Proc. Natl. Acad. Sci. USA
Vol. 94, pp. 13487–13492, December 1997
Biochemistry

B12-dependent ribonucleotide reductases from deeply rooted eubacteria are structurally related to the aerobic enzyme from *Escherichia coli*

(deoxyribonucleotide synthesis/allosteric regulation/evolution/thermophilic)

ALBERT JORDAN*†, EDUARD TORRENTS*†, CHRISTIAN JEANTHON‡, ROLF ELIASSON*, ULF HELLMAN§, CHRISTER WERNSTEDT§, JORDI BARBÉ†, ISIDRE GIBERT†, AND PETER REICHARD*¶

*Department of Biochemistry I, Medical Nobel Institute, MBB, Karolinska Institute, S-17177 Stockholm, Sweden; †Department of Genetics and Microbiology, Faculty of Sciences, Autonomous University of Barcelona, Bellaterra, E-08193 Barcelona, Spain; ‡Station Biologique, UPR9042, Centre National de la Recherche Scientifique and Université Pierre et Marie Curie, F-29680 Roscoff, France; and §Ludwig Institute for Cancer Research, Biomedical Center, Box 595, S-75124 Uppsala, Sweden

Contributed by Peter Reichard, September 24, 1997

ABSTRACT **The ribonucleotide reductases from three ancient eubacteria, the hyperthermophilic *Thermotoga maritima* (TM), the radioresistant *Deinococcus radiodurans* (DR), and the thermophilic photosynthetic *Chloroflexus aurantiacus,* were found to be coenzyme-B_{12} (class II) enzymes, similar to the earlier described reductases from the archaebacteria *Thermoplasma acidophila* and *Pyrococcus furiosus*. Reduction of CDP by the purified TM and DR enzymes requires adenosylcobalamin and DTT. dATP is a positive allosteric effector, but stimulation of the TM enzyme only occurs close to the temperature optimum of 80–90°C. The TM and DR genes were cloned by PCR from peptide sequence information. The TM gene was sequenced completely and expressed in *Escherichia coli*. The deduced amino acid sequences of the two eubacterial enzymes are homologous to those of the archaebacteria. They can also be aligned to the sequence of the large protein of the aerobic *E. coli* ribonucleotide reductase that belongs to a different class (class I), which is not dependent on B_{12}. Structure determinations of the *E. coli* reductase complexed with substrate and allosteric effectors earlier demonstrated a 10-stranded β/α-barrel in the active site. From the conservation of substrate- and effector-binding residues we propose that the B_{12}-dependent class II enzymes contain a similar barrel.**

All living cells produce the four deoxyribonucleotides required for DNA replication and repair by reduction of ribonucleotides. Three classes of ribonucleotide reductases exist (1, 2). All use free radical chemistry for catalysis but differ in the way in which they produce a protein radical required for the activation of the substrate. Class I enzymes, with the aerobic *Escherichia coli* enzyme as prototype, are $\alpha_2\beta_2$ proteins. The large α protein (R1) harbors catalytic and allosteric sites, whereas β contains a diferric center and a stable tyrosyl radical. They occur in eukaryotes and some aerobic eubacteria. Class II enzymes, with the *Lactobacillus leichmannii* enzyme as prototype, have an α or α_2 structure and use adenosylcobalamin as radical generator. They occur in aerobic and anaerobic bacteria. Class III enzymes finally, with the anaerobic *E. coli* enzyme as prototype, have an $\alpha_2\beta_2$ structure and use *S*-adenosylmethionine to generate a stable glycyl radical. They occur in anaerobic bacteria.

The prototypes for the three classes were not homologous and it seemed possible that they had arisen separately during evolution (3). However, functional aspects are in favor of a common root (1, 2). Members of the three classes reduce ribonucleotides by an identical mechanism and critical, functionally involved cysteine residues are found in corresponding positions in the class I and II enzymes (4, 5). Furthermore the allosteric regulation of the substrate specificity of the enzymes is highly similar, suggesting that the different classes may in part have closely related tertiary structures in spite of the large divergence of their primary structures (1, 2).

For a long time the only characterized class II enzyme was the prototype from *L. leichmannii,* a eubacterium (6, 7). However, two recent studies reported the sequences and some properties of class II reductases from two archaebacteria, *Pyrococcus furiosus* (8) and *Thermoplasma acidophilum* (9). Both sequences show considerable homology not only to the *L. leichmannii* enzyme but also to class I reductases. In addition, at the N terminus ≈100 amino acids were homologous to the corresponding sequence of the *E. coli* class III reductase. This result is persuasive evidence that all three classes originated from a common ancestor.

Do the negative results with the *Lactobacillus* enzyme mean that its sequence has evolved so far that a common origin no longer is recognized? Knowledge of eubacterial class II ribonucleotide reductases is limited. We have now studied class II enzymes from some deeply rooted eubacteria (10, 11) and find that their sequences show a greater kinship to the enzymes from archaebacteria than to that from *L. leichmannii.* From a combination of sequence data of the class II enzymes with recent structural results from complexes between the *E. coli* class I reductase and allosteric effectors (12) we hypothethize that the two classes share an appreciable part of the tertiary structure involved in effector binding.

To facilitate the presentation and discussion of our work we suggest the name *nrdJ* for the gene of class II ribonucleotide reductases.

MATERIALS AND METHODS

Materials. *Thermotoga maritima* (TM) strain MSB8 (DSM 3109 from the Deutsche Sammlung von Mikroorganismen und Zellkulturen, Braunschweig, Germany) and *Deinococcus radiodurans* R1 (ATCC 13939) (DR) were used for enzyme purification and PCR amplification of genomic DNA. Extracts from *Chloroflexus aurantiacus* J-10-fl (ATCC 29366), were used for enzyme assays. *E. coli* DH5αF′, and plasmids pBlueScript SK (+) (pBSK, Stratagene), pGEM-T (Promega) and pET22b (Novagen) were used for recombinant DNA

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: TM, *Thermotoga maritima*; DR, *Deinococcus radiodurans*; TIGR, The Institute for Genomic Research.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. Y12877).

