Transcriptional regulation of NAD metabolism in bacteria: genomic reconstruction of NiaR (YrxA) regulon

Dmitry A. Rodionov, Xiaoqing Li, Irina A. Rodionova, Chen Yang, Leonardo Sorci, Etienne Dervyn, Dariusz Martynowski, Hong Zhang, Mikhail S. Gelfand, Andrei L. Osterman

HAL Id: hal-02661572
https://hal.inrae.fr/hal-02661572

Submitted on 30 May 2020

Transcriptional regulation of NAD metabolism in bacteria: genomic reconstruction of NiaR (YrxA) regulon

Dmitry A. Rodionov[1,2,*], Xiaoqing Li[1], Irina A. Rodionova[1], Chen Yang[1], Leonardo Sorci[1], Etienne Dervyn[3], Dariusz Martynowski[4], Hong Zhang[4], Mikhail S. Gelfand[2] and Andrei L. Osterman[1,5]

[1]Burnham Institute for Medical Research, La Jolla, CA 92037, USA, [2]Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow 127994, Russia, [3]Génétique Microbienne, INRA, 78352 Jouy-en-Josas, France, [4]Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX 75390 and [5]Fellowship for Interpretation of Genomes, Burr Ridge, IL 60527, USA

ABSTRACT

A comparative genomic approach was used to reconstruct transcriptional regulation of NAD biosynthesis in bacteria containing orthologs of *Bacillus subtilis* gene *yrxA*, a previously identified niacin-responsive repressor of NAD *de novo* synthesis. Members of YrxA family (re-named here NiaR) are broadly conserved in the *Bacillus/Clostridium* group and in the deeply branching Fusobacteria and Thermotogales lineages. We analyzed upstream regions of genes associated with NAD biosynthesis to identify candidate NiaR-binding DNA motifs and assess the NiaR regulon content in these species. Representatives of the two distinct types of candidate NiaR-binding sites, characteristic of the Firmicutes and Thermotogales, were verified by an electrophoretic mobility shift assay. In addition to transcriptional control of the *nadABC* genes, the NiaR regulon in some species extends to niacin salvage (the *pncAB* genes) and includes uncharacterized membrane proteins possibly involved in niacin transport. The involvement in niacin uptake proposed for one of these proteins (re-named NiaP), encoded by the *B. subtilis* gene *yceI*, was experimentally verified. In addition to bacteria, members of the NiaP family are conserved in multicellular eukaryotes, including human, pointing to possible NaiP involvement in niacin utilization in these organisms. Overall, the analysis of the NiaR and NrtR regulons (described in the accompanying paper) revealed mechanisms of transcriptional regulation of NAD metabolism in nearly a hundred diverse bacteria.

INTRODUCTION

Nicotinamide adenine dinucleotide (NAD) plays an important role in cellular metabolism because it functions as a cofactor of numerous redox reactions and is utilized as a consumable substrate by several classes of NAD-dependent enzymes such as DNA ligase, protein deacetylase and a variety of ADP-ribosyltransferases (1–3). The biochemistry, enzymology and genetics of NAD metabolism were studied in a number of model species as reflected in numerous overviews [e.g. see (4–9)]. In many bacteria, NAD is synthesized *de novo* from aspartic acid by the consecutive action of L-aspartate oxidase NadB, quinolinate synthetase NadA, quinolinate phosphorybosyltransferase NadC, nicotinate mononucleotide (NaMN) adenylyltransferase NadD and NAD synthetase NadE (Figure 1A). Mononucleotide intermediates, NaMN and its amidated analog (NMN), can be generated via a variety of salvage/recycling pathways from nicotinamide (Nam), nicotinic acid (NA) or ribosyl nicotinamide (RNam). Both Nam and NA (often indiscriminately referred to as vitamin B_3 or niacin) can originate from exogenous sources via relatively poorly studied uptake mechanisms, as well as via the breakdown of endogenous NAD by NAD-consuming enzymes. The most widely distributed niacin salvage pathway is a two-step conversion of Nam to NaMN via nicotinamidase PncA and nicotinate phosphoribosyltransferase PncB. Alternatively, Nam can be converted to NMN by nicotinamide phosphoribosyltransferase NadV. Another salvage
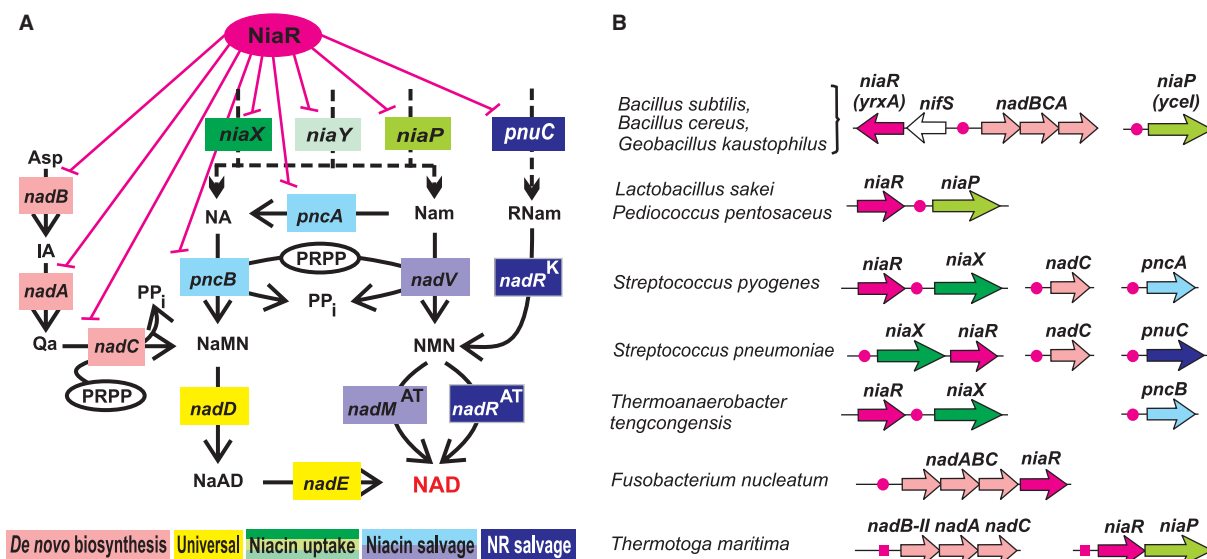

Figure 1. NAD biogenesis and its transcriptional regulation in bacteria containing NiaR regulator. (**A**) A simplified pathway diagram showing biochemical transformations (arrows) of precursors and intermediates (shown by the same abbreviations as in the text). Dashed arrows reflect uptake of exogenous precursors. Enzymes and transporters are shown using the respective gene names in boxes whose colors correspond to various pathways of NAD metabolism. Magenta-colored lines point to regulated genes (NrtR regulon members). (**B**) Genomic organization of the *niaR*-containing loci and NiaR-regulated genes in representative bacterial genomes. Genes are shown by arrows colored as in Figure 1A. Red circles and squares mark positions of candidate NiaR-binding sites of types I and II, respectively. (**C**) Sequence logos for the NiaR-binding motifs in the *Bacillus/Clostridium* group and the Thermotogales. The colors of arrows representing reactions in (A) and genes in (B) correspond to cellular functions listed in the legend.

pathway proceeding via NMN intermediate includes RNam uptake via the PnuC transporter and phosphorylation by the RNam kinase domain of the multifunctional NadR protein (10). The ultimate conversion of NMN intermediate to NAD is performed by the NMN adenylyltransferase domain of the NadR or NadM family [as in *Haemophilus influenzae* (11), or *Synechocystis* sp. (6), respectively].

In contrast with a rather detailed picture of NAD-related biosynthetic pathways in diverse bacteria, the knowledge of transcriptional regulation of these pathways until recently was limited to Enterobacteria. The NadR protein of *Escherichia coli* and *Salmonella* was shown to function as a NAD-responsive repressor of NAD *de novo* and salvage pathways (12–14). An optional N-terminal HTH-domain of a multifunctional NadR protein mediates specific recognition of a palindromic 18-bp operator with the consensus sequence TGTTTA-N$_6$-TAAACA. A comparative genomic analysis revealed conserved NadR-binding DNA motifs and allowed us to reconstruct the NadR regulon (including autoregulation of the *nadR* gene) in the Enterobacteria (15). NadR orthologs in bacterial lineages beyond the Enterobacteria lack the N-terminal DNA-binding domain, and they apparently function only as RNam salvage enzymes and not as transcriptional regulators.

A novel transcriptional regulator of NAD biosynthesis was recently discovered in *Bacillus subtilis* (16). This regulator encoded by the *yrxA* gene was shown to be an NA-responsive transcriptional repressor of two divergently transcribed operons, *nadBCA* and *nifS-yrxA*, with overlapping promoters (17). In this study, we used a

comparative genomic approach to reconstruct this novel regulon (re-named here NiaR, for <u>nia</u>cin <u>r</u>epressor) in many species of the *Bacillus/Clostridium* group and in the deep-branched Fusobacteria and Thermotogales lineages that contain orthologs of the *niaR* (*yrxA*) genes in their genomes. The approach combining metabolic reconstruction with analysis of genome context and conserved upstream regions was successfully applied by us and other research groups to expand known regulons to a broad range of related species as well as to predict regulons *ab initio* [for a recent review see (18)]. Moreover, such analysis generates tentative functional assignments of previously uncharacterized genes that in many cases have been already confirmed by further experimental studies. Reconstruction of regulons associated with metabolic pathways is a particularly efficient strategy for predicting new transporter specificities (19).

In this study we tentatively identified two distinct NiaR-binding DNA motifs: type I (with the TGT-N$_4$-ACA core) characteristic of the entire *Bacillus/Clostridium* group and the Fusobacteria lineage; and type II (with an ACA-N$_5$-TGT core), found exclusively in Thermotogales. Predicted DNA sites were experimentally verified by electrophoretic mobility shift assays (EMSAs) for two sites of each type I (in *B. subtilis*) and II (in *Thermotoga maritima*) genomes, using purified recombinant proteins YrxA (*bs*NiaR) and TM1602 (*tm*NiaR) respectively. In addition to NAD *de novo* synthesis genes, the NiaR regulon in these divergent species includes conserved homologous genes encoding uncharacterized membrane proteins (YceI in *B. subtilis* and TM1603 in *T. maritima*) of the Major Facilitator Superfamily (MFS). The hypothesized role of this

subfamily of transporters (named NiaP) in niacin uptake was confirmed by a series of growth experiments using *B. subtilis yceI* knockout mutant and recombinant strains of *E. coli* overexpressing this gene. Overall, the results of this study provide insights in two relatively poorly explored aspects of bacterial NAD metabolism, transcriptional regulation and vitamin B_3 transport.

