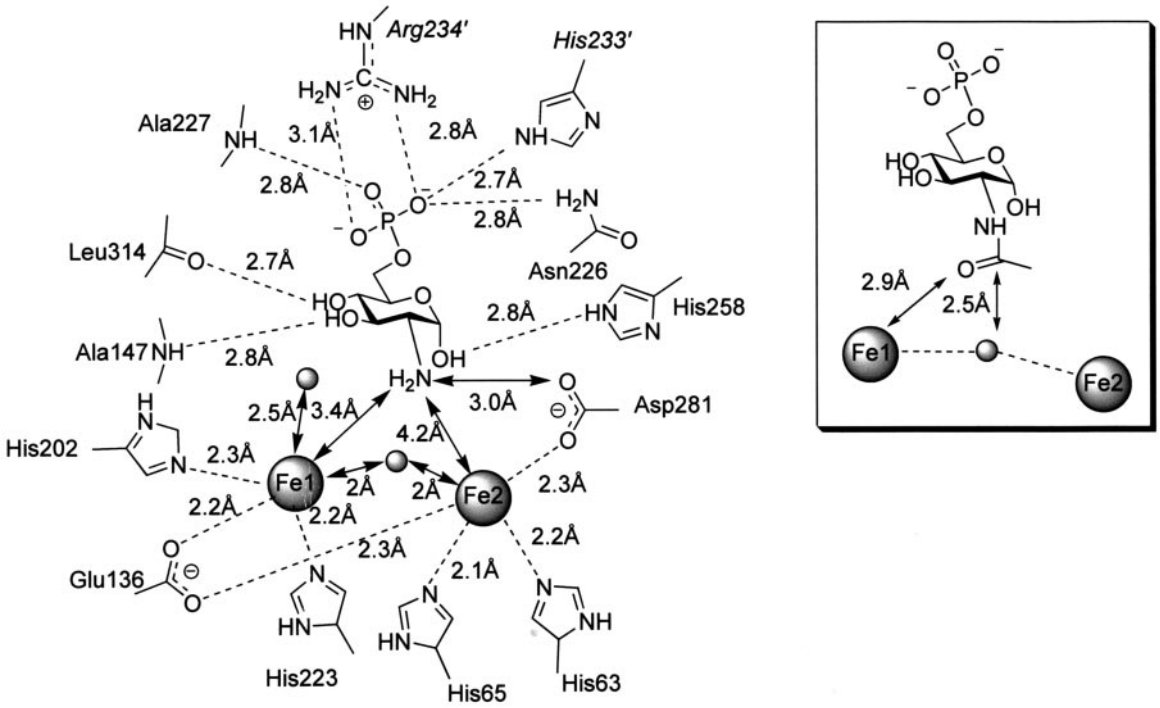

FIG. 4. **Schematic diagram of the active center interactions of NagA.** Water molecules are shown as *small shaded spheres* and the two Fe atoms as *labeled large spheres*. The *inset* shows the interactions of the partially occupied substrate molecule. This figure was drawn with CHEMDRAW (CambridgeSoft Corporation, Cambridge, MA).

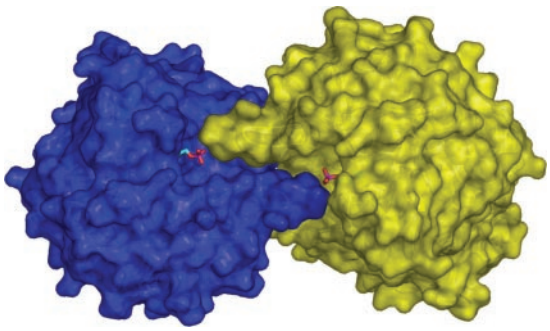

FIG. 5. **Surface of the NagA dimer.** The 6-phospho group of the ligand (shown in *liquorice*) can just be seen protruding from the surface and interacting with the second molecule that forms the dimer. This figure was drawn using PyMOL.