¶To whom reprint requests should be addressed.

techniques. Lysyl endopeptidase was from Wako Chemicals (Neuss, Germany).

Bacterial Growth and Enzyme Extraction. Batch cultures of TM were grown under nitrogen for 12–14 h at 80°C in basal medium containing sodium sulfide (13), with the NaCl concentration adjusted to 20 g/liter and thiosulfate used as electron acceptor at 20 mM. After rapid cooling on ice the cells were sedimented at 8,000 × *g* and stored at −20°C under nitrogen until processed. Extraction of 8 g of packed cells was made in air at +4°C with 12 ml of 50 mM Tris·HCl, pH 7.5/50 mM KCl/1 mM EDTA/10 mM DTT/1 mM phenylmethanesulfonyl fluoride and 1 mg/ml egg white lysozyme by six consecutive cycles of freezing/thawing. A clear extract was obtained by centrifugation (45,000 rpm for l h in a Ty65 Beckman centrifuge).

DR was grown aerobically at 30°C in TGY medium (14) for 8 h until the end of the logarithmic growth phase. The cells were centrifuged, extracted in a French press in the buffer described for TM cells lacking KCl, and centrifuged at 45,000 rpm for 1 h to give a clear extract. *Chloroflexus* was grown anaerobically with light (900 lux) at 55°C in DG medium (15) for 48 h. The cells were extracted in air by sonication and centrifuged as described for DR.

Enzyme Purification. The TM extract was precipitated with streptomycin (final concentration, 1%). After removal of the precipitate the TM reductase was precipitated with ammonium sulfate (final saturation, 45%), centrifuged and dialyzed against buffer A (50 mM Tris·HCl, pH 7.5/10 mM DTT/1 mM EDTA/0.1 mM phenylmethanesulfonyl fluoride). The enzyme was chromatographed on a Mono Q HR 5/5 column on a Pharmacia fast protein liquid chromatography (FPLC) machine with a 0–0.4 M KCl gradient in buffer A. The enzyme eluted at around 0.3 M KCl. The sample was concentrated by ultrafiltration in Centriprep 30 (Amicon) tubes and adsorbed to a column of dATP-Sepharose (0.2 ml/mg protein) in 30 mM Tris·HCl, pH7.5/10 mM $CaCl_2$/2 mM DTT (buffer B)/0.5 M KCl. The column was washed successively with buffer B, buffer B/1 mM ATP, buffer B again, and finally 0.5 M ammonia. The last eluate contained the reductase. The enzyme was immediately concentrated and ammonia was removed by centrifugation in a Centricon 30 tube.

The DR reductase was purified by a similar procedure omitting the FPLC chromatography. The dATP-Sepharose step was slightly different in that the enzyme was eluted with 1 mM dATP instead of ammonia.

Enzyme Assays. Under standard conditions the TM reductase was incubated aerobically at 90°C for 20 min with 0.5 mM [^{3}H]CDP or [^{3}H]CTP (19 cpm/pmol), 20–40 mM DTT, 15 μM adenosylcobalamin, 0.3 mM dATP, 10 mM $MgCl_2$ or $CaCl_2$, and 50 mM Tris·HCl (pH 8). The DR and *Chloroflexus* enzymes were assayed similarly except for a temperature of 30°C for DR and 50°C for *Chloroflexus*. After acid hydrolysis the amount of [^{3}H]dCMP formed was determined (16). One enzyme unit corresponds to 1 nmol dCMP/min. Specific activity is units/mg protein.

Isolation of *nrdJ* from TM by PCR and Single Specific Primer-PCR. On gel electrophoresis the TM enzyme gave a single band at 90 kDa, the DR enzyme gave several bands, but with a dominating, rather diffuse band at 85 kDa. These bands were used for the determination of the N-terminal sequence as well as internal sequences from peptides obtained by in-gel digestion with lysyl endopeptidase (17) (Table 1). TM peptides 1 and 4 were used to design the following two degenerate primers according to the TM codon usage (18): PrA [5′-GA(C/T)AG(A/G)TA(C/T)TT(C/T)ATGAA-3′; peptide 1] and PrB [5′-TC(C/T)CTGTA(A/C)AC(A/G/C/T)GT(A/G/T)AT-3′; peptide 4, antisense]. Genomic DNA from TM (19) was amplified by PCR with the two primers by *Taq* polymerase at 50°C for annealing. The 1.8-kb amplification product was cloned in pGEM-T, and by sequencing it was found to contain the central part of *nrdJ*.

The cloning of the two extremes of the gene was made with single specific primer-PCR (20). After Southern blot hybridization of the TM genome with the 1.8-kb fragment as a probe we deduced that the 5′ extreme would be included in a 1.4-kb *Xho*I fragment and the 3′ extreme in a 1.6-kb *Cla*I fragment. To clone the 5′ end *Xho*I-digested chromosomal DNA was ligated to *Xho*I-digested pBSK plasmid and PCR amplified with primer PrC (5′-CTTCGAAAATCTCTTCAATGC-3′) as specific primer and PrFor (pUC/M13 forward) as generic primer (see Fig. 1). Similarly, the 3′ extreme was obtained by amplifying *Cla*I-digested chromosomal DNA ligated to *Cla*I-digested pBSK plasmid with PrD (5′-TACGTGAACCAGGT-3′) as specific and PrRev (pUC/M13 reverse) as generic primers. The resulting products were cloned into pGEM-T and sequenced in both strands from several independent clones. From these sequences and their overlap with the 1.8-kb sequence a definite sequence of 3,256 bp between the restriction sites *Xho*I and *Cla*I (Fig. 1) was obtained and deposited in the GenBank database (accession no Y12877).

