Chimia 56 (2002) 270–273

ISSN 0009–4293

Phosphoketolase, a Neglected Enzyme of Microbial Carbohydrate Metabolism

Lukas M. Rohr*, Michael Teuber, and Leo Meile

Abstract: Phosphoketolases are thiamine diphosphate (ThDP) dependent enzymes of the phosphoketolase (PK) pathway of heterofermentative and facultatively homofermentative lactic acid bacteria and of the fructose 6-phosphate shunt of bifidobacteria. PK activity was also measured in protein extracts of other microorganisms including yeasts. The dual substrate xylulose 5-phosphate/fructose 6-phosphate phosphoketolase (Xfp) from the 'probiotic' *Bifidobacterium lactis* was purified, and its encoding gene (*xfp*) was cloned and sequenced. Comparisons with public databases revealed an unexpectedly wide spread of more than 30 homologous *xfp* sequences in the kingdom of the bacteria, but not of the archaea. We assigned amino acid motifs typically found in PKs. Two of them (G-P-G-H-G and E-G-G-E-L-G-Y, respectively) discriminate PKs from transketolases, which have at least one ThDP binding site in common. On the basis of further comparative analyses we conclude that the PK prevalence among diverse organisms is due to longitudinal and not horizontal gene transmission.

Keywords: Bifidobacteria · Carbohydrate metabolism · Lactic acid bacteria · Phosphoketolase · Thiamine diphosphate

1. Introduction

Phosphoketolases (EC 4.1.2.9, EC 4.1.2.22) are key enzymes of the phosphoketolase (PK) pathway of heterofermentative and facultatively homofermentative lactic acid bacteria and of the D-fructose 6-phosphate shunt of bifidobacteria, which are used as 'probiotic' supplements in food [1]. PK activity has been sporadically reported in other microorganisms including eukaryotic yeasts. The study of their presence, specificities and roles in other microbes has been hampered by the fact that the substrate D-xylulose 5-phosphate, as long as it was commercially available, was quite expensive, but as of the beginning of 2001, it is no longer on the market.

PKs catalyse an irreversible thiamine diphosphate (ThDP) dependent phosphorolytic reaction splitting D-xylulose 5-phosphate (EC 4.1.2.9) or D-fructose 6-phosphate (EC 4.1.2.22) in the presence of inorganic phosphate to yield acetyl phosphate and glycerol 3-phosphate or erythrose 4-phosphate, respectively (Fig. 1) [2]. The exact interaction mechanism of the ThDP moiety with its binding partners has not been investigated so far, but it was assumed to be similar to that of transketolases with respect to the formation of a Schiff's base between the ThDP molecule and the carbohydrate substrate on the carbonyl carbon at which cleavage occurs [3]. For anaerobic bacteria, the enzyme product acetyl phosphate represents a good and quickly available energy source by producing either acetate plus ATP from ADP or S-acetyl-CoA from CoASH. It was proposed by Buckel [4] that two distinct enzymes catalyse these reactions, namely a fructose 6-phosphate dependent phosphoketolase-1 and a xylulose 5-phosphate phosphoketolase-2, respectively. Among bifidobacteria, former work which had related mostly on enzyme activities from crude or partially purified extracts revealed some evidence for this hypothesis, but also to the contradictory conclusion presenting dual substrate specificity for a single PK enzyme [5]. Since bifidobacteria are supposed to metabolise certain sugars (so-called prebiotic substrates) in the human intestinal tract, the key degradation enzymes should be investigated by modern techniques. Therefore, due to the lack of any pure enzyme and known gene encoding PK, we approached the gaps at the molecular level.

Purification of a PK from a *Bifidobacterium* species and subsequent cloning of its gene has opened the way to look for homologous genes in microorganisms whose genomes have been or are presently being sequenced [6]. Surprisingly, homologous sequences were detected in publicly available databases for many microorganisms other than the lactic acid bacteria or bifidobacteria. In this communication, we present the phylogenetic distribution of PK encoding genes in the bacterial and eukaryotic kingdoms proving the potential of the chemical structure (sequence) of DNA to search for genetic information for a known enzyme in biochemically not investigated species.