MATERIALS AND METHODS

Bioinformatic analysis of bacterial genomes and reconstruction of NiaR regulon

Bacterial genome sequences were taken mainly from GenBank (20). Some draft genomes were downloaded from the web sites of J. Craig Venter Institute (http://www.tigr.org), the Wellcome Trust Sanger Institute (http://www.sanger.ac.uk) and the DOE Joint Genome Institute (http://jgi.doe.gov). Functional annotations of genes involved in NAD metabolism and related pathways in bacterial genomes were from the collection of metabolic subsystems in the SEED comparative genomic database (21). Functional coupling of genes via clustering on the chromosome and distribution of genes in the genomes were analyzed using the integrated SEED tools. Multiple sequence alignments of protein sequences were produced by the Clustal series of program (22). The PHYLIP package was used for construction of maximum likelihood phylogenetic tree for the NiaR and NiaP protein families including bootstrapping with 100 replicates and drawing of a consensus tree (23). The Protein Families database (http://www.sanger.ac.uk/Software/Pfam/) (24) and the Clusters of Orthologous Groups database (http://www.ncbi.nlm.nih.gov/COG/) (25) were used to identify conserved functional domains. Transmembrane segments (TMSs) in proteins were predicted using TMpred (http://www.ch.embnet.org/software/TMPRED_form.html).

To identify candidate DNA-binding motifs of NiaR(YrxA) regulators, we started from a set of upstream regions of potentially coregulated genes in the *Bacillus*/*Clostridium* group. The training set included potential regulatory regions of genes implicated in NAD biosynthesis and/or genes positionally clustered with the *niaR* genes. An iterative motif detection procedure implemented in the program SignalX was used to identify a common regulatory DNA motif (type I) in the training set and to construct a motif recognition profile [for details see (26)]. The constructed recognition profile was used to scan a subset of the *Bacillus*/*Clostridium* genomes encoding NiaR orthologs. A divergent NiaR-binding motif (type II) characteristic of Thermotogales was determined by comparing upstream regions of the *nadB* and *yrxA* genes from four available *Thermotoga* genomes, *T. maritima*, *Thermotoga lettingae*, *Thermotoga neapolitana* and *Thermotoga petrophila*, and the DNA-recognition profile was constructed using the SignalX program. Positional nucleotide weights in the recognition profile and *Z*-scores of candidate sites were calculated as the sum of the respective positional nucleotide weights as previously described ref. (27). Genome scanning for additional candidate NiaR-binding sites was performed using the

Genome Explorer software (28). The threshold for the site search was defined as the lowest score observed in the training set. Sequence logos for the derived DNA-binding motifs were constructed using the WebLogo package v.2.6 (http://weblogo.berkeley.edu/) (29).

Expression and purification of recombinant NiaR proteins

The gene *yrxA* was PCR-amplified from *B. subtilis* genomic DNA using the primers: 5′-ctcggatccTTGACCGAAGAATTAAAGCTAATGGGCGCC (forward) and 5′-ctcctcgagTTAATTAAAAATGCCGGCTTCTTCGAGGGC (reverse). This gene was then cloned into the pHisparellel1 vector (30) between the BamHI and XhoI sites. The encoded recombinant protein (*bs*NiaR) was overexpressed in *E. coli* BL21(DE3) in the LB medium in shake flasks with induction by 0.8 mM IPTG at $OD_{600}\sim0.6$ followed by overnight shaking at 18°C, 175 r.p.m. The cells were lyzed, and the protein was extracted by 20 mM HEPES pH 7.2, 100 mM NaCl, 0.03% Brij35 buffer containing protease inhibitor cocktail (Sigma) and purified by Ni–NTA column chromatography. The His_6-tag was removed by treatment with TEV protease overnight at 4°C in a dialysis bag (50 mM HEPES pH. 8.0 50 mM NaCl buffer). The processed bsNiaR was recovered from the flow-through of the second Ni–NTA column, concentrated and further purified by gel-filtration on a Superdex75 16/60 column using the AKTA FPLC system. The protein eluted as a dimer with the apparent $M_r\sim43$ kDa. It was homogenous by SDS–PAGE (Figure S1 in Supplementary Data available online) and formed diffractable crystals that were further used for structural studies (Zhang,H., unpublished data).

An *E. coli* DL41 strain carrying *T. maritima* gene encoding *tm*NiaR (*TM1602*) in pMH2T7 vector under arabinose-inducible promoter (31) was a kind gift from Dr Scott Lesley at the Joint Center for Structural Genomics (www.jcsg.org). Protein expression was performed in TB media (24 g/l yeast extract, 12 g/l tryptone) containing 1% glycerol and 50 mM Mops, pH 7.6. The cells were induced by 0.15% arabinose and harvested after 3 h shaking at 37°C. Purification of the recombinant *tm*NiaR containing an N-terminal His_6-tag was performed using rapid Ni–NTA agarose (Qiagen) minicolumn protocol as described (32). Briefly, harvested cells were resuspended in 20 mM HEPES buffer pH 7 containing 100 mM NaCl, 0.03% Brij 35 and 2 mM β-mercaptoethanol supplemented with 2 mM phenylmethylsulfonyl fluoride and a protease inhibitor cocktail (Sigma–Aldrich). Lysozyme was added to 1 mg/ml, and the cells were lyzed by freezing–thawing followed by sonication. After centrifugation at 18 000 r.p.m., the Tris–HCl buffer (pH 8) was added to the supernatant (50 mM, final concentration), and it was loaded onto a Ni–NTA 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. The buffer was then changed to 20 mM HEPES containing 2 mM DTT, 0.5 mM EDTA and 150 mM NaCl by using Bio-Spin columns (Bio-Rad). This simplified procedure allowed us to obtain protein of >90% purity

by SDS–PAGE (Supplementary Figure S1) in an amount sufficient (∼2 mg) for further functional studies.

Analysis of NiaR–DNA interactions by EMSA

Interaction of purified recombinant proteins *bs*NiaR and *tm*NiaR with their cognate DNA motifs was assessed by the EMSA technique using the following dsDNA segments obtained by PCR amplification or by custom synthesis of both complementary strands, annealing and purification (Allele Inc., CA, USA), as indicated below. The PCR products were purified with the PCR purification kit (Promega, WI, USA), and their concentration was determined spectrophotometrically (Nanodrop ND-1000, DE, USA).

(i) *bs_nadB_87*, an 87-bp segment of *B. subtilis* from the upstream region of *nadB* gene, identical to the one experimentally verified in previously published study (16), was obtained by PCR amplification with primers 5′-CCTCCTGTTGTTTACACCTGTCT (forward) and 5′-CTTCCATCCGTTCTCCATAAAA (reverse) and used as a positive control of *bs*NiaR binding;

(ii) *bs_niaP_87*, a synthetic 87-bp fragment corresponding to positions (−13 to −99) in the upstream region of *B. subtilis niaP* gene (*yceI*), was used to test *bs*NiaR binding;

(iii) *tm_nadB_68*, a synthetic 68-bp DNA segment corresponding to positions (−3 to −65) in the upstream region of *T. maritima nadB* gene (TM1643), was used to test *tm*NiaR binding;

(iv) *tm_niaR_68*, a synthetic 68-bp DNA segment corresponding to positions (−6 to −62) in the upstream region of *T. maritima niaR* gene (TM1602), was used to test *tm*NiaR binding;

(v) *tm_pncA_111*, a 111-bp segment from the upstream region of *T. maritima* ortholog of *pncA* gene (*TM0475*), was obtained by PCR amplification using *T. maritima* MSB8 genomic DNA and primers 5′-GCTAAGGAGCCGGTCTTGAA (forward) and 5′-CACGTTTTCACCCCCTAAAA (reverse). This fragment lacking a NiaR-binding site was used as negative control (nonspecific competitor) in all binding studies. A size difference between this and other fragments was sufficient to monitor differential binding and gel-shift in a mixture of two DNA fragments without labeling (see below).

A typical reaction mixture for binding studies contained two DNA fragments (including the negative control *tm_pncA_111*), each at 20 nM, in a total volume of 10 μl of 20 mM Tris–HCl (pH 8.0), 5 mM MgCl$_2$, 100 mM KCl, 10% (v/v) glycerol, 0.5 mM DTT with or without 0.3 mM NA [as in (16)]. Recombinant proteins corresponding to the tested DNA segments were added: *bs*NiaR (20 μM) or *tm*NiaR (in a range 0–40 μM), and incubated for 30 min at ambient temperature or at 75°C, respectively. The complex formation was monitored by the electorphoresis in 10% PAAG (with or without 0.3 mM NA) at 90 V using 0.5× TBE buffer (45 mM Tris base, 45 mM boric acid, 10 mM EDTA). The gels were subsequently stained with Syber Green I according to the instructions of the supplier (Sigma–Aldrich, USA).

Testing the proposed role of *bs*NiaP in niacin uptake by growth experiments

Cloning of B. subtilis gene yceI (niaP). The gene was PCR-amplified from *B. subtilis* genomic DNA using the primers 5′-gaagaccATGGGAAAACAACAGCC (forward), and 5′-cCTATTCTAGCTCAGTTTGTTTCGT TTC (reverse) and cloned in pBAD TOPO TA expression vector (Invitrogen) under control of the arabinose-inducible promoter.

Growth studies of E. coli (nadA$^-$) overexpressing the B. subtilis gene yceI on limiting concentrations of Nam. The expression plasmid described above was transformed into the *E. coli* BW25113 (ΔnadA) knockout mutant with disrupted NAD *de novo* synthesis pathway from the Keio collection (a kind gift by Dr H. Mori, Keio University, Japan) (33). The same host strain containing the expression vector without any insert was used as a negative control. Growth curves of both strains were obtained in the minimal M9 growth media containing 20 mM glucose, 50 mg/l ampicillin and 0.15% of L-arabinose and supplemented with limiting concentrations of Nam. Starter cultures were grown overnight in LB medium. Cells were harvested, washed resuspended in the medium (OD$_{600}$∼0.03), and used in equal amounts to start growth experiments in a series of duplicates with increasing concentrations of Nam in a range of 0–0.8 μM. The growth was monitored by OD$_{600}$ measurement using a microplate reader DTX800 (Beckman).

Growth studies of E. coli (nadA$^+$) overexpressing the B. subtilis gene yceI in the presence of 6-Amino-Nam. The same expression plasmid was transformed into the *E. coli* BW25113 (ΔxylF) knockout mutant from the Keio collection. This host strain was chosen to have an undisrupted *de novo* NAD pathway in a similar genetic background as in the Nam experiments. The same *E. coli* strain containing the expression vector without any insert was used as a negative control. Growth curves for both strains were obtained using the same approach and in the same medium without Nam in a series of increasing concentrations of its toxic analog, 6-Amino-Nam (6ANam) in the range of 0–150 μM.

Growth studies of B. subtilis yceI knockout strain in the presence of 6ANam. Growth curves were obtained for *B. subtilis (yceI$^-$)* and compared to *B. subtilis (ypaA$^-$)* strain used as an isogenic negative control. Both knockout strains were from the collection constructed by the joint effort of the Japanese and European Bacillus Subtilis Functional Analysis programs. They were obtained by a standard single crossover-based protocol using PCR-amplified fragments of target genes cloned in pMUTIN2 vector as previously described (34). Both strains were grown at 37°C in chemically defined medium containing glucose (4 g/l), tryptophan (50 mg/l), glutamine (2 g/l), K$_2$HPO$_4$ (10 g/l), KH$_2$PO$_4$ (6 g/l), sodium citrate (1 g/l), MgSO$_4$ (0.2 g/l), K$_2$SO$_4$ (2 g/l), FeCl$_3$ (4 mg/l) and MnSO$_4$ (0.2 mg/l) overnight in the presence of 0.5 mg/l erythromycin (pMUTIN2 marker). These cultures were diluted ∼10–20-fold to the same cell density (OD$_{600}$ = 0.05) in the

Table 1. Occurrence and features of genes involved in NAD metabolism in representative bacterial genomes

Genome	NiaR (YrxA)	De novo synthesis			Universal		Niacin salvage			Uptake Transporters			
		NadB	NadA	NadC	NadD	NadE	PncA	PncB	NadR	PnuC	NiaP	NiaX	NiaY
• Bacillales													
Bacillus subtilis, B.licheniformis	⊕	⊕	⊕	⊕	+	+	+	+		⊕			
Bacillus cereus, G.kaustophilus	⊕	⊕	⊕	⊕	+	+	+	+		⊕			
Bacillus halodurans, B. clausii	⊕	⊕	⊕	⊕	+	+	+	+					⊕
Exiguobacterium sp.	⊕	⊕	⊕	⊕	+	+	+	+					⊕
Listeria monocytogenes	⊕	⊕	⊕	⊕	+	+	+	+					
Oceanobacillus iheyensis	0				+	+	+	+					
Staphylococcus aureus	0				+	+	+	+					
• Lactobacillales													
Enterococcus faecalis	+				+	+	+	+				⊕	
Lactobacillus plantarum	⊕				+	+	+	+ ⊕		+	+		
Lactobacillus casei	+				+	+	+	+			⊕		
L.sakei, Pediococcus pentosaceus	⊕				+	+	+	+			⊕		
L.acidophilus, L. johnsonii	0				+	+	+	+					
Leuconostoc mesenteroides	0				+	+		+	+	+	+		
• Streptococcaceae													
Lactococcus lactis	0				+ +	+	+	+	+			+	
Streptococcus agalactiae	+				+	+	+	+				⊕	
Streptococcus mutans	+				+	+	+	+		⊕		⊕	
Streptococcus pneumoniae	⊕			⊕	+	+	+	+		⊕		⊕	
S.thermophilus, S.suis	⊕				+	+	+	+	+			⊕	
Streptococcus pyogenes	+			⊕	+	+	⊕	+				⊕	
Streptococcus equi	+			⊕	+	+	⊕	+				⊕	
• Clostridiales													
Alkaliphilus metalliredigens	⊕	⊕	⊕	⊕	+	+	+	+					⊕
Clostridium acetobutylicum	+	⊕	⊕	⊕	+ +	+	+ +	+ +					
Clostridium difficile	+	⊕	⊕	⊕	+	+	+	+					⊕
C.botulinum, C.beijerincki	+	⊕	⊕	⊕	+	+	+	+				⊕	
Clostridium tetani	+				+	+	⊕	+				⊕	
Clostridium perfringens	0	+	+	+	+	+	+	+					
Thermoanaerobacter tengcongensis	+				+	+	+	⊕				⊕	
• Fusobacteria													
Fusobacterium nucleatum	⊕	⊕	⊕	⊕	+	+		+					+
• Thermotogales													
Thermotoga maritima	⊕	⊕II	⊕	⊕	+	+	+	+			⊕		
T. petrophila, T. neapolitana	⊕	⊕II	⊕	⊕	+	+	+	+			⊕		
Thermotoga lettingae	⊕				+	+	+	+			⊕		

The *Bacillus/Clostridium*, *Thermotogales* and *Fusobacteria* lineages that include species having the NiaR regulon are shown. The presence (+) or absence (0) of the NiaR regulator is denoted in the second column. Genes clustered on the chromosome are marked by the same background color. Candidate NiaR regulon members are circled. II denotes a nonorthologous substitution of the aspartate oxidase NadB by the aspartate dehydrogenase NadB-II in Thermotogales.

fresh medium and grown in duplicate series containing increasing concentrations of 6ANam $(0-75\,\mu M)$ in duplicates. A similar experiment was performed using another toxic niacin analog, 6-aminonicotinic acid (6ANA), instead of 6ANam.

RESULTS

Comparative genomic reconstruction of NiaR regulon in bacteria

The main goal of this study was to refine and expand the knowledge of transcriptional regulation of NAD synthesis from *B. subtilis*, where the function of NiaR (YrxA) as NA-responsive repressor was discovered (16), to all bacteria that encode NiaR orthologs in their genomes.

This was done using the subsystems-based approach to comparative genome analysis as supported by the SEED genomic annotation platform (21), in combination with specialized software tools for the identification of regulatory DNA motifs, SignalX and Genome Explorer (28).

Taxonomic distribution and structural features of the NiaR (YrxA) protein family. Members of the NiaR family show a mosaic distribution among Gram-positive bacteria from the *Bacillus/Clostridium* (Firmicutes) group (see Table 1, Supplementary Table S1 and 'NAD regulation' subsystem in the SEED database available online at http://theseed.uchicago.edu/FIG/index.cgi for more details). Among other taxonomic groups, NiaR was only found in two deeply diverged bacterial lineages, the Thermotogales and the Fusobacteria. Like several other
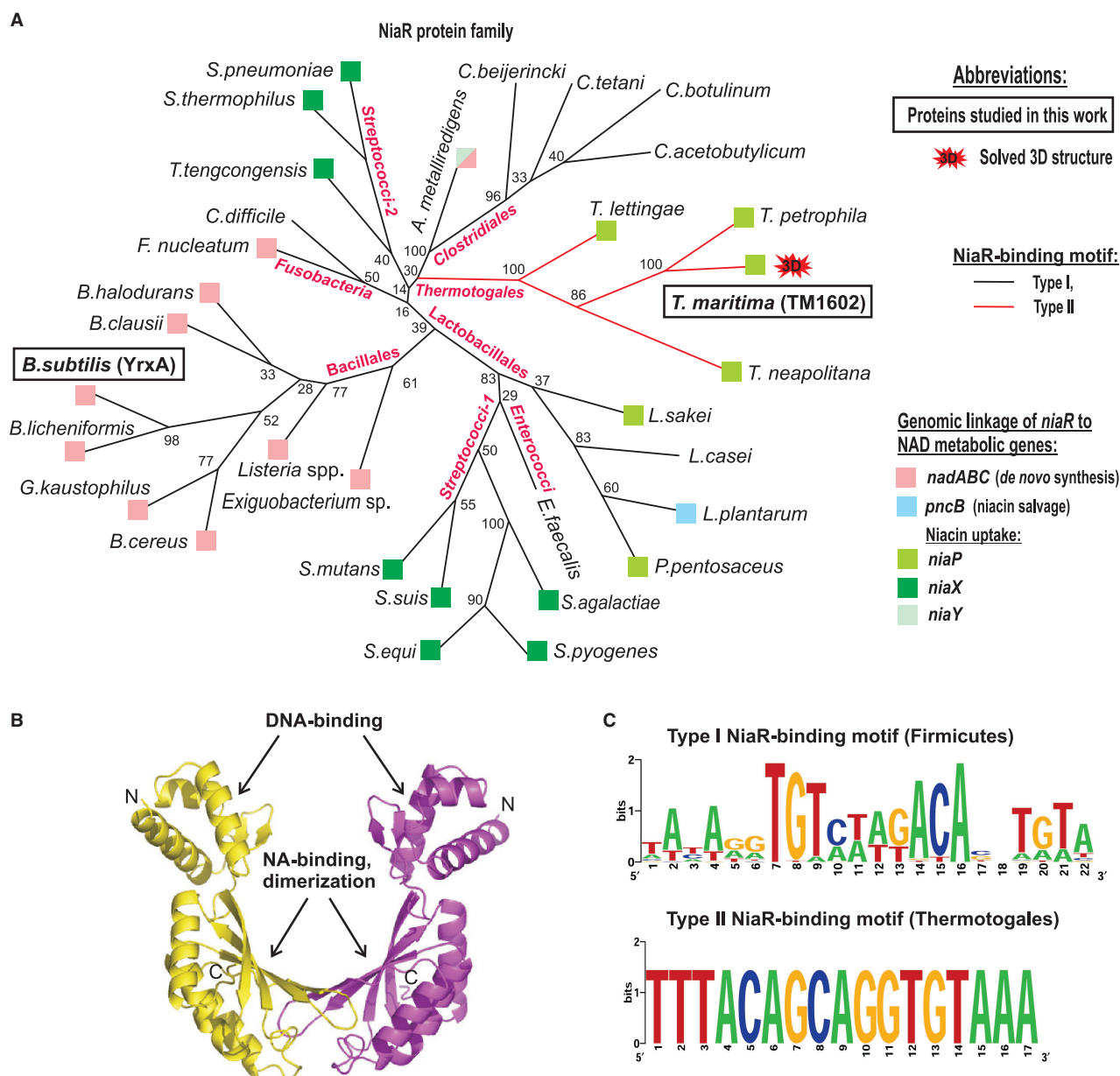
Figure 2. Structural and functional features of the NiaR family. (**A**) The maximum-likelihood tree of the NiaR family of transcriptional regulators. Genome context (chromosomal clustering) of the *niaR* genes is reflected by the color of squares with colors as in Figure 1A. The NiaR proteins from *T. maritima* and *B. subtilis* studied in this work are outlined. The NiaR proteins recognizing type I and type II DNA motifs are shown by black and red lines, respectively. The numbers indicate the number of bootstrap replications, out of 100, that support each node on the tree. (**B**) The arrangement of DNA-binding (HTH) and putative niacin-binding (3H) domains in the 3D structure *tm*NiaR (PDB accession number 1J5Y); (**C**) The NrtR recognition DNA motifs of type I and type II are shown by their respective sequence logos.

transcription factors, e.g. NrdR (35) and ArgR (36), regulators from the NiaR family always occur as a single copy pointing to a likely functional homogeneity of this family. Although a maximum-likelihood tree of the entire NiaR family is in good overall correlation with the phylogenetic tree of bacterial species, some deviations were observed (Figure 2A). For example, the NiaR proteins from two *Streptococcus* species, *Streptococcus thermophilus* and *Streptococcus pneumoniae*, branch together with Clostridiales, whereas all other Streptococci occur on a distinct Lactobacillales branch.

The first 3D structure from the NiaR family was solved and reported for *tm*NiaR, the TM1602 protein from *T. maritima* (PDB accession number 1J5Y) by the Joint Center for Structural Genomics [www.jcsg.org, (37)]. The TM1602 monomer consists of two structural domains (Figure 2B), an N-terminal domain similar to the 'winged helix' DNA-binding domain of the biotin repressor BirA, and a C-terminal 3H domain (so-called to reflect the conserved three-histidine signature), a widely distributed small-molecule-binding domain found in many signal transduction proteins and enzymes such as TrkA-module

proteins or 2-phosphoglycerate kinase (38). Significant amino acid sequence conservation of both domains, including three highly conserved histidine residues in the putative effector-binding 3H domain (His85, His153 and His155 in *bs*NiaR), supports the anticipated conservation of function within the entire NiaR family (see multiple alignment in Supplementary Figure S2). The analysis of genome context and the reconstruction of the NiaR regulons (see below) provided additional strong evidence of the proposed role of the NiaR proteins in transcriptional regulation of NAD synthesis in all analyzed bacterial species.

Genome context of the niaR genes. Clustering of genes on the prokaryotic chromosome, a major technique of genome context analysis, provides the key evidence of functional coupling of genes in pathways (39). Particularly, bacterial transcriptional factors often occur close on the chromosome to the genes that they regulate (40,41). Indeed, the *B. subtilis yrxA* (*niaR*) gene is located next to its target operon, *nadBAC*. The same chromosomal arrangement is retained in other Bacillales genomes except for the *Staphylococcus* and *Oceanobacillus* species that lack both *niaR* and the *de novo* NAD biosynthetic genes. Among *niaR*-containing species, the Lactobacillales and Streptococcaceae also lack the *de novo* NAD biosynthesis genes, and in many of them *niaR* is found clustered on the chromosome with the niacin salvage gene *pncB* and with hypothetical transporter genes (Figure 1B). In *Fusobacterium nucleatum* and *Alkaliphilus metalliredigens*, *niaR* is apparently cotranscribed with the upstream *nadABC* operon, whereas in the Thermotogales it forms an operon with an ortholog of the *B. subtilis yceI* (*niaP*) gene (tentatively assigned a role in niacin uptake, see below). This analysis further confirmed functional coupling between the NiaR regulator and NAD metabolism (Figure 2A) and defined a training gene set to analyze conserved upstream DNA sequences and to derive putative NiaR-binding motifs.

NiaR-binding sites and regulon members. A comparative genomic approach was efficiently applied for prediction of DNA-binding motifs for a variety of transcription factors in a broad range of bacterial genomes (18,26,27,35,36,42). Using this approach in this study, we tentatively identified NiaR-binding sites in 35 bacterial species with completely sequenced genomes (see Table 1 and Supplementary Table S2). For the Firmicutes group, the training set included upstream regions of the *nadABC* operon from 12 genomes. The SignalX program applied to this training set derived a candidate 22-bp NiaR-binding motif with the imperfect inverted repeat symmetry (Figure 2C). To assess significance of this motif (designated here type I) we defined the level of nucleotide conservation within the respective DNA segments in related bacteria by a phylogenetic footprinting method (43). Multiple alignments of the *nadABC* operon upstream regions from the *Bacillus* and *Clostridium* species revealed a high level of conservation in predicted NiaR-binding sites (Supplementary Figure S3). In *B. subtilis*, the predicted NiaR-binding site overlaps the −10 promoter box and the

transcriptional start site of the *nadB gene* (17), consistent with the proposed role of NiaR as a transcriptional repressor.

A recognition profile capturing the identified NiaR-binding sites was constructed and used to detect their additional occurrences and new members of the NiaR regulon in all analyzed genomes (see Supplementary Table S2). A threshold in the site scanning was set at the lowest score observed within the training set. In addition to the genes of the *de novo* pathway, candidate NiaR-binding sites were found upstream of genes involved in salvage of niacin (*pncA* and *pncB*) and uptake of RNam (*pnuC*). The second occurrence of an NiaR-binding site in the genome of *B. subtilis* and several other Bacillales was detected upstream of the uncharacterized gene *yceI* (*niaP*) encoding a transmembrane protein of the MFS transporter family. This provided additional support for the inferred role of the NiaP protein family in niacin utilization (see next section for a more detailed analysis). Due to the important functional implications of this inference, we chose to experimentally assess binding of *bs*NiaR to both DNA target sites identified in *B. subtilis*.

Two additional nonorthologous families of putative transporters were identified among NiaR regulon members in other genomes of the *Bacillus*/*Clostridium* group, represented by *BH3254* (named *niaY*) in *B. halodurans*, and Spy1425 (*niaX*) in *Streptococcus pyogenes* (see Table 1, Figure 1B and next section for a more detailed analysis). The identified candidate NiaR-binding sites show a significant conservation in closely related species (as illustrated by multiple alignments in Supplementary Figure S3). An observed overlap between these sites and putative promoter elements of downstream genes is consistent with the negative (repressor) mode of NiaR regulation, as previously proposed for the *nadABC* operon in *B. subtilis* (16).

Apart from the *Bacillus*/*Clostridium* group, candidate NiaR-binding sites identified in Fusobacterium lineage generally conform to the type I DNA motif defined above. An imperfect inverted repeat 5′-TTTATT<u>GTCTTGTCA</u> ATAAA-3′ in the promoter region of the *F. nucleatum nadABC-niaR* operon has a conserved core (underlined) and divergent AT-rich flanking regions. At the same time, no similar DNA motifs could be detected in the available genomes of the Thermotogales. We used the phylogenetic footprinting approach (43) to identify the type II NiaR-binding motif in this taxonomically distant group of bacteria. In the *de novo* NAD synthesis of *T. maritima*, the common bacterial aspartate oxidase (NadB) is replaced by a nonorthologous aspartate dehydrogenase (NadBII) (44). Multiple alignment of DNA regions upstream of the *nadBII-nadC-nadA* operon (TM1643–TM1645 in *T. maritima*), which is present in three compared *Thermotoga* genomes revealed a highly conserved 17-bp imperfect inverted repeat 5′-<u>TTTACAGCAGGTGTA</u> <u>AA</u>-3′ (symmetrical nucleotides are underlined) located between the candidate −35 and −10 promoter boxes (Figure 2C and Supplementary Figure S3). Remarkably, the identical DNA sequence was detected upstream of the *niaR-niaP* operon (*TM1602-TM1603* in *T. maritima*). This operon and the respective DNA site are conserved in

all four compared *Thermotoga* genomes including *T. lettingae*. The latter species lacks the *de novo* biosynthesis operon, and it appears to be fully dependent on the niacin salvage pathway for NAD synthesis. Although the genes composing this pathway (*pncAB*) do not belong to the NiaR regulon in *Thermotoga* species, niacin salvage in these organisms may still be regulated since the expression of the putative NiaP transporter depends on NiaR. Based on these observations, we hypothesized that the identified DNA signal represents the type II NiaR-binding motif, involved in coregulation of the *de novo* NAD biosynthetic operon and the *niaR-niaP* genes in Thermotogales. This prediction was experimentally tested by the EMSA approach using the purified recombinant *tm*NiaR (as described below). Despite the apparent difference in the sequence of NiaR-binding motifs in Firmicutes and Thermotogales, one may notice a mirror arrangement of the conserved triplets, TGT-(N_5)-ACA versus ACA-(N_4)-TGT, in the two motifs (Figure 2C). The latter may reflect the difference in a relative orientation of the NiaR subunits that are expected to interact with DNA in the dimeric form.

Candidate niacin transporters. Utilization of the exogenous Nam includes its uptake followed by a deamidating (PncA/PncB-dependent) or nondeamidating (NadV-dependent) conversion to NaMN or NMN precursors of NAD (Figure 1A). The salvage of NA, a deamidated form of vitamin B_3, is performed by a single enzyme (PncB) upon NA uptake mediated by the same or different transporters. Notwithstanding the importance of the salvage pathways, especially for bacterial species that lack *de novo* biosynthesis (e.g. many pathogens), relatively little is known about niacin uptake. Although several studies in various bacterial species pointed to the existence of active transport (45,46), no specific transporter genes have been identified so far. A high-affinity NA permease identified in *Saccharomyces cerevisiae* belongs to the yeast-specific *DAL5* family and does not have homologs in bacteria (47). As already mentioned, regulon and genome context analysis was previously used to predict novel transporter genes in a number of metabolic subsystems [as reviewed in (19)]. The analysis of the NiaR regulon in this study implicated three different families of uncharacterized membrane proteins (named NiaP, NiaX, and NiaY) as putative niacin transporters in three different groups of analyzed bacterial species.

The proteins of the NiaP family contain 10 predicted TMSs and belong to the MFS superfamily of transporters with various specificities to sugars and carboxylic acids (Pfam accession number PF00083). NiaP displays a detectable sequence similarity to the benzoate transporter BenK of *Acinetobacter* sp. (48) and 4-hydroxybenzoate permease PcaK from *Pseudomonas putida* (49) (see multiple alignment in Supplementary Figure S4). Importantly, these two aromatic compounds are quite similar to NA (Figure 3A). The observed functional coupling of the NiaP family with NAD biosynthesis is strengthened by the evidence coming independently from two divergent groups of bacteria, the *Bacillus* and *Thermotoga* species. As discussed above, the *niaP* (*yceI*) gene in *B. subtilis* is

a predicted member of the NiaR regulon, whereas its ortholog in *T. maritima* (TM1603) forms a self-regulated operon with the upstream *niaR* regulator gene (Table 1 and Figure 1B). Notably, all analyzed *niaP*-containing genomes also contain niacin salvage genes, whereas in some of the Lactobacillales species that lack a *de novo* NAD biosynthesis pathway, *niaP* is the only NiaR-regulated gene. NiaP orthologs are also found in bacterial species that do not contain NiaR (e.g. in the Proteobacteria and Actinobacteria) and even in the Metazoa (see the phylogenetic tree in Figure 3B and the distribution of genes in Supplementary Table S1). A likely conservation of the NiaP function in a broad range of bacteria not having the NiaR regulon, is additionally supported by: (i) co-occurrence of the *niaP* (uptake) and *pncB* (salvage) genes in most genomes; (ii) chromosomal clustering of the *niaP* and *pncB* genes in a divergent *Deinococcus* branch; and (iii) regulation by a NAD-related transcriptional regulator of the NrtR family, as in *Acinetobacter* sp. [see the accompanying paper (50)].

The second candidate transporter family NiaX (represented by *Spy1425* in *S. pyogenes*) is present in all Streptococcaceae, in four Clostridiales species, and in *Enterococcus faecalis* (Table 1). In many genomes, the *niaX* gene is located adjacent to *niaR* on the chromosome (Figure 1B). NiaX is an integral membrane protein with four predicted TMSs showing no detectable similarity with any known or unknown protein in any other genome beyond this group. Most NiaX-containing species (among them many pathogens) lack the *de novo* NAD biosynthesis but possess the PncA–PncB salvage pathway. The NiaR regulon in two *Streptococcus* species and *Clostridium tetani* contains the *niaX* and *pncA* genes, whereas in *Thermoanaerobacter tengcongensis* it includes the *niaX* and *pncB* genes. These observations support possible involvement of NiaX in niacin transport.

The third novel gene in the NiaR regulon (represented by *BH3254* in *B. halodurans*) encodes a hypothetical transmembrane protein with nine predicted TMSs. This protein family, named NiaY, is present in a limited number of bacterial species and shows no detectable similarity with any proteins of known function. In *Bacillus clausii*, *Exiguobacterium* sp. and *Clostridium difficile*, the *niaY* gene is preceded by a strong NiaR-binding site (Table 1), whereas in *A. metalliredigens* it belongs to the NiaR-regulated *nadABC-niaR-niaY* gene cluster (Figure 1B). In addition to these species, members of the NiaY family were identified only in *Staphylococcus saprophyticus*, *Mannheimia succiniciproducens*, *Treponema denticola* and *F. nucleatum*. Notably, *niaY* genes do not occur in genomes that contain other candidate niacin transporter genes (*niaP*, *niaX*), whereas they always co-occur with the salvage gene *pncB*. This phyletic pattern provides an additional support for the proposed functional assignment of the NiaY protein family.

In summary, the bioinformatics analysis performed in this study allowed us to tentatively identify two types of DNA motifs that correspond to the binding sites of transcriptional regulators of the NiaR family. Genomic
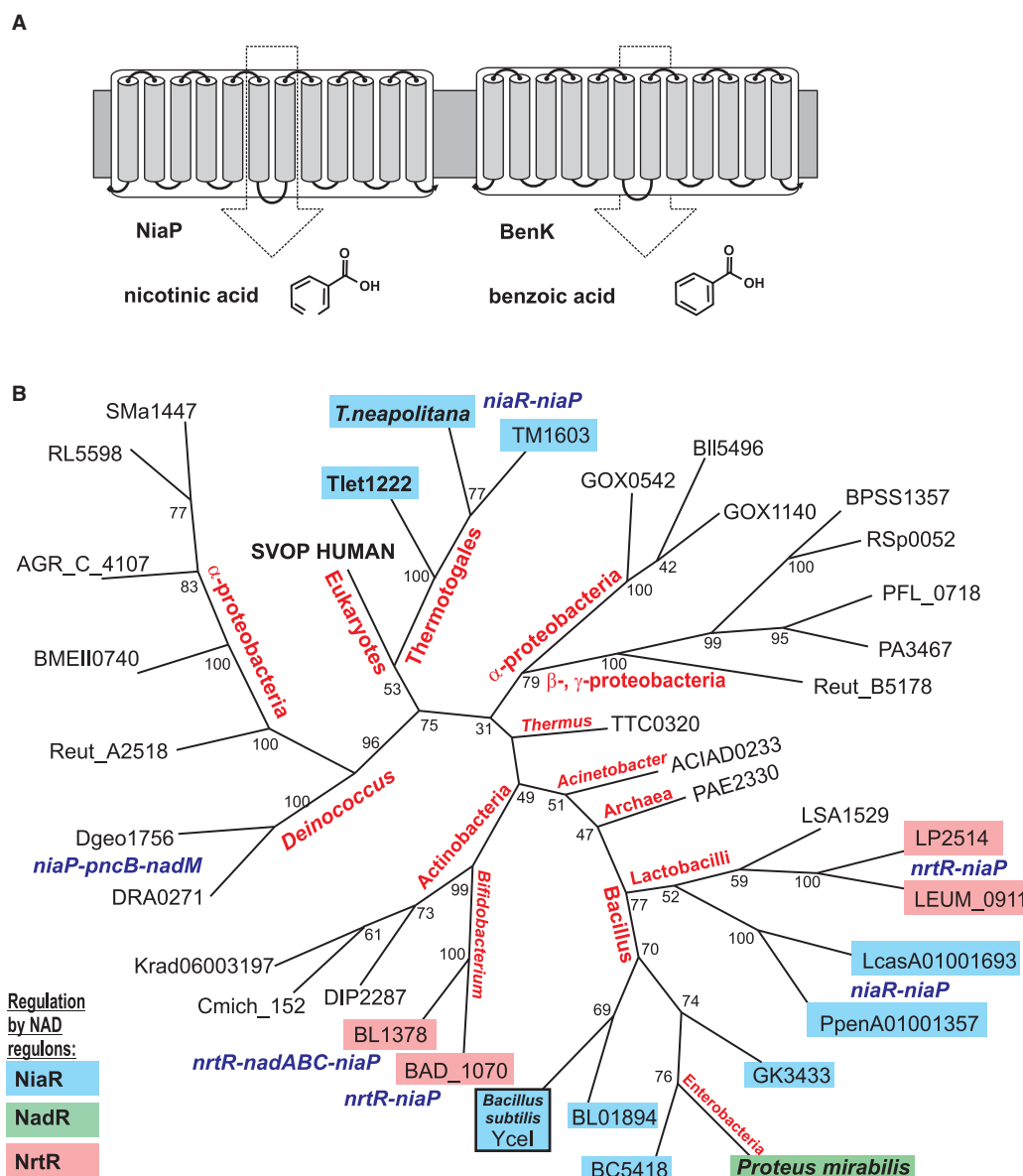
Figure 3. Structural features of the NiaP family. (**A**) Topology prediction for a niacin (NiaP) and a related benzoate (BenK) transporters from the MFS superfamily. *Bacillus subtilis* NiaP and *Acinetobacter* sp. BenK share 27% sequence identity and are composed of 12 predicted transmembrane helices shown by cylinders. (**B**) Maximum-likelihood tree of the NiaP family of niacin transporters. Names of taxonomic groups are given in red. Genes predicted to be regulated by one of the three NAD regulons (NadR, NrtR or NiaR) are highlighted. The chromosomal clusters of the *niaP* genes with NAD regulatory and metabolic genes are given in blue text. The numbers indicate the number of bootstrap replications, out of 100, that support each node on the tree. Multiple alignment of NiaP proteins, as well as of BenK and other homologous transporters is provided in the Supplementary Figure S4. The NiaP protein from *B. subtilis* studied in this work is outlined.

reconstruction and comparative analysis of the NiaR regulons in 35 diverse bacterial genomes led to the prediction of candidate genes for previously uncharacterized niacin transporters. The following key inferences were chosen for the experimental testing: (i) NrtR-binding sites detected in *B. subtilis* and *T. maritima* were assessed using EMSA with recombinant purified *bs*NiaR and *tm*NiaR proteins; (ii) the predicted role of the NiaP family in niacin uptake was tested by growth studies in *B. subtilis* *yceI* gene knockout and heterologous expression of *yceI* in *E. coli*.

Experimental testing of NiaR binding to predicted DNA targets in *T. maritima* and *B. subtilis*

Interaction of the NiaR proteins with DNA operator regions containing candidate NiaR-binding sites was assessed *in vitro* by EMSA. For validation experiments, we have chosen two representatives of the NiaR family, YrxA from *B. subtilis* (*bs*NiaR) and TM1602 from *T. maritima* (*tm*NiaR). The choice of these two organisms was dictated by the following considerations: (i) conservation of regulatory function between highly divergent
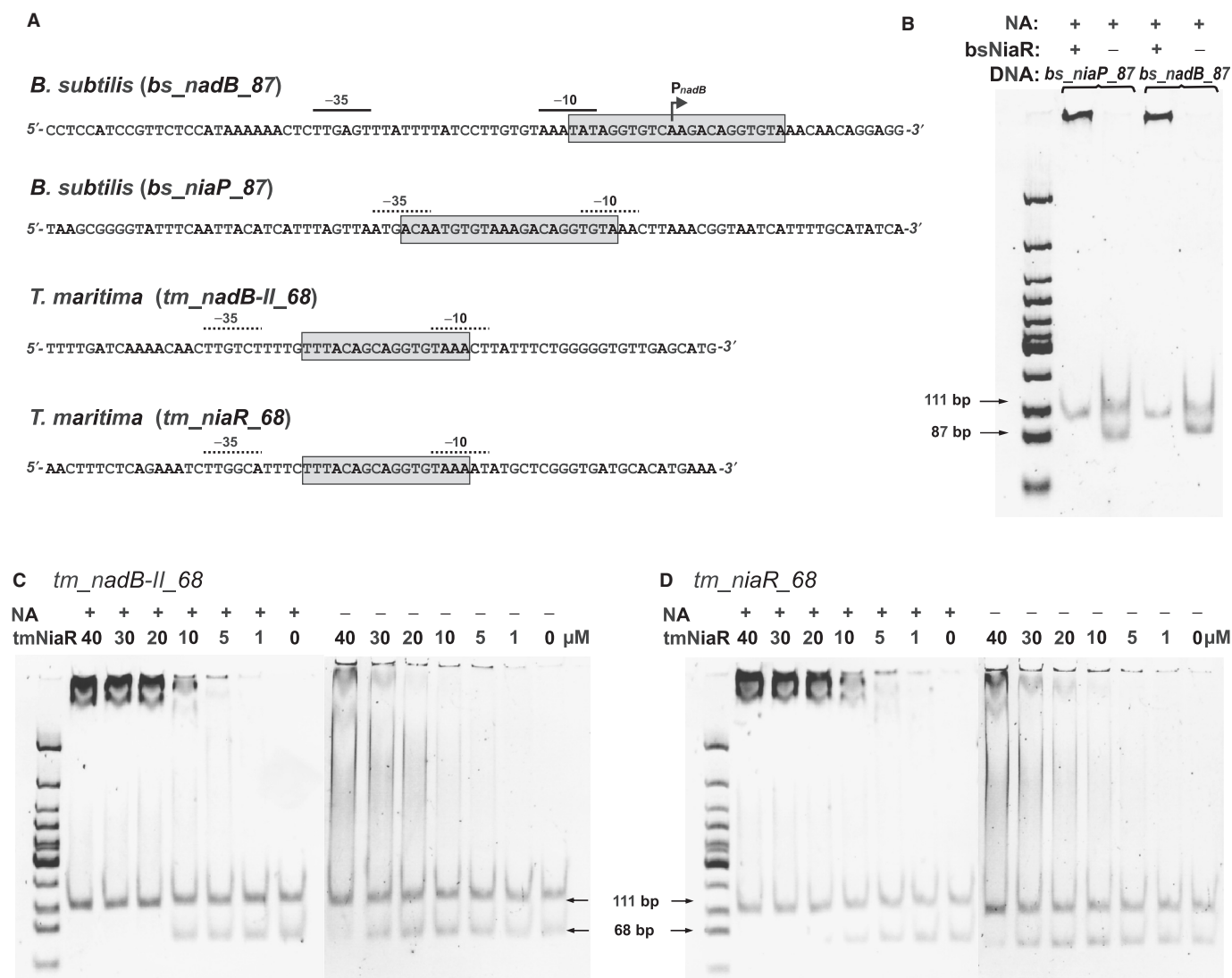
Figure 4. Electrophoretic mobility shift assays (EMSA) to assess the interactions of NiaR proteins from *B. subtilis* (*bs*NiaR) and *T. maritima* (*tm*NiaR) with their cognate DNA signals. (**A**) Sequences of DNA fragments used in EMSA assays. Candidate NiaR-binding sites are shown in gray boxes, and promoter elements are shown by solid lines for the known *B. subtilis nadB* promoter (17) and by dotted lines for the predicted promoters. (**B**) Target DNA segments *bs_yceI_87* and *bs_nadB_87* (20 nM) were incubated for 30 min at 37°C with or without 20 μM of bsNiaR in the presence of 0.3 mM NA and analyzed by EMSA. (**C**) and (**D**) EMSA of DNA segments *tm_niaR_68* and *tm_nadB_68* (20 nM) after incubation for 30 min at 75°C with increasing concentrations of *tm*NiaR (0–40 μM) in the presence or absence of the effector, 0.3 mM NA. A negative control DNA segment *tm_pncA_111* (without NrtR motif) was added to all incubation mixtures as a nonspecific competitor.

representatives of the NiaR protein family, especially at such a large phylogenetic distance as between the Bacillales and Thermotogales lineages, would provide an additional support for the functional assignment of the entire family; (ii) predicted NiaR-binding sites in these species represent the two distinct DNA motifs characteristic of the NiaR family; (iii) the previously described *bs*NiaR binding with one of the two target sites in *B. subtilis* (16) provided us with a positive control for EMSA experiments; and (iv) the availability of the recently solved 3D structure of *tm*NiaR (37) as well as of diffractable crystals of *bs*NiaR (Zhang,H., unpublished data) sets the stage for structure-functional studies.

Incubation of the purified recombinant *bs*NiaR protein with its predicted cognate DNA segments (*bs_nadB_87*

and *bs_niaP_87*) at 37°C, in the presence of 0.3 mM NA and nonspecific competitor (negative control DNA, *tm_pncA_111*) followed by PAGE and Syber Green staining, confirmed the specific *bs*NiaR binding to both target sites (Figure 4A and B). These results demonstrated that in addition to the previously described regulation of the *yrxA/nifS–nadBCA* divergon (16), the NiaR regulon in *B. subtilis* includes the remote *yceI* (*niaP*) gene with a suggested role in niacin uptake.

The EMSA results obtained upon incubation of *tm*NiaR with each of the two predicted target DNA segments (*tm_nadB_68* and *tm_niaR_68*) from the upstream regions of the *T. maritima nadBII-nadA-nadC* and *niaR-niaP* operons, respectively, are illustrated in Figure 4C and D. The incubation was performed at 75°C,

close to the optimal physiological temperature for *T. maritima*, as only weak binding was observed at 37°C (data not shown). For both targets a clear gel mobility shift was observed in the presence of 0.3 mM NA (added to both the incubation mixture and the gel). Under these experimental conditions, the shift of the specific 68-bp DNA band was essentially complete at *tm*NiaR concentration above 10 μM, whereas no change in the 111-bp nonspecific competitor DNA band (*tm_pncA_111*) could be detected at up to 40 μM *tm*NiaR. Exclusion of NA from the incubation and PAGE buffers led to substantial changes in the *tm*NiaR–DNA interactions: (i) the shift of either segment remained incomplete even at 40 μM *tm*NiaR and (ii) even at a high concentration of *tm*NiaR, a substantial fraction of shifted DNA appeared as a smear, likely due to the microheterogeneity and/or instability of the complex in the absence of the NA effector. This is in contrast to a sharp binary distribution between nonshifted and fully shifted bands observed in the presence of NA. These observations are similar to those previously reported, where the presence of the NA effector led to both, increased DNA-binding affinity and possible change in the stoichiometry of the *bs*NiaR–DNA complex, as deduced from the change in the mobility of a shifted band (referred to as *supershift*) (16).

Overall, this analysis of two representative NiaR proteins provided a sufficient experimental confirmation of the target NiaR-binding sites and NiaR-regulated genes tentatively identified by comparative genomic techniques (as described in the previous section). We believe that in addition to specific findings related to the NiaR regulon in *B. subtilis* and *T. maritima*, these results constitute an additional supportive evidence of the similar assignments made for the members of the NiaR family in other species.

Experimental assessment of the predicted NiaP transporter from *B. subtilis*

To test a proposed involvement of the NiaP protein family in niacin transport, we cloned the gene *yceI* (*niaP*) from *B. subtilis* and assessed the effect of its overexpression on the growth rate of an *E. coli* BW25113 (Δ*nadA*) mutant in the minimal medium supplemented with limiting amount of Nam. This strain, made available to us from the Keio collection of *E. coli* single-gene deletion mutants (33), requires Nam or NA to grow on the minimal medium as it has a disrupted *de novo* synthesis making it fully dependent on the salvage of exogenous niacin (via the PncA-PncB pathway, see Figure 1A). Although *E. coli* has an intrinsic ability to uptake niacin via a yet unknown mechanism (46), we speculated that overexpression of the *B. subtilis yceI* gene in the *E. coli* Δ*nadA* mutant could increase the efficiency of Nam uptake and accelerate the strain growth at suboptimal concentrations of Nam. Indeed, a substantial (up to ~6fold) growth rate acceleration was observed in this engineered strain compared to the control strain carrying an empty vector plasmid at 0.16 μM Nam (Figure 5). The difference between the growth curves was smaller at higher concentrations of niacin, becoming negligible at above 1 μM niacin (data not shown).
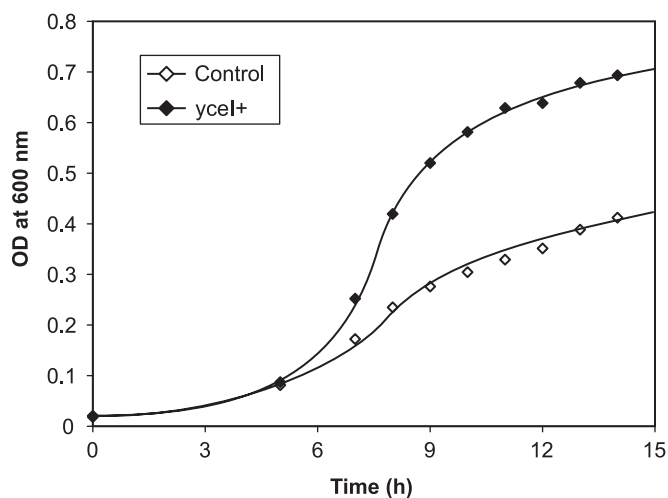

Figure 5. An accelerated growth of *E. coli* Δ*nadA* mutant carrying *yceI* gene in the minimal media supplemented by a limiting concentration of nicotinamide (0.16 μM). The plasmid containing the *B. subtilis yceI* gene was introduced into the *E. coli* Δ*nadA* mutant lacking *de novo* NAD biosynthesis, whereas the same strain with an empty vector was used as a negative control (black and open squares, respectively).

An additional verification of the *bs*NiaP functional assignment was obtained using a different approach based on growth inhibition by 6ANam, a toxic analog of niacin. This compound is known to suppress bacterial growth by entering the regular PncA-PncB salvage pathway and ultimately converting to NAD antagonist (6-amino-NAD) (51,52). Assuming that NiaP would facilitate the uptake of the niacin analog, we expected to observe the increased sensitivity of *E. coli* containing the same *yceI* expression plasmid to growth suppression by 6ANam. This experiment was performed in the strain of *E. coli* containing an undisrupted *de novo* pathway transformed with the same expression plasmid containing the *B. subtilis yceI* gene. A significant difference in the growth rate between this strain and an isogenic control strain transformed with an empty vector was observed in a narrow range of 6ANam concentration 7.5–15 μM (Figure 6A). The growth of both strains was completely suppressed at 6ANam concentrations above 150 μM, whereas no difference in their growth rates was observed in the absence of 6ANam.

A similar approach was used to test whether *bs*NiaP contributes to niacin uptake in its natural host. It was accomplished by comparing the growth of *B. subtilis* knockout mutant with disrupted *yceI* (*niaP*) gene versus a control strain with intact *yceI* in the presence of 6ANam. The expectation was that inactivation of the *yceI* gene would cause resistance to the toxic niacin analogs. However, the growth of this knockout mutant was rather efficiently suppressed by 6ANam pointing to the existence of the NiaP-unrelated niacin uptake mechanism in *B. subtilis*. At the same time, a substantial difference in the growth rate of the two strains was indeed observed in the presence of 0.75–1.5 μM 6ANam (Figure 6B), whereas no difference in the respective growth curves could be detected in the absence of 6ANam. In the presence of another toxic niacin analog, 6-amino-NA (6ANA),
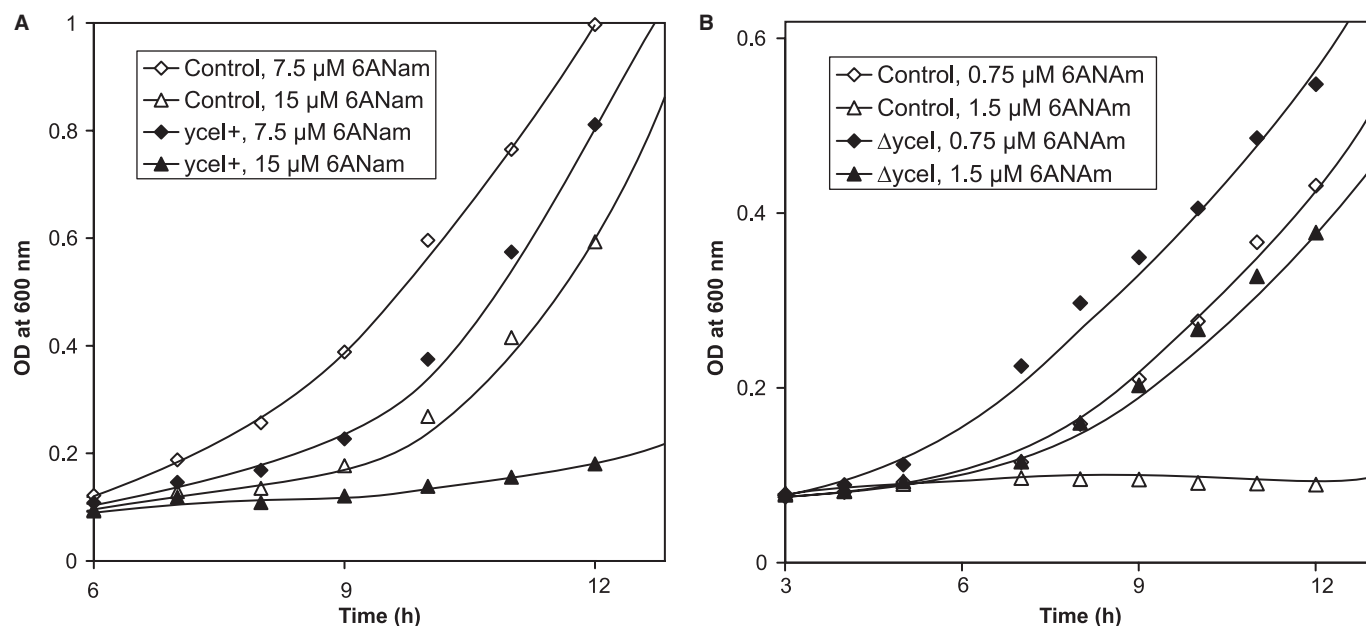
Figure 6. Growth studies in the presence of 6ANam, a toxic analog of Nam. (**A**) An enhanced 6ANam sensitivity (growth suppression) of *E. coli* strain overexpressing the *B. subtilis yceI* gene is observed when grown in minimal medium supplemented by 6ANam compared to the control strain containing empty vector. Both strains have an intact *de novo* NAD biosynthesis. (**B**) In similar growth studies the *B. subtilis* (*yceI⁻*) mutant displayed a decreased 6ANam sensitivity (faster growth) compared to the *B. subtilis* strain with an intact *yceI* gene (carrying the same transposon insert in the unrelated gene ypaA as a negative control).

the largest difference in the growth rate of the two strains was observed at 1–2.5 μM 6ANA. These observations suggest that *bs*NiaP contributes to the uptake of both amidated and deamidated forms of niacin. The obtained results also revealed that, in addition to NiaP, *B. subtilis* must have another, hitherto-unknown transport system for niacin uptake.

Overall, the cumulative evidence generated in the three different types of experiments described above provided strong support for the tentative functional assignment proposed for the NiaP family by comparative genomic analysis and reconstruction of the NiaR regulon. To our knowledge, this is the first report implicating a protein family in niacin uptake in bacteria.

DISCUSSION

The comparative genomic approach was applied to characterize a novel NiaR regulon for transcriptional control of NAD metabolic genes in bacteria. The niacin-responsive transcription factor NiaR (YrxA) was originally identified and characterized as a nicotinic acid-responsive repressor of the *de novo* NAD biosynthesis operon in *B. subtilis* (16). NiaR orthologs were found in 30 out of 45 species from the *Bacillus/Clostridium* group (Firmicutes), as well as in the deeply diverged groups, the Fusobacteria and Thermotogales (Supplementary Table S1). Two distinct types of DNA-binding sites of NiaR were identified, type I operator characteristic of the Firmicutes and Fusobacteria, and type II in the Thermotogales (Figure 2A and C). Despite the apparent overall dissimilarity, both types of operators share

palindromic signature sequences, TGT and ACA, albeit in opposite orientation (Figure 2A and C). Comparison of closely related species showed high conservation of candidate NiaR-binding sites in the promoter regions of target genes (Supplementary Figure S3).

To validate the predicted NiaR-binding sites of both types, the NiaR proteins from two organisms, *B. subtilis* (*bs*NiaR) and *T. maritima* (*tm*NiaR), were cloned, purified, and assessed in EMSA assays for their ability to bind DNA fragments containing their target DNA sites. The results of these experiments (Figure 4) confirmed both the second type I operator site predicted in *B. subtilis* and the predicted DNA sites of type II in *T. maritima*. They also confirmed that, as previously shown for *bs*NiaR (16), NA is a likely effector (corepressor) for the phylogenetically divergent *tm*NiaR. Comparative 3D structural studies are underway to assess mechanistic details of the interactions with DNA and the effector in these two distant representatives of the NiaR family.

A combination of comparative genomic techniques, including positional gene clustering, phylogenetic footprinting of regulatory sites and metabolic reconstruction, helped us establish the details of the NiaR regulon organization in 35 bacterial species (Figure 1A and Table 1). In 18 genomes including species from the Firmicutes, Fusobacteria and Thermotogales groups, NiaR regulates *de novo* biosynthesis operon *nadABC*. Less frequent is the NiaR regulation of the niacin salvage genes *pncB* (in *Lactobacillus plantarum*) and/or *pncA* (in *S. pyogenes*, *Streptococcus equi*, and *C. tetani*) and the RNam salvage transporter *pnuC* (in *S. pneumoniae* and *Streptococcus mutans*). In addition, the NiaR regulon

includes three previously uncharacterized gene families encoding membrane proteins of unknown function, to which we here tentatively assigned a role in niacin uptake. Of those, the most abundant NiaP family is present in 10 NiaR-containing species (Bacilli, Lactobacilli and Thermotogales) as well as in a number of species that do not contain the NiaR regulator. Among two other families, NiaX is present in twelve genomes (Streptococci and Clostridia), and NiaY is present in five genomes (Bacilli and Clostridia). The proposed involvement of these gene families in niacin uptake is supported by several lines of genomic evidence: (i) predicted coregulation with NAD biosynthesis and niacin salvage genes; (ii) clustering on the chromosome with the *niaR* genes; (iii) nonoverlapping (complementary) phyletic patterns of the genome occurrence; (iv) regulation by NiaR in species lacking the *nadABC* genes and fully dependent on niacin salvage for NAD synthesis; and (v) co-occurrence with the niacin salvage genes *pncB-pncA* (Table 1).

Although niacin uptake has been studied in many species, both bacterial, such as *E. coli* (46) and *E. faecalis* (45), and eukaryotic, such as yeasts (47) and mammals (53,54), the knowledge of the respective transport systems and mechanisms is quite limited. The only NA-specific permease was identified in yeast (47), and the absence of its orthologs in other lineages suggests that other niacin transporters are yet to be identified.

The predicted niacin transporter NiaP, a member of the NiaR regulon in Firmicutes and Thermotogales, belongs to other NAD-related regulons in other bacterial species (Figure 3B). Particularly, in two Actinobacteria and two Lactobacilli, NiaP is a member of the NrtR regulon, which is described in the accompanying paper (50). Furthermore, an NiaP ortholog detected in a enterobacterium *Proteus mirabilis* is presumably regulated by binding of the NadR repressor to its characteristic operator site (TGTTTAGTATACTAAACA) in the upstream region of the respective gene. The NiaP transporters belong to the MFS superfamily of transporters that catalyze uniport, solute:cation symport and solute:cation antiport of a great variety of metabolites (55). NiaP is most similar to the aromatic acid/H^+ symporters PcaK from *Pseudomonas* spp. and BenK from *Acinetobacter* spp. involved in 4-hydroxybenzoate and benzoate uptake, respectively (48,49).

We selected a representative of the NiaP family encoded by the *yceI* gene in *B. subtilis* for the experimental testing of its proposed involvement in niacin uptake. Three types of growth experiments were performed to test for: (i) the growth acceleration of *E. coli* (*nadA$^-$*) mutant strain in minimal medium supplemented with limiting concentrations of Nam, as a result of overexpression of the *B. subtilis* gene *yceI*; (ii) the increased sensitivity of *E. coli* (*nadA$^+$*) to the growth suppression by a toxic analog of Nam (6ANam), as a result of overexpression of the *B. subtilis* gene *yceI* and (iii) the increased resistance of the *B. subtilis* (*yceI$^-$*) mutant to growth suppression by toxic niacin analogs (6ANam and 6ANA). The results of these experiments (Figures 5 and 6) allowed us to conclude that *bs*NiaP is involved in niacin transport in *B. subtilis*,

although there exist additional, unknown mechanism for niacin uptake. It is important to emphasize that although the performed experiments provided sufficient support for the predicted functional assignment of the NiaP family, additional studies are necessary to establish its actual mechanism, energy dependence and kinetic parameters, as well as to elucidate the mechanism of NiaP-independent niacin uptake.

High-scoring homologs of NiaP are widely distributed among Gram-positive and Gram-negative bacteria (Supplementary Table S1), and they are also present in eukaryotic genomes including nematodes, insects and mammals showing substantial (up to 30%) sequence identity with their bacterial counterparts (Supplementary Figure S4). Although the mechanism of niacin transport was characterized in human intestinal epithelial and liver cells, a gene encoding the respective transporter is yet unknown. At the same time, the human homolog of NiaP, a synaptic vesicle 2-related protein (SVOP) expressed in brain and endocrine cells, was extensively studied but its physiological and molecular function remained obscure (56). It is tempting to speculate that human SVOP and its orthologs in other mutlicellular eukaryotes constitute a plausible candidate for a missing niacin transporter in these organisms. This hypothesis, however farfetched it may seem due to the evolutionary distance and well-anticipated differences between bacterial and eukaryotic membrane transport systems, is supported by a precedent in the field of vitamin transport. Indeed, the human homolog of the bacterial pantothenate (vitamin B$_5$) transporter of the Solute:Sodium Symporter family (e.g. PanF in *E. coli*) was implicated in uptake of B$_5$ (and possibly some other vitamins) in human cells (57).

This work and the accompanying study (50) demonstrated a significant variability of regulatory strategies for the transcriptional control of NAD metabolism in different bacterial species (see Supplementary Table S1 that summarizes the distribution of NAD metabolic and regulatory genes across genomes). This variability parallels substantial variations in the topology of NAD biosynthetic pathways [as illustrated in the 'NAD and NADP biosynthesis' subsystem provided by the SEED database (http://theseed.uchicago.edu/) and briefly discussed in (58)] and, more importantly, a variety of bacterial habitats and lifestyles. Although the NiaR regulatory mechanism bears some analogy with NadR regulation previously studied in Enterobacteria (e.g. suppression of the *de novo* biosynthesis enhanced by interactions with a corepressor), fundamental differences between these two regulatory strategies are apparent. Whereas NadR blocks all biosynthetic routes to NAD as long as NAD concentration is above threshold (NAD operates is an effector of NadR), NiaR tends to switch the organism from the *de novo* biosynthetic route to niacin salvage as long as niacin is available at the level sufficient to operate as an NiaR corepressor. On the other hand, niacin may also regulate its own uptake to maintain NAD homeostasis at the desired level as the genes of the downstream biosynthetic machinery (NadD and NadE) appear to be constitutively expressed, or at least not regulated by NiaR. Regulation of the niacin salvage

enzymes (PncA and PncB) by NiaR is relatively rare, and happens only in species that lack a *de novo* biosynthetic route. Both NadR and NiaR regulatory strategies are dramatically different from the NrtR regulatory network described in the accompanying paper (50). The latter regulon includes the entire NAD biosynthetic machinery, including the downstream enzymes. The derepression appears to be triggered by the accumulation of NAD degradation products (e.g. ADP ribose that operates as NrtR antirepressor) likely interpreted by the cell as depletion of the NAD pool which thus needs to be replenished.

In summary, this study provided a comprehensive bioinformatic analysis of the NiaR regulon that constitutes a system of transcriptional regulation of NAD synthesis in several groups of Gram-positive bacteria. The key conjectures delivered by this analysis, two types of NiaR-binding DNA motifs and one of the three predicted families of niacin transporters, were experimentally validated providing additional support for these specific findings as well as for the reconstruction of the entire NiaR regulon. Despite the implications of this and the accompanying study (50), the mechanisms and the specific genes involved in the regulation of NAD synthesis as well as in the uptake of its precursors in many groups of bacteria, remain to be elucidated.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Conflict of interest statement. None declared.

REFERENCES

1. Marmorstein,R. (2004) Structure and chemistry of the Sir2 family of NAD+-dependent histone/protein deactylases. *Biochem. Soc. Trans.*, **32**, 904–909.
2. Wilkinson,A., Day,J. and Bowater,R. (2001) Bacterial DNA ligases. *Mol. Microbiol.*, **40**, 1241–1248.
3. Ziegler,M. (2000) New functions of a long-known molecule. emerging roles of NAD in cellular signaling. *Eur. J. Biochem.*, **267**, 1550–1564.
4. Begley,T.P., Kinsland,C., Mehl,R.A., Osterman,A. and Dorrestein,P. (2001) The biosynthesis of nicotinamide adenine dinucleotides in bacteria. *Vitam. Horm.*, **61**, 103–119.
5. Belenky,P., Bogan,K.L. and Brenner,C. (2007) NAD+ metabolism in health and disease. *Trends Biochem. Sci.*, **32**, 12–19.
6. Gerdes,S.Y., Kurnasov,O.V., Shatalin,K., Polanuyer,B., Sloutsky,R., Vonstein,V., Overbeek,R. and Osterman,A.L. (2006) Comparative genomics of NAD biosynthesis in cyanobacteria. *J. Bacteriol.*, **188**, 3012–3023.
7. Katoh,A. and Hashimoto,T. (2004) Molecular biology of pyridine nucleotide and nicotine biosynthesis. *Front. Biosci.*, **9**, 1577–1586.
8. Magni,G., Amici,A., Emanuelli,M., Raffaelli,N. and Ruggieri,S. (1999) Enzymology of NAD+ synthesis. *Adv. Enzymol. Relat. Areas Mol. Biol.*, **73**, 135–182.
9. Penfound,T. and Foster,J. (1996) *Biosynthesis and Recycling of NAD*, 2nd edn. American Society for Microbiology, Washington, D.C.
10. Kurnasov,O.V., Polanuyer,B.M., Ananta,S., Sloutsky,R., Tam,A., Gerdes,S.Y. and Osterman,A.L. (2002) Ribosylnicotinamide kinase domain of NadR protein: identification and implications in NAD biosynthesis. *J. Bacteriol.*, **184**, 6906–6917.
11. Singh,S.K., Kurnasov,O.V., Chen,B., Robinson,H., Grishin,N.V., Osterman,A.L. and Zhang,H. (2002) Crystal structure of Haemophilus influenzae NadR protein. A bifunctional enzyme endowed with NMN adenyltransferase and ribosylnicotinimide kinase activities. *J. Biol. Chem.*, **277**, 33291–33299.
12. Grose,J.H., Bergthorsson,U. and Roth,J.R. (2005) Regulation of NAD synthesis by the trifunctional NadR protein of Salmonella enterica. *J. Bacteriol.*, **187**, 2774–2782.
13. Penfound,T. and Foster,J.W. (1999) NAD-dependent DNA-binding activity of the bifunctional NadR regulator of Salmonella typhimurium. *J. Bacteriol.*, **181**, 648–655.
14. Raffaelli,N., Lorenzi,T., Mariani,P.L., Emanuelli,M., Amici,A., Ruggieri,S. and Magni,G. (1999) The Escherichia coli NadR regulator is endowed with nicotinamide mononucleotide adenylyltransferase activity. *J. Bacteriol.*, **181**, 5509–5511.
15. Gerasimova,A.V. and Gelfand,M.S. (2005) Evolution of the NadR regulon in Enterobacteriaceae. *J. Bioinform. Comput. Biol.*, **3**, 1007–1019.
16. Rossolillo,P., Marinoni,I., Galli,E., Colosimo,A. and Albertini,A.M. (2005) YrxA is the transcriptional regulator that represses de novo NAD biosynthesis in Bacillus subtilis. *J. Bacteriol.*, **187**, 7155–7160.
17. Sun,D. and Setlow,P. (1993) Cloning, nucleotide sequence, and regulation of the Bacillus subtilis nadB gene and a nifS-like gene, both of which are essential for NAD biosynthesis. *J. Bacteriol.*, **175**, 1423–1432.
18. Rodionov,D.A. (2007) Comparative genomic reconstruction of transcriptional regulatory networks in bacteria. *Chem. Rev.*, **107**, 3467–3497.
19. Gelfand,M.S. and Rodionov,D.A. (2007) Comparative genomics and functional annotation of bacterial transporters. *Phys. Life Rev.*, **4**, j.plrev.2007.2010.2003.
20. Benson,D.A., Karsch-Mizrachi,I., Lipman,D.J., Ostell,J. and Wheeler,D.L. (2007) GenBank. *Nucleic Acids Res.*, **35**, D21–D25.
21. 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. *et al.* (2005) The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res.*, **33**, 5691–5702.
22. Chenna,R., Sugawara,H., Koike,T., Lopez,R., Gibson,T.J., Higgins,D.G. and Thompson,J.D. (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.*, **31**, 3497–3500.
23. Felsenstein,J. (1996) Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods. *Methods Enzymol.*, **266**, 418–427.
24. Finn,R.D., Mistry,J., Schuster-Bockler,B., Griffiths-Jones,S., Hollich,V., Lassmann,T., Moxon,S., Marshall,M., Khanna,A., Durbin,R. *et al.* (2006) Pfam: clans, web tools and services. *Nucleic Acids Res.*, **34**, D247–D251.
25. Tatusov,R.L., Galperin,M.Y., Natale,D.A. and Koonin,E.V. (2000) The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.*, **28**, 33–36.
26. Gelfand,M.S., Koonin,E.V. and Mironov,A.A. (2000) Prediction of transcription regulatory sites in Archaea by a comparative genomic approach. *Nucleic Acids Res.*, **28**, 695–705.
27. Mironov,A.A., Koonin,E.V., Roytberg,M.A. and Gelfand,M.S. (1999) Computer analysis of transcription regulatory patterns in completely sequenced bacterial genomes. *Nucleic Acids Res.*, **27**, 2981–2989.

28. Mironov,A.A., Vinokurova,N.P. and Gelfand,M.S. (2000) [Software for analyzing bacterial genomes]. *Mol. Biol.*, **34**, 253–262.

29. Crooks,G.E., Hon,G., Chandonia,J.M. and Brenner,S.E. (2004) WebLogo: a sequence logo generator. *Genome Res.*, **14**, 1188–1190.

30. Sheffield,P., Garrard,S. and Derewenda,Z. (1999) Overcoming expression and purification problems of RhoGDI using a family of "parallel" expression vectors. *Protein Expr. Purif.*, **15**, 34–39.

31. Lesley,S.A., Kuhn,P., Godzik,A., Deacon,A.M., Mathews,I., Kreusch,A., Spraggon,G., Klock,H.E., McMullan,D., Shin,T. *et al.* (2002) Structural genomics of the Thermotoga maritima proteome implemented in a high-throughput structure determination pipeline. *Proc. Natl Acad. Sci. USA*, **99**, 11664–11669.

32. Osterman,A.L., Lueder,D.V., Quick,M., Myers,D., Canagarajah,B.J. and Phillips,M.A. (1995) Domain organization and a protease-sensitive loop in eukaryotic ornithine decarboxylase. *Biochemistry*, **34**, 13431–13436.

33. Baba,T., Ara,T., Hasegawa,M., Takai,Y., Okumura,Y., Baba,M., Datsenko,K.A., Tomita,M., Wanner,B.L., Mori,H. *et al.* (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.*, **2**, 2006 0008.

34. Vagner,V., Dervyn,E. and Ehrlich,S.D. (1998) A vector for systematic gene inactivation in Bacillus subtilis. *Microbiology*, **144**(Pt 11), 3097–3104.

35. Rodionov,D.A. and Gelfand,M.S. (2005) Identification of a bacterial regulatory system for ribonucleotide reductases by phylogenetic profiling. *Trends Genet.*, **21**, 385–389.

36. Makarova,K.S., Mironov,A.A. and Gelfand,M.S. (2001) Conservation of the binding site for the arginine repressor in all bacterial lineages. *Genome Biol.*, **2**, RESEARCH0013.

37. Weekes,D., Miller,M.D., Krishna,S.S., McMullan,D., McPhillips,T.M., Acosta,C., Canaves,J.M., Elsliger,M.A., Floyd,R. *et al.* (2007) Crystal structure of a transcription regulator (TM1602) from Thermotoga maritima at 2.3 Å resolution. *Proteins*, **67**, 247–252.

38. Anantharaman,V., Koonin,E.V. and Aravind,L. (2001) Regulatory potential, phyletic distribution and evolution of ancient, intracellular small-molecule-binding domains. *J. Mol. Biol.*, **307**, 1271–1292.

39. Overbeek,R., Fonstein,M., D'Souza,M., Pusch,G.D. and Maltsev,N. (1999) The use of gene clusters to infer functional coupling. *Proc. Natl Acad. Sci. USA*, **96**, 2896–2901.

40. Korbel,J.O., Jensen,L.J., von Mering,C. and Bork,P. (2004) Analysis of genomic context: prediction of functional associations from conserved bidirectionally transcribed gene pairs. *Nat. Biotechnol.*, **22**, 911–917.

41. Kolesov,G., Wunderlich,Z., Laikova,O.N., Gelfand,M.S. and Mirny,L.A. (2007) How gene order is influenced by the biophysics of transcription regulation. *Proc. Natl Acad. Sci. USA*, **104**, 13948–13953.

42. Gelfand,M.S. (1999) Recognition of regulatory sites by genomic comparison. *Res. Microbiol.*, **150**, 755–771.

43. Florea,L., McClelland,M., Riemer,C., Schwartz,S. and Miller,W. (2003) EnteriX 2003: visualization tools for genome alignments of Enterobacteriaceae. *Nucleic Acids Res.*, **31**, 3527–3532.

44. Yang,Z., Savchenko,A., Yakunin,A., Zhang,R., Edwards,A., Arrowsmith,C. and Tong,L. (2003) Aspartate dehydrogenase, a novel enzyme identified from structural and functional studies of TM1643. *J. Biol. Chem.*, **278**, 8804–8808.

45. Neujahr,H.Y. and Varga,Z. (1966) Transport of B-vitamins in microorganisms. VII. The uptake of 14 C-niacinamide by non-proliferating cells and by protoplasts of Streptococcus faecalis. *Acta Chem. Scand.*, **20**, 1529–1534.

46. Rowe,J.J., Lemmon,R.D. and Tritz,G.J. (1985) Nicotinic acid transport in Escherichia coli. *Microbios*, **44**, 169–184.

47. Llorente,B. and Dujon,B. (2000) Transcriptional regulation of the Saccharomyces cerevisiae DAL5 gene family and identification of the high affinity nicotinic acid permease TNA1 (YGR260w). *FEBS Lett.*, **475**, 237–241.

48. Collier,L.S., Nichols,N.N. and Neidle,E.L. (1997) benK encodes a hydrophobic permease-like protein involved in benzoate degradation by Acinetobacter sp. strain ADP1. *J. Bacteriol.*, **179**, 5943–5946.

49. Nichols,N.N. and Harwood,C.S. (1997) PcaK, a high-affinity permease for the aromatic compounds 4-hydroxybenzoate and protocatechuate from Pseudomonas putida. *J. Bacteriol.*, **179**, 5056–5061.

50. Rodionov,D.A., De Ingeniis,J., Mancini,C., Cimadamore,F., Zhang,H., Osterman,A.L. and Raffaelli,N. (2008) Transcriptional regulation of NAD metabolism in bacteria: NrtR family of Nudix-related regulators. *Nucleic Acids Res.*, **36**, 2047–2059.

51. Cobb,J.R., Pearcy,S.C. and Gholson,R.K. (1977) Metabolism of 6-aminonicotinic acid in Escherichia coli. *J. Bacteriol.*, **131**, 789–794.

52. Hughes,K.T., Cookson,B.T., Ladika,D., Olivera,B.M. and Roth,J.R. (1983) 6-Aminonicotinamide-resistant mutants of Salmonella typhimurium. *J. Bacteriol.*, **154**, 1126–1136.

53. Nabokina,S.M., Kashyap,M.L. and Said,H.M. (2005) Mechanism and regulation of human intestinal niacin uptake. *Am. J. Physiol. Cell Physiol.*, **289**, C97–C103.

54. Said,H.M., Nabokina,S.M., Balamurugan,K., Mohammed,Z.M., Urbina,C. and Kashyap,M.L. (2007) Mechanism of nicotinic acid transport in human liver cells: studies with HepG2 cells and primary hepatocytes. *Am. J. Physiol. Cell Physiol.*, **293**, C1773–C1778.

55. Pao,S.S., Paulsen,I.T. and Saier,M.H. Jr. (1998) Major facilitator superfamily. *Microbiol. Mol. Biol. Rev.*, **62**, 1–34.

56. Janz,R., Hofmann,K. and Sudhof,T.C. (1998) SVOP, an evolutionarily conserved synaptic vesicle protein, suggests novel transport functions of synaptic vesicles. *J. Neurosci.*, **18**, 9269–9281.

57. Wang,H., Huang,W., Fei,Y.J., Xia,H., Yang-Feng,T.L., Leibach,F.H., Devoe,L.D., Ganapathy,V. and Prasad,P.D. (1999) Human placental Na + -dependent multivitamin transporter. Cloning, functional expression, gene structure, and chromosomal localization. *J. Biol. Chem.*, **274**, 14875–14883.

58. Osterman,A. and Begley,T.P. (2007) *A Subsystems-Based Approach to the Identification of Drug Targets in Bacterial Pathogens,* Birkhauser, Basel.