Table 1. Peptide sequences

No.	Location	Sequence
T. maritima		
1	N terminal	M^{1}KLSDLISRWIDVEPSKNAQIILRDRYFMK^{30}
2	Internal	K^{158}VGGGVGSNFSELRPK^{173}
3	Internal	K^{523}EPLLYVNQVLREK^{536}
4	Internal	K^{597}TINMPQSATVDDVLNVYLEALRTNVRxIT-VYRDGSLQTQVL^{638}
D. radiodurans		
1	N terminal	x^{2}xAPDRSA^{9}
2	Internal	K^{15}saQHFDENAQH^{26}
3	Internal	K^{382}xFPVPQEGEYAGTFPEL^{399}
4	Internal	k^{1012}FVELLESYPAPK^{1024}
5	Internal	K^{1111}SAYEEAYRTGxK^{1123}
6	Internal	K^{15}xxQHFDENAQ^{25}
7	Internal	K^{118}LALVTK^{124}

The numbers at the beginning and at the end of each sequence indicate their positions in the predicted translated sequence of the TM *nrdJ* gene obtained in this work and the *nrdJ* gene sequence found in the DR genome database. Lowercase letters are used when discrepancies exist. Peptides 1–5 from DR have been obtained from the 85-kDa band, and peptides 6 and 7 from the 105-kDa band. The intein present in the DR-translated product is included in the sequence; the total size of the unprocessed protein is 1,358 amino acids. No peptides were found corresponding to the intein region (positions 465–830).

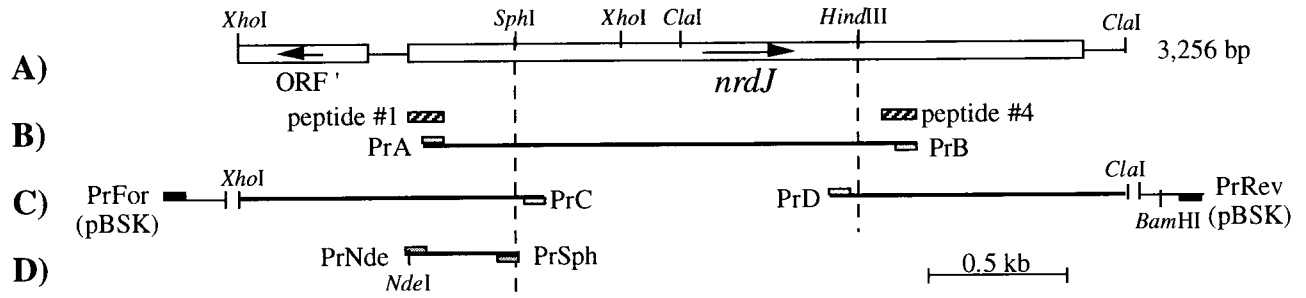

FIG. 1. (*A*) Physical map of the 3,256-bp sequenced fragment containing the TM *nrdJ* gene with restriction sites mentioned in the text. (*B*) PCR amplification of an internal fragment with primers derived from peptides 1 and 4 (Table 1). (*C*) Single specific primer-PCR for cloning the 5′ and 3′ extremes of *nrdJ*. (*D*) Generation of an *Nde*I restriction site in the translation start codon. The entire gene was obtained by joining the different fragments via the *Sph*I and *Hin*dIII sites.

Expression of TM *nrdJ* in *E. coli*. An *Nde*I site and the entire reconstructed gene were introduced into pET22b plasmid, downstream the T7 promoter and ribosome binding site. Primers PrNde (5′-AGGAGGGAGCATATGAATTGTCCG-3′, *Nde*I site underlined, start codon in bold letters) and PrSph (5′-CGAAGCATGCGGAGAGCATGTGAAGATG-3′, antisense, *Sph*I site underlined) were used to amplify the 5′ extreme of the gene. The 418-bp product was used together with the fragments generated with the amplifications PrA-PrB and PrD-PrRev to reconstruct the entire gene by using the unique *Sph*I and *Hin*dIII sites (Fig. 1). The entire gene was obtained by digestion with *Nde*I–*Bam*HI and cloned into pET22b. The resulting plasmid (pUA724) was transformed into *E. coli* BL21(DE3). *nrdJ* was expressed by addition of 1 mM isopropyl β-D-thiogalactoside to an exponential culture (OD_{550} of 0.6) and cells were harvested after 4 h. For denaturation of mesophilic host cells extracts were heated to 85°C for 15 min.

PCR Amplification of an Internal *nrdJ* Fragment of DR. DR peptide 3 (Table 1) was used for the construction of primer 5′-GGCGA(A/G)TACGCCGGCAC(G/C)TTCCC-3′ and peptide 4 for primer 5′-CTTGGG(G/C)GC(G/C)GGG-TAGCT(T/C)TC-3′ (antisense) for PCR amplification of a plate colony of DR with an annealing temperature of 64°C. An amplification product of 1.9 kb was obtained, cloned, and sequenced.

RESULTS

Enzyme Purification. When CTP reduction was determined in extracts from TM, DR, and *Chloroflexus* under the conditions given in *Materials and Methods,* the extracts gave specific activities of 1.8, 0.6, and 0.4, respectively. Omission of adenosylcobalamin completely abolished the reaction with the TM and DR enzymes but left some activity in the *Chloroflexus* extract. Enzyme activity also depended on adenosylcobalamin when enzyme extraction and assay were made anaerobically excluding the additional activity of a class III enzyme in the anaerobes TM and *Chloroflexus*. The enzymes from TM and DR were then purified aerobically, as described in *Materials and Methods*. The purified TM enzyme had a specific activity of 1,100 with CDP as substrate. On gel electrophoresis it gave a single band of ≈90 kDa. During chromatography on a column of Superdex 200 at room temperature the protein eluted very close to the void volume, corresponding to a molecular mass of more than 500 kDa, indicating aggregation. The purified DR enzyme was not homogeneous on gel electrophoresis. It gave a minor sharp band at 105 kDa and a strong, rather diffuse band at 85 kDa, as well as additional minor bands of lower molecular mass. The 105- and 85-kDa bands were used for peptide analyses (Table 1). The deduced amino acid sequence [The Institute for Genomic Research (TIGR); personal communication] suggests a molecular mass of 107-kDa and the 85-kDa band may therefore correspond to degradation products. The specific enzyme activity with CDP as substrate was 150–200.