*Correspondence: L.M. Rohr
Laboratory of Food Microbiology
Institute of Food Science and Nutrition
ETH Zentrum, LFO G
Schmelzbergstrasse 9
CH–8092 Zurich
Tel.: +41 1 632 33 63
Fax: +41 1 632 12 66
E-Mail: lukas.rohr@alumni.ethz.ch

2. Results and Discussion

2.1. *Cloning and Sequencing of a Bifidobacterial Phosphoketolase*

Starting from a crude cell extract, the main fructose 6-phosphate cleaving enzyme from the 'probiotic' *Bifidobacterium lactis* was purified to homogeneity [6]. Its apparent molecular weight was 550,000 when estimated by gel filtration chromatography. On sodium dodecyl sulfate polyacrylamide gel electrophoresis, the purified PK displayed a single band with an apparent molecular mass of approximately 90,000 Da. Therefore, the native enzyme is likely to be a homohexamer. Enzyme assays with a purified protein preparation revealed that *B. lactis* possesses a dual substrate PK combining the substrate specificities of both EC 4.1.2.9 and EC 4.1.2.22. The corresponding gene, named *xfp* (*x*ylulose 5-phosphate/*f*ructose 6-phosphate *p*hosphoketolase), was identified on the *B. lactis* chromosome by using molecular biological techniques. A cloned 4.1-kb DNA fragment contained the *xfp* gene. The molecular mass of the Xfp polypeptide was calculated to be 92,529 Da, a value coincident with that estimated by gel electrophoresis. In the crude extract of *Escherichia coli* harbouring the gene on a plasmid, we found a strong PK activity, whereas no activity could be detected in the absence of the plasmid [7]. These results definitely confirmed that the cloned *xfp* gene codes for the purified enzyme.

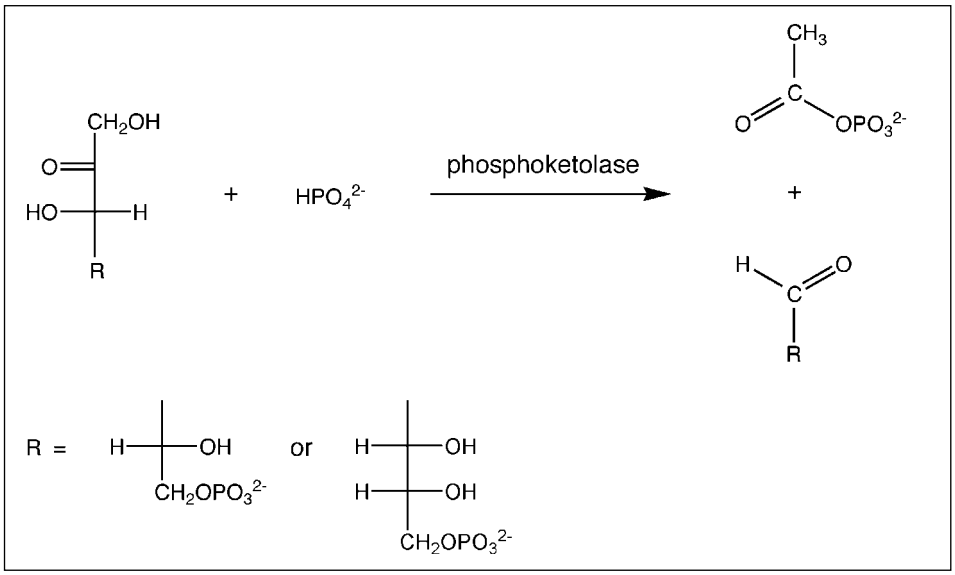

Fig. 1. Reaction of thiamine diphosphate (ThDP) dependent phosphoketolases.

2.2. *Distribution of Phosphoketolase Gene Homologues in Microbial Genomes*

The enormous advances in the sequencing of whole genomes from a broad range of organisms (currently around 550 completed or ongoing projects) and the availability of the respective data enables the prevalence of a particular nucleotide or amino acid sequence to be investigated. Comparison of the *B. lactis* PK amino acid sequence with public databases revealed homologous putative proteins in an unexpectedly large number of different microorganisms. Similar sequences with up to 55% amino acid identity were identified in the genomes of three fungi (*Aspergillus fumigatus*, *Neurospora crassa* (2 copies), and *Schizosaccharomyces pombe*) and 25 bacteria, but not in the genomes of archaea, plants or animals. Especially the presence of homologues in the N_2-fixing cyanobacteria (*Synechocystis* sp. PCC6803, *Anabaena* sp. PCC7120 (2 copies), and *Nostoc punctiforme*) as well as in chemoautotrophic species (*Nitrosomonas europaea* and *Acidithiobacillus ferrooxidans*) is noteworthy. A multiple alignment of these putative PK sequences shows several highly conserved regions. The consensus of a segment near the N-terminus, calculated from 31 amino acid sequences, is depicted in Fig. 2. This section, in which 26% out of 157 amino acids are identical in all the sequences, contains a pattern matching closely a signature sequence common for ThDP binding enzymes ([LIVMF]-G-X_5-P-X_4-V-X-I-X-G-D-G-[GSAC], Prosite acc. no. PS00187). Two further consensus patterns, G-P-G-H-G and E-G-G-E-L-G-Y, respectively, are located near the cofactor binding site. Their functional role is still unknown, but they do not seem to be present in any other proteins, therefore they are proposed as signature sequences for the new family of Xfp phosphoketolases. (Since a slightly homologous sequence in the gen-ome of *Pseudomonas aeruginosa* does not contain a proper ThDP binding site, it was excluded from the above analysis.)

To elucidate the wide spread of the *xfp* gene, a phylogenetic tree was constructed based on the 16S rRNA sequences of bacterial organisms possessing a putative *xfp* gene and of species showing no *xfp* homologues in their completely sequenced genomes (Fig. 3). Contrary to the hitherto assumption that the PK and the fructose 6-phosphate shunts were specific for a small group of bacteria and yeasts, the enzyme seems to be prevalent in the bacterial world, though not ubiquitous.

... - G(63) - H - [FW] - G - [ST] - X_6 - [LIMF] - X_6 - N - X_2 - I - X_7 - [LIVMF] - X - [LIVM] - X - G - P - G - H - G - X_{26-27} - G - [LIM] - X_2 - [LIF] - X_3 - F - S - X_3 - G - [IVFY] - X - S - H - X_5 - P - X_2 - [LI] - X - E - G - G - E - L - G - Y - X - [LIV] - X_4 - G - X_3 - D - X - P - X_2 - [LIVF] - X_3 - [LIVM] - X - G - D - G - E - [AS] - E - T - X_6 - W - X_4 - [LFY] - [LIV] - X - P - X_3 - G - X - [IV] - [LI] - P - [LIV] - [LV] - X - [LIV] - N - G - X - K - I(219) - ...

Fig. 2. Consensus of an N-terminal 157-amino acid segment, calculated from 31 microbial phosphoketolase (PK) homologues. The shaded parts mark a ThDP binding site known from transketolases (framed) and two proposed PK signature sequences, respectively. Indicated positions refer to the *B. lactis* Xfp protein.

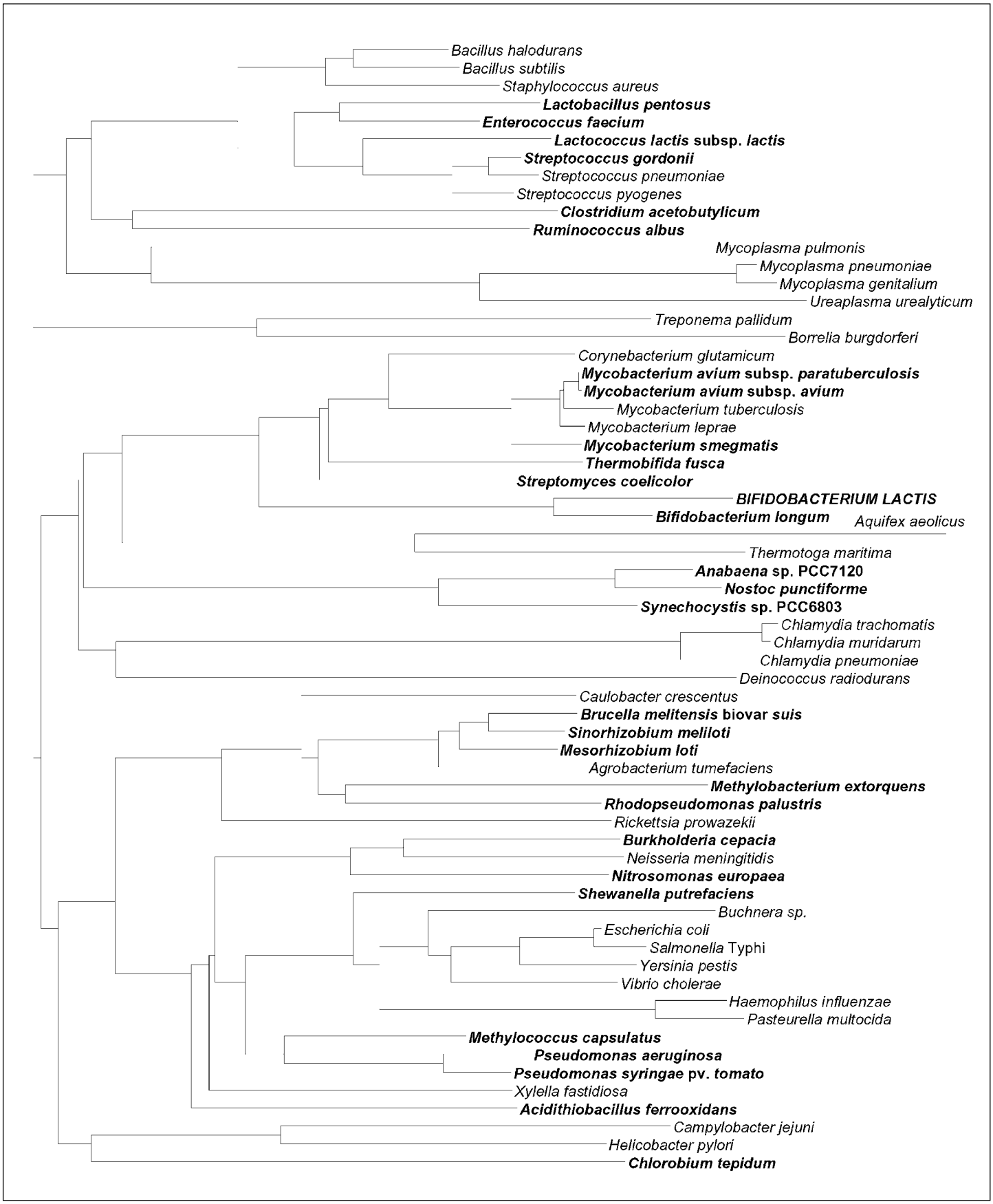

Fig. 3. Phylogenetic tree based on the 16S rRNA sequences of bacteria possessing putative phosphoketolase genes (bold) and of species showing no *xfp* homologues in their completely sequenced genomes (data from the Ribosomal Database Project II)

2.3. *Horizontal Gene Transfer or Evolutionary Gene Loss*

The wide phylogenetic distribution within the bacteria provokes the question of the evolutionary history of the *xfp* gene. The PK could be a very ancient enzyme having played an important role in the metabolism of most bacteria in the past. During evolution, new enzymes were developed and therefore more advantageous pathways could be established. The now obsolete *xfp* gene was hereafter rejected by several organisms, and so would be a useless relic in most genomes today. A contrary hypothesis to explain the described findings would assume that PK originated from a small group of bacteria from which the beneficial *xfp* gene was spread *e.g.* by means of mobile genetic elements or natural transformation by horizontal transfer into many other species. An approach which allows one hypothesis to be favoured over the other is the analysis of the G+C content, the molar ratio of guanosine and cytosine in the nucleic acid. Owing to the codon preference during gene transcription, the G+C content in the genome is typical for a particular species and normally consistent from gene to gene. Assuming an evolutionary recently acquired gene, its G+C content is likely to differ from the genomic G+C content. We calculated the G+C content of the various *xfp* genes and compared it with that of the whole genomes or of the so far accessible nucleotide sequences of the same species. As depicted in Fig. 4, the G+C contents of all identified putative PK genes are very similar to that of their respective genomes (sequences from eukaryotes included). These results lead to the conclusion that the ancient *xfp* gene was transmitted longitudinally, getting lost in a number of organisms, but was not subject to horizontal gene transfer and must therefore be regarded as an 'ur-protein'.

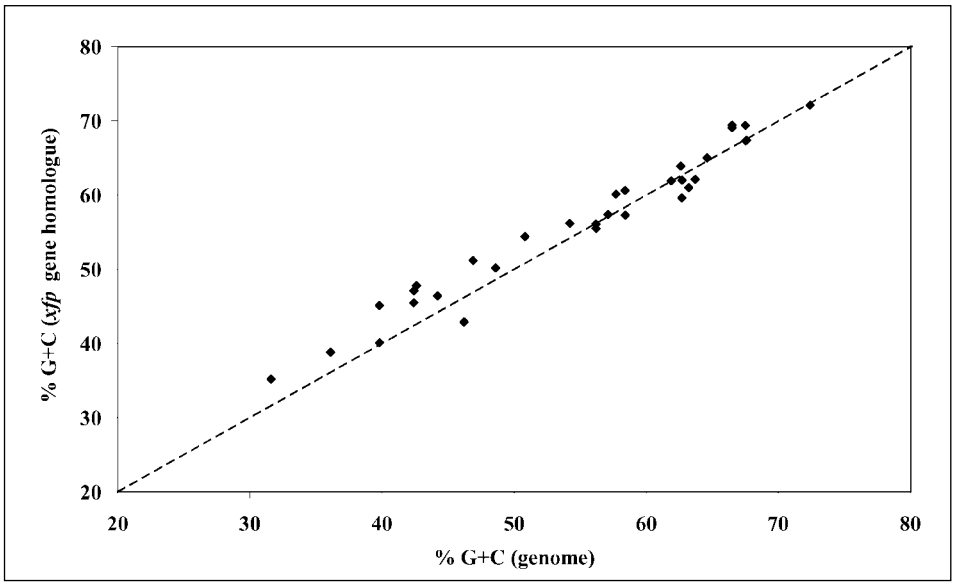

Fig. 4. Comparison of the G+C contents of 31 microbial phosphoketolase gene (*xfp*) homologues with those of the whole genomes or of the so far accessible nucleotide sequences of the respective species. The dashed line marks equality of the nucleic acid compositions. In addition to the species containing *xfp* that are named in Fig. 3 the three fungi *Schizosaccharomyces pombe, Neurospora crassa*, and *Aspergillus fumigatus* were included in the analysis.

3. Outlook

In this study we presented the dual substrate phosphoketolase of *Bifidobacterium lactis* and revealed the new family of Xfp phosphoketolases, which turned out to be ancient enzymes widespread in microorganisms. The performed heterologous expression of the *B. lactis* PK gene in *Escherichia coli* enables the thorough investigation of the native structure and the reaction mechanism of the enzyme by crystallisation studies. It will furthermore be interesting to clarify the substrate specificity and the biochemical significance of the PK within the metabolism of the different species harbouring its gene.

4. Methods

Preliminary molecular sequence data was obtained from The Institute for Genomic Research (TIGR) website at *http://www.tigr.org* as well as from databases provided by the National Center for Biotechnology Information (NCBI, *http://www.ncbi.nlm.nih.gov)*. For multiple sequence alignments the ClustalX program (version 1.81) was employed [8]. The construction of phylogenetic trees was performed using the data and algorithms (default parameters) of the Ribosomal Database Project II [9]. Data on nucleic acid sequence composition (G+C content) was retrieved from the Codon Usage Database [10].

Received: May 22, 2002

[1] L. Meile, *Lebensmittel-Technologie* **1998**, *31*, 68.
[2] U. Schörken, G.A. Sprenger, *Biochim. Biophys. Acta* **1998**, *1385*, 229.
[3] A. Schellenberger, *Biochim. Biophys. Acta* **1998**, *1385*, 177.
[4] W. Buckel, in 'Biology of the prokaryotes', Eds. J.W. Lengeler, G. Drews, H. G. Schlegel, Thieme-Verlag, Stuttgart, **1999**, p. 278.
[5] B. Sgorbati, G. Lenaz, F. Casalicchio, *Antonie Van Leeuwenhoek* **1976**, *42*, 49.
[6] L. Meile, L.M. Rohr, T.A. Geissmann, M. Herensperger, M. Teuber, *J. Bacteriol.* **2001**, *183*, 2929.
[7] Unpublished results.
[8] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, D.G. Higgins, *Nucl. Acids Res.* **1997**, *25*, 4876.
[9] B.L. Maidak, J.R. Cole, T.G. Lilburn, C.T.J. Parker, P.R. Saxman, R.J. Farris, G.M. Garrity, G.J. Olsen, T.M. Schmidt, J.M. Tiedje, *Nucl. Acids Res.* **2001**, *29*, 173.
[10] Y. Nakamura, T. Gojobori, T. Ikemura, *Nucl. Acids Res.* **2000**, *28*, 292.