native crystal grown from lithium acetate was used, which revealed the presence of phosphorus, sulfur, chlorine, potassium, calcium, manganese, iron, copper, and zinc. Because this sample had no additional sulfur in the buffer, the number of atoms for these elements per protein molecule (*i.e.* per 12 sulfur atoms) could be calculated and averaged from individual point spectra: iron 1.3 (±0.2); zinc 0.21 (±0.03); manganese 0.17 (±0.02); potassium 0.11 (±0.02); and copper 0.05 (±0.02). Because iron was the predominant metal atom found in the purified protein (grown from *E. coli* in LB medium), iron was modeled in the electron density. The two Fe ions are situated in the bottom of the active site. One is more buried (Fe-2) and makes four interactions with protein side chains (His-63 and His-65 on β_1, Glu 136 from the end of β_3 and the active site Asp-281), and the other (Fe-1) is held in position by His-223, His-202, and Glu-136. The two metal atoms are 3.5 Å apart and refine with *B*-factors of 26 Å^2 for Fe-2 and 21 Å^2 for Fe-1 in each molecule of the asymmetric unit.

A comparison of NagA protein sequences from Gram-negative, Gram-positive, and archeal bacteria highlights other residues important for NagA structure and function (Fig. 6). There are four residues in the NagA structure whose backbone dihedral angles fall outside of the normal range as defined by a Ramachandran plot (36). A *cis*-proline (Pro-138) is conserved and occurs in a Gly-Pro motif immediately after Glu-136 at the end of β_3, which bridges the two Fe ions. The metal-binding His-223 is held at an unusual angle (ψ = 26.1, φ = 72.2) by the side chain of Glu-251, and Tyr-225 (ψ = 86.7, φ = −27.5) also displays a high energy conformation, which allows the substrate-binding Asn-226 side chain to turn back toward the substrate. Ser-344 (ψ = −152.1, φ = −59.3) mediates a tight turn between αA and αB. The substrate-binding residue His-258 is held in position by the side chain of Asp-255, which is conserved. The barrel is completed by an interaction between the conserved residues Thr-280 (β_8) and Asp-61 on strand β_1. The alignment also highlights some potential differences between NagA sequences of *B. subtilis* and *E. coli*. The intersubunit salt linkage, which is conserved between *B. subtilis* and *T. maritima* NagA proteins, appears to be absent in the same position of the *E. coli* sequence. Although the phosphate-binding residue Arg-234 is conserved in all of the three sequences, its partner, His-233, is not present in *E. coli* NagA. Perhaps the biggest difference is that Gln and Asn residues in *E. coli* NagA replace the residues His-63 and His-65, which bind Fe-2 and coordinate the active site Asp-281.

DISCUSSION

The structural superfamily of metal-dependent hydrolases includes enzymes with highly diverse substrates. This family has been coined an "evolutionary treasure" (26). The trove has expanded with structural characterization of predicted family members (27) including isoaspartyldipeptidase (37), *N*-acyl-D-aminoacylase (38), renal dipeptidase (39), and now NagA. A recurrent feature in the family is the post-translational modification of a lysine residue on strand β_4 to a carboxylated form, which serves to bridge two metal-ions in the catalytic site. In an interesting variation, both NagA and renal dipeptidases have exploited glutamate carboxylates in a similar spatial position,

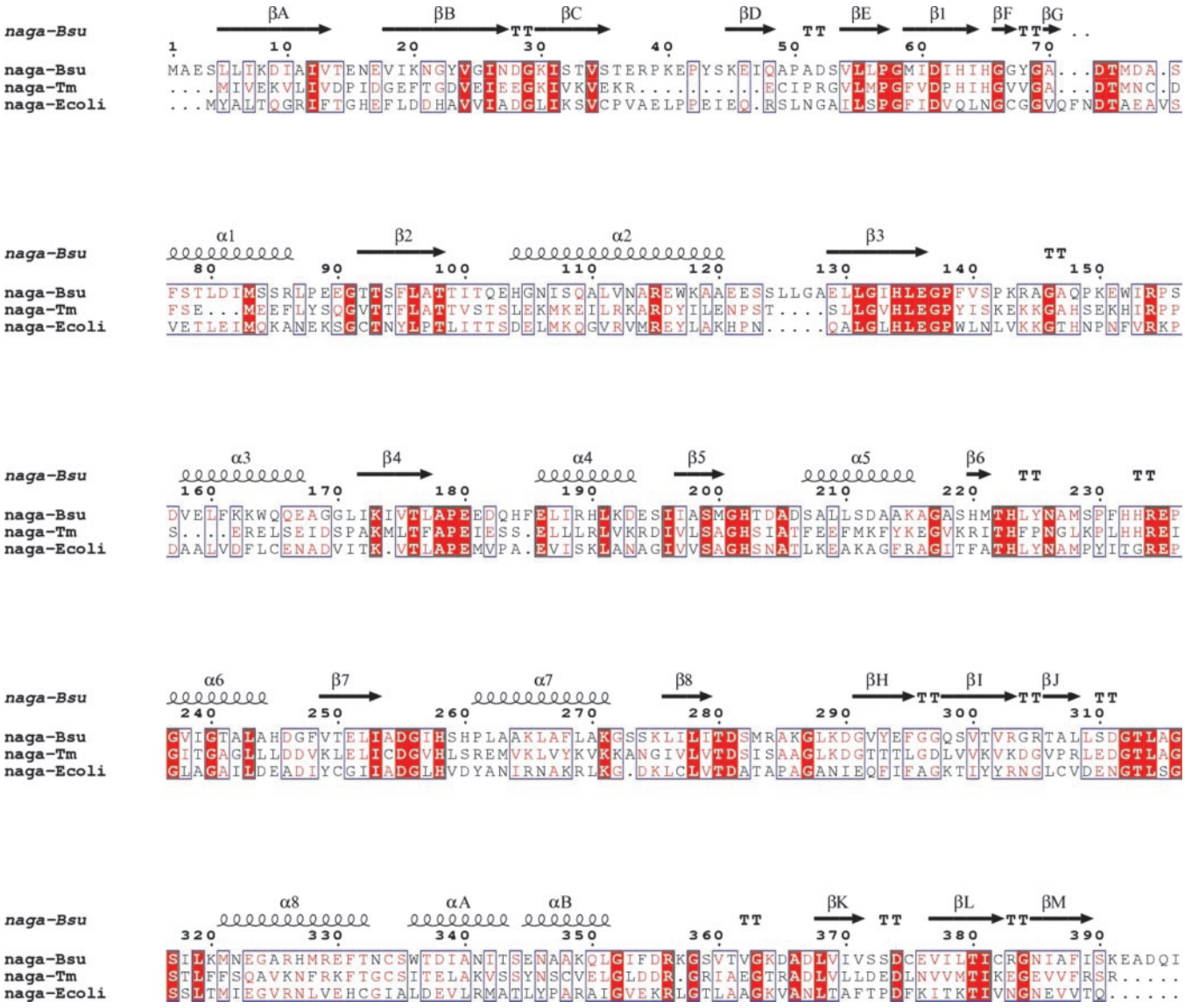

FIG. 6. **Sequence alignment of representative NagA proteins.** The primary sequence of NagA proteins from *B. subtilis* (*Bsu*), *T. maritima* (*Tm*, ORF TM0184), and *E. coli* are aligned. Some bacteria have multiple NagA paralogs. The *E. coli* sequence given here is the gene product that has been crystallized (3). Secondary structure elements labeled as in Fig. 2*b* are drawn above the alignment. The figure was generated using MultAlin (prodes.toulouse.inra.fr/multalin/multalin.html).

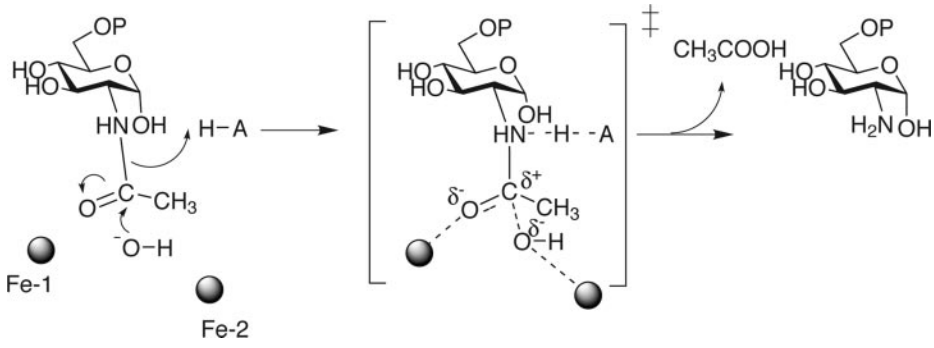

FIG. 7. **Putative reaction mechanism for NagA**. Attack of a metal-bound hydroxide is enhanced through stabilization of the carbonyl by a second metal and proton donation to the leaving group. A candidate for this general acid is Asp-281, although other mechanisms have been proposed for related enzymes (for example, see Refs. 44–48).

which are instead donated from strand β_3 of the β barrel to serve the same role, which may be one reason that NagA was not predicted as a family member by the original grouping based on primary sequence (26).