Properties of the TM Reaction. The temperature optimum for the reaction was ≈80°C. Both Mg^{2+} and Ca^{2+} stimulated 3- to 4-fold, with Ca^{2+} giving a slightly higher temperature optimum (data not shown). The homogeneous enzyme was active with both CDP and CTP. Comparing the two substrates in a time curve (Fig. 2) we found that CTP was essentially not used as substrate during the first minutes, whereas CDP was reduced immediately. This finding suggests that the activity of CTP depended on its dephosphorylation to CDP at the high temperature of the reaction. In support of this we found a time-dependent transformation of CTP to CDP during incubation (Fig. 2).

The reduction of CDP is influenced by allosteric effectors, with dATP and ATP stimulating and dTTP and dGTP inhibiting the reaction (data not shown). The extent of stimulation by dATP depended on temperature. At both 80°C and 90°C dATP strongly stimulated CDP reduction, whereas it had a marginal effect at 70°C (Fig. 3).

Properties of the DR Reaction. All experiments were done at 30°C. Again, both Mg^{2+} and Ca^{2+} stimulated the reaction and the enzyme reduced both CTP and CDP. Also in this case a time curve suggested CDP to be the preferred substrate (data not shown). The allosteric regulation was similar to that of the

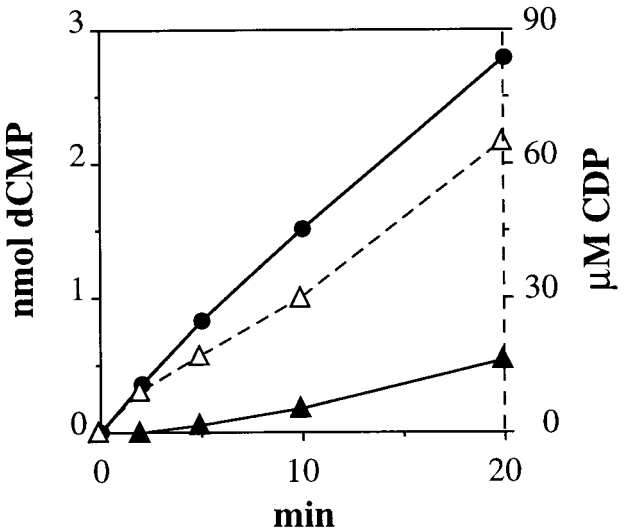

FIG. 2. Comparison of CDP and CTP as substrates for the TM reductase. Assays were made under standard conditions with $MgCl_2$ and 0.13 μg of enzyme and the amount of dCMP formed was determined at the indicated time points (CDP, ●; CTP, ▲). In the CTP experiment, the concentration of CDP (△) caused by degradation of CTP was also determined.

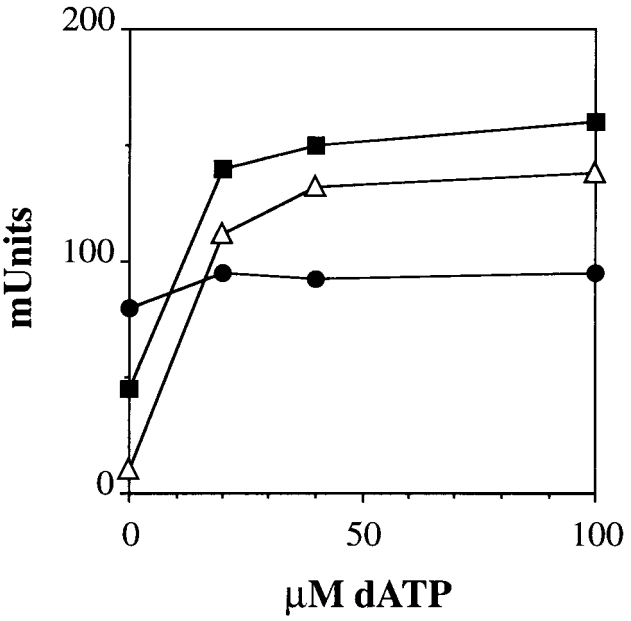

FIG. 3. Dependence of the allosteric stimulation of the TM enzyme (0.13 μg) by dATP on temperature (70°C ●; 80°C, ■; 90°C △) in the presence of $MgCl_2$.

TM enzyme with dATP and ATP stimulating and dTTP and dGTP inhibiting the reaction (data not shown).

Cloning and Sequencing of the TM and DR *nrdJ* Genes. The complete TM *nrdJ* gene was cloned by PCR and single specific primer-PCR as described in *Materials and Methods*. It is composed of 2,484 bp, encoding a putative protein of 827 residues with a predicted molecular mass of 94,018, lacking inteins. A hypothetical ribosome binding site (AGGAGG) complementary to the 3′ end of the TM 16S rRNA (21) is located 7 nt upstream of the ATG start codon for the first methionine. The methionine is not processed from the mature protein (Table 1). No definitive promoter or terminator sequences were identified. The 5′ end (467 bp) of an hypothetical ORF with a transcriptional sense opposed to the *nrdJ* gene is present upstream the *nrdJ* gene. It shows significant homologies with two ORFs of unknown function present in the *Methanococcus jannaschii* (22) and *Synechocystis* sp. (23) genomes.

The complete TM *nrdJ* gene was cloned into pET22b to give plasmid pUA724 as described in *Materials and Methods*. After transformation of *E. coli* BL21(DE3) with pUA724 expression of the *nrdJ* gene could be induced with isopropyl β-D-thiogalactoside resulting in a moderate production of the NrdJ protein. Gels of extracts from induced bacteria did not show overproduction of a specific band, but enzyme assays demonstrated the presence of an adenosylcobalamin-dependent, heat-stable reductase activity. The specific activity of the recombinant enzyme in the crude extract was increased from 5 to 150 by heating at 85°C. The poor overproduction is probably caused by differences in the codon usage of TM and *E. coli* (18). Nevertheless pUA724 should be a useful tool for further work.

The DR *nrdJ* gene was only partially cloned because in the course of our work the TIGR group released the provisional and uncompleted sequence of the DR genome, including the *nrdJ* gene. Our work is presented here briefly, because it confirms and slightly extends the other results. The amino acid sequence deduced from the nucleotide sequence of the 1.9-kb product cloned by us was in complete agreement with the reported sequence. All the DR peptides of Table 1 obtained from either the 85-kDa and 105-kDa bands are present in the sequence. The sequence of peptide 1 suggests that the N-terminal methionine is processed. The sequence contains one intein that is located in the same position as the second intein of the *P. furiosus* enzyme (8) and also is of similar length. In the position of the first *Pyrococcus* intein, the DR sequence contains an insertion relative to the TM sequence that does not possess the elements of an intein (Fig. 4).

DISCUSSION

Species belonging to three of the most deeply rooted eubacteria groups (10, 11) are here shown to use a class II reductase for deoxyribonucleotide synthesis. The genome of DR contains sequences characteristic of both a class Ib and a class II reductase (TIGR, personal communication). Our work shows that it is the latter that provides an active enzyme under normal growth conditions. At least two of the investigated enzymes were found to be ribonucleoside *di*phosphate reductases, as are two earlier investigated class II enzymes from archaebacteria (8, 9). They differ from the *L. leichmannii* enzyme that reduces triphosphates (6). Clearly a distinction between class I and II enzymes can no longer be made on the basis of the level of phosphorylation of the substrate. TM is an extreme thermophile (25) and the TM reductase accordingly requires

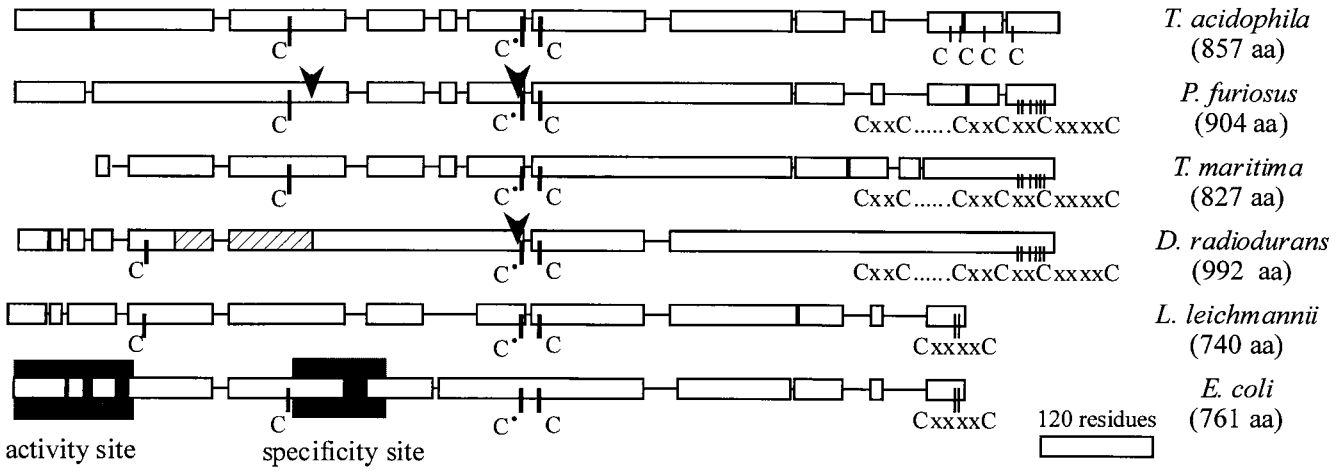

FIG. 4. Schematic alignment of class II reductases with the class I *E. coli* R1 protein. The alignment was generated with the CLUSTAL W version 1.7 (24) program. Single lines show larger gaps introduced in the sequences. The lowest bar gives the length of 120 residues. The arrows show the positions of inteins in the *P. furiosus* and DR enzymes. The striped area of the DR structure indicates the insertion referred to in the text. The shaded areas show the locations of the two allosteric sites of the *E. coli* enzyme identified by crystallography (12).

temperatures above 60°C for activity, with an optimum between 80°C and 90°C. All ribonucleotide reductases are allosteric enzymes (1, 2, 5) whose substrate specificity is regulated by binding of effectors to a specific site (specificity site). Reduction of CDP is stimulated by binding of ATP or dATP to this site. A group of class I reductases, among them the *E. coli* enzyme, has a second site (activity site). Binding of dATP there strongly inhibits all enzyme activity. Reduction of CDP by the reductases from the two ancient eubacteria investigated here are stimulated, not inhibited, by dATP, suggesting the absence of an activity site. With the TM enzyme a clear stimulation by dATP is only seen at high temperatures, even though the nucleotide binds to the protein in the cold room (unpublished data). This result suggests that the protein at low temperature does not have the required flexibility to transmit the message from the allosteric to the catalytic site.

Fig. 4 gives an overview of the general construction of the TM and DR enzymes, comparing them with the two archaeal class II reductases, the *Lactobacillus* enzyme and the R1 protein of the *E. coli* class Ia reductase. The latter two enzymes use three active cysteines with specific functions for the same free radical-based chemistry (4, 5). In Fig. 4 these cysteines can be identified in all sequences. Although located at slightly different positions, their relative order is the same. The first redox-active cysteine is located closer to the N terminus in the *Lactobacillus* and *Deinococcus* enzymes than in the other organisms. This result created difficulties in sequence alignments. The *Deinococcus* sequence was, however, quite homologous with the *Thermotoga* sequence, which could form a bridge to the other sequences. The redox-active cysteine pair at the C terminus of the *E. coli* and *Lactobacillus* enzymes have been shown to transfer electrons from thioredoxin to the cysteine pair in the catalytic site (4, 5). The *Thermoplasma* enzyme lacks such a pair, whereas the *Pyrococcus*, TM and DR enzymes have three. It is not clear how this affects transthiolation. The TM enzyme lacks the first 100 amino acids present in the others and has instead an extended C terminus.

Fig. 5 shows a computer alignment of the sequences of the TM reductase with those of the *Pyrococcus* and *E. coli* enzymes. The alignment of the *Deinococcus* enzyme whose preliminary sequence was reported by the TIGR group cannot be included, but it also shows retention of crucial residues. The three catalytically active cysteines are found in identical positions, which gives confidence to the correctness of the alignments and also suggests that the tertiary structure of the enzymes around the catalytic site is similar. In the *E. coli* enzyme the activity site is located at the N terminus of R1, the specificity site at the interface between the two subunits of R1 after C225 in the primary structure. Specific amino acids involved in effector binding have been identified and are shaded in Fig. 5 (12). B_{12} enzymes do not contain an activity site. The *Thermotoga* enzyme lacks the first 60 residues at the N terminus of the *E. coli* enzyme responsible for the activity site, which explains the absence of this site. The signature sequence VXKRDG at the N terminus involved in effector binding is found in *Pyrococcus* and *Thermoplasma*, but not in *Deinococcus*. However, all three sequences lack most other binding residues of the activity site (Fig. 5). In the specificity site, D232, H-bonded to the 3′-OH group of the effector, and R262, binding the γ-phosphate, are two key amino acids present in all class I enzymes. These residues are also present in the TM and *Pyrococcus* class II enzymes (Fig. 5). The other

```
              +         +         +         +         +          +         +         +         +         +
Pfur  ~MAVEKVMKRDGRIVPFDESRIRWAVQRAMWEVGIRDEKKLDEVVKSIVQ RINELYDGKIPHIENIQDIVELELMRAGLFEVAKAYILYRKKKAEIREEK  99
Tmar  ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~ ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~MKLS.   4
Ecol  MNQNLLVTKRDGSTERI......................NLDKIHR.... VLDWAAEGLHN....VSISQVELRSHIQFYDGIK...TSDIHETIIKAA  66
      -------------------------------------------------- --------------------------------------------------

Pfur  KRILNKKELDEIDKRF....SINALRVLASR.YLKKDENGNIVESP.REL FERVAILAVIPDLLYDERVFDKNGNY.SQDLKRVEYYLEHFEEFDRKYSI 192
Tmar  .DLISR....WIDVEP....SKNAQIILRDR.YFMKDLDGNYLETKWEDV ARRVARVVATAELL........NPSY.KKN...........EKLDR....  70
Ecol  ADLISR...DAPDYQYLAAR.......LAI.FHLRKKAYGQF.EP..PAL YDHVVKMVEMGK..YDNHLLE...DYTEEEFKQMDTFIDHDRDMTFSYAA 147
      -----------iD-------s-na---L--r-y--Kd--Gn--E------ --rVa-------ll-----------Y--------------e--dr----

Pfur  GKYKLNKYHFERMVN.LYKELAEQGKMKVSIDEFLAMLEKGEFNEYEKEI NEYFRLMTNQIFMPNTPALINSG.........RP................ 266
Tmar  ..........................................IKEWE... DIFFRVLKARLFIPNSPTLFNAGLGVKHDLLWKPIDQMTLEDYEEIYRSR 125
Ecol  VKQLEGKYLVQNRVTGEIYESAQFLYILVAACLFSNYPRETRLQ....YV KRFYDAVST..FKISLPTPIMSGVR.......TPTRQ............ 221
      -------------------------------------------e-e--- ---fr------F-pn-P-l-n-G----------P----------------
                   ▼ ▼                          ▼      ▼         ▼       ▼              ▼
Pfur  ..LGMLSACFVVPIEDDMESIMKAAHDVAMIQKMGGGTGLNFSKLRPEGD IVGTTTGAASGPVSFMHLIDAVSDVIKQGGVRR.GANMGILEIWHPDIEK 363
Tmar  NHLHMLSACFVVPVGDSIEEIFEAVKEYALITKVGGGVGSNFSELRPKGS FVAGTHGKASGPVSFMHVFNSAISVVKQGYRRR.GALMGILNINHPDIEE 224
Ecol  .....FSSCVLIECGDSLDSINATSSAIVKYVSQRAGIGINAGRIRALGS PIRGGEAFHTGCIPFYKHFQTAVKSCSQGG.VRGGAATLFYPMWHLEVES 315
      --l-mlSaCfvvp--D--e-I--a----a-i-k-ggG-G-Nfs-lRp-G- -v--t-g-asGpvsFmh-------v-kQG--rR-GA-mgil-i-Hpdie-
             225    232 234                         262     268   275      281          292
Pfur  FIHAKEKNIGTNVLSNFNISVGIWEDFWEALK...EGKKYPLINP..... ......................RTGEVVK.EVDPKTLFEELAYMAWAKADPG 434
Tmar  FIDAKKENTGEAVLNFFNLSVGFPMDKKEILKLYEEDGELELSHP..... ......................RS.TIRK.KVKIRELFRKIATNAWKSGDPG 297
Ecol  LLVLK.NNRGVEGNRVRHMDYGVQINKL.MYTRLLKGEDITLFSPSDVPG LYDAFFADQEEFERLYTKYEKDDSIRKQRVKAVELFSLMMQERASTGRIY 413
      fi-aK--N-G--vl--fn-svG---d--e-lk---e-----L--P----- ------------------------r-----K--V----LF---a--aw---dpg

Pfur  VIFFDVINRRNVLKKAKGGPIRATNPCGEEPLY............EYESC NLASINLAKFVKYDENGKPYFDWDEYAYVIQKVAKYLDNSIDVNKFPLPE 522
Tmar  LAFLGEMNKYYPLYPHR..KINSTNPCGEIGLS............DYEAC NLGSIDVAKF..YN.NG..FVDLEALQELVQIAVRFLDNVIDVNVFPIDK 378
Ecol  IQNVDHCNTHSPFDPA.IAPVRQSNLCLEIALPTKPLNDVNDENGEIALC TLSAFNLGA..........INNLDELEELAILAVRALDALLDYQDYPIPA 502
      --f---------l-------i--tNpCgE--L-----------------C nL-si--akf--y--ng----d---------q-----LDn-iDvn-fP---
                              439                    462
Tmar  ITKAVKESRRLGLGIMGFADLLYKLEIPYNSQEARDFAANLMAFIALHAH RTSYELGKEKGNFPLLEISRYRTEDNFVPFAMGMSNYDDEIREVMKMTKE 478
Tmar  FRRNVALLTIAPTGSISNIADTSSGLEPNFLLAYTRFVTKEDGTKEPLLY VNQVLREKLNPEILKRIEKELIEKGSLKDIPDVPEKIKKVFVVALDIDPM 578
Tmar  DHLLMQDAFQRYVDNNISKTINMPQSATVDDVLNVYLEALRTNVRGITVY RDGSLQTQVLTKALKTPEAPKVQFFVVDEKLKLHPRPRKDTLRSVTRKYK 678
Tmar  RPDGTTYITISFDDTGEAVEIFISNGSEMAEAIGRLSSIALRAGVSIDEI VEQLSKVKGEYCKGLAEEIKKALEDFAKLWLRTGEEAPESEEEPIEREKF 778
Tmar  IVAHNLRWQSGYYVDDEGNVYCPVCLSKNSLIKQEGCVSCKNCGWSKCE*                                                  #827#
```

