# Comparative Genomics and Experimental Characterization of N-Acetylglucosamine Utilization Pathway of Shewanella oneidensis<sup>\*S</sup>

Received for publication, May 26, 2006, and in revised form, July 20, 2006 Published, JBC Papers in Press, July 20, 2006, DOI 10.1074/jbc.M605052200

Chen Yang<sup>‡1</sup>, Dmitry A. Rodionov<sup>‡§1,2</sup>, Xiaoqing Li<sup>‡</sup>, Olga N. Laikova<sup>¶</sup>, Mikhail S. Gelfand<sup>§∥</sup>, Olga P. Zagnitko<sup>\*\*</sup>, Margaret F. Romine<sup>‡‡3</sup>, Anna Y. Obraztsova<sup>§§</sup>, Kenneth H. Nealson<sup>§§3</sup>, and Andrei L. Osterman<sup>‡\*\*</sup>

From the <sup>‡</sup>Burnham Institute for Medical Research, La Jolla, California 92037, <sup>§</sup>Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow 127994, Russia, <sup>¶</sup>State Scientific Center GosNIIGenetika, Moscow 117545, Russia, <sup>∥</sup>Department of Bioengineering and Bioinformatics, Moscow State University, Moscow 119992, Russia, \*\*Fellowship for Interpretation of Genomes, Burr Ridge, Illinois 60527, <sup>‡‡</sup>Pacific Northwest National Laboratory, Richland, Washington 99352, and <sup>§§</sup>Department of Earth Sciences, University of Southern California, Los Angeles, California 90089

SEED annotation environment to reconstruct the chitin and GlcNAc utilization subsystem and regulatory network in most proteobacteria, including 11 species of Shewanella with completely sequenced genomes. Comparative analysis of candidate regulatory sites allowed us to characterize three different GlcNAc-specific regulons, NagC, NagR, and NagQ, in various proteobacteria and to tentatively assign a number of novel genes with specific functional roles, in particular new GlcNAc-related transport systems, to this subsystem. Genes SO3506 and SO3507, originally annotated as hypothetical in Shewanella oneidensis MR-1, were suggested to encode novel variants of GlcN-6-P deaminase and GlcNAc kinase, respectively. Reconstitution of the GlcNAc catabolic pathway in vitro using these purified recombinant proteins and GlcNAc-6-P deacetylase (SO3505) validated the entire pathway. Kinetic characterization of GlcN-6-P deaminase demonstrated that it is the subject of allosteric activation by GlcNAc-6-P. Consistent with genomic data, all tested Shewanella strains except S. frigidimarina, which lacked representative genes for the GlcNAc metabolism, were capable of utilizing GlcNAc as the sole source of carbon and energy. This study expands the range of carbon substrates utilized by Shewanella spp., unambiguously identifies several genes involved in chitin metabolism, and describes a novel variant of the classical three-step biochemical conversion of GlcNAc to fructose 6-phosphate first described in Escherichia coli.

We used a comparative genomics approach implemented in the

The genus *Shewanella* is composed of a number of species and strains, all of which are Gram-negative  $\gamma$ -proteobacteria (order Alteromonadales). These bacteria are widespread in aquatic environments and abundant in environments where nutrient levels are high and redox interfaces are found. *Shewanella oneidensis* MR-1 was originally isolated from freshwater sediments of Oneida Lake, NY (1), whereas other species have been isolated from marine sediments, marine waters, and a variety of other environments (2). Because of their respiratory versatility (many can respire 10 or more different electron acceptors), they are thought to play an important role in the cycling of organic carbon and in the geobiology and ecology of sedimentary environments. In contrast to their versatility with regard to electron acceptors, *Shewanella* tend to be rather limited with regard to energy sources utilized.

Thus, it was of great interest to perform the genomic analysis of *Shewanella* to see if this limited nutritional versatility is a feature of the organism and not a result of our limited knowledge of its biology. To this end, a significant effort was allocated to the genomic and functional analysis of this interesting group of species (3). As of now, the genomes of 11 strains of *Shewanella* have been sequenced (4), and one of these strains, *S. oneidensis* MR-1, became a subject of careful genomic annotation and in-depth functional genomics analysis (5–7). The wealth of accumulated data, together with features enabling physiological and genetic studies (facultative anaerobe, robust growth in defined medium, etc.) and practical importance, make *S. oneidensis* a very attractive model system representing a genuine environmental (aquatic) lifestyle.

Chitin, the second-most abundant organic polymer in nature after cellulose, is composed of GlcNAc residues that play an important role in supplying carbon and energy in a variety of organisms (8). Chitin is present in the cell walls of fungi as well as in cuticles and exoskeletons of worms, mollusks, and arthropods, and it constitutes a natural source of GlcNAc for many bacteria in respective ecosystems, such as large aquatic reservoirs. An extremely efficient utilization of chitin in the ocean is evidenced by the fact that its presence in marine sediments is nearly undetectable. Marine chitinolytic bacteria, such as Vibrionales, are presumed to play a key role in this process (9). Chitin catabolism in Vibrios is initiated by detection of chitin oligosaccharides and/or GlcNAc by chemotactic sensors and is followed by the up-regulation of a suite of genes that mediate chitin depolymerization at the cell surface, uptake of resulting monomers across the outer membrane, further degradation

•<del>OCTO</del>BER 6, 2006

<sup>\*</sup> This work was supported in part by NIAID Grant 1-R01-Al059146-01A2 from the National Institutes of Health (to A. L. O.), Howard Hughes Medical Institute Grant 55005610 (to M. S. G.), Russian Academy of Sciences Program "Molecular and Cellular Biology" (to D. A. R. and M. S. G.), Russian Fund of Basic Research Grant 04-04-49361 (to D. A. R.), and INTAS Grant 05-8028 (to M. S. G. and D. A. R.). 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.

S The on-line version of this article (available at http://www.jbc.org) contains supplemental Tables S1, S2, S3, and S4.

<sup>&</sup>lt;sup>1</sup> Both authors contributed equally to this work.

<sup>&</sup>lt;sup>2</sup> To whom correspondence should be addressed: Burnham Institute for Medical Research, 10901 N. Torrey Pines Rd., La Jolla, CA 92037. Tel.: 858-646-3100; Fax: 858-713-9949; E-mail: rodionov@burnham.org.

<sup>&</sup>lt;sup>3</sup> Supported by the Dept. of Energy Office of Biological and Environmental Sciences under the Genomics-GTL Program via the *Shewanella* Federation consortium.

into GlcNAc, and subsequent transport across the inner membrane to the cytoplasm where it is finally transformed to the common metabolic intermediate fructose 6-phosphate (Fru-6-P)<sup>4</sup> (10, 11). Enzymology and regulation of the chitinolytic machinery have also been studied in some aquatic representatives of the Alteromonadales order, such as *Alteromonas* sp. (12) and *Saccharophagus degradans* (formerly *Microbulbifer degradans*) (13), but so far not in *Shewanella* spp.

The transformation of GlcNAc to Fru-6-P is catalyzed via consecutive phosphorylation, deacetylation, and isomerization-deamination reactions (Fig. 1). This three-step biochemical pathway (*i.e.* the NAG pathway) appears to be conserved both in chitinolytic bacteria (such as *Vibrio cholerae*) and in non-chitinolytic species such as *Escherichia coli* (11, 14). In *E. coli*, genes involved in the NAG pathway are organized in the *nagBACD* operon (Fig. 2) and controlled by the transcriptional regulator NagC (15). The same regulator controls expression of the *chb* operon, whose products enable *E. coli* to utilize chitobiose (Fig. 1), a disaccharide form of GlcNAc (16).

Because the initial genomic survey of S. oneidensis MR-1 only identified the presence of one of the three canonical NAG pathway genes encoding the GlcNAc-6-P deacetylase NagA (5), the occurrence of a NAG catabolic pathway in MR-1 is not obvious. We used a subsystems-based comparative genomic analysis as implemented in the SEED platform (17) to explore the GlcNAc and chitin utilization machinery in several species of Shewanella with completely sequenced genomes. The analysis of the subsystems was used to project gene assignments from model species to other genomes and to identify which pathway variants are likely implemented in groups of species (18). It also revealed gaps in the existing knowledge, such as missing genes for enzymatic steps inferred by metabolic reconstruction. The analysis of genome context, e.g. conserved operons and regulons, was then used to predict gene candidates amenable to focused experimental verification (19).

The Journal of Biological Chemistry

By using this approach, we have identified and experimentally validated a novel functional variant of the NAG pathway present in most Shewanella spp. Two of three enzymes involved in this pathway, GlcNAc kinase (SO3507 in S. oneidensis, named NagK) and an alternative GlcN-6-P deaminase (SO3506, named NagB-II), that were not homologous to the respective components of the E. coli NAG pathway, were characterized in more detail. In addition, we have tentatively identified a novel transcriptional regulator (SO3516, named NagR) and a permease (SO3503, named NagP) associated with the NAG pathway in *Shewanella* and several related species, as well as several other genes likely associated with the utilization of chitin and/or chito-oligosaccharides. Many of these genes occur in operons that form a novel predicted regulon controlled by NagR. Although most of these functional predictions have yet to be validated experimentally, the overall pattern of genes tentatively associated with the NAG pathway (biochemically

## N-Acetylglucosamine Metabolism in Shewanella

characterized in this study) provides the first strong evidence of an elaborate chitin utilization machinery, which is likely to be conserved in many *Shewanella* spp.

#### **EXPERIMENTAL PROCEDURES**

# Genomic Reconstruction of GlcNAc Utilization and Related Pathways in Proteobacteria

Genomes and Bioinformatics Tools-Complete and nearly complete genomes of proteobacteria analyzed in this study were uploaded from GenBank<sup>TM</sup>, the Institute for Genomic Research, the Department of Energy Joint Genome Institute, and Welcome Trust Sanger Institute as listed in supplemental Table S1. Most of these and other (supporting) genomes used for comparative analysis are integrated in the SEED genomic data base as listed on line. For analysis of regulatory DNAbinding motifs and their genomic context, we used the integrated GenomeExplorer software, including the SignalX program for identification of conserved regulatory motifs (20). Sequence logos for regulatory motifs were depicted using the WebLogo package version 2.6 (21). In addition we used ClustalX to construct multiple protein alignments and trees (22); Psi-BLAST (23) to conduct long range similarity searches; the PFAM (24) and Conserved Domain data bases (25) to identify conserved functional domains; and TMpred (26) to predict transmembrane segments.

Subsystem Encoding, Annotations, and Genome Context Analysis—A set of subsystems-based genomic annotations and metabolic reconstruction tools implemented in the SEED platform were used to capture the existing knowledge of GlcNAc utilization pathways and to tentatively project it to a broader collection of bacteria with completely sequenced genomes. The general approach was originally described by Overbeek et al. (17), and its application was recently illustrated for the analysis of NAD metabolism in cyanobacteria (27). Briefly, a subsystem is initiated by defining a list of functional roles (enzymes, transporters, and regulators) immediately associated with a biological process under study. This information is obtained by review of the literature related to NAG metabolism and pathway-reaction-compound information available in public resources such as KEGG and MetaCyc. In this study we have focused on the conversion of GlcNAc to Fru-6-P placed in a broader functional context with feeding (chitin hydrolysis and uptake) and regulatory pathways. The subsystem phylogenetic boundaries were limited to the proteobacteria lineage. Most initial information about functional roles and respective genes was derived from the studies in E. coli K12 (14) and a group of chitinolytic bacteria, including Vibrio spp. (11) and Saccharophagus degradans (13). A subsystem is encoded as a spreadsheet, where rows correspond to sequenced genomes and columns correspond to functional roles. Individual genes (proteins) are connected to cells in the spreadsheet via internally consistent annotations. Functional roles encoded by nonorthologous groups of genes are represented by sets of two (or more) separate columns (alternative forms). Subsystem expansion beyond model species and their closest relatives is accomplished by addition of increasingly distant genomes (rows) and orthology-based projection of gene annotations. The analysis of genome context (e.g.



<sup>&</sup>lt;sup>4</sup> The abbreviations used are: Fru-6-P, fructose 6-phosphate; GlcNAc-6-P, N-acetylglucosamine-6-phosphate; NAG pathway, N-acetylglucosamine catabolic pathway; PTS, phosphoenolpyruvate-carbohydrate phosphotransferase system; GlcN-6-P, glucosamine 6-phosphate; ManNAc-6-P, N-acetylmannosamine 6-phosphate; Ni-NTA, nickel-nitrilotriacetic acid; PIPES, 1,4-piperazinediethanesulfonic acid.

# Downloaded from www.jbc.org at Burnham Institute for Medical Research LIBRARY on November 7, 2006

# N-Acetylglucosamine Metabolism in Shewanella

clustering on the chromosome) and *functional context (e.g.* pathway requirements) significantly improves the quality and consistency of tentative functional assignments. This analysis also allows us to do the following: (i) recognize gene patterns corresponding to major functional variants of the subsystem (18); (ii) map inconsistencies and gaps in the existing knowledge, *e.g.* missing genes for required functional roles; and (iii) identify additional protein families (with known or hypothetical functions) possibly related to the subsystem, such as potential transporters and regulators implicated by genome context analysis.

Regulatory Signals and Regulons-An iterative motif detection procedure implemented in the program SignalX was used to identify common regulatory DNA motifs in a set of upstream gene fragments and to construct the motif recognition profiles as described previously (28). For the NagC regulon, we started from a training set of the upstream regions of known NagC targets in E. coli and their orthologs in other NagC-encoding genomes (15, 29, 30). For predicted novel transcriptional regulators (termed NagR and NagQ), the training sets included the upstream regions of potentially co-regulated genes in the chitin utilization and NAG pathways. The constructed recognition rules were used to scan against a subset of genomes of proteobacteria having a respective regulator. Positional nucleotide weights in the recognition profile and Z-scores of candidate sites were calculated as the sum of the respective positional nucleotide weights (31). The threshold for the site search was defined as the lowest score observed in the training set. This analysis produced gene sets with candidate regulatory sites in the upstream regions.

*Functional Predictions*—Previously uncharacterized genes detected within conserved clusters and/or as conserved members of corresponding regulons (as described above) are considered to be potentially involved (functionally coupled) with the subsystem under study. Depending on the functional context in a subset of genomes containing such genes, they may constitute candidates either for previously identified missing genes or for additional functional roles not included in the initial subsystem encoding. The analysis of long range homologs (*e.g.* by Psi-BLAST) and conserved domains in proteins (*e.g.* by Conserved Domain and PFAM data base search) often suggests a general class function providing additional support for specific functional predictions amenable to focused experimental verification, as demonstrated previously (19) and in this study.

#### Bacterial Strains, Plasmids, and Reagents

The Journal of Biological Chemistry

ibc

*E. coli* strains DH5 $\alpha$  (Invitrogen) and BL21/DE3 (Stratagene, La Jolla, CA) were used for gene cloning and protein overexpression, respectively. For expression of all genes in *E. coli*, a pET-derived vector containing the T7 promoter, His<sub>6</sub> tag, and tobacco etch virus-protease cleavage site (32) was used. Enzymes for PCR and DNA manipulations were from New England Biolabs Inc. (Beverly, MA). Plasmid purification kits were from Promega (Madison, WI). PCR purification kits and nickel-nitrilotriacetic acid resin were from Qiagen Inc. (Valencia, CA). Oligonucleotides for PCR and sequencing were synthesized by Sigma-Genosys (Woodlands, TX). All other chemicals, including the assay components GlcNAc, GlcNAc-6-P,

GlcN-6-P, NADH, NADP<sup>+</sup>, ATP, phosphoenolpyruvate, lactate dehydrogenase, pyruvate kinase, phosphoglucose isomerase, glucose-6-phosphate dehydrogenase, and *p*-dimethylaminobenzaldehyde, were purchased from Sigma.

#### PCR Amplification and Cloning

Three protein-coding genes from S. oneidensis MR-1 (SO3505, SO3506, and SO3507) were amplified by using three sets of primers shown below. Introduced restriction sites (BspHI for the 5'-end and SalI and XhoI for the 3'-end) are shown in boldface; nucleotides not present in the original sequence are shown in lowercase. For GlcNAc-6-P deacetylase (SO3505), no mutations were introduced, 5'-ggcgctcATGAAAT-TCACTTTAATTGCCGAGCAATTATTTG and 3'-gacgcgtcgacTTACGGGCGAACATAGACCGCATTTCC; for predicted GlcN-6-P deaminase (SO3506), no mutations were introduced, 5'-ggcgctcATGACAAACACTATTATGGAACA-AGAAGCG and 3'-gacgcctcgagCTACAGCGTTTGAGT-GACTTTTTTCAG; for predicted GlcNAc kinase (SO3507), the second codon was mutated from GGA to ACA (Gly  $\rightarrow$  Thr), 5'-ggcgctcATGACATTAGTCCAGACAAATGATCAACAAC and 3'-gacgcgtcgacTTAAACTGTTGCTGAATTAAATTG-CTG. PCR amplification was performed using S. oneidensis MR-1 genomic DNA. PCR fragments were cloned into the expression vector, which was cleaved by NcoI and SalI. Selected clones were confirmed by DNA sequence analysis.

Protein overexpression and purification were performed from 50-ml cultures using rapid Ni-NTA-agarose minicolumn protocol as described (33). Briefly, cells were grown in LB medium to  $A_{600} = 0.8$  at 37 °C, induced by 0.2 mM isopropyl 1-thio- $\beta$ -D-galactopyranoside, and harvested after 12 h of shaking at 20 °C. Harvested cells were resuspended in 20 mM HEPES buffer (pH 7) containing 100 mM NaCl, 0.03% Brij-35, and 2 mM  $\beta$ -mercaptoethanol supplemented with 2 mM phenylmethylsulfonyl fluoride and a protease inhibitor mixture (Sigma). Lysozyme was added to 1 mg/ml, and the cells were lysed by freezing-thawing followed by sonication. After centrifugation at 18,000 rpm, the Tris-HCl buffer (pH 8) was added to the supernatant (50 mM, final concentration), and it was loaded onto a Ni-NTA-agarose minicolumn (0.2 ml). After washing with the starting buffer containing 1 M NaCl and 0.3% Brij-35, bound proteins were eluted with 0.3 ml of the starting buffer containing 250 mM imidazole. Protein size, expression level, distribution between soluble and insoluble forms, and extent of purification were monitored by SDS-PAGE. In all three cases, soluble proteins were obtained with high yield (>1 mg from 50-ml culture) purified to >90% by SDS-PAGE (Fig. 3).