Another unusual feature of NagA is that across species there are inconsistencies in the gene order and chromosomal operon structure. In *E. coli*, *nagA* is in an operon with components of the import GlcNAc phophoenol pyruvate-phosphotransferase system. Although the same is true in some Gram-positive organisms, such as *Bacillus sphaericus* (40), many Bacillus *nagA* genes are in a short operon with *nagB*, which encodes a GlcN-6P deaminase/epimerase. In *B. subtilis*, the two genes overlap by 2 bp and in an extreme case by 13 bp in *Bacillus halodurans*, strongly suggesting that they are co-transcribed and translated. Indeed, there may be some real differences between the GlcNAc deacetylase/GlcN epimerase systems

across the bacterial order divides. The archaebacteria *T. maritima* lacks a phosphotransferase system, and its *nagA* gene is not genetically linked to *nagB*. The Gram-negative NagA proteins, which have been characterized, appear to be functional as homotetramers (1, 2). The *E. coli* GlcN-6P epimerase is a hexamer that displays allosteric activation (41) and has greater sequence identity to NagB from eukaryotic sources than to the *B. subtilis* NagB, which is a monomer.[3] It is clear that the role and regulation of the NagA/NagB systems is a complex area that demands further investigation.

A Putative Reaction Mechanism for NagA—There is little consensus in the literature concerning the reaction mechanism of urease superfamily members. Even for urease itself, a number of different mechanisms have been proposed (42–45), none of which is compatible with the active center of NagA. What is clear is that the active site must provide at least two functions: (*a*) a catalytic (Brønsted) acid to protonate the leaving group ($pK_a > 30$) to facilitate its departure and (*b*) a water molecule, which must in some way be activated through deprotonation, or hydroxide-ion to perform the hydrolytic attack at the carbonyl carbon of the substrate. For those superfamily members acting on acetyl or *N*-acetyl substrates, it is likely that some means of stabilizing the developing negative charge on the carbonyl oxygen must be provided in a manner analogous to the "oxyanion hole" of serine proteases. There does seem general consensus as to the latter function. Ni-1 of ureases analogous to Fe-1 in NagA (44, 45) seems perfectly positioned to stabilize the carbonyl oxygen. Indeed, it is likely to further enhance the electrophilicity of the carbon through withdrawal of electrons along the C=O bond. In the partial substrate complex described above, Fe-1 lies 2.9 Å from the substrate carbonyl oxygen. Consistent with such proposals is the observation that those superfamily members, such as adenosine and bacterial cytosine deaminases (46–48), that do not act on carbonyl-substituted substrates lack this metal-ion and perform catalysis via a single metal center, although stabilization of the oxyanion in the bi-zinc enzyme renal dipeptidase is proposed to occur via two histidine side chains and not through the structurally analogous metal (39).

The contentious mechanistic issues are the identification of the catalytic acid and the position and direction of the putative attacking nucleophile, and here there is no consistent proposal. For ureases, perhaps the most discussed homologs whose study dates back to Sumners crystallization of urease in 1926 (eloquently reviewed in Ref. 45), two proposals exist: a "reverse protonation" mechanism in which the catalytic acid is provided by the side chain of His-320 (45) and an unusual mechanism (given the pK_a of superoxide of ~21) in which "the proton needed. . . . could easily be provided by the hydroxide itself" (44). Neither mechanism is consistent with the NagA structure, although this may not be surprising because the resonance stabilization of urea through its double amide serves to make the carbonyl less electrophilic and suggests that strategies for its hydrolysis are unrepresentative of the superfamily. There is no direct spatial equivalent to His-320 of the *Klebsiella aerogenes* urease in NagA as this space is instead occupied by an Ala and Phe side-chain from elsewhere in the structure, nor is there any group in this vicinity that could function as a proton donor. Although it is possible that "late" proton donation somehow comes in an intramolecular manner via the departing acetate group (pK_a ~4.6), a more likely acid is provided by the side chain of Asp-281, which lies exactly 3 Å from the glucosamine amine group in perfect geometry for proton donation (Fig. 4). This residue is strictly invariant in all of the superfamily members, which may reflect a common role as proton donor, although strict conservation may equally well reflect its role in binding the second metal-ion. The partially occupied substrate suggests that the most probable attacking nucleophile is the water/hydroxide that bridges the two metals as suggested by Benini *et al.* (44) for the *B. pasteurii* urease (44). Therefore, we propose a putative mechanism for catalysis by NagA (Fig. 7), one that would equally well be appropriate for other members of the urease superfamily involving (*a*) nucleophilic attack by the bridging hydroxide-ion and (*b*) proton-donation to the amine leaving group via Asp-281 and stabilization of the tetrahedral transition state through interaction of Fe-1 with the polarized carbonyl oxygen. We hope that further studies with inhibitors and putative transition-state mimics will lead to a greater understanding of catalysis in this unusual enzyme.

Acknowledgments—We thank the European Synchrotron Radiation Source (Grenoble, France) for the provision of x-ray data collection facilities and Professor A. Karplus (Oregon State) for insightful discussions. The release of coordinates by the structural genomics community is appreciated.

[3] F. Vincent, G. J. Davies, and J. A. Brannigan, unpublished observations.

REFERENCES

1. Souza, J. M., Plumbridge, J. A., and Calcagno, M. L. (1997) *Arch. Biochem. Biophys.* **340,** 338–346
2. Yamano, N., Matsushita, Y., Kamada, Y., Fujishima, S., and Arita, M. (1996) *Biosci. Biotechnol. Biochem.* **60,** 1320–1323
3. Ferreira, F. M., Mendoza-Hernandez, G., Calcagno, M. L., Minauro, F., Delboni, L. F., and Oliva, G. (2000) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **56,** 670–672
4. Park, J. T. (2001) *J. Bacteriol.* **183,** 3842–3847
5. Ronning, D. R., Klabunde, T., Besra, G. S., Vissa, V. D., Belisle, J. T., and Sacchettini, J. C. (2000) *Nat. Struct. Biol.* **7,** 141–146
6. Coggins, B. E., Li, X., McClerren, A. L., Hindsgaul, O., Raetz, C. R. H., and Zhou, P. (2003) *Nat. Struct. Biol.* **10,** 645–651
7. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25,** 3389–3402
8. Prates, J. A. M., Tarbouriech, N., Charnock, S. J., Fontes, C., Ferreira, L. M. A., and Davies, G. J. (2001) *Structure* **9,** 1183–1190
9. Ghosh, D., Sawicki, M., Lala, P., Erman, M., Pangborn, W., Eyzaguirre, J., Gutiérrez, R., Jörnvall, H., and Thiel, D. J. (2001) *J. Biol. Chem.* **276,** 11159–11166
10. Vincent, F., Charnock, S. J., Verschueren, K. H. G., Turkenburg, J. P., Scott, D. J., Offen, W. A., Roberts, S., Pell, G., Gilbert, H. J., Davies, G. J., and Brannigan, J. A. (2003) *J. Mol. Biol.* **330,** 593–606
11. Mølgaard, A., Kauppinen, S., and Larsen, S. (2000) *Structure* **8,** 373–383
12. Jenkins, J., Mayans, O., Smith, D., Worboys, K., and Pickersgill, R. W. (2001) *J. Mol. Biol.* **305,** 951–960
13. Whittington, D. A., Rusche, K. M., Shin, H., Fierke, C. A., and Chistianson, D. W. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 8146–8150
14. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
15. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50,** 760–763
16. Navaza, J. (1994) *Acta Crystallogr. Sect. A* **50,** 157–163
17. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **53,** 240–255
18. Perrakis, A., Harkiolaki, M., Wilson, K. S., and Lamzin, V. S. (2001) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **57,** 1445–1450
19. Brünger, A. T. (1992) *Nature* **355,** 472–475
20. Laskowski, R. A., Macarthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291
21. Grime, G., Dawson, M., Marsh, M., McArthur, I. C., and Watt, F. (1991) *Nucl. Instr. Methods Phys. Res. B* **54,** 52–63
22. Johansson, S., Campbell, J. L., and Malmqvist, Q. G. (1995) *Particle Induced X-ray Emission Spectrometry*, John Wiley & Sons, Inc., New York
23. Kabsch, W., and Sander, C. (1983) *Biopolymers* **22,** 2577–2637
24. Holm, L., and Sander, C. (1995) *Trends Biochem. Sci.* **20,** 478–480
25. Murzin, A. G., Brenner, S. E., Hubbard, T., and Chothia, C. (1995) *J. Mol. Biol.* **247,** 536–540
26. Holm, L., and Sander, C. (1997) *Proteins* **28,** 72–82
27. May, O., Habenicht, A., Mattes, R., Syldatk, C., and Siemann, M. (1998) *Biol. Chem.* **379,** 743–747
28. Jones, S., and Thornton, J. M. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 13–20
29. Oldfield, T. J. (2001) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **57,** 82–94
30. White, R. J., and Pasternak, C. A. (1975) *Methods Enzymol.* **41,** 497–502
31. Bell, J. K., Goetz, D. H., Mahrus, S., Harris, J. L., Fletterick, R. J., and Craik, C. S. (2003) *Nat. Struct. Biol.* **10,** 527–534
32. Hink-Schauer, C., Estebanez-Perpina, E., Kurschus, F. C., Bode, W., and Jenne, D. E. (2003) *Nat. Struct. Biol.* **10,** 535–540
33. Burling, F. T., Kniewel, R., Buglino, J. A., Chadha, T., Beckwith, A., and Lima, C. D. (2003) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **59,** 73–76
34. Mao, C., Cook, W. J., Zhou, M., Koszalka, G. W., Krenitsky, T. A., and Ealick,

S. E. (1997) *Structure* **5,** 1373–1383
35. Garman, E. (1999) *Structure Fold Des.* **7,** 291–299
36. Ramakrishnan, C., and Ramachandran, G. N. (1965) *Biophys. J.* **5,** 909–933
37. Thoden, J. B., Marti-Arbona, R., Raushel, F. M., and Holden, H. M. (2003) *Biochemistry* **42,** 4874–4882
38. Liaw, S. H., Chen, S. J., Ko, T. P., Hsu, C. S., Chen, C. J., Wang, A. H., and Tsai, Y. C. (2003) *J. Biol. Chem.* **278,** 4957–4962
39. Nitanai, Y., Satow, Y., Adachi, H., and Tsujimoto, M. (2002) *J. Mol. Biol.* **321,** 177–184
40. Alice, A. F., Perez-Martinez, G., and Sanchez-Rivas, C. (2003) *Microbiology* **149,** 1687–1698
41. Oliva, G., Fontes, M. R., Garratt, R. C., Altamirano, M. M., Calcagno, M. L., and Horjales, E. (1995) *Structure* **3,** 1323–1332
42. Benini, S., Rypniewski, W. R., Wilson, K. S., Ciurli, S., and Mangani, S. (2001) *J. Biol. Inorg. Chem.* **6,** 778–790
43. Pearson, M. A., Park, I. S., Schaller, R. A., Michel, L. O., Karplus, P. A., and Hausinger, R. P. (2000) *Biochemistry* **39,** 8575–8584
44. Benini, S., Rypniewski, W. R., Wilson, K. S., Miletti, S., Ciurli, S., and Mangani, S. (1999) *Structure Fold Des.* **7,** 205–216
45. Karplus, P. A., Pearson, M. A., and Hausinger, R. P. (1997) *Acc. Chemical Res.* **30,** 330–337
46. Ireton, G. C., McDermott, G., Black, M. E., and Stoddard, B. L. (2002) *J. Mol. Biol.* **315,** 687–697
47. Wang, Z. M., and Quiocho, F. A. (1998) *Biochemistry* **37,** 8314–8324
48. Cooper, B. F., Sideraki, V., Wilson, D. K., Dominguez, D. Y., Clark, S. W., Quiocho, F. A., and Rudolph, F. B. (1997) *Protein Sci.* **6,** 1031–1037
49. Esnouf, R. M. (1997) *J. Mol. Graphics* **15,** 133–138
50. Read, R. J. (1986) *Acta Crystallogr. Sect. A* **42,** 140–149