FIG. 5. Alignment of the N-terminal 378 residues of the *T. maritima* ribonucleotide reductase (Tmar) with the reductase from *P. furiosus* (Pfur) (8) and the R1 protein from *E. coli* (Ecol) (26). The remaining TM sequence is shown by itself. The cysteines corresponding to those in the active site of *E. coli* (C225, C439, and C462) as well as the C-terminal cysteines hypothetically involved in transthiolation are shaded lightly. Residues involved in effector binding at the activity (N terminal) and specificity sites of *E. coli* are shaded as well as the conserved counterparts of the class II enzymes. The specificity site residues also carry a ▼. Consensus among the two class II proteins is shown by lowercase letters; consensus among all three enzymes is shown by capital letters. The numbering of the three sequences is given at the end of each line. The alignment was generated with the PILEUP program of the GCG (Genetic Computer Group, University of Wisconsin, version 9.0-UNIX) package.

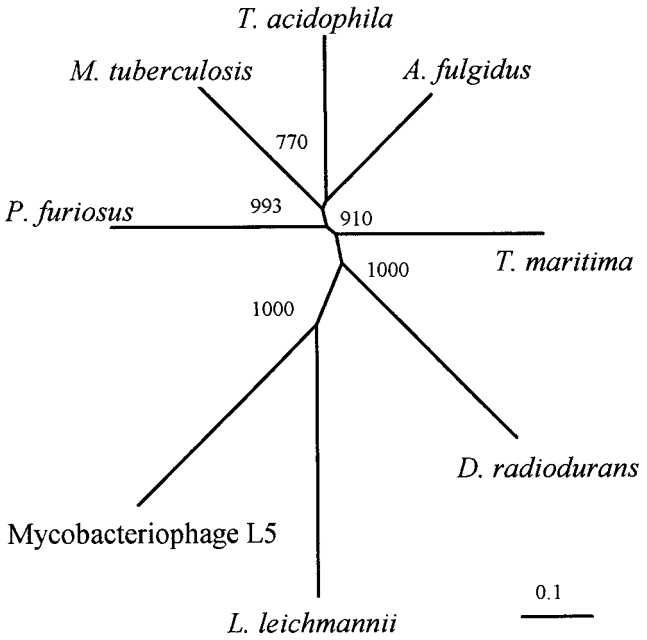

FIG. 6. Unrooted phylogenetic tree of the deduced amino acid sequences from *T. maritima*, *D. radiodurans* (TIGR, personal communication), *P. furiosus* (8), *T. acidophila* (9), *A. fulgidus* (TIGR), *M. tuberculosis* (TIGR; partial sequence), *L. leichmannii* (7), and the mycobacteriophage L5 (27). The CLUSTAL W program was used for sequence alignments and to compute the phylogenetical tree. The significance of the branching order was evaluated by bootstrap analysis of 1,000 computer-generated trees. Bootstrap values are indicated. (Bar = 0.1 changes per site.)

binding residues vary slightly between various class I reductases. Thus L234 is I in the mouse and M in class Ib, I268 is V in *Synechocystis* (23) and C292 is V or I in some eukaryotes. In all but one case the corresponding residues of the B_{12} enzymes fall within these variations. Also the enzymes from *Thermoplasma* and *Deinococcus*, as well as two sequences from *Archaeoglobus fulgidus* and *Mycobacterium tuberculosis* reported by genome sequencing (TIGR, personal communication), conform completely to this pattern (data not shown). Thus the binding residues for the allosteric effector in the specificity site are identical in class I reductases and six B_{12} enzymes, three from eubacteria and three from archaea. This finding together with retention of the catalytically active cysteines strongly suggests that the tertiary structure of the catalytic and allosteric sites, i.e., the 10-stranded β/α-barrel, is present also in class II reductases.

The allosteric regulation of the *L. leichmannii* enzyme is identical to that of the other B_{12} enzymes (unpublished data), indicating that also this enzyme may contain the barrel structure. One complication is that the *Lactobacillus* enzyme is believed to be a monomer (6), whereas the allosteric *E. coli* site is located between the polypeptide chains of a dimer.

At present the sequences of 8 class II enzymes are available. Fig. 6 shows an unrooted phylogenetic tree of the deduced amino acid sequences. The main finding is that the TM and DR enzymes are closely related to the archeal reductases, whereas the *L. leichmannii* enzyme is more distant. Interestingly, also the *M. tuberculosis* class II sequence is more related to the archeal enzymes. In contrast to the DR case, a class Ib reductase is the active enzyme in extracts from *M. tuberculosis* (28).

This work and the appearance of additional genome sequences show that a large number of class II sequences are more related to the two reported archeal proteins than to the until now considered prototype from *Lactobacillus*. In the *Lactobacillus* sequence it is more difficult to recognize the structure of the specificity site, mainly because of the large distance existing between the cysteine residues homologous to C225 and C439. In DR where this distance also is large, one can recognize by comparison with the TM sequence a fragment that may have been inserted more recently during evolution. One can speculate that an intein originally was present in this position, as is the case in the *Pyrococcus* enzyme (Fig. 4).

We acknowledge TIGR for availability of sequence data before publication, and K. W. Minton, R. Sirevag, and J. Shiozawa for providing us with *D. radiodurans* and *C. aurantiacus* strains. A.J. was supported by fellowships from Margit and Folke Pehrzon Foundation and Axel Wenner-Grens Stiftelse, E.T. by a predoctoral fellowship from the Direccio General d'Universitats de la Generalitat de Catalunya and the European Molecular Biology Organization, I.G. and J.B. by grants from the Spanish Dirección General de Investigación Cientifica y Técnica and from the Comissionat per Universitats in Recerca de la Generalitat de Catalunya, P.R. by a grant from the Swedish Medical Research Council, and C.J. by grants from Centre National de la Recherche Scientifique and Fonds Structurel Européen (FEDER 5b).

1. Reichard, P. (1993) *Science* **260,** 8383–8386.
2. Reichard, P. (1997) *Trends Biochem. Sci.* **22,** 81–85.
3. Benner, S. A., Ellington, A. D. & Tauer, A. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 7054–7058.
4. Stubbe, J. & van der Donk, W. A. (1995) *Chem. Biol.* **2,** 793–801.
5. Sjöberg, B.-M. (1997) *Struct. Bonding* **88,** 139–173.
6. Panagou, D., Orr, M. D., Dunstone, J. R. & Blakley, R. L. (1972) *Biochemistry* **11,** 2378–2388.
7. Booker, S. & Stubbe, J. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 8352–8356.
8. Riera, J., Robb, F. T., Weiss, R. & Fontecave, M. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 475–478.
9. Tauer, A. & Benner, S. A. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 53–58.
10. Woese, C. R. (1987) *Microbiol. Rev.* **51,** 221–271.
11. Stetter, K. O. (1996) *FEMS Microbiol. Rev.* **18,** 149–158.
12. Eriksson, M., Uhlin, U., Ramaswamy, S., Ekberg, M., Regnström, K., Sjöberg, B.-M. & Eklund, H. (1997) *Structure* **5,** 1077–1092.
13. Ravot, G., Ollivier, B., Magot, M., Patel, B. C., Crolet, J.-L., Fardeau, M.-L. & Garcia, J.-L. (1995) *Appl. Environ. Microbiol.* **61,** 2053–2055.
14. Carrol, J. D., Daly, M. J. & Minton, K. W. (1996) *J. Bacteriol.* **178,** 130–135.
15. Pierson, B. K. & Castenholz, R. W. (1991) in *The Prokaryotes: A Handbook on the Biology of Bacteria,* eds. Barlows, A., Truper, H. G., Dworkin, M., Harder, W. & Schleifer, K.-H. (Springer, Berlin), pp. 3754–3774.
16. Thelander, L., Sjöberg, B.-M. & Eriksson, S. (1978) *Methods Enzymol.* **51,** 227–237.
17. Hellman, U. (1997) in *Protein Structure Analysis: Preparation, Characterization, and Microsequencing,* eds. Kamp, R. M., Choli-Papadopoulou, T. & Wittmann-Liebold, B. (Springer, Berlin), pp. 97–104.
18. Kim, C. W., Markiewicz, P., Lee, J. J., Schierle, C. F. & Miller, J. H. (1993) *J. Mol. Biol.* **231,** 960–981.
19. Jeanthon, C., Reysenbach, A.-L., L'Haridon, S., Gambacorta, A., Pace, N. & Prieur, D. (1995) *Arch. Microbiol.* **164,** 91–97.
20. Shyamala, V. & Ames, G. F.-L. (1993) *Methods Mol. Biol.* **15,** 339–348.
21. Achenbach-Richter, L., Gupta, R., Stetter, K. O. & Woese, C. R. (1987) *Syst. Appl. Microbiol.* **9,** 34–39.
22. Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., *et al.* (1996) *Science* **273,** 1058–1073.
23. Kaneko, T., Sato S., Kotani, H., Tanaka, A., Asamizu, E., *et al.* (1996) *DNA Res.* **3,** 109–136.
24. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680.
25. Huber, R., Langworthy, T. A., Konig, H., Thomm, M., Woese, C. R., Sleytr, U. B. & Stetter, K. O. (1986) *Arch. Microbiol.* **144,** 324–333.
26. Nilsson, O., Åberg, A., Lundqvist, T. