



# Ribonucleotide Reductases from Bifidobacteria Contain Multiple Conserved Indels Distinguishing Them from All Other Organisms: *In Silico* Analysis of the Possible Role of a 43 aa Bifidobacteria-Specific Insert in the Class III RNR Homolog

Seema Alnajar, Bijendra Khadka and Radhey S. Gupta\*

Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, ON, Canada

## OPEN ACCESS

### Edited by:

Manuel Espinosa,  
Centro de Investigaciones Biológicas  
(CSIC), Spain

### Reviewed by:

Paloma López,  
Centro de Investigaciones Biológicas  
(CSIC), Spain  
Paul Meyers,  
University of Cape Town, South Africa

### \*Correspondence:

Radhey S. Gupta  
gupta@mcmaster.ca

### Specialty section:

This article was submitted to  
Evolutionary and Genomic  
Microbiology,  
a section of the journal  
Frontiers in Microbiology

Received: 17 April 2017

Accepted: 11 July 2017

Published: 31 July 2017

### Citation:

Alnajar S, Khadka B and Gupta RS  
(2017) Ribonucleotide Reductases  
from Bifidobacteria Contain Multiple  
Conserved Indels Distinguishing  
Them from All Other Organisms: *In  
Silico* Analysis of the Possible Role  
of a 43 aa Bifidobacteria-Specific  
Insert in the Class III RNR Homolog.  
Front. Microbiol. 8:1409.  
doi: 10.3389/fmicb.2017.01409

Bifidobacteria comprises an important group/order of bacteria whose members have widespread usage in the food and health industry due to their health-promoting activity in the human gastrointestinal tract. However, little is known about the underlying molecular properties that are responsible for the probiotic effects of these bacteria. The enzyme ribonucleotide reductase (RNR) plays a key role in all organisms by reducing nucleoside di- or tri- phosphates into corresponding deoxyribose derivatives required for DNA synthesis, and RNR homologs belonging to classes I and III are present in either most or all *Bifidobacteriales*. Comparative analyses of these RNR homologs have identified several novel sequence features in the forms of conserved signature indels (CSIs) that are exclusively found in bifidobacterial RNRs. Specifically, in the large subunit of the aerobic class Ib RNR, three CSIs have been identified that are uniquely found in the *Bifidobacteriales* homologs. Similarly, the large subunit of the anaerobic class III RNR contains five CSIs that are also distinctive characteristics of bifidobacteria. Phylogenetic analyses indicate that these CSIs were introduced in a common ancestor of the *Bifidobacteriales* and retained by all descendants, likely due to their conferring advantageous functional roles. The identified CSIs in the bifidobacterial RNR homologs provide useful tools for further exploration of the novel functional aspects of these important enzymes that are exclusive to these bacteria. We also report here the results of homology modeling studies, which indicate that most of the bifidobacteria-specific CSIs are located within the surface loops of the RNRs, and of these, a large 43 amino acid insert in the class III RNR homolog forms an extension of the allosteric regulatory site known to be essential for protein function. Preliminary docking studies suggest that this large CSI may be playing a role in enhancing the stability of the RNR dimer complex. The possible significances of the identified CSIs, as well as the distribution of RNR homologs in the *Bifidobacteriales*, are discussed.

**Keywords:** novel features of ribonucleotide reductases, probiotic bacteria, *Bifidobacteriales*, conserved signature inserts and deletions, homology modeling and protein docking studies, extended allosteric site, phylogenetic analysis

## INTRODUCTION

The *Bifidobacteriales* constitute an important order of bacteria within the phylum Actinobacteria (Ventura et al., 2007; Zhi et al., 2009; Gao and Gupta, 2012). While some species belonging to this order are pathogenic (Smith et al., 1992; Bradshaw et al., 2006; Alves et al., 2014; Kenyon and Osbak, 2014) many *Bifidobacteriales* species belonging to the genus *Bifidobacterium* are known for their beneficial health-promoting effects in humans and other mammals (Gibson et al., 1995; Leahy et al., 2005; Masco et al., 2005; Ventura et al., 2009; Cronin et al., 2011). These probiotic bifidobacteria form a significant constituent in the microbiota of the human colon, and exert their effects as commensal microorganisms (Biavati et al., 2000; Turroni et al., 2008, 2009; Mills et al., 2011; Milani et al., 2014; Ventura et al., 2014). As a result, these bacteria are frequently exploited by the food industry to create consumable products that increase their relative proportion in the gut (Gibson et al., 1995; Sanders, 1998; Masco et al., 2005; Oberg et al., 2011; Ventura et al., 2014). Bifidobacteria are Gram-positive, anaerobic, saccharolytic organisms with a unique metabolic pathway known as the “bifid shunt” (Palframan et al., 2003; Biavati and Mattarelli, 2006; Milani et al., 2015). While many characteristics are known about this important group of bacteria, the biochemical and molecular properties contributing toward their probiotic effects, and adaptability in their respective environments, remain elusive (Ventura et al., 2009; Turroni et al., 2014).

The present study focuses on the enzyme ribonucleotide reductase (RNR), the sole enzyme capable of reducing nucleoside di- or tri- phosphates (NDPs or NTPs) into deoxyribonucleotides (dNDPs or dNTPs) (Eklund et al., 2001; Nordlund and Reichard, 2006; Torrents, 2014). There are currently three recognized classes of RNRs, named classes I, II, and III, sharing no more than 10% sequence identity across their lengths, which are distributed in different organisms (Logan et al., 1999; Sintchak et al., 2002; Torrents et al., 2002). Class I RNR is further divided into three subclasses viz. Ia, Ib, Ic (Jordan et al., 1996; Jiang et al., 2007; Bollinger et al., 2008). The distributions of these different classes of RNRs within the bacterial domain does not follow any specific pattern that can be correlated with the phylogenies of the bacterial phyla (Torrents et al., 2002; Lundin et al., 2009). However, since the different classes of RNR employ different mechanisms of action and require differing environmental prerequisites to function, we explore their distribution in bifidobacteria in an attempt to identify any unique characteristics that may distinguish them.

Each RNR is capable of reducing all four ribonucleotides into their corresponding deoxyribonucleotides by exhibiting a tightly regulated allosteric substrate specificity site, and employing a convoluted mechanism involving radical chemistry that ultimately results in the removal of a hydrogen from the 3' carbon of the substrate (Brown and Reichard, 1969; Reichard, 1993, 2010; Eriksson et al., 1997; Eliasson et al., 1999). Some RNRs have an additional overall activity site, made possible by the existence of an ATP cone domain at the N-terminus (Thelander and Reichard, 1979). Class I RNRs use NDPs as their substrate, and are aerobic tetramers consisting of one large (R1) and one

small (R2) homodimers (Nordlund and Reichard, 2006). The R2 dimer harbors a dinuclear metallocofactor where the radical is formed and subsequently transferred to the active site located at the R1 subunit. Classes Ia, Ib, and Ic differ in the type of metallocofactor in R2 (manganese and/or iron), as well as the different cofactors required for enzymatic function (Pettersson et al., 1980; Nordlund and Reichard, 2006; Jiang et al., 2007; Bollinger et al., 2008; Torrents, 2014). Most bifidobacteria harbor a class Ib RNR (Lundin et al., 2009), whose large and small subunit are encoded by the *nrdE* and *nrdF* genes, respectively, and require NrdH as a reductant and NrdI as a cofactor; in contrast, classes Ia and Ic utilize thioredoxin and/or glutaredoxin as reductants, do not require additional cofactors, and their large and small subunits are encoded by *nrdA* and *nrdB*, respectively (Jordan et al., 1997; Cotruvo and Stubbe, 2008; Roca et al., 2008; Crona et al., 2011). Class II RNRs are not oxygen sensitive, use either NDPs or NTPs as their substrate, and are the only monomeric class of RNR (Tamao and Blakley, 1973; Larsson et al., 2010), however, their structural topology mimics a dimer (Sintchak et al., 2002). No known bifidobacteria harbor a class II homolog, but all bifidobacteria possess a class III RNR. Class III RNRs are encoded by *nrdD* and *nrdG* genes and function under strictly anaerobic conditions, with NTPs as their sole substrates (Garriga et al., 1996; Torrents et al., 2001). They consist of a large R1 subunit (NrdD) that is a dimer in its native state, and works concomitantly with a small activase (NrdG) which generates the radical utilizing a [4Fe-4S] cluster (Eliasson et al., 1992; Sun et al., 1995; Logan et al., 1999). This is unlike class I RNRs, where radical formation by the small subunit is required to induce dimer formation of the large subunit (Ollagnier et al., 1996; Torrents, 2014). Despite the described differences in the properties of the different classes of RNRs, the remarkable structural similarities seen across the three main RNRs strongly suggest a common evolutionary origin of them (Poole et al., 2002; Sintchak et al., 2002; Torrents et al., 2002). In all three RNRs, allosteric regulation involves binding of the dNDP/dNTP products at a 4-helix bundle, involving two helices from each monomeric subunit, at the dimer interface of the enzyme (Uhlen and Eklund, 1994; Larsson et al., 2001). The allosteric regulation causes conformational changes at a highly conserved 10 stranded  $\alpha/\beta$  barrel where the active site “finger loop” structure resides in its center, or is brought to its center upon activation (Aurelius et al., 2015).

Although previous studies have significantly contributed to the current understanding of the structure and function of the different RNRs, in the present work we focus on the specific biochemical/molecular properties of the RNRs from bifidobacteria that may shed light on their unique physiological effects. Our earlier work describes a number of conserved signature indels (CSIs) in the homologs of many important proteins from that are uniquely found in all *Bifidobacteriales* (Zhang et al., 2016). These CSIs represent vertically transferred genetic changes that are indicated to have occurred in a common ancestor of the group in which they are found, thus asserting their value as highly specific molecular markers. In the present work, we have performed similar comparative genomic studies that have led to the identification of several novel CSIs in class Ib

and III RNR homologs that are shared by all genome-sequenced *Bifidobacteriales* species that contain the respective protein homolog(s), but are absent in all other bacteria. We also describe the results of protein modeling which illustrate the structural location of these CSIs, as well as the results of preliminary *in silico* docking studies which suggest that one of the large CSIs [a 43 amino acid (aa) insertion in class III RNR] may be playing a role in NrdD complex stability.

## MATERIALS AND METHODS

### Identification of Conserved Signature Indels

The approach used to identify CSIs in RNR was as described in earlier work (Gupta, 2014; Zhang et al., 2016). Multiple sequence alignments (MSAs) were initially created using the Clustal\_X 2.1 (Larkin et al., 2007; Goujon et al., 2010) program for the protein sequences of NrdE, NrdF, NrdH, NrdD, and NrdG homologs from about 10–15 *Bifidobacteriales* species, as well as 8–10 species from other groups/phyla of bacteria. These sequence alignments were examined for the presence of conserved indels that are limited to the *Bifidobacteriales* homologs and are flanked on both sides by at least five conserved residues in the neighboring 30–40 aa. A detailed Blastp search (Altschul et al., 1997) was then conducted on the sequence region containing the potential conserved indels to investigate the species-specificities of the identified indels. The indels that were not flanked by conserved regions were not further investigated in our work. The signature files shown here were created using SIG\_CREATE and SIG\_STYLE from the GLEANS.net program as described in earlier work (Gupta, 2014; Zhang et al., 2016). Unless otherwise indicated, all of the reported CSIs are specific for the *Bifidobacteriales* homologs and similar CSIs were not observed in homologs from any other bacterial species within the top 500–1000 blast hits examined.

### Phylogenetic Tree Construction

In this study we have constructed three separate phylogenetic trees: (i) based on NrdE (large subunit of class Ib RNR) sequences, (ii) based on NrdD (large subunit of class III RNR) sequences, and (iii) based on the large subunit sequences from class I (NrdA, NrdE), II (NrdJ), and III (NrdD) RNRs. For these studies, NrdE and NrdD homologs from all genome sequenced bifidobacterial species were obtained from the NCBI GenBank sequence database (Benson et al., 2017). The species represented in the tree based on NrdD sequences included 49 of 58 validly published *Bifidobacterium* species, the two known *Scardovia* species, all three *Alloscardovia* species, and the single species known from the *Parascardovia* and *Gardnerella* genera. The tree based on NrdE sequences similarly included the subset of these *Bifidobacteriales* species where the protein homolog was detected. Sequences from members of the *Bifidobacteriales* genera *Aeriscardovia* and *Pseudoscardovia* were not available at the present time and were not included in our study. For each tree, a MSA of RNR homologs was created using the Clustal\_X 2.1 (Larkin et al., 2007; Goujon et al., 2010) program. For each of

these trees, we have additionally included a number of outgroup species (20 species for the NrdD tree, 23 species for the NrdE tree) from other orders in the Actinobacteria phylum, as well as Firmicutes species. For the tree concerning the sequences from large subunits of all RNR classes, we have included <10 NrdE and <10 NrdD sequences from representative *Bifidobacteriales*, in addition to several species across various bacterial phyla in order to depict the evolutionary history of RNR classes. The MEGA 6 program (Tamura et al., 2013) was used to construct a maximum likelihood (ML) tree based on 1000 bootstrap replicates for each alignment employing the Whelan and Goldman model substitution method (Whelan and Goldman, 2001). Gaps and regions with missing data from the sequence alignments were completely removed. In each case, a discrete Gamma distribution was used to model evolutionary rate differences among sites (five categories) and the Jones–Taylor–Thornton substitution method was used to compute the initial trees for the heuristic search using the Neighbor-joining method with a matrix of pairwise distances (Jones et al., 1992).

### Homology Modeling of RNR Homologs and Structural Analysis of CSIs

The approach used to model the CSIs involves homology modeling based on previously crystallized class Ib and III RNR proteins. A Position-Specific-Iterated Blastp search (Altschul et al., 1997) was performed on *Bifidobacterium longum* NrdE (Accession no. EPE39971) and NrdD (Accession no. KXS29127) sequences against the PDB database which revealed that the class Ib RNR from *Salmonella typhimurium* (PDB ID: 1PEQ) (Uppsten et al., 2003) and class III RNR from Enterobacteria phage T4 (PDB ID: 1H7B) (Larsson et al., 2001) exhibited the highest sequence similarity to the *Bifidobacteriales* homologs and provided suitable templates for homology modeling of the RNR isoforms of *B. longum*. A conserved domain search (CD-Search) (Marchler-Bauer and Bryant, 2004) was conducted on the *B. longum* sequences. Homology modeling was performed using MODELLER v9.11 (Eswar et al., 2007) and the top 500 models were initially created and ranked on the basis of their Discrete Optimized Protein Energy (DOPE) scores (Shen and Salí, 2006). The selected models of RNR homologs with the highest DOPE score were then submitted to the GalaxyRefine server (Heo et al., 2013; Lee et al., 2016) to obtain atomic-level energy minimization and to improve the stereochemical quality of the model. The secondary structure elements in the regions containing CSIs were examined and compared with results of the PSIPRED and CONCORD analyses to ensure their reliability (Jones, 1999; Wei et al., 2011; Buchan et al., 2013). The stereochemical properties of the final models were assessed using four independent servers: RAMPAGE, ERRAT, PROSA and Verify3D (Bowie et al., 1991; Luthy et al., 1992; Colovos and Yeates, 1993; Sippl et al., 1999; Lovell et al., 2003; Wiederstein and Sippl, 2007). These validation tools utilize a dataset of highly refined solved structures to evaluate the statistical significance of models based on the conformation, location, and the environment of individual amino acids in the protein sequence, as well as the model's overall structural stability. The structural alignments of the models with the respective templates were carried out using PyMOL

**TABLE 1** | Results of protein–protein docking studies for the *Bifidobacterium longum* RNR (NrdD subunit) either containing or lacking the large 43 aa CSI.

Docking server	Protein models of the NrdD		
	CSI-containing (Extended helix)	CSI-containing (based on PSIPRED/CONCORD)	
		CSI lacking	
ClusPro	–1297.200	–1190.900	–953.000
ClusPro + ROSIE	–1092.851	–1110.039	–1042.091
PatchDock	*	16192.000	11234.000
PatchDock + ROSIE	*	–1109.473	–1033.714
“ROSIE”	–1105.550	–1114.689	–1043.235

The results of docking studies are shown for two different models of the CSI-containing protein (Extended helix and based on PSIPRED/CONCORD) and compared with those for the model for CSI-lacking protein. The monomers based on the indicated models were submitted to ClusPro and PatchDock servers. Results from the ClusPro and PatchDock servers were also submitted to the ROSIE (RosettaDock) server for local refinement. The results shown under the heading “ROSIE” were obtained by submitting model monomers that were superimposed onto the crystallized 1H7B (Larsson et al., 2001) template dimer. For the ClusPro and ROSIE servers, the lower binding energy (more negative values) is indicative of the increase in stability of the docking complexes, while for the PatchDock server higher (positive) scores indicate stronger binding affinity (see Materials and Methods section for further details). Asterisks (\*) indicate that no biologically relevant complexes were obtained from the indicated docking servers.

Version 1.8 (Schrödinger, 2016) in order to analyze the location and the structural features of the CSIs in the protein structure. This procedure was followed to create the homology models of both of the RNR homologs found in bifidobacteria. In addition, a structural model of class III RNR was also generated using I-TASSER, an online server that uses threading to predict three dimensional protein structure (Zhang, 2008; Roy et al., 2010; Yang et al., 2015).

## Protein–Protein Docking Analysis of the Class III RNR Homologs

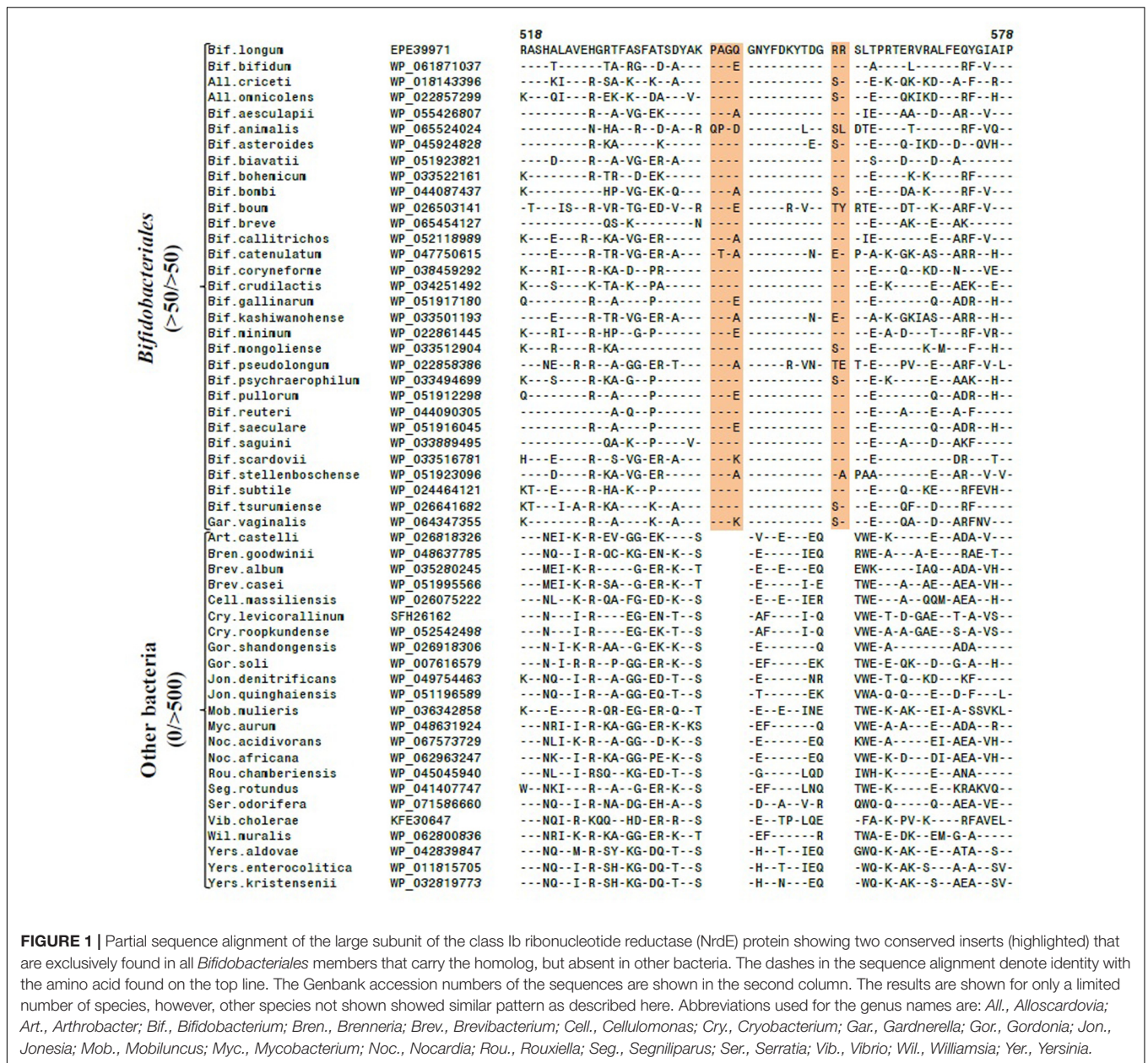
Protein–protein docking studies were performed in order to assess the possible role of a large CSI in the formation or stabilization of the dimeric structure of class III RNR in bifidobacteria. A structural model of the class III RNR from *B. longum* was created by removing the CSI residues from its primary sequence, using the methods described above for other RNR homologs. An additional structural model of RNR was generated with the CSI region constructed as an extended helix. In this structural model, the CSI has a slightly different secondary structure than those of the models that followed PSIPRED/CONCORD, or I-TASSER predictions. This was done in an attempt to be inclusive of multiple possible structural conformations of the CSI. Four structures of the anaerobic RNR monomer viz. PSIPRED/CONCORD based model, I-TASSER generated model, model with extended helix, and the CSI-lacking model, were submitted to two independent web-based protein–protein docking programs using default parameters, viz. PatchDock Version B 1.3 (Schneidman-Duhovny et al., 2005) and ClusPro Version 2.0 (Comeau et al., 2004). PatchDock is an efficient molecular docking algorithm that employs a geometry-based shape complementarity approach which aims

to yield refined atomic contacts of protein–protein complexes. Its scoring function takes into consideration both geometric fit and atomic desolvation energy (Duhovny et al., 2002). On the other hand, ClusPro utilizes PIPER, a rigid body docking program, which is based on a novel Fast-Fourier Transform (FFT) docking approach with pairwise potential. Its scoring function is thus based on pairwise interaction potentials (Comeau et al., 2004; Kozakov et al., 2006). The resulting top scoring dimer complex models of RNR from each server (if any) were then refined using the RosettaDock (ROSIE) server (Lyskov and Gray, 2008; Chaudhury et al., 2011; Lyskov et al., 2013). For the docking scores of ClusPro and RosettaDock, the lower (negative) binding energy value indicates improved stability of the docking complexes. In the case of PatchDock, the geometry shape complementarity score was utilized to determine rank, and higher (positive) scores indicate stronger binding affinity (see also notes in Table 1). In addition, the monomeric forms of each of the four models were structurally aligned with the established biological assembly of the RNR dimer, and the resulting dimer orientations were utilized as additional inputs for submission to the ROSIE server. The resulting refined structure from ROSIE with the lowest total score, maximum cluster size and the smallest RMSD with respect to the solved structure of RNR complex, was chosen as a representative structure for detail interface interaction analysis. To analyze the dimer interface, this class III RNR dimeric output structure was submitted to the PDBePISA Version 1.48 server, using default parameters (Krissinel and Henrick, 2007).

## RESULTS

### Identification of Conserved Signature Indels in Class I and Class III RNR Homologs and Their Phylogenetic Implications

Comparative analysis of the *Bifidobacteriales* genomes indicated that all of the sequenced species from this order contain an anaerobic class III RNR homolog. In addition, an aerobic class I RNR belonging to the class Ib group was also found in most species from this order except *B. adolescentis*, *B. angulatum*, *B. dentium*, *B. gallicum*, *B. cuniculi*, *B. lemurum*, *B. merycicum*, *B. moukalabense*, *B. ruminantium*, and members of the genera *Parascardovia* and *Scardovia*. The sequences of class Ib and III RNRs were examined for the presence of any CSIs that are specific for bifidobacteria. The results of these studies have identified three CSIs in the large subunit of the class Ib RNR (NrdE) homologs, which are specifically found in the bifidobacterial enzyme. Sequence information for two of these CSIs, which are comprised of 4 and 2 aa inserts in a conserved region of the NrdE protein, is shown in Figure 1. As seen in the figure, both of these CSIs are flanked on either side by conserved regions and while they are commonly shared by all of the bifidobacteria harboring the NrdE homolog, they are not present in any other bacterial species in the top 500 blast hits. Sequence information for one additional CSI in the NrdE protein, consisting of a 1 aa



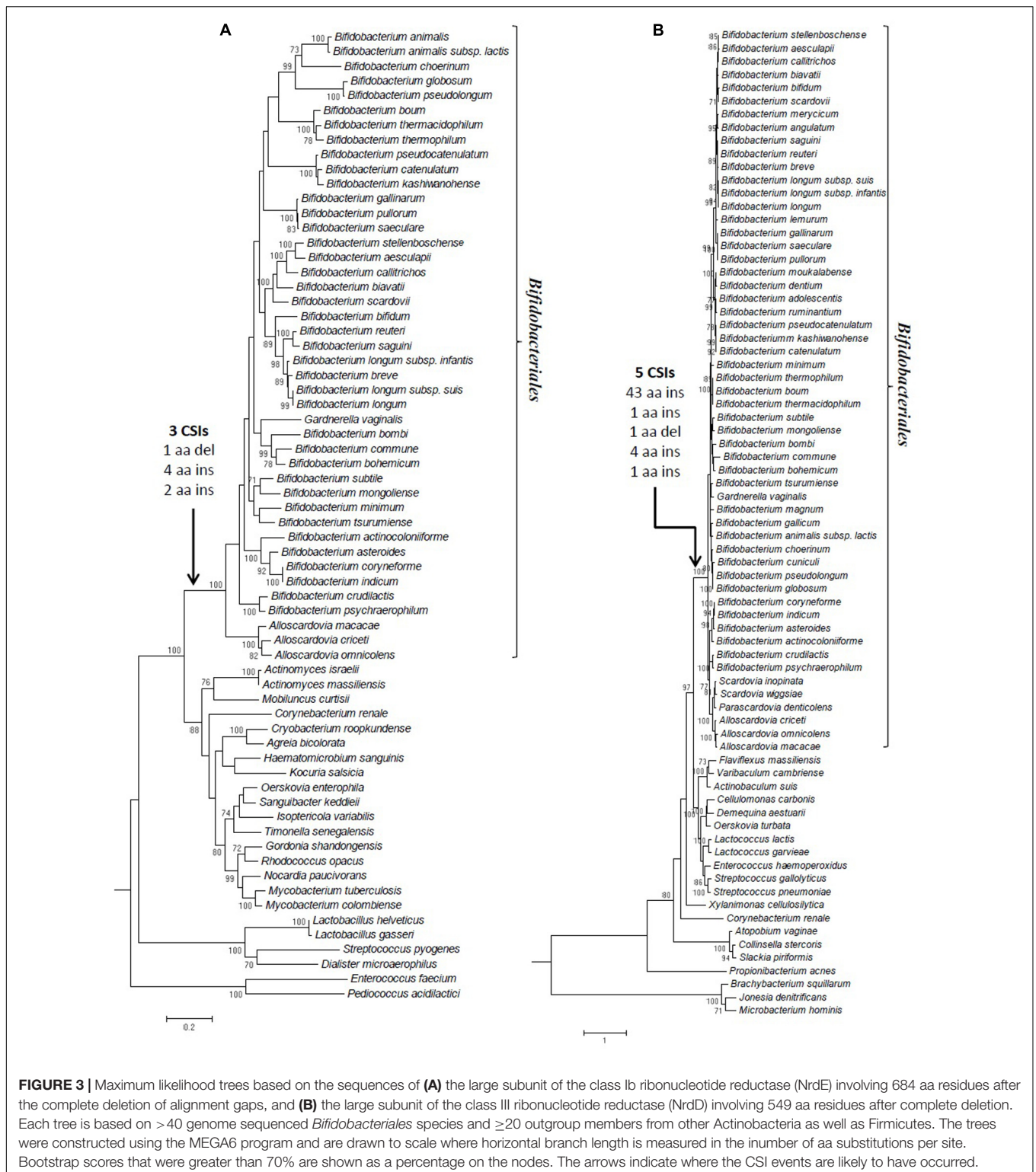
**FIGURE 1** | Partial sequence alignment of the large subunit of the class Ib ribonucleotide reductase (NrdE) protein showing two conserved inserts (highlighted) that are exclusively found in all *Bifidobacteriales* members that carry the homolog, but absent in other bacteria. The dashes in the sequence alignment denote identity with the amino acid found on the top line. The Genbank accession numbers of the sequences are shown in the second column. The results are shown for only a limited number of species, however, other species not shown showed similar pattern as described here. Abbreviations used for the genus names are: *All.*, *Alloscardovia*; *Art.*, *Arthrobacter*; *Bif.*, *Bifidobacterium*; *Bren.*, *Brenneria*; *Brev.*, *Brevibacterium*; *Cell.*, *Cellulomonas*; *Cry.*, *Cryobacterium*; *Gar.*, *Gardnerella*; *Gor.*, *Gordonia*; *Jon.*, *Jonesia*; *Mob.*, *Mobiluncus*; *Myc.*, *Mycobacterium*; *Noc.*, *Nocardia*; *Rou.*, *Rouxiella*; *Seg.*, *Segniliparus*; *Ser.*, *Serratia*; *Vib.*, *Vibrio*; *Wil.*, *Williamsia*; *Yer.*, *Yersinia*.

deletion also specific for all *Bifidobacteriales* strains that harbor the protein, is presented in Supplementary Figure 1, and is once again absent in other bacteria.

Similarly, analysis of the sequences from the NrdD homolog has also led to the identification of five CSIs that are specific for the *Bifidobacteriales* homologs. Sequence information for one large CSI, a 43 aa insertion, that is specifically found in the NrdD homolog from bifidobacteria, is presented in Figure 2. Sequence information for four other CSIs, consisting of a 1 aa deletion, two 1 aa insertions, and a 4 aa insertion, which are also either exclusively or mainly found in the *Bifidobacteriales* NrdD homologs, are presented in Supplementary Figures 2–5. Of these other CSIs, the 1 aa deletion (Supplementary Figure 3) is also present in *Coriobacteriales* species, which are also

anaerobic and saccharolytic bacteria. Additionally, one of the CSIs consisting of a 1 aa insert is also shared by *Lactobacillus* species (Supplementary Figure 2), and the other single aa insert is shared by few other species from the Actinobacteria phylum (Supplementary Figure 5). Besides from these cases, the CSIs were not found in the additional 500–1000 bacterial outgroups examined. The locations of the different identified CSIs over the lengths of the NrdE and NrdD proteins and their respective domains are presented in Supplementary Figures 6, 7. All of the CSIs in the NrdE and NrdD subunits are located in the RNR domain of the respective proteins. In contrast to the NrdE and NrdD proteins, no specific CSIs were found in the NrdF, NrdH, or NrdG proteins. Due to the exclusive presence of most of these CSIs in the RNR homologs from bifidobacteria, they





from the Firmicutes species were used to root the trees. The sequences from *Bifidobacteriales* species formed strongly supported monophyletic clades in both trees. Because the sequence alignments used for construction of these phylogenetic

trees did not contain any sequence gaps, the observed branching pattern was not influenced by the presence of the identified CSIs. Therefore, the distinct branching of bifidobacteria observed in both trees supports the notion that the reported CSIs in the

NrdD and NrdE proteins most likely first occurred in a common ancestor of the order *Bifidobacteriales*, and were inherited by descendants due to incurring an evolutionary advantage. In addition to these trees, we have also constructed a tree based on the sequences of the large subunit from the three main RNR classes (Supplementary Figure 8). The three classes of RNR formed distinct clades in the tree, which were separated from each other by long branches. Based on the midpoint rooting of the tree, the sequences from the class III RNR, which function under strictly anaerobic conditions, were found to form a sister clade to sequences from the classes I and II. The observed branching of the class III RNR in the tree is in agreement with earlier work (Reichard, 1993; Logan et al., 1999; Larsson et al., 2001; Poole et al., 2002; Sintchak et al., 2002; Torrents et al., 2002) suggesting that this class of RNR represents the ancestral form of the reductase. Although phylogenetic analysis can shed light on the evolutionary history of the three types of RNR, it does not explain the variable distribution of these classes in different organisms. As an important protein, there is at least one type of RNR in every organism. However, the combination of different RNRs which are found in various organisms is unpredictable and it does not show any correlation with the evolutionary histories of the organisms (Torrents et al., 2000; Lundin et al., 2009, 2010; Torrents, 2014).

## Locations of the CSIs in the Ribonucleotide Reductase Homologs Structures

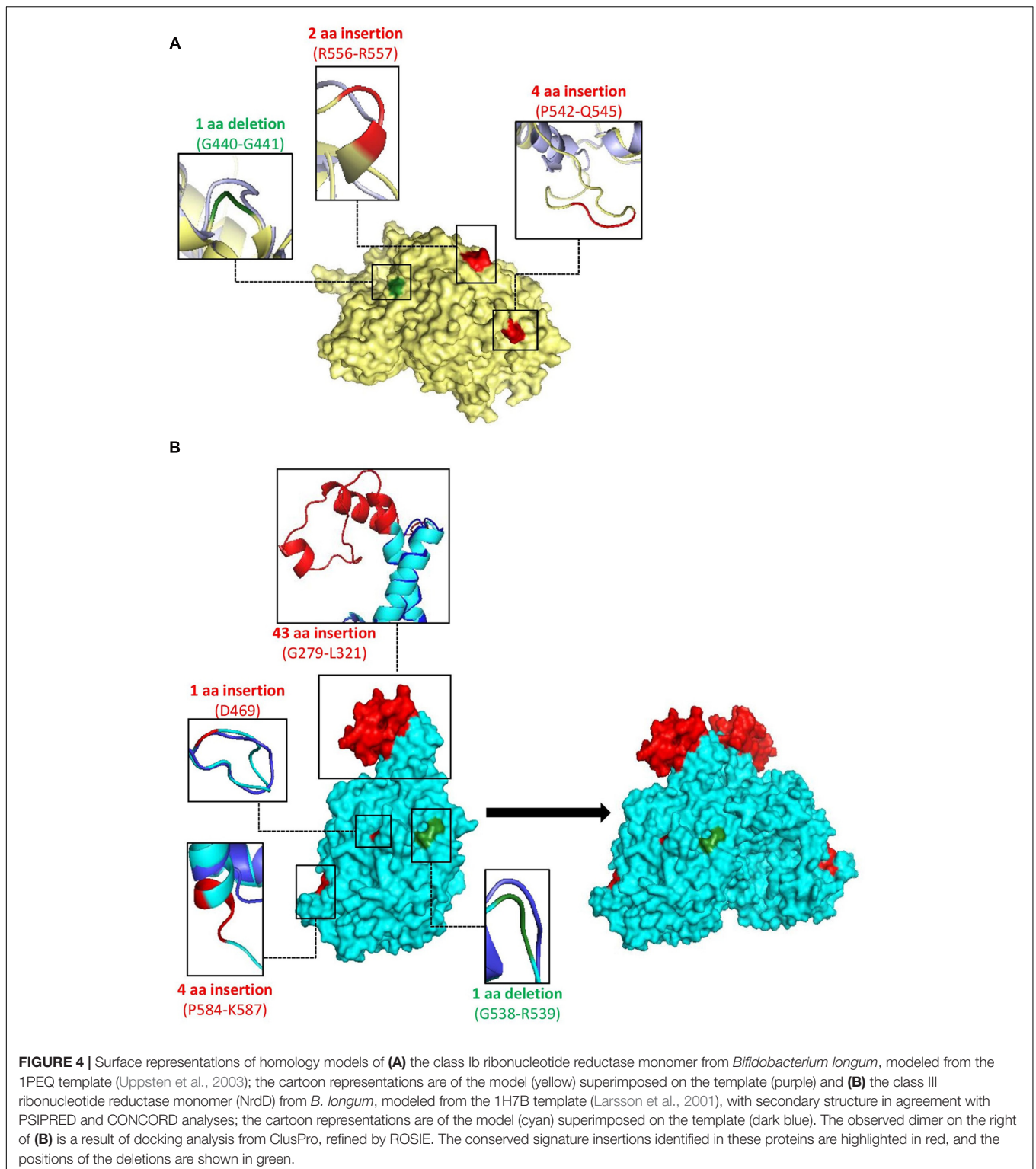
To gain insights into the possible significance of the identified CSIs, homology models for the class Ib and III RNRs from *B. longum* were constructed (see Materials and Methods section) based on previously crystallized template structures of class Ib and III RNR proteins (Larsson et al., 2001; Uppsten et al., 2003). After the validation of the homology models using a variety of tools described under section “Materials and Methods,” a superimposition of the final selected models with the template structures was carried out using PyMOL to determine the locations of the identified CSIs in the structures of the class Ib and III proteins. The locations of the three CSIs identified in class Ib RNR in the modeled structure of the NrdE protein is shown in **Figure 4A**. As seen, all three CSIs in the NrdE homolog were located within the surface loops of the protein (**Figure 4A**). However, of these CSIs, the 2 aa insert also appears to extend a helix. The current model is in agreement with secondary structure analyses (PSIPRED/CONCORD), and yielded reassuring measurements by ERRAT, Verify3D, RAMPAGE and PROSA. The locations of these CSIs in the structure of the *B. longum* NrdE subunit indicate that they are topologically distant from the dimer interface/allosteric regulatory site, as well as the active site as seen in Supplementary Figure 9A (Uppsten et al., 2003, 2006). However, the locations of these CSIs within surface loops in the NrdE structure indicate that they could be involved in mediating novel protein–protein interactions (Cherkasov et al., 2006; Singh and Gupta, 2009; Gupta, 2016; Zhang et al., 2016; Khadka and Gupta, 2017).

Of the five CSIs found in the NrdD homologs (Class III) of bifidobacteria, the structural locations of four indels could be determined and are illustrated on the modeled structure (**Figure 4B**). Structural location of one of the CSIs present near the C-terminal end could not be determined as the structural information for the corresponding region was absent from the template structure (PDB ID: 1H7B) used for homology modeling (Larsson et al., 2001). Similar to the indels found in the NrdE structure, most of the CSIs in the NrdD structure are also found on surface loops that are structurally distant from the active site (**Figure 4B** and Supplementary Figure 9B). The 4 aa insert, also located on the surface exposed region, appears to form a loop and elongate a helix. The large 43 aa insert in NrdD exists as an elongation of the allosteric regulatory region, in between two helices that form the 4-helix bundle in the NrdD dimer (Uhlin and Eklund, 1994; Logan et al., 1999; Larsson et al., 2001). Since the 43 aa insert did not correspond to a characterized domain or motif, elucidating its structure was a challenging task. We present a model that is in agreement with secondary structure predictions from the *B. longum* primary sequence (PSIPRED/CONCORD) (**Figure 4B**), and an additional model according to the prediction made *in silico* by the I-TASSER server (Supplementary Figure 10B). As seen in Supplementary Figure 10B, in the I-TASSER model, the 43 aa insert appears to form two helices that are connected to one another by a loop, and each are also connected to the two existing helices by loops. The orientation of the insert is such that it folds back toward the bulk of the protein, and the dimer interface is relatively uninfluenced. In the case of the model generated based on the secondary structure predictions, the CSI forms an extension of the two helices, along with two small helices connected by loops in between (**Figure 4B**). In order to be inclusive of all reasonable possibilities, we also modeled the NrdD homolog based on the hypothesis that, instead of small helices with breaks induced by loops according to the PSIPRED/CONCORD model, perhaps the loops connecting the main helices to the small helices may be extended helices without loop-induced breaks (Supplementary Figure 10A). This extended helix hypothesis is a corollary to the observation that the elongated helix in class III compared to class I has important functional significance regarding allosteric binding, and also influences dimer packing (Larsson et al., 2001). All three models were refined and validation scores were maximized in these CSI-containing regions.

## Analyzing the Possible Functional Significance of the Large Conserved Insert

To determine the possible role of the large 43 aa insert in dimer formation or complex stability, we have performed a series of docking studies to reveal dimerization potentials of the models compared to an additional model of bifidobacterial NrdD that lacks the CSI. These four models (I-TASSER, PSIPRED/CONCORD, extended helix hypothesis and CSI-lacking) were submitted to two online servers viz. ClusPro and PatchDock. The complexes obtained from ClusPro and PatchDock were refined and scored using ROSIE, and an





additional dimerization measurement was performed by superimposing the models with the available experimentally solved structure of the NrD dimer (Larsson et al., 2001). The docking scores of the dimer complexes from the protein–protein docking studies are summarized in **Table 1**. The PatchDock

server did not yield any dimer complex for the modeled proteins containing extended helix that was plausible with the known biological assembly. However, the docking scores for the CSI-containing protein model based on PSIPRED/CONCORD were consistently improved (or superior) for all servers in

comparison to the protein model lacking the CSI. Thus, the model generated according to the PSIPRED/CONCORD prediction is more likely to approximate the true structure of the class III RNR in *B. longum*. For all docking servers, the I-TASSER generated protein model did not form a plausible dimer complex and hence its results are not shown. This may be due to the fact that in the I-TASSER model of NrdD, the CSI was found to protrude away from the dimeric interface (Supplementary Figure 10B).

To examine if any of the residues from the CSI are involved in dimer interaction, the structural coordinate file for the dimeric form of the PSIPRED/CONCORD model was submitted to the PDBePISA server (Krissinel and Henrick, 2007). Analysis of the results obtained suggests that two residues from the large CSI (viz. Gly279, Met280 for *B. longum* NrdD) are present at the protein dimer interface. However, both of these residues, which are present at the N-terminal end of this large CSI, are not conserved in other *Bifidobacteriales* homologs. Thus, it is difficult to infer with any degree of confidence the possible role of these two residues in protein dimerization. Besides from these two residues, the remaining CSI residues are partly or fully solvent accessible near the interface.

## DISCUSSION

The *Bifidobacteriales* are an important group of bacteria that consist of both pathogenic species and health-promoting commensal microorganisms that are frequently exploited in the food industry as probiotics (Gibson et al., 1995; Sanders, 1998; Leahy et al., 2005; Masco et al., 2005; Ventura et al., 2009; Cronin et al., 2011). However, the use of these bacteria as probiotics faces several challenges as very little is understood about the mechanism responsible for the beneficial effects exerted by bifidobacteria (Oberge et al., 2011). In the present work, we have identified several novel signatures in the form of CSIs in the sequences of both classes I and III RNR homologs that differentiate the RNR homologs from *Bifidobacteriales* from all those found in all other organisms. Earlier work on CSIs (including 1–2 amino acid indels) in several important proteins (e.g., GroEL, DnaK, GyrB, PIP5K, etc.) provides evidence that the CSIs play important functional roles in CSI-containing organisms, and deletion or other changes in the CSIs adversely impact cell growth or other critical functions (Chatterji et al., 2000; Singh and Gupta, 2009; Schoeffler et al., 2010; Clarke and Irvine, 2013; Gupta et al., 2017). In this context, the results reported here that both classes I and III RNR homologs harbor multiple CSIs that are uniquely present in all bifidobacteria is of much interest. These CSIs serve to clearly differentiate the bifidobacterial RNR homologs from those found in all other organisms, and they could function in conjunction with each other to impart certain novel functional characteristic(s) that is only shared by the classes I and III RNR homologs from bifidobacteria.

The distribution of RNR homologs in the *Bifidobacteriales* reveals that while all bifidobacteria species contain a class III

anaerobic RNR homolog (NrdDG), the class Ib aerobic RNR homologs (NrdeF) were not found (or detected) in a number of *Bifidobacterium* species (viz. *B. adolescentis*, *B. angulatum*, *B. dentium*, and *B. gallicum*, *B. cuniculi*, *B. lemurum*, *B. merycicum*, *B. moukalabense*, *B. ruminantium*) as well as in members of the genera *Parascardovia* and *Scardovia* (Lundin et al., 2009). The bifidobacteria species lacking the class Ib RNR homologs do not show any specific branching pattern, or belong to any specific clade(s) of bifidobacteria, but they appear to be distributed sporadically within the order *Bifidobacteriales* (Figure 3). Thus, it is likely that the genes for both classes I and III RNR were present in the common ancestor of all *Bifidobacteriales* and subsequently some species have lost the genes for the class I RNR. Since the different identified CSIs in the classes I and III RNR homologs are present in all bifidobacteria, the most likely explanation for this fact is that the genetic changes responsible for the observed CSIs occurred in a common ancestor of the order *Bifidobacteriales*, presumably at the time of divergence of this group of bacteria from other organisms, similar to the CSIs in many other proteins that are uniquely found in the members of this order (Zhang et al., 2016).

The biological significance of the wide-spread presence of an aerobic RNR (class Ib) in bifidobacterial species, which are generally regarded as anaerobic organisms, remains to be understood. It is known that bifidobacterial species exhibit varying levels of aerotolerance which affects their viability outside of their natural habitats (e.g., gastrointestinal tract, mouth and vagina of mammals) upon exposure to an oxidative environment (Biavati and Mattarelli, 2006; Oberge et al., 2011). Under these conditions, RNRs are pivotal in maintaining a pool of dNDPs/dNTPs to compensate for DNA and protein damage by reactive oxygen species (ROS), ultimately alleviating the detrimental effects of superoxide stress. Thus, a reasonable assumption is that the class I aerobic RNR may play an important role in the tolerance of bifidobacteria to oxidative environment. Although how bifidobacteria manage oxidative exposure remains to be understood, in few studies that have explored the gene expression of both NrdeFHI and NrdDG systems in *Bifidobacterium* species, including data available from the Gene Expression Omnibus<sup>1</sup>, reveal that, under oxidative stress, the expression of both gene systems was rapidly induced by the altered environment; first, the class Ib system was upregulated, followed by the class III system and other proteins, including proteolytic enzymes (Larsson et al., 2001; Edgar et al., 2002; Xiao et al., 2011; Zuo and Chen, 2016). However, further work is necessary to understand the biological and physiological roles of the class I RNR in bifidobacteria under normal conditions and during oxidative stress.

Earlier work on CSIs in protein structures provides evidence that most of the studied CSIs are located in the surface loops of proteins (Geszvain et al., 2004; Akiva et al., 2008; Singh and Gupta, 2009; Gupta et al., 2017; Khadka and Gupta, 2017). The results of our modeling of the identified CSIs in the NrdE and NrdD protein structures also show that all of described

<sup>1</sup><http://www.ncbi.nlm.nih.gov/geo/>

CSIs in the classes I and III RNR protein subunits are present on the surface loops of these proteins (**Figure 4**). Based on extensive earlier work, the surface loops in protein structures often serve as platforms for facilitating novel protein–protein or protein–ligand interactions that are specific for the CSI-containing organisms, without affecting the core functions of the target proteins (Geszvain et al., 2004; Akiva et al., 2008; Singh and Gupta, 2009; Schoeffler et al., 2010; Gupta et al., 2017; Khadka and Gupta, 2017). In a number of cases, the surface loops formed by the CSIs have also been shown to play important role in determining the oligomeric state of the proteins (Itzhaki et al., 2006; Akiva et al., 2008; Hashimoto and Panchenko, 2010). Based on these studies, it is expected that the identified CSIs in the RNR homologs of bifidobacteria will also play novel and functionally important roles that are specific for bifidobacteria. There is no information available at present regarding the biochemical properties of the RNR homologs from bifidobacteria or whether they exhibit any novel functional characteristics. However, the identified CSIs provide highly specific tools for the genetic and biochemical exploration of novel functional characteristics of the RNRs from bifidobacteria.

Of the five different CSIs identified within the NrdD homologs, one of these CSIs is a large 43 aa insert, located proximal to the allosteric regulatory site in the protein monomer (**Figure 4B** and Supplementary Figure 9B) (Larsson et al., 2001). This regulatory site consists of a 4-helix bundle where each monomer contributes two helices. These helices are significantly longer in class III RNRs compared to those of class I (~21 aa), resulting in an altered dimer packing, mode of effector binding and consequential conformational changes to the active site (Larsson et al., 2001). Due to the presence of the 43 aa CSI in the class III bifidobacterial RNRs in this region, this helical region is further elongated in bifidobacteria, which is suggestive of a feature uniquely shared by these bacteria. In the absence of any information regarding the structure of the NrdD protein from bifidobacteria, or the conformation of this large CSI in the protein structure, it is difficult to predict the precise role that this CSI may play in the RNR structure and/or function. However, the results of our preliminary *in silico* protein–protein

docking using three separate docking servers suggest that when the CSI is oriented in such a way that it can interact with the other monomer, the stability of the dimer complex is improved in the presence of the 43 aa CSI, in comparison to the similar docking studies carried out with the protein lacking the CSI (**Figure 4B** and **Table 1**). The results from our *in silico* analyses are broadly suggestive of one possible function of this large CSI. However, a clearer understanding of the functional significances of the identified CSIs should emerge from future biochemical and structural studies on classes I and III RNRs. Nevertheless, the results presented here highlight the many unique sequence features of the bifidobacterial RNRs, whose further investigations could provide important insights into novel functional aspects of these enzymes in bifidobacteria.

## AUTHOR CONTRIBUTIONS

SA identified some of the conserved inserts, carried out phylogenetic and structural work and wrote the draft manuscript. BK was involved in the structural analysis of the identified conserved inserts and read and commented on the draft manuscript. RG identified the initial CSIs and conceived and directed the entire project and edited the final version of the manuscript.

## FUNDING

This work was supported by the research grant number 249924 from the Natural Sciences and Engineering Research Council of Canada.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2017.01409/full#supplementary-material>

## REFERENCES

- Akiva, E., Itzhaki, Z., and Margalit, H. (2008). Built-in loops allow versatility in domain-domain interactions: lessons from self-interacting domains. *Proc. Natl. Acad. Sci. U.S.A.* 105, 13292–13297. doi: 10.1073/pnas.0801207105
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389
- Alves, P., Castro, J., Sousa, C., Cereija, T. B., and Cerca, N. (2014). *Gardnerella vaginalis* outcompetes 29 other bacterial species isolated from patients with bacterial vaginosis, using an *in vitro* biofilm formation model. *J. Infect. Dis.* 210, 593–596. doi: 10.1093/infdis/jiu131
- Aurelius, O., Johansson, R., Bagenholm, V., Lundin, D., Tholander, F., Balhuizen, A., et al. (2015). The crystal structure of *Thermotoga maritima* class III ribonucleotide reductase lacks a radical cysteine pre-positioned in the active site. *PLoS ONE* 10:e0128199. doi: 10.1371/journal.pone.0128199
- Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., et al. (2017). GenBank. *Nucleic Acids Res.* 45, D37–D42. doi: 10.1093/nar/gkw1070
- Biavati, B., and Mattarelli, P. (2006). “The family *Bifidobacteriaceae*,” in *The Prokaryotes*, eds M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer, and E. Stackebrandt (New York, NY: Springer), 322–382.
- Biavati, B., Vescovo, M., Torriani, S., and Bottazzi, V. (2000). Bifidobacteria: history, ecology, physiology and applications. *Ann. Microbiol.* 50, 117–132.
- Bollinger, J. M. Jr., Jiang, W., Green, M. T., and Krebs, C. (2008). The manganese(IV)/iron(III) cofactor of *Chlamydia trachomatis* ribonucleotide reductase: structure, assembly, radical initiation, and evolution. *Curr. Opin. Struct. Biol.* 18, 650–657. doi: 10.1016/j.sbi.2008.11.007
- Bowie, J. U., Luthy, R., and Eisenberg, D. (1991). A method to identify protein sequences that fold into a known three-dimensional structure. *Science* 253, 164–170. doi: 10.1126/science.1853201
- Bradshaw, C. S., Tabrizi, S. N., Fairley, C. K., Morton, A. N., Rudland, E., and Garland, S. M. (2006). The association of *Atopobium vaginae* and *Gardnerella vaginalis* with bacterial vaginosis and recurrence after oral metronidazole therapy. *J. Infect. Dis.* 194, 828–836. doi: 10.1086/506621
- Brown, N. C., and Reichard, P. (1969). Role of effector binding in allosteric control of ribonucleoside diphosphate reductase. *J. Mol. Biol.* 46, 39–55. doi: 10.1016/0022-2836(69)90056-4

- Buchan, D. W., Minneci, F., Nugent, T. C., Bryson, K., and Jones, D. T. (2013). Scalable web services for the PSIPRED protein analysis workbench. *Nucleic Acids Res.* 41, W349–W357. doi: 10.1093/nar/gkt381
- Chatterji, M., Unniraman, S., Maxwell, A., and Nagaraja, V. (2000). The additional 165 amino acids in the B protein of *Escherichia coli* DNA gyrase have an important role in DNA binding. *J. Biol. Chem.* 275, 22888–22894. doi: 10.1074/jbc.M001047200
- Chaudhury, S., Berrondo, M., Weitzner, B. D., Muthu, P., Bergman, H., and Gray, J. J. (2011). Benchmarking and analysis of protein docking performance in Rosetta v3.2. *PLoS ONE* 6:e22477. doi: 10.1371/journal.pone.0022477
- Cherkasov, A., Lee, S. J., Nandan, D., and Reiner, N. E. (2006). Large-scale survey for potentially targetable indels in bacterial and protozoan proteins. *Proteins* 62, 371–380. doi: 10.1002/prot.20631
- Clarke, J. H., and Irvine, R. F. (2013). Evolutionarily conserved structural changes in phosphatidylinositol 5-phosphate 4-kinase (PI5P4K) isoforms are responsible for differences in enzyme activity and localization. *Biochem. J.* 454, 49–57. doi: 10.1042/BJ20130488
- Colovos, C., and Yeates, T. O. (1993). Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci.* 2, 1511–1519. doi: 10.1002/pro.5560020916
- Comeau, S. R., Gatchell, D. W., Vajda, S., and Camacho, C. J. (2004). ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics* 20, 45–50. doi: 10.1093/bioinformatics/btg371
- Cotruvo, J. A. Jr., and Stubbe, J. (2008). NrdI, a flavodoxin involved in maintenance of the diferric-tyrosyl radical cofactor in *Escherichia coli* class Ib ribonucleotide reductase. *Proc. Natl. Acad. Sci. U.S.A.* 105, 14383–14388. doi: 10.1073/pnas.0807348105
- Crona, M., Torrents, E., Rohr, A. K., Hofer, A., Furrer, E., Tomterm, A. B., et al. (2011). NrdH-redoxin protein mediates high enzyme activity in manganese-reconstituted ribonucleotide reductase from *Bacillus anthracis*. *J. Biol. Chem.* 286, 33053–33060. doi: 10.1074/jbc.M111.278119
- Cronin, M., Ventura, M., Fitzgerald, G. F., and van Sinderen, D. (2011). Progress in genomics, metabolism and biotechnology of bifidobacteria. *Int. J. Food Microbiol.* 149, 4–18. doi: 10.1016/j.ijfoodmicro.2011.01.019
- Duhovny, D., Nussinov, R., and Wolfson, H. J. (2002). “Efficient unbound docking of rigid molecules,” in *Proceedings of the Second International Workshop on Algorithms in Bioinformatics (WABI) 2452* (Berlin: Springer-Verlag), 185–200. doi: 10.1007/3-540-45784-4\_14
- Edgar, R., Domrachev, M., and Lash, A. E. (2002). Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30, 207–210. doi: 10.1093/nar/30.1.207
- Eklund, H., Uhlin, U., Farnegardh, M., Logan, D. T., and Nordlund, P. (2001). Structure and function of the radical enzyme ribonucleotide reductase. *Prog. Biophys. Mol. Biol.* 77, 177–268. doi: 10.1016/S0079-6107(01)00014-1
- Eliasson, R., Pontis, E., Fontecave, M., Gerez, C., Harder, J., Jörnvall, H., et al. (1992). Characterization of components of the anaerobic ribonucleotide reductase system from *Escherichia coli*. *J. Biol. Chem.* 267, 25541–25547.
- Eliasson, R., Pontis, E., Jordan, A., and Reichard, P. (1999). Allosteric control of three B12-dependent (class II) ribonucleotide reductases. Implications for the evolution of ribonucleotide reduction. *J. Biol. Chem.* 274, 7182–7189. doi: 10.1074/jbc.274.11.7182
- Eriksson, M., Uhlin, U., Ramaswamy, S., Ekberg, M., Regnström, K., and Sjöberg, B. M. (1997). Binding of allosteric effectors to ribonucleotide reductase protein R1: reduction of active-site cysteines promotes substrate binding. *Structure* 5, 1077–1092. doi: 10.1016/S0969-2126(97)00259-1
- Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., and Shen, M. Y. (2007). Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein Sci.* 15:5.6.1–5.6.30. doi: 10.1002/0471140864.ps0209s50
- Gao, B., and Gupta, R. S. (2012). Phylogenetic framework and molecular signatures for the main clades of the phylum *Actinobacteria*. *Microbiol. Mol. Biol. Rev.* 76, 66–112. doi: 10.1128/MMBR.05011-11
- Garriga, X., Eliasson, R., Torrents, E., Jordan, A., Gibert, I., and Reichard, P. (1996). nrdD and nrdG genes are essential for strict anaerobic growth of *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 229, 189–192. doi: 10.1006/bbrc.1996.1778
- Geszvain, K., Gruber, T. M., Mooney, R. A., Gross, C. A., and Landick, R. (2004). A hydrophobic patch on the flap-tip helix of *E. coli* RNA polymerase mediates sigma(70) region 4 function. *J. Mol. Biol.* 343, 569–587. doi: 10.1016/j.jmb.2004.08.063
- Gibson, G. R., Beatty, E. R., Wang, X., and Cummings, J. H. (1995). Selective stimulation of bifidobacteria in the human colon by oligofructose and inulin. *Gastroenterology* 108, 975–982. doi: 10.1016/0016-5085(95)90192-2
- Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., et al. (2010). A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res.* 38, W695–W699. doi: 10.1093/nar/gkq313
- Gupta, R. S. (2014). “Chapter 8 - Identification of conserved indels that are useful for classification and evolutionary studies,” in *Methods in Microbiology: New Approaches to Prokaryotic Systematics*, eds M. Goodfellow, I. Sutcliffe, and J. Chun (Cambridge, MA: Academic Press), 153–182.
- Gupta, R. S. (2016). Impact of genomics on the understanding of microbial evolution and classification: the importance of Darwin’s views on classification. *FEMS Microbiol. Rev.* 40, 520–553. doi: 10.1093/femsre/fuw011
- Gupta, R. S., Nanda, A., and Khadka, B. (2017). Novel molecular, structural and evolutionary characteristics of the phosphoketolases from bifidobacteria and *Coriobacteriales*. *PLoS ONE* 12:e0172176. doi: 10.1371/journal.pone.0172176
- Hashimoto, K., and Panchenko, A. R. (2010). Mechanisms of protein oligomerization, the critical role of insertions and deletions in maintaining different oligomeric states. *Proc. Natl. Acad. Sci. U.S.A.* 107, 20352–20357. doi: 10.1073/pnas.1012999107
- Heo, L., Park, H., and Seok, C. (2013). GalaxyRefine: protein structure refinement driven by side-chain repacking. *Nucleic Acids Res.* 41, W384–W388. doi: 10.1093/nar/gkt458
- Itzhaki, Z., Akiva, E., Altuvia, Y., and Margalit, H. (2006). Evolutionary conservation of domain-domain interactions. *Genome Biol.* 7:R125. doi: 10.1186/gb-2006-7-12-r125
- Jiang, W., Yun, D., Saleh, L., Barr, E. W., Xing, G., Hoffart, L. M., et al. (2007). A manganese(IV)/iron(III) cofactor in *Chlamydia trachomatis* ribonucleotide reductase. *Science* 316, 1188–1191. doi: 10.1126/science.1141179
- Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* 292, 195–202. doi: 10.1006/jmbi.1999.3091
- Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* 8, 275–282. doi: 10.1093/bioinformatics/8.3.275
- Jordan, A., Aslund, F., Pontis, E., Reichard, P., and Holmgren, A. (1997). Characterization of *Escherichia coli* NrdH. A glutaredoxin-like protein with a thioredoxin-like activity profile. *J. Biol. Chem.* 272, 18044–18050. doi: 10.1074/jbc.272.29.18044
- Jordan, A., Pontis, E., Aslund, F., Hellman, U., Gibert, I., and Reichard, P. (1996). The ribonucleotide reductase system of *Lactococcus lactis*. Characterization of an NrdEF enzyme and a new electron transport protein. *J. Biol. Chem.* 271, 8779–8785. doi: 10.1074/jbc.271.15.8779
- Kenyon, C. R., and Osbak, K. (2014). Recent progress in understanding the epidemiology of bacterial vaginosis. *Curr. Opin. Obstet. Gynecol.* 26, 448–454. doi: 10.1097/GCO.0000000000000112
- Khadka, B., and Gupta, R. S. (2017). Identification of a conserved 8 aa insert in the PIP5K protein in the *Saccharomycetaceae* family of fungi and the molecular dynamics simulations and structural analysis to investigate its potential functional role. *Proteins* doi: 10.1002/prot.25306 [Epub ahead of print].
- Kozakov, D., Brenke, R., Comeau, S. R., and Vajda, S. (2006). PIPER: an FFT-based protein docking program with pairwise potentials. *Proteins* 65, 392–406. doi: 10.1002/prot.21117
- Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372, 774–797. doi: 10.1016/j.jmb.2007.05.022
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404
- Larsson, K. M., Andersson, J., Sjöberg, B. M., Nordlund, P., and Logan, D. T. (2001). Structural basis for allosteric substrate specificity regulation in anaerobic ribonucleotide reductases. *Structure* 9, 739–750. doi: 10.1016/S0969-2126(01)00627-X

- Larsson, K. M., Logan, D. T., and Nordlund, P. (2010). Structural basis for adenosylcobalamin activation in AdoCbl-dependent ribonucleotide reductases. *ACS Chem. Biol.* 5, 933–942. doi: 10.1021/cb1000845
- Leahy, S. C., Higgins, D. G., Fitzgerald, G. F., and van Sinderen, D. (2005). Getting better with bifidobacteria. *J. Appl. Microbiol.* 98, 1303–1315. doi: 10.1111/j.1365-2672.2005.02600.x
- Lee, G. R., Heo, L., and Seok, C. (2016). Effective protein model structure refinement by loop modeling and overall relaxation. *Proteins* 84(Suppl. 1), 293–301. doi: 10.1002/prot.24858
- Logan, D. T., Andersson, J., Sjöberg, B. M., and Nordlund, P. (1999). A glycol radical site in the crystal structure of a class III ribonucleotide reductase. *Science* 283, 1499–1504. doi: 10.1126/science.283.5407.1499
- Lovell, S. C., Davis, I. W., Arendall, W. B. III, de Bakker, P. I., Word, J. M., Prisant, M. G., et al. (2003). Structure validation by Alpha geometry: phi, psi and Cbeta deviation. *Proteins* 50, 437–450. doi: 10.1002/prot.10286
- Lundin, D., Gribaldo, S., Torrents, E., Sjöberg, B. M., and Poole, A. M. (2010). Ribonucleotide reduction - horizontal transfer of a required function spans all three domains. *BMC Evol. Biol.* 10:383. doi: 10.1186/1471-2148-10-383
- Lundin, D., Torrents, E., Poole, A. M., and Sjöberg, B. M. (2009). RNRdb, a curated database of the universal enzyme family ribonucleotide reductase, reveals a high level of misannotation in sequences deposited to Genbank. *BMC Genomics* 10:589. doi: 10.1186/1471-2164-10-589
- Luthy, R., Bowie, J. U., and Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature* 356, 83–85. doi: 10.1038/356083a0
- Lyskov, S., Chou, F. C., Conchuir, S. O., Der, B. S., Drew, K., Kuroda, D., et al. (2013). Serverification of molecular modeling applications: the Rosetta Online Server that Includes Everyone (ROSIE). *PLoS ONE* 8:e63906. doi: 10.1371/journal.pone.0063906
- Lyskov, S., and Gray, J. J. (2008). The RosettaDock server for local protein-protein docking. *Nucleic Acids Res.* 36, W233–W238. doi: 10.1093/nar/gkn216
- Marchler-Bauer, A., and Bryant, S. H. (2004). CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.* 32, W327–W331. doi: 10.1093/nar/gkh454
- Masco, L., Huys, G., De Brandt, E., Temmerman, R., and Swings, J. (2005). Culture-dependent and culture-independent qualitative analysis of probiotic products claimed to contain bifidobacteria. *Int. J. Food Microbiol.* 102, 221–230. doi: 10.1016/j.ijfoodmicro.2004.11.018
- Milani, C., Lugli, G. A., Duranti, S., Turrioni, F., Bottacini, F., Mangifesta, M., et al. (2014). Genomic encyclopedia of type strains of the genus *Bifidobacterium*. *Appl. Environ. Microbiol.* 80, 6290–6302. doi: 10.1128/AEM.02308-14
- Milani, C., Turrioni, F., Duranti, S., Lugli, G. A., Mancabelli, L., Ferrario, C., et al. (2015). Genomics of the genus *Bifidobacterium* reveals species-specific adaptation to the glycan-rich gut environment. *Appl. Environ. Microbiol.* 82, 980–991. doi: 10.1128/AEM.03500-15
- Mills, S., Stanton, C., Fitzgerald, G. F., and Ross, R. P. (2011). Enhancing the stress responses of probiotics for a lifestyle from gut to product and back again. *Microb. Cell Fact.* 10(Suppl. 1):S19. doi: 10.1186/1475-2859-10-S1-S19
- Nordlund, P., and Reichard, P. (2006). Ribonucleotide reductases. *Annu. Rev. Biochem.* 75, 681–706. doi: 10.1146/annurev.biochem.75.103004.142443
- Oberg, T. S., Steele, J. L., Ingham, S. C., Smeianov, V. V., Briczinski, E. P., Abdalla, A., et al. (2011). Intrinsic and inducible resistance to hydrogen peroxide in *Bifidobacterium* species. *J. Ind. Microbiol. Biotechnol.* 38, 1947–1953. doi: 10.1007/s10295-011-0983-y
- Ollagnier, S., Mulliez, E., Gaillard, J., Eliasson, R., Fontecave, M., and Reichard, P. (1996). The anaerobic *Escherichia coli* ribonucleotide reductase. Subunit structure and iron sulfur center. *J. Biol. Chem.* 271, 9410–9416. doi: 10.1074/jbc.271.16.9410
- Palframan, R. J., Gibson, G. R., and Rastall, R. A. (2003). Carbohydrate preferences of *Bifidobacterium* species isolated from the human gut. *Curr. Issues Intest. Microbiol.* 4, 71–75.
- Petersson, L., Graslund, A., Ehrenberg, A., Sjöberg, B. M., and Reichard, P. (1980). The iron center in ribonucleotide reductase from *Escherichia coli*. *J. Biol. Chem.* 255, 6706–6712.
- Poole, A. M., Logan, D. T., and Sjöberg, B. M. (2002). The evolution of the ribonucleotide reductases: much ado about oxygen. *J. Mol. Evol.* 55, 180–196. doi: 10.1007/s00239-002-2315-3
- Reichard, P. (1993). From RNA to DNA, why so many ribonucleotide reductases? *Science* 260, 1773–1777.
- Reichard, P. (2010). Ribonucleotide reductases: substrate specificity by allostery. *Biochem. Biophys. Res. Commun.* 396, 19–23. doi: 10.1016/j.bbrc.2010.02.108
- Roca, I., Torrents, E., Sahlin, M., Gibert, I., and Sjöberg, B. M. (2008). NrdI essentiality for class Ib ribonucleotide reduction in *Streptococcus pyogenes*. *J. Bacteriol.* 190, 4849–4858. doi: 10.1128/JB.00185-08
- Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* 5, 725–738. doi: 10.1038/nprot.2010.5
- Sanders, M. E. (1998). Development of consumer probiotics for the US market. *Br. J. Nutr.* 80, S213–S218.
- Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., and Wolfson, H. J. (2005). PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.* 33, W363–W367. doi: 10.1093/nar/gki481
- Schoeffler, A. J., May, A. P., and Berger, J. M. (2010). A domain insertion in *Escherichia coli* GyrB adopts a novel fold that plays a critical role in gyrase function. *Nucleic Acids Res.* 38, 7830–7844. doi: 10.1093/nar/gkq665
- Schrödinger, L. (2016). *The PyMOL Molecular Graphics System, Version 1.8*. Available at: <https://www.pymol.org/>
- Shen, M. Y., and Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. *Protein Sci.* 15, 2507–2524. doi: 10.1110/ps.062416606
- Singh, B., and Gupta, R. S. (2009). Conserved inserts in the Hsp60 (GroEL) and Hsp70 (DnaK) proteins are essential for cellular growth. *Mol. Genet. Genomics* 281, 361–373. doi: 10.1007/s00438-008-0417-3
- Sintchak, M. D., Arjara, G., Kellogg, B. A., Stubbe, J., and Drennan, C. L. (2002). The crystal structure of class II ribonucleotide reductase reveals how an allosterically regulated monomer mimics a dimer. *Nat. Struct. Biol.* 9, 293–300. doi: 10.1038/nsb774
- Sippl, M. J., Lackner, P., Domingues, F. S., and Koppensteiner, W. A. (1999). An attempt to analyse progress in fold recognition from CASP1 to CASP3. *Proteins Suppl.* 3, 226–230. doi: 10.1002/(SICI)1097-0134(1999)37:3+<226::AID-PROT29>3.0.CO;2-Z
- Smith, S. M., Ogbara, T., and Eng, R. H. (1992). Involvement of *Gardnerella vaginalis* in urinary tract infections in men. *J. Clin. Microbiol.* 30, 1575–1577.
- Sun, X., Eliasson, R., Pontis, E., Andersson, J., Buist, G., Sjöberg, B. M., et al. (1995). Generation of the glycol radical of the anaerobic *Escherichia coli* ribonucleotide reductase requires a specific activating enzyme. *J. Biol. Chem.* 270, 2443–2446. doi: 10.1074/jbc.270.6.2443
- Tamao, Y., and Blakley, R. L. (1973). Direct spectrophotometric observation of an intermediate formed from deoxyadenosylcobalamin in ribonucleotide reduction. *Biochemistry* 12, 24–34. doi: 10.1021/bi00725a005
- Tamura, K., Stecher, G., Peterson, D., Filipinski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197
- Thelander, L., and Reichard, P. (1979). Reduction of ribonucleotides. *Annu. Rev. Biochem.* 48, 133–158. doi: 10.1146/annurev.bi.48.070179.001025
- Torrents, E. (2014). Ribonucleotide reductases: essential enzymes for bacterial life. *Front. Cell Infect. Microbiol.* 4:52. doi: 10.3389/fcimb.2014.00052
- Torrents, E., Aloy, P., Gibert, I., and Rodriguez-Trelles, F. (2002). Ribonucleotide reductases: divergent evolution of an ancient enzyme. *J. Mol. Evol.* 55, 138–152. doi: 10.1007/s00239-002-2311-7
- Torrents, E., Eliasson, R., Wolpher, H., Graslund, A., and Reichard, P. (2001). The anaerobic ribonucleotide reductase from *Lactococcus lactis*. Interactions between the two proteins NrdD and NrdG. *J. Biol. Chem.* 276, 33488–33494. doi: 10.1074/jbc.M103743200
- Torrents, E., Jordan, A., Karlsson, M., and Gibert, I. (2000). Occurrence of multiple ribonucleotide reductase classes in gamma-proteobacteria species. *Curr. Microbiol.* 41, 346–351. doi: 10.1007/s002840010147
- Turrioni, F., Ribbera, A., Foroni, E., van Sinderen, D., and Ventura, M. (2008). Human gut microbiota and bifidobacteria: from composition to functionality. *Antonie Van Leeuwenhoek* 94, 35–50. doi: 10.1007/s10482-008-9232-4
- Turrioni, F., van Sinderen, D., and Ventura, M. (2009). Bifidobacteria: from ecology to genomics. *Front. Biosci.* 14:4673–4684. doi: 10.2741/3559
- Turrioni, F., Ventura, M., Butto, L. F., Duranti, S., O'Toole, P. W., Motherway, M. O., et al. (2014). Molecular dialogue between the human gut microbiota and the host: a *Lactobacillus* and *Bifidobacterium* perspective. *Cell Mol. Life Sci.* 71, 183–203. doi: 10.1007/s00018-013-1318-0

- Uhlin, U., and Eklund, H. (1994). Structure of ribonucleotide reductase protein R1. *Nature* 370, 533–539. doi: 10.1038/370533a0
- Uppsten, M., Farnegardh, M., Domkin, V., and Uhlin, U. (2006). The first holocomplex structure of ribonucleotide reductase gives new insight into its mechanism of action. *J. Mol. Biol.* 359, 365–377. doi: 10.1016/j.jmb.2006.03.035
- Uppsten, M., Farnegardh, M., Jordan, A., Eliasson, R., Eklund, H., and Uhlin, U. (2003). Structure of the large subunit of class Ib ribonucleotide reductase from *Salmonella typhimurium* and its complexes with allosteric effectors. *J. Mol. Biol.* 330, 87–97. doi: 10.1016/S0022-2836(03)00538-2
- Ventura, M., Canchaya, C., Tauch, A., Chandra, G., Fitzgerald, G. F., Chater, K. F., et al. (2007). Genomics of *Actinobacteria*: tracing the evolutionary history of an ancient phylum. *Microbiol. Mol. Biol. Rev.* 71, 495–548. doi: 10.1128/MMBR.00005-07
- Ventura, M., O'Flaherty, S., Claesson, M. J., Turrone, F., Klaenhammer, T. R., van Sinderen, D., et al. (2009). Genome-scale analyses of health-promoting bacteria: probiogenomics. *Nat. Rev. Microbiol.* 7, 61–71. doi: 10.1038/nrmicro2047
- Ventura, M., Turrone, F., Lugli, G. A., and van Sinderen, D. (2014). Bifidobacteria and humans: our special friends, from ecological to genomics perspectives. *J. Sci. Food Agric.* 94, 163–168. doi: 10.1002/jsfa.6356
- Wei, Y., Thompson, J., and Floudas, C. A. (2011). CONCORD: a consensus method for protein secondary structure prediction via mixed integer linear optimization. *Proc. R. Soc. A Math. Phys. Eng. Sci.* 468, 831–850. doi: 10.1098/rspa.2011.0514
- Whelan, S., and Goldman, N. (2001). A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.* 18, 691–699. doi: 10.1093/oxfordjournals.molbev.a003851
- Wiederstein, M., and Sippl, M. J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* 35, W407–W410. doi: 10.1093/nar/gkm290
- Xiao, M., Xu, P., Zhao, J., Wang, Z., Zuo, F., Zhang, J., et al. (2011). Oxidative stress-related responses of *Bifidobacterium longum* subsp. *longum* BBMN68 at the proteomic level after exposure to oxygen. *Microbiology* 157, 1573–1588. doi: 10.1099/mic.0.044297-0
- Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. *Nat. Methods* 12, 7–8. doi: 10.1038/nmeth.3213
- Zhang, G., Gao, B., Adeolu, M., Khadka, B., and Gupta, R. S. (2016). Phylogenomic analyses and comparative studies on genomes of the *Bifidobacteriales*: identification of molecular signatures specific for the order *Bifidobacteriales* and its different subclades. *Front. Microbiol.* 7:978. doi: 10.3389/fmicb.2016.00978
- Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9:40. doi: 10.1186/1471-2105-9-40
- Zhi, X. Y., Li, W. J., and Stackebrandt, E. (2009). An update of the structure and 16S rRNA gene sequence-based definition of higher ranks of the class Actinobacteria, with the proposal of two new suborders and four new families and emended descriptions of the existing higher taxa. *Int. J. Syst. Evol. Microbiol.* 59, 589–608. doi: 10.1099/ijs.0.65780-0
- Zuo, F., and Chen, S. (2016). Recently performed Next generation sequencing based transcriptomic analysis of oxidative stress response in *Bifidobacterium longum* BBMN68 (data accessible at NCBI GEO database (Edgar et al., 2002), accession GSE65320). NCBI GEO database. Available at: <http://www.ncbi.nlm.nih.gov/geo/>

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer PL and handling Editor declared their shared affiliation, and the handling Editor states that the process nevertheless met the standards of a fair and objective review.

Copyright © 2017 Alnajjar, Khadka and Gupta. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.