#### Enzyme Assays

GlcNAc kinase activity was assayed by two methods as described (34). In the end point assay A, the consumption of GlcNAc substrate was monitored using *p*-dimethyl-aminobenzaldehyde reagent after the specific removal of GlcNAc-6-P product by the addition of  $ZnSO_4$  and  $Ba(OH)_2$  solutions. In the continuous assay B, the conversion of ATP to ADP was enzymatically coupled to the oxidation of NADH to NAD<sup>+</sup> and monitored at 340 nm. Briefly, 0.2–0.4  $\mu$ g of purified GlcNAc kinase was added to 200  $\mu$ l of reaction mixture containing



50 mM Tris buffer (pH 7.5), 10 mM MgSO<sub>4</sub>, 1.2 mM ATP, 1.2 mM phosphoenolpyruvate, 0.3 mM NADH, 1.2 units of pyruvate kinase, 1.2 units of lactate dehydrogenase, and 1 mM GlcNAc. For determination of the apparent  $k_{cat}$  and the  $K_m$  values for GlcNAc, its concentration was varied in the range of 0.1–5.0 mM in the presence of 3.2 mM (saturating) ATP.

GlcNAc-6-P deacetylase activity was tested by adding 2  $\mu$ g of purified GlcNAc-6-P deacetylase to 500  $\mu$ l of reaction mixture containing 0.5 mM GlcNAc-6-P and 0.1 M Tris (pH 7.5) at 37 °C. The amount of GlcNAc-6-P consumed after 0–3 min was determined by using *p*-dimethylaminobenzaldehyde reagent and reading the absorbance at 585 nm (35). Determination of the  $k_{cat}$  and  $K_m$  values was performed by varying the concentration of GlcNAc-6-P in the range of 0.1–5.0 mM.

GlcN-6-P deaminase activity was assayed by coupling the formation of Fru-6-P to the reduction of NADP<sup>+</sup> to NADPH and detection at 340 nm as described (35). Briefly, 1.5–3.0  $\mu$ g of purified GlcN-6-P deaminase was added to 200  $\mu$ l of reaction mixture containing 40 mM sodium phosphate buffer (pH 7.5), 0.2 mM NADP<sup>+</sup>, 8 units of phosphoglucose isomerase, 3 units of glucose-6-phosphate dehydrogenase, and 1 mM GlcN-6-P. No activity was detected in a control experiment without phosphoglucose isomerase, which verified the production of Fru-6-P but not glucose 6-phosphate. Determination of  $k_{cat}$  and  $K_m$  values was performed by varying the concentration of GlcN-6-P in the range of 0.1–10.0 mM. The allosteric activation of GlcN-6-P deaminase by GlcNAc-6-P was detected and assessed by performing the same assay in the presence of 0.05–1.00 mM GlcNAc-6-P.

All these assays were performed at 37 °C. The change in NADH or NADPH absorbance was monitored at 340 nm using a Beckman DTX-880 multimode microplate reader. An NADH or NADPH extinction coefficient of 6.22 mm<sup>-1</sup> cm<sup>-1</sup> was used to determine initial rates using Multimode Detection software (Beckman). Further kinetic analysis was performed with the software XLfit 4 (IDBS). A standard Michaelis-Menten model was used to determine the apparent  $k_{cat}$  and  $K_m$  values of Glc-NAc kinase and GlcNAc-6-P deacetylase, and the Hill equation was used to fit the data of GlcN-6-P deaminase.

#### In Vitro Reconstitution of NAG Pathway

A three-step conversion of GlcNAc to Fru-6-P was assessed by coupling the formation of Fru-6-P to the reduction of NADP<sup>+</sup> and detection of NADPH at 340 nm (as in GlcN-6-P deaminase activity assay described above). The reaction mixture contained 1 mM GlcNAc in 50 mM Tris (pH 7.5) with 20 mм MgCl<sub>2</sub>, 1 mм GlcNAc, 0.2 mм NADP<sup>+</sup>, 8 units/ml phosphoglucose isomerase, 3 units/ml glucose-6-phosphate dehydrogenase, and a mixture of three purified recombinant proteins, GlcNAc kinase, GlcNAc-6-P deacetylase, and GlcN-6-P deaminase (2  $\mu$ M each). All possible binary combinations of these enzymes were used in control runs to confirm that the presence of all three enzymes is strictly required. The order of enzymatic steps was additionally verified by replacing GlcNAc with 1 mM GlcNAc-6-P and incubating with a binary mixture of 2 µM GlcNAc-6-P deacetylase and GlcN-6-P deaminase. To assess potential rate-limiting steps in the NAG pathway, the concentration of each enzyme in a series of experiments was

# N-Acetylglucosamine Metabolism in Shewanella

decreased to 1 and 0.5  $\mu{\rm M},$  although the concentration of the other two enzymes was kept constant (2  $\mu{\rm M}).$ 

#### Growth Conditions for Phenotypic Analysis of Shewanella spp.

The following seven Shewanella strains were used in the phenotypic analysis: S. oneidensis MR-1 (culture collection deposit number ATCC BAA-1096); Shewanella amazonensis SB2B (ATCC BAA-1098); Shewanella putrefaciens CN-32 (ATCC BAA-454); Shewanella sp. PV-4 (ATCC BAA-1088); Shewanella denitrificans OS217 (ATCC BAA-1090); Shewanella baltica OS155 (ATCC BAA-1091); and Shewanella frigidimarina NCIMB400 (ATCC BAA-1089). Cells were aerobically grown at room temperature in 5-ml test tubes on defined medium of the following composition (g/liter): NH<sub>4</sub>Cl (0.3), KCl (0.1), NaCl (1.75), PIPES (0.91), NaOH (0.3), and NaH<sub>2</sub>PO<sub>4</sub> (0.6). 10 ml of  $100 \times$  stock of trace elements, vitamins, and amino acids solutions (36) were also added into the medium. pH of the medium was adjusted to 7.0-7.2. The ability to grow with 20 mM lactate as sole carbon and energy source was compared with the ability to grow on 10 mM GlcNAc. Growth was monitored by measuring absorbance at 600 nm with a visible light spectrophotometer (Spec 20).

#### RESULTS

*Comparative Genomics of GlcNAc and Chitin Utilization in Proteobacteria*—We used a subsystems-based approach to assess metabolic potential and possible regulatory mechanisms associated with utilization of GlcNAc and of its polymeric or oligomeric precursors in *S. oneidensis* and related species with completely sequenced genomes. This approach combines a number of comparative genomics techniques to analyze genes of the target organism(s) in a broader context of related pathways and species. The detailed results of this analysis are captured in the SEED subsystem available on line and in supplemental Table S1. The key results are illustrated in Fig. 1 and Tables 1 and 2 and contain both previously known and novel features predicted using the genome context analysis.

Although an overall topology of this subsystem and its role in utilizing exogenous sources of GlcNAc for feeding Fru-6-P to central carbon metabolism is preserved in most of the analyzed species, nearly all aspects of its implementation are associated with remarkable variations (functional variants). As illustrated in Fig. 1, only one gene of the NAG pathway (nagA) is shared between S. oneidensis and E. coli. Major species-to-species variations in the key functional modules include the following: (i) existence of three alternative transcriptional regulators and regulatory signals; (ii) presence or absence of chitinolytic machinery and variations therein; (iii) alternative mechanisms of sugar uptake; and (iv) nonorthologous gene displacements for two of three enzymatic steps of the NAG pathway (Table 1 and Table S1). Some of these variations are briefly highlighted in the below where we focus mostly on novel findings and conjectures.

#### Transcriptional Regulation

*NagC Regulon*—The GlcNAc-6-P-responsive transcription factor NagC was initially characterized in enterobacteria as a repressor of the *nag* operon encoding NAG pathway genes (15). NagC was further shown to additionally regulate the *chb* 





FIGURE 1. Reconstruction of the GlcNAc and chitin utilization pathways in proteobacteria. Functional roles and metabolic blocks (pathways) present in *S. oneidensis* and *E. coli* are shown on *green* and *yellow backgrounds*, respectively. Those present in other proteobacteria (but not in *S. oneidensis* and *E. coli*) are shown on a *gray background. Solid arrows* denote enzymatic reactions, and *dotted arrows* denote transport. *Broken arrows* denote links to other catabolic subsystems (utilization of chitobiose, sialic acids, and glucosamine) that are not analyzed in this study.

operon responsible for utilization of chitobiose, (GlcNAC)<sub>2</sub>, and the glmUS genes involved in the GlcNAC biosynthetic pathway in E. coli (16, 29). NagC belongs to the ROK (Repressor, Open reading frame, Kinase) protein family and has an N-terminal DNA-binding domain and a C-terminal sugar kinase-like domain (37). Although NagC orthologs are present in the genomes of enterobacteria, Vibrionales, and Pasteurella multocida (Table 1), the NagC-binding motif and regulons have not been systematically analyzed. The analysis of upstream regions of NagC-controlled genes and their orthologs in these taxonomic groups resulted in identification of the NagC-binding motif, which was used for identification of additional NagC targets in the analyzed microbial genomes (supplemental Table S2). The obtained consensus logo for NagC-binding sites in  $\gamma$ -proteobacteria (Fig. 2) is in accordance with the consensus experimentally determined for E. coli NagC (30). The most

conserved part of the NagC regulon includes the NAG pathway genes *nagA* and *nagB* and PTS<sup>GlcNAc</sup> *nagE*. The largest extension of the NagC regulon observed in chitinolytic Vibrionales species (Table 1) includes genes involved in chitin sensing and degradation (chitinases,  $\beta$ -*N*-acetylglucosaminidase, chitodextrinase, and CBP (chitin-binding protein)), chemotaxis toward chitin (MCP (methyl-accepted chemotaxis protein), uptake of chitin degradation products across the outer and inner membranes (chitoporins and permeases), and desulfation (Table S2). These findings agree with microarray data on the induction of gene expression by GlcNAc and chito-oligosaccharides in *V. cholerae* (11).

Many other proteobacteria, beyond Enterobacteriales and Vibrionales, having the NAG pathway do not contain NagC orthologs. Based on the analysis of conserved chromosomal clusters (operons), we have tentatively identified previously uncharacterized members of the LacI and GntR families as alternative transcription regulators of the NAG pathway in two other groups of proteobacteria.

Novel NagR Regulon-A transcriptional regulator of LacI type (named here *nagR*) was identified in a chromosomal cluster with the nag genes conserved in Alteromonadales, including Shewanella spp., and Xanthomonadales (Table 1). The length (18 bp) and the central CG pair of the derived palindromic consensus aaTGACArCGyTGTCAtt (named here NAG-box, Fig. 2) are characteristic of DNA-binding sites of LacI family regulators (38). Predicted members of the novel NagR regulon (supplemental Table S3) include enzymes that are likely involved in chitin degradation (ChiA, CBP, CdxA, and Hex) as well as a number of new genes implicated in uptake and utilization of GlcNAc via a new version of the NAG pathway (see below). The presence of NAG-boxes upstream of the nagR gene in Shewanella and Colwellia species suggests possible autoregulation of its expression. Colwellia psychrerythraea has an additional gene locus, CPS2360-CPS2383, with multiple NAGboxes in upstream regions of genes encoding two additional NagR paralogs, a TonB-dependent outer membrane receptor and a number of putative sugar hydrolases and sulfatases that may be involved in degradation/desulfation of a complex sulfated GlcNAc polysaccharide. High sequence similarity of NagR paralogs and the presence of similar NAG-boxes in their regulatory regions suggest a recent duplication of *nagR* genes in the C. psychrerythraea genome accompanied by an extension of the NagR regulon to co-regulate the NAG and CPS2360-CPS2383 genes.

Novel NagQ Regulon—Another transcription factor of the GntR type (named here NagQ) was inferred based on chromosomal clustering with NAG pathway genes in  $\alpha$ - and  $\beta$ -proteobacteria and in several representatives of  $\gamma$ -proteobacteria (Table 1). The deduced binding motif consists of two or three imperfect repeats of TGGTATT separated by 4 bp (Fig. 2). Such operator structure is typical for regulators from the GntR family (39). Candidate NagQ sites occur in upstream regions of all *nag* operons containing the *nagQ* gene. Additional sites were detected upstream of chitinolytic genes in *Chromobacterium violaceum, Caulobacter crescentus*, and *Burkholderia* spp. (supplemental Table S4). Interestingly, regulation of GlcNAc and chitin utilization in a Gram-positive bacterium *Streptomyces* 

#### TABLE 1

The Journal of Biological Chemistry

ibc

#### Occurrence and features of genes involved with GlcNAc and chitin utilization subsystem

Species with completely sequenced genomes in several taxonomic groups of proteobacteria are shown as rows. The presence of genes for the respective functional roles and metabolic blocks (columns) is shown by capital letters corresponding to the three identified regulons. C, NagC regulon ( $\beta$ - and  $\alpha$ -proteobacteria). Other genes that were not identified within the above GlcNAc regulons are marked by "U." Genes clustered on the chromosome (operons) are outlined by matching background colors. Tentatively predicted functional roles are marked by asterisks. The complete version of this table with gene identifications is available in Table S1 in supplementary data.

Touronamia amoun	NAG pathway				GIcNAc-related transport				Regulators			Chitin Degradation						
(total number of genomes /	Na	ıgK		Na	ıgВ	Inn	er Me	embra	ne	Por	ins						es	
species with NAG pathway) Organism	I*	*	NagA	I	II*	PTS	ABC*	NagP*	NagX*	ChiP-II*	ChiP-III*	NagC	NagQ'	NagR*	НЕХ	CdxA	Chitinas	CBP
Alteromonadales (11/15)																		
Shewanella oneidensis MR-1	R		R		R			R	R	R				R	R		U	R
S. putrefaciens CN-32	R		R		R			R	R	R				R	R			
S. amazonensis SB2B	R		R		R			R	R	R				R	R			
Shewanella sp. PV-4	R		R	R				R	R	R				U	R			
Shewanella baltica OS155	R		R		R			R	R	R				R	R	R	R	R
S. denitrificans OS-217	R		R		R			R	R	R				R	R		R	R
Shewanella sp. ANA-3 <sup>1</sup>	R		R		R			R	R	R				R	R		R	R
Shewanella frigidimarina																		
Saccharophagus degradans	R		R		R			R	R	R				U	R	R	R	
Colwellia psychrerythraea	R		R	R				R	R	R				R	R			
Xanthomonadales (5/5)																		
Xanthomonas axonopodis		R	Q		Q			Q	R	R			Q	Q	U			
Xylella fastidiosa		R	R		R			R						R	U			
Stenotrophomonas maltophilia		R	R		R			R	R	R				R	R		R	U
Enterobacteriales (11/12)																		
Escherichia coli			С	С		С					С	С					С	
Serratia marcescens			С	С		С					С	С			С		С	С
Erwinia carotovora			С	С		С						С						
Photorhabdus luminescens			С	С		С						С			С		С	С
Yersinia pestis			С	С		С					С	С			С			
Pasteurellales (1/6)																		
Pasteurella multocida			С	С		С						U						
Vibrionales (5/5)																		
Vibrio cholerae			С	С		С						С			С	С	С	С
Photobacterium profundum			С	С		С					С	С			С	С	С	С
Pseudomonadales (2/5)																		
Pseudomonas aeruginosa			Q		Q	Q							Q					
eta -proteobacteria (9/24)																		
Burkholderia pseudomallei			Q		Q	Q							Q		Q		U	U
Chromobacterium violaceum	U		Q		Q	Q	Q				Q		Q		Q	Q	Q	Q
lpha -proteobacteria (14/42)				I														
Brucella melitensis	Q		Q		Q		Q						Q					
Mesorhizobium loti	Q		Q		Q		Q						Q					
Gluconobacter oxydans	U		U	U				U			-				U			
Caulobacter crescentus			Q		Q	Q			Q	Q			Q		Q			

<sup>1</sup>Closely related *Shewanella* species ANA-3, MR-4, and MR-7 have the same set of genes and genome context.

*coelicolor* is also mediated by a transcriptional factor from the GntR family (DasR or SCO5231) (40).

#### Chitin Degradation and Uptake to the Periplasm

The analysis of operons and regulons associated with GlcNAc and chitin utilization subsystem allowed us to accurately annotate

and map a number of previously uncharacterized components of chitinolytic machinery in *Shewanella* spp. and other proteobacteria (supplemental Table S1). Most genes encoding chitinases (*chi*), chitin-binding proteins (*cbp*), chitodextrinase (*cdxA*), and  $\beta$ -*N*-acetylglucosaminidase (*hex*) appear to be co-regulated with the NAG pathway (Table 1 and Fig. 2).

#### TABLE 2

#### Predicted functional roles in the chitin and GlcNAc utilization pathway

I O	Phylogenetic distribution			
NagB-II SO3506 GlcN-6P deaminase (validated)	Alteromonadales, <i>Pseudomonas</i> , $\alpha$ - and $\beta$ -proteobacteria			
NagK-I SO3507 GlcNAc kinase (validated)	Alteromonadales, $\alpha$ -proteobacteria, Chromobacterium violaceum			
NagK-II XF1460 GlcNAc kinase, the glucokinase family	Xanthomonadales			
NagR SO3516 NAG regulator, the LacI family	Alteromonadales, Xanthomonadales			
NagQ CV0555 NAG regulator, the GntR family	α-, β-Proteobacteria; <i>Pseudomonas</i>			
NagP SO3503 GlcNAc permease	Alteromonadales, Xanthomonadales, Gluconobacter oxydans			
NagX SO3504 GlcNAc-related transporter	Alteromonadales, Xanthomonadales, Caulobacter crescentus			
ChiP-II SO3514 TonB-dependent outer membrane transporter for chito-oligosaccharides	Alteromonadales, Xanthomonadales, C. crescentus			
ChiP-III YbfM_EC Outer membrane porin for chito-oligosaccharides	Enterobacteriales, C. violaceum, Photobacterium profundum			
ABC <sup>GlcNAc</sup> M14771–4779 GlcNAc ABC transporter	α-Proteobacteria, <i>C. violaceum</i>			
MCP SO3510 Chemotaxis toward chitin	Alteromonadales, Vibrionales			
YdeMN B1497-B1498 GlcNAc-related sulfatase	Enterobacteriales			



FIGURE 2. Genome context of genes associated with GlcNAc and chitin utilization subsystem in Shewanella spp. and related proteobacteria. The presence of regulatory sites upstream of the respective genes and operons is shown by symbols explained in the boxed inset (together with "logo" representing corresponding consensus regulatory DNA motifs). Genes predicted by genome context analysis are marked by *asterisks*. Homologous genes are marked by matching colors.



In Vibrionales, chito-oligosaccharides are known to enter the periplasmic space via a specific chitoporin encoded by the (Glc- $NAc)_{2}$ -induced gene *chiP* (11, 41), a predicted member of the NagC regulon. Based on the genome context analysis, we predicted the involvement of two other types of outer membrane proteins, ChiP-II and ChiP-III, in the uptake of GlcNAc and/or its oligosaccharides.

(i) ChiP-II belongs to a family of TonB-dependent outer membrane receptors, many of which are known to mediate high affinity binding and energy-dependent uptake of specific substrates into the periplasm (42). The chiP-II gene was identified as a member of the NagR regulon in Shewanella spp. (SO3514 in S. oneidensis) and other Alteromonadales and Xanthomonadales. It is also positionally clustered with hex and genes of the NAG pathway (Table 1 and Fig. 2). Some species (such as Xanthomonas and Caulobacter) contain multiple copies of the chiP-II genes that are mostly controlled by NAG-dependent regulators. One of 67 hypothetical TonB-dependent receptors encoded in the C. crescentus genome was shown to be involved in the TonB-ExbBD-dependent uptake of maltodextrins (43). This allowed us to suggest a similar mechanism for the ChiP-II-mediated uptake of chito-oligosaccharides.

(ii) A hypothetical outer membrane porin (named here ChiP-III) was mapped in the NagC regulons of 10  $\gamma$ -proteobacteria (Tables S1 and S2). ChiP-III belongs to the OprD porin family, many representatives of which are involved in uptake of various nutrients through the outer membrane in Pseudomonas aeruginosa (44). In contrast, the previously characterized chitoporin ChiP from Vibrio spp. belongs to the OmpC porin family. The conserved co-localization of the chiP-III and hex genes in the chromosomes of the C. violaceum, Yersinia, and Serratia species suggests involvement of ChiP-III in the outer membrane uptake of oligosaccharides that are substrates of the periplasmic  $\beta$ -N-acetylglucosaminidase Hex. In non-chitinolytic Escherichia and Salmonella species that have no Hex hydrolase, ChiP-III could be involved in uptake of chitobiose through the outer membrane supplying the PTS transporter for this disaccharide.

#### GlcNAc Transport across the Inner Membrane

PTS System—The GlcNAc-specific component NagE of the PTS transport system associated with the NAG pathway was described in *E. coli*. This signature gene is a conserved member of nag operons and respective regulons in Enterobacteriales and Vibrionales and in some  $\alpha$ - and  $\beta$ -proteobacteria (Table 1 and Table S1). Another committed PTS system for transport of chitobiose in *E. coli* and other enterobacteria is a part of the *chb* operon controlled by NagC and an additional regulator ChbR (16, 45).

ABC Transport System-An alternative system of GlcNAc transport via a committed ABC cassette was originally described in Streptomyces olivaceoviridis (46). An analogous ABC system is clustered on the chromosome with other nag genes in Silicibacter spp. and Rhizobiales (e.g. mll4771mll4779 in Mesorhizobium loti). These species appear to have a functional variant of the NAG pathway but lacking any other candidate genes for GlcNAc transport, and this phyletic distribution supports the suggested tentative specificity assignment

of the identified ABC cassette. On the other hand, Vibrionales use an ABC-type transporter (VC0620-VC0616 in V. cholerae) for the uptake of chitobiose (10), which is induced by chitooligosaccharides (11) without being a member of the NagC regulon.

Novel Permease NagP-A novel GlcNAc transporter, named here nagP (SO3503 in S. oneidensis), was tentatively identified in the genomes of Alteromonadales and Xanthomonadales. The NagP proteins contain 12 predicted transmembrane segments and are homologous to the fucose permease FucP from E. coli (~25% identity). This functional assignment is supported by clustering on the chromosome and by predicted coregulation (via upstream NagR-binding sites) with other nag genes (Table 1 and Fig. 2). The presence of NagP orthologs only in species containing a functional variant of the NAG pathway (see below) but lacking GlcNAc-specific PTS or ABC systems is more evidence in favor of the NagP proposed functional role. Another family of uncharacterized membrane proteins with 11 predicted transmembrane segments, named here NagX (SO3504 in S. oneidensis), may also be involved in uptake and transport of GlcNAc or other products of chitin degradation. In addition to the evidence from operons and regulons (Table 1 and Fig. 2), NagX has a perfect co-occurrence profile with ChiP-III as well as good correlation with occurrence of NagP.

#### Three-step NAG Pathway

Step 1: From GlcNAc to GlcNAc-6-P-In enterobacteria and some other lineages of proteobacteria, GlcNAc uptake and phosphorylation are coupled to a PTS transport system containing a GlcNAc-specific NagE component (Fig. 1). The absence of nagE orthologs in many other microbial genomes (including Shewanella species) presumably containing the NAG pathway and alternative (nonphosphorylating) transporters pointed to the existence of a specific GlcNAc kinase (EC 2.7.1.59). The S. oneidensis gene SO3507, which is present in the conserved chromosomal cluster containing the nag genes in Alteromonadales and some  $\alpha$ -proteobacteria (Fig. 2), encodes a motif typical for the ATP-binding domain of sugar kinases (PFAM domain PF01869). Therefore, it was deemed a candidate for the missing GlcNAc kinase role (named here NagK-I). This functional prediction was supported by significant sequence similarity ( $\sim$ 24%) with the recently characterized mammalian Glc-NAc kinase (47). Although the observed homology alone would not be sufficient to assign the sugar specificity of the respective bacterial enzymes, taken together with the genome and functional context analysis, it provided strong evidence, sufficient for experimental testing as described below.

In five analyzed Xanthomonadales species having a similar variant of the subsystem (supplemental Table S1), NagK-I appears to be functionally replaced by a nonorthologous enzyme from the glucokinase family (PFAM domain PF02685). A candidate gene for this novel GlcNAc kinase (XF1460 in Xylella fastidiosa; named here NagK-II) was identified using the genome context evidence as follows: chromosomal clustering, assignment to the NagR regulon, and a characteristic phyletic occurrence profile (Table 1 and Fig. 2).

The Journal of Biological Chemistry

Step 2: From GlcNAc-6-P to GlcN-6-P—GlcNAc-6-P deacetylase (EC 3.5.1.25), NagA, is the only invariant component of the NAG pathway. Orthologs of this enzyme, originally characterized in *E. coli* (48), are conserved over a large phylogenetic distance from bacteria to mammals (*e.g.* NP\_057028 in *Homo sapiens*). In bacteria, *nagA* is commonly clustered on the chromosome with other genes associated with the NAG pathway (Table 1), making it a nearly perfect signature gene. However, the presence of NagA is necessary but not sufficient to infer the complete pathway. For example, in some bacteria (*e.g. Haemophilus influenzae*), NagA is involved in sialic acid metabolism, where GlcNAc-6-P is supplied by epimerization of Man-NAc-6-P by NanE.

Step 3: From GlcN-6-P to Fru-6-P—Glucosamine-6-phosphate deaminase (EC 3.5.99.6) is required for the last committed step of the NAG pathway, simultaneous isomerization and deamination of GlcN-6-P yielding Fru-6-P, a key intermediate in the central carbon metabolism of many species. This enzyme, a product of the *nagB* gene in *E. coli*, was characterized in detail (49), and its orthologs are present in many bacteria and mammals (*e.g.* human gene GNPDA, NP\_005462). However, analysis of the GlcNAc utilization subsystem showed that this gene is missing in *S. oneidensis* as well as in most of other proteobacteria except Enterobacteriales and Vibrionales.

We have tentatively identified a candidate gene for an alternative GlcN-6-P deaminase (termed here NagB-II to distinguish from the orthologs of E. coli NagB, which we will further refer to as NagB-I) based on its chromosomal clustering and occurrence profile (Table 1 and Fig. 2). NagB-II (SO3506 in S. oneidensis) belongs to the sugar isomerase protein family, and it is most similar to the C-terminal domain of GlcN-6-P synthase (GlmS in E. coli). The latter catalyzes the first committed step in the biosynthesis of UDP-N-acetylmuramate from Fru-6-P (via GlcNAc-6-P), providing a critical intermediate in the biogenesis of peptidoglycan in most bacteria (50). Although this step, glutamine-dependent transamination-isomerization of Fru-6-P to GlcN-6-P, is formally a reversal of a catabolic reaction catalyzed by NagB-I, these enzymes are structurally and mechanistically unrelated to each other. On the other hand, NagB-II shares the C-terminal domain with GlmS but does not contain an equivalent of its N-terminal domain responsible for the utilization of glutamine as a nitrogen source and strictly required for the biosynthetic activity of GlmS. Based on detailed mechanistic studies of GlmS and its individual domains (51), it was reasonable to expect that NagB-II may have a catabolic, deaminase-isomerase activity (as in NagB-I) but not the biosynthetic activity of GlmS. This conjecture was captured earlier by the on-line annotation of a NagB-II ortholog from Thermotoga maritima (TM0813) proposed in the context of structural studies at the Joint Center for Structural Genomics (Protein Data Bank code 1J5X). When this study was nearly complete, another group published the experimental characterization of an NagB-II ortholog from the hyperthermophilic archaeon Thermococcus kodakaraensis, where it is encoded within the chitin utilization gene cluster (52). A functional context (pathway) of this enzyme in Archaea is different from the bacterial NAG pathway and has yet to be elucidated. In S. oneidensis and other bacteria, the nagB-II gene is consistently clustered with



FIGURE 3. **Purified recombinant enzymes of NAG pathway of** *S. oneidensis.* Proteins (2 µg each) purified by Ni-NTA mini-column. *Lane 1*, SO3505, GlcNAc-6-P deacetylase (NagA); *lane 2*, SO3506, predicted GlcN-6-P deaminase (NagB-II); *lane 3*, SO3507, GlcNAc kinase (NagK-I).

*nagA* and other *nag* genes (Fig. 2). Remarkably, in a single strain of *Shewanella* species (PV-4) and in *C. psychrerythraea, nagB-II* is functionally replaced by a standalone *nagB-I* gene preceded by a candidate NagR-binding site (Fig. 2).

In summary, we used the subsystem reconstruction and the genome context analysis to infer the existence of a novel functional variant of the NAG pathway of *S. oneidensis* and a number of other proteobacteria. In the second part of this study, we have performed experimental validation of the predicted pathway and characterized the enzymes involved in this pathway, focusing on its novel components, NagK-I and NagB-II, nonhomologous to known *E. coli* genes with the same function.

#### Experimental Characterization of S. oneidensis NAG Pathway

To validate a novel variant of the NAG pathway and to characterize its individual components, three genes of the *S. oneidensis* NAG operon, *SO3507* (*nagK-II*), *SO3505* (*nagA*), and *SO3506* (*nagB-II*), were cloned in an expression vector with an N-terminal His<sub>6</sub> tag and overexpressed in *E. coli* BL21/DE3. Recombinant proteins were partially purified (>90% by SDS-PAGE; see Fig. 3) using a rapid Ni-NTA mini-column protocol.

#### Properties of Individual Enzymes

Expected enzymatic activities of all three proteins were verified and characterized using the specific assays described under "Experimental Procedures." The steady-state kinetic parameters are provided in Table 3.

The Journal of Biological Chemistry

ibc

ch

#### TABLE 3

Kinetic constants for GlcNAc kinase, GlcNAc-6P deacetylase, and GlcN-6P deaminase

	Substrate	$k_{\rm cat}$	$K_m$	$k_{\rm cat}/K_m$
		$s^{-1}$	тм	$M^{-1} s^{-1}$
se	GlcNAc	$14.1 \pm 0.6$	$0.24\pm0.01$	$6.0 imes10^4\pm0.4 imes10^4$
eacetylase	GlcNAc-6P	$12.0 \pm 0.9$	$0.43 \pm 0.01$	$2.8 imes10^4\pm0.2 imes10^4$
minase <sup>a</sup>	GlcN-6P	$0.9 \pm 0.1$	$0.35\pm0.01$	$2.5 imes10^3\pm0.2 imes10^3$
	se leacetylase minase <sup>a</sup>	Substrate           se         GlcNAc           leacetylase         GlcNAc-6P           minase <sup>a</sup> GlcN-6P	Substrate $k_{cat}$ seGlcNAc $14.1 \pm 0.6$ leacetylaseGlcNAc-6P $12.0 \pm 0.9$ minase <sup>a</sup> GlcN-6P $0.9 \pm 0.1$	Substrate $k_{cat}$ $K_m$ $s^{-1}$ $mM$ se         GlcNAc         14.1 ± 0.6         0.24 ± 0.01           leacetylase         GlcNAc-6P         12.0 ± 0.9         0.43 ± 0.01           minase <sup>a</sup> GlcN-6P         0.9 ± 0.1         0.35 ± 0.01

<sup>*a*</sup> Data were obtained in the presence of a saturating concentration (1 mM) of the allosteric effector GlcNAc-6P.



value of NagB-II (0.9  $s^{-1}$ ) remains much lower than the reported  $k_{cat}$ value of NagB-I (168 s<sup>-1</sup>) (49). The overall catalytic efficiency of NagB-II ( $k_{cat}/K_m$ ) is 1–1.5 orders of magnitude lower compared with NagK-I and NagA of S. oneidensis (Table 3), suggesting that (given the expected comparable expression levels of all three enzymes encoded within a single operon) the former may be a rate-limiting step of the NAG pathway in this organism. A similar conclusion was derived from the in vitro pathway reconstitution experiments described below.

NagB-I (0.55 mM), although the  $k_{cat}$ 

FIGURE 4. **Allosteric activation of** *S. oneidensis* **GlcN-6-P deaminase by GlcNAc-6-P.** *A*, plot of initial velocity against substrate concentration obtained at different GlcNAc-6-P concentrations. Data were fit to the Hill equation. GlcNAc-6-P concentrations were as follows: 0 ( $\bigcirc$ ), 0.05 mM ( $\square$ ), 0.2 mM ( $\blacksquare$ ), and 1 mM ( $\blacksquare$ ). *B*, apparent steady-state parameters  $k_{cat}$  ( $\blacksquare$ ) and  $K_m$  values ( $\bigcirc$ ) for GlcN-6-P plotted as a function of GlcNAc-6-P concentration.

GlcNAc kinase NagK-II, the proposed first enzyme of the NAG pathway in *S. oneidensis*, has kinetic parameters similar to those of its recently described eukaryotic homologs. The  $k_{cat}$  of GlcNAc kinase is 14.1 s<sup>-1</sup> (Table 3). The apparent  $K_m$  value for GlcNAc (0.24 mM) at a saturating concentration of ATP (1.2 mM) is comparable with the respective values reported for the enzymes from *Candida albicans* (0.37 mM) or from mouse (0.20 mM) (47, 53). To our knowledge, this is the first documented report of bacterial GlcNAc kinase functionally replacing GlcNAc phosphorylation by PTS system, such as in the NAG pathway of *E. coli*.

GlcNAc-6-P deacetylase NagA of *S. oneidensis* catalyzing the second step of the NAG pathway, which is broadly conserved in bacteria and eukaryotes, has a  $K_m$  value (0.43 mM) similar to NagA from *E. coli* (0.4 mM) (35). The  $k_{cat}$  of Glc-NAc-6-P deacetylase (12.0 s<sup>-1</sup>) is comparable with that of GlcNAc kinase (Table 3).

GlcN-6-P deaminase NagB-II, a predicted last step of the S. oneidensis NAG pathway, displayed a surprisingly low specific activity in the initial tests. Consequently, we have shown that its activity can be strongly enhanced by the addition of an in-pathway intermediate GlcNAc-6-P. A further detailed analysis of this allosteric activation (Fig. 4) revealed that GlcNAc-6-P decreases the apparent K<sub>m</sub> value for GlcN-6-P without modifying the maximal velocity. Remarkably, a very similar allosteric regulation by GlcNAc-6-P was described previously for a nonhomologous counterpart of this enzyme (NagB-I) of E. coli (54). A conservation of the allosteric regulation in the absence of any notable homology at the sequence level between NagB-I and NagB-II is quite unexpected and argues for the functional importance of this regulatory mechanism. At a saturating concentration of GlcNAc-6-P, the apparent  $K_m$  values reach a similar level for both S. oneidensis NagB-II (0.35 mM) and E. coli

#### In Vitro Reconstitution of the NAG Pathway

We successfully validated the inferred S. oneidensis NAG pathway by in vitro reconstitution of the three-step biochemical conversion of GlcNAc to Fru-6-P. The reaction mixture contained GlcNAc and ATP and equimolar amounts of the three purified recombinant enzymes NagK-I, NagA, and NagB-II from S. oneidensis. The production of Fru-6-P was monitored by enzymatic coupling with two consecutive reactions catalyzed by phosphoglucose isomerase and glucose-6phosphate dehydrogenase and led to a chromogenic conversion of NADP to NADPH (Fig. 5A). The conversion of GlcNAc to Fru-6-P could be detected only in the presence of all three recombinant enzymes (Fig. 5B). In the same series of experiments, intermediates GlcNAc-6-P and GlcN-6-P could be converted to the final product Fru-6-P in the presence of two enzymes (NagA and NagB-II) and one enzyme (NagB-II), respectively, confirming the order of enzymatic steps in the NAG pathway (Fig. 5B).

To assess rate-controlling reactions in the GlcNAc utilization pathway, we determined the change in the overall flux (a net rate of Fru-6-P formation) resulting from a change in the concentration of each enzyme in the mixture (Fig. 6). The most significant concentration dependence of the overall flux was observed for NagB-II consistent with the possible (at least partially) rate-limiting nature of the last step of the NAG pathway.

# Phenotypic Characterization of GlcNAc Utilization in Shewanella spp.

Seven *Shewanella* strains were tested for the ability to respire GlcNAc, and all except *S. frigidimarina* NCIMB400 were positive for NAG respiration. The same seven strains were tested for ability to grow aerobically using either lac-



FIGURE 5. *In vitro* reconstitution of NAG pathway using purified *S. oneidensis* enzymes. *A*, the Fru-6-P formation was monitored by change in absorbance at 340 nm due to enzymatic coupling to NADP-NADPH conversion via phosphoglucose isomerase and glucose-6-phosphate dehydrogenase. *B*, all samples contained 1 mM of substrate and 2 µM of each enzyme. Formation of Fru-6-P was observed by incubating GlcNAc with all the three enzymes, GlcNAc-6-P with GlcNAc-6-P deacetylase and GlcN-6-P deaminase, and GlcN-6-P with GlcNAc-6-P deaminase, and the corresponding production rates of Fru-6-P were determined.



FIGURE 6. Evaluation of the possible rate-limiting steps in the *in vitro* reconstituted NAG pathway. Monitoring of the overall rate of conversion of GlcNAc to Fru-6-P was performed as described in Fig. 5. The initial value of the overall rate (flux) was obtained at 1:1:1 molar ratio of three enzymes. Consecutive measurements were performed while reducing the concentration of each enzyme (one at a time) to 50 and 25% of its initial value, and the concentration of the other two enzymes was kept constant. The observed overall flux was normalized to the initial value (at 1:1:1 ratio of all enzymes). All values are means  $\pm$  S.D. of three independent experiments.

tate, a universal substrate for most *Shewanella* strains, or GlcNAc as a sole carbon and energy source. The results of the phenotype screening showed that the cells of *S. oneidensis* MR-1, *S. putrefaciens* CN-32, *Shewanella* sp. PV-4, *S. denitrificans* OS-217, *S. baltica* OS155, and *S. amazonensis* SB2B grew equally well on both substrates achieving OD from 0.4 to 0.8. In contrast, *S. frigidimarina* NCIMB400 showed no growth on GlcNAc but grew well on lactate as a sole carbon and energy source. This observation is consistent with the absence of the GlcNAc utilization genes in the genome of *S. frigidimarina*.

In summary, the experimental results obtained in this study provide an experimental validation of the novel variant of the NAG pathway in *S. oneidensis* and reveal a role of GlcNAc utilization in biomass production and respiration in all strains of *Shewanella* with completely sequenced genomes with the notable exception of *S. frigidimarina*. Kinetic analysis of the individual enzymes of the pathway revealed a remarkable conservation of the allosteric regulatory mechanism for the last, possibly rate-limiting, step of the pathway.

The Journal of Biological Chemistry

#### DISCUSSION

The Journal of Biological Chemistry

 $\dot{b}c$ 

The comparative genomics analysis and metabolic reconstruction of the chitin and GlcNAc metabolic subsystem across a broad range of proteobacteria, including *S. oneidensis* and related species from the Alteromonodales family, revealed a remarkable pattern of conserved and variable features. The key conserved aspects of this subsystem are as follows.

1) An overall catabolic strategy is implemented in a number of common aspects of the subsystem topology (Fig. 1), including chemotaxis, depolymerization of chitin by an array of extracellular hydrolases, uptake of chito-oligosaccharides to the periplasm, further degradation to GlcNAc, and transport across the inner membrane feeding GlcNAc to the biochemical conversion to Fru-6-P.

2) The latter module, the NAG pathway, is present in both chitinolytic (*e.g. S. oneidensis* or *V. cholerae*) and nonchitinolytic (*e.g. E. coli* or *P. aeruginosa*) proteobacteria (Table 1). Within this module, one enzyme, GlcNAc-6-P deacetylase (NagA), appears to be invariantly conserved in all bacteria and eukaryotes containing one or another variant of the NAG pathway.

3) Expression of nearly all components of the subsystem is tightly coordinated and regulated as revealed by the extended conserved operons and regulons in all groups of proteobacteria.

4) An allosteric regulatory mechanism is remarkable conserved for both nonhomologous isozymes (NagB-I and NagB-II) involved with the last step of the NAG pathway.

At the same time, all these aspects are also associated with significant variations such as the following.

1) Different groups of species (and even closely related species) differ in their ability to utilize GlcNAc from a variety of natural reservoirs, such as chitin, chitobiose, and GlcNAc. Moreover, most of the pathogenic Pasteurellales (*e.g. H. influenzae*) use only the two last steps of the NAG pathway to catabolize GlcNAc-6-P produced by the sialic acid utilization pathway (55).

2) At least three nonorthologous types of transcription regulators appear to control the expression of the *nag* genes in various groups of proteobacteria. Among them, only NagC, characteristic of Enterobacteriales and Vibrionales species, was described previously. Additional regulators of the LacI type (named here NagR, characteristic of Alteromonadales, including *Shewanella* spp.) and of the GntR type (named here NagQ, characteristic of  $\alpha$ - and  $\beta$ -proteobacteria) were predicted in this study (Table 1 and Fig. 2).

3) A similarly high level of variations and nonorthologous displacements is observed for the components of transport machinery (Fig. 2). Most notably, the PTS-mediated transport of GlcNAc characteristic of Enterobacteriales and Vibrionales appears to be functionally replaced by a specific permease (NagP in *Shewanella* and other Alteromonadales and Xantomonadales) or an ABC cassette (as in some  $\alpha$ -proteobacteria) in conjunction with a GlcNAc kinase, see below.

4) Notably, all but one (NagA) enzymatic components of the most universal NAG pathway occur in a number of alternative forms. At least two nonorthologous kinases, NagK-I (as in *Shewanella* spp. and other Alteromonadales) and NagK-II (in

Xanthomonadales), were inferred by the genome context analysis in this study. The former, a distant homolog of the eukaryotic GlcNAc kinase, was experimentally validated in this study. Of note, yet another (presently unknown or "missing") GlcNAc kinase is expected to be present in some bacterial species beyond the order of Proteobacteria described in this study. For example, T. maritima contains an extended nag operon but lacks the PTS system as well as obvious orthologs of NagK-I or NagK-II.<sup>5</sup> Moreover, GlcNAc kinase activity was reported in crude extracts of *E. coli*, where it could be needed for the utilization of one molecule of GlcNAc, which is formed, along with one molecule of GlcNAc-6-P, as a by-product of the chitobiose utilization pathway (56). Although it is tempting to speculate that such activity could be displayed by NagC, a member of the ROK family that contains a number of established sugar kinases, we were unable to detect any GlcNAc activity for the recombinant NagC proteins from either E. coli or T. maritima (TM0808, clustered with the nag operon). Therefore, GlcNAc kinase remains a missing gene in some bacteria. Multiple nonorthologous displacements appear to be characteristic of many small molecule kinases, even beyond sugar-kinase families (57). For example, at least four nonorthologous forms of pantothenate kinase are associated with functional variants of the universal coenzyme A biosynthetic pathway (18).

5) The canonical form of the GlcN-6-P deaminase (NagB-I) present in Enterobacteriales and Vibrionales is conserved in eukaryotes but replaced by a nonorthologous isozyme NagB-II in *S. oneidensis* (characterized in this study) and all other proteobacteria, as well as in many distant groups of bacteria, including *T. maritima*.<sup>6</sup>

A subsystems-based approach (17) applied in this study to the analysis of chitin and GlcNAc utilization in S. oneidensis and related species has allowed us to significantly improve the quality of genomic annotations and to accurately infer metabolic and regulatory networks for relatively unexplored organisms. For example, before this study very little was known about the chitin utilization pathways in Shewanella species, although some of the strains were shown previously to have this catabolic potential, especially important in the aquatic environment enriched by this polysaccharide (9). Among Gram-negative bacteria, degradation and utilization of chitin have been studied in some species of Vibrionales, Alteromonas sp., S. degradans, and S. marcesens, whereas the intracellular GlcNAc catabolism was mostly characterized in a non-chitinolytic bacterium, E. coli. For example, many components of the chitin utilization system of V. cholerae were mapped previously by genome analysis, microarray expression, and biochemical studies (11). Another example is an insightful analysis of S. degradans, which revealed several secreted and periplasmic chitinolytic enzymes by initial genomic analysis and further enzymatic assays (13). We applied comparative genomics techniques, including the analysis of conserved operons and regulons, to

Downloaded from www.jbc.org at Burnham Institute for Medical Research LIBRARY on November 7, 2006

<sup>&</sup>lt;sup>5</sup> No GlcNAc kinase activity was detected for TM1280, a remote homolog of NagK-I from *T. maritima* (C. Yang and A. L. Osterman, unpublished data).

<sup>&</sup>lt;sup>6</sup> We have confirmed the respective enzymatic activity for TM0813 (C. Yang and A. L. Osterman, unpublished data).

project this knowledge to *Shewanella* species and other proteobacteria with completely sequenced genomes.

Among the important outcomes of this analysis is our ability to assert the existence and the biochemical details of the NAG pathway in all analyzed species. For example, all but one of the analyzed *Shewanella* species is capable of utilizing GlcNAc, as inferred by the genome analysis and confirmed by phenotype screening. Several strains (namely *S. oneidensis, S. baltica, S. denitrificans*, and *Shewanella* spp. ANA-3, MR-4, and MR-7) are expected to have the ability to utilize chitin or chito-oligosaccharides. Although all enzymes of the *S. oneidensis* NAG pathway were experimentally characterized in this study, many other functional predictions (Table 2) have yet to be challenged by focused experiments, which are now underway.

Acknowledgments—We are grateful to Ross Overbeek and other members of The SEED development and annotation team at Fellowship for Interpretation of Genomes. Pacific Northwest National Laboratory is operated for the Department of Energy by Battelle Memorial Institute under Contract DE-AC06–76RLO 1830.

#### REFERENCES

- 1. Myers, C. R., and Nealson, K. H. (1988) Science 240, 1319-1321
- Nealson, K. H., and Scott, J. (2003) in *The Prokaryotes: An Evolving Electronic Resource for the Microbiological Community* (Dworkin, M., ed) 3rd Ed., Springer-Verlag Inc., New York
- Fredrickson, J. K., and Romine, M. F. (2005) Curr. Opin. Biotechnol. 16, 269–274
- Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., and Wheeler, D. L. (2006) *Nucleic Acids Res.* 34, D16–D20
- Heidelberg, J. F., Paulsen, I. T., Nelson, K. E., Gaidos, E. J., Nelson, W. C., Read, T. D., Eisen, J. A., Seshadri, R., Ward, N., Methe, B., Clayton, R. A., Meyer, T., Tsapin, A., Scott, J., Beanan, M., Brinkac, L., Daugherty, S., DeBoy, R. T., Dodson, R. J., Durkin, A. S., Haft, D. H., Kolonay, J. F., Madupu, R., Peterson, J. D., Umayam, L. A., White, O., Wolf, A. M., Vamathevan, J., Weidman, J., Impraim, M., Lee, K., Berry, K., Lee, C., Mueller, J., Khouri, H., Gill, J., Utterback, T. R., McDonald, L. A., Feldblyum, T. V., Smith, H. O., Venter, J. C., Nealson, K. H., and Fraser, C. M. (2002) Nat. Biotechnol. 20, 1118–1123
- Qiu, X., Daly, M. J., Vasilenko, A., Omelchenko, M. V., Gaidamakova, E. K., Wu, L., Zhou, J., Sundin, G. W., and Tiedje, J. M. (2006) *J. Bacteriol.* 188, 1199–1204
- Liu, Y., Gao, W., Wang, Y., Wu, L., Liu, X., Yan, T., Alm, E., Arkin, A., Thompson, D. K., Fields, M. W., and Zhou, J. (2005) *J. Bacteriol.* 187, 2501–2507
- Alvarez-Anorve, L. I., Calcagno, M. L., and Plumbridge, J. (2005) J. Bacteriol. 187, 2974–2982
- Keyhani, N. O., and Roseman, S. (1999) Biochim. Biophys. Acta 1473, 108–122
- 10. Li, X., and Roseman, S. (2004) Proc. Natl. Acad. Sci. U. S. A. 101, 627-631
- Meibom, K. L., Li, X. B., Nielsen, A. T., Wu, C. Y., Roseman, S., and Schoolnik, G. K. (2004) Proc. Natl. Acad. Sci. U. S. A. 101, 2524–2529
- Orikoshi, H., Nakayama, S., Miyamoto, K., Hanato, C., Yasuda, M., Inamori, Y., and Tsujibo, H. (2005) *Appl. Environ. Microbiol.* 71, 1811–1815
- Howard, M. B., Ekborg, N. A., Taylor, L. E., Weiner, R. M., and Hutcheson, S. W. (2003) *J. Bacteriol.* 185, 3352–3360
- 14. Peri, K. G., Goldie, H., and Waygood, E. B. (1990) *Biochem. Cell Biol.* 68, 123–137
- 15. Vogler, A. P., and Lengeler, J. W. (1989) Mol. Gen. Genet. 219, 97-105
- 16. Plumbridge, J., and Pellegrini, O. (2004) Mol. Microbiol. 52, 437-449
- Overbeek, R., Begley, T., Butler, R. M., Choudhuri, J. V., Chuang, H. Y., Cohoon, M., de Crecy-Lagard, V., Diaz, N., Disz, T., Edwards, R., Fonstein, M., Frank, E. D., Gerdes, S., Glass, E. M., Goesmann, A., Hanson, A., Iwata-Reuyl, D., Jensen, R., Jamshidi, N., Krause, L., Kubal, M., Larsen, N.,

Linke, B., McHardy, A. C., Meyer, F., Neuweger, H., Olsen, G., Olson, R., Osterman, A., Portnoy, V., Pusch, G. D., Rodionov, D. A., Ruckert, C., Steiner, J., Stevens, R., Thiele, I., Vassieva, O., Ye, Y., Zagnitko, O., and Vonstein, V. (2005) *Nucleic Acids Res.* **33**, 5691–5702

- Ye, Y., Osterman, A., Overbeek, R., and Godzik, A. (2005) *Bioinformatics* Suppl. 11, **21**, 478 – 486
- 19. Osterman, A., and Overbeek, R. (2003) Curr. Opin. Chem. Biol. 7, 238-251
- Mironov, A. A., Vinokurova, N. P., and Gelfand, M. S. (2000) *Mol. Biol.* (*Moscow*) 34, 222–231
- 21. Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) *Genome Res.* **14**, 1188–1190
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997) Nucleic Acids Res. 25, 4876 – 4882
- McGinnis, S., and Madden, T. L. (2004) Nucleic Acids Res. 32, W20-W25
- Finn, R. D., Mistry, J., Schuster-Bockler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R., Eddy, S. R., Sonnhammer, E. L., and Bateman, A. (2006) *Nucleic Acids Res.* 34, D247–D251
- Marchler-Bauer, A., Anderson, J. B., Cherukuri, P. F., DeWeese-Scott, C., Geer, L. Y., Gwadz, M., He, S., Hurwitz, D. I., Jackson, J. D., Ke, Z., Lanczycki, C., Liebert, C. A., Liu, C., Lu, F., Marchler, G. H., Mullokandov, M., Shoemaker, B. A., Simonyan, V., Song, J. S., Thiessen, P. A., Yamashita, R. A., Yin, J. J, Zhang, D., and Bryant, S. H. (2005) *Nucleic Acids Res.* 33, D192–D196
- 26. Hofmann, K., and Stoffel, W. (1993) Biol. Chem. Hoppe-Seyler 347, 166
- Gerdes, S. Y., Kurnasov, O. V., Shatalin, K., Polanuyer, B., Sloutsky, R., Vonstein, V., Overbeek, R., and Osterman, A. L. (2006) *J. Bacteriol.* 188, 3012–3023
- Gelfand, M. S., Koonin, E. V., and Mironov, A. A. (2000) Nucleic Acids Res. 28, 695–705
- 29. Plumbridge, J. (1995) EMBO J. 14, 3958-3965
- 30. Plumbridge, J. (2001) Nucleic Acids Res. 29, 506-514
- Mironov, A. A., Koonin, E. V., Roytberg, M. A., and Gelfand, M. S. (1999) Nucleic Acids Res. 27, 2981–2989
- Osterman, A., Grishin, N. V., Kinch, L. N., and Phillips, M. A. (1994) Biochemistry 33, 13662–13667
- Osterman, A. L., Lueder, D. V., Quick, M., Myers, D., Canagarajah, B. J., and Phillips, M. A. (1995) *Biochemistry* 34, 13431–13436
- 34. Datta, A. (1975) *Methods Enzymol.* **42**, 58–62
- 35. White, R. J., and Pasternak, C. A. (1975) Methods Enzymol. 41, 497-502
- 36. Myers, C. R., and Nealson, K. H. (1990) J. Bacteriol. 172, 6232-6238
- Titgemeyer, F., Reizer, J., Reizer, A., and Saier, M. H., Jr. (1994) *Microbiology* 140, 2349–2354
- Gelfand, M. S., and Laikova, O. N. (2003) in *Frontiers in Computational Genomics* (Koonin, E. V., and Galperin, M. Y., eds) pp. 195–216, Caister Academic Press, Wymondham, UK
- Rigali, S., Derouaux, A., Giannotta, F., and Dusart, J. (2002) J. Biol. Chem. 277, 12507–12515
- Rigali, S., Schlicht, M., Hoskisson, P., Nothaft, H., Merzbacher, M., Joris, B., and Titgemeyer, F. (2004) *Nucleic Acids Res.* 32, 3418–3426
- Keyhani, N. O., Li, X. B., and Roseman, S. (2000) J. Biol. Chem. 275, 33068–33076
- Chimento, D. P., Kadner, R. J., and Wiener, M. C. (2003) J. Mol. Biol. 332, 999–1014
- Neugebauer, H., Herrmann, C., Kammer, W., Schwarz, G., Nordheim, A., and Braun, V. (2005) *J. Bacteriol.* 187, 8300 – 8311
- 44. Tamber, S., Ochs, M. M., and Hancock, R. E. (2006) J. Bacteriol. 188, 45–54
- Uchiyama, T., Kaneko, R., Yamaguchi, J., Inoue, A., Yanagida, T., Nikaidou, N., Regue, M., and Watanabe, T. (2003) *J. Bacteriol.* 185, 1776–1782
- Xiao, X., Wang, F., Saito, A., Majka, J., Schlosser, A., and Schrempf, H. (2002) *Mol. Genet. Genomics* 267, 429–439
- Hinderlich, S., Berger, M., Schwarzkopf, M., Effertz, K., and Reutter, W. (2000) *Eur. J. Biochem.* 267, 3301–3308
- 48. Souza, J. M., Plumbridge, J. A., and Calcagno, M. L. (1997) Arch. Biochem.

Mast

The Journal of Biological Chemistry

F

Biophys. 340, 338-346

- 49. Bustos-Jaimes, I., Sosa-Peinado, A., Rudino-Pinera, E., Horjales, E., and Calcagno, M. L. (2002) J. Mol. Biol. 319, 183-189
- 50. Dutka-Malen, S., Mazodier, P., and Badet, B. (1988) Biochimie (Paris) 70, 287-290
- 51. Teplyakov, A., Obmolova, G., Badet, B., and Badet-Denisot, M. A. (2001) J. Mol. Biol. 313, 1093-1102
- 52. Tanaka, T., Takahashi, F., Fukui, T., Fujiwara, S., Atomi, H., and Imanaka, T. (2005) J. Bacteriol. 187, 7038-7044
- 53. Yamada-Okabe, T., Sakamori, Y., Mio, T., and Yamada-Okabe, H. (2000) Eur. J. Biochem. 268, 2498-2505
- 54. Calcagno, M., Campos, P. J., Mulliert, G., and Suastegui, J. (1984) Biochim. Biophys. Acta 787, 165-173
- 55. Vimr, E. R., Kalivoda, K. A., Deszo, E. L., and Steenbergen, S. M. (2004) Microbiol. Mol. Biol. Rev. 68, 132-153
- 56. Asensio, C., and Ruiz-Amil, M. (1966) Methods Enzymol. 9, 421-425
- 57. Cheek, S., Ginalski, K., Zhang, H., and Grishin, N.V. (2005) BMC Struct. *Biol.* 5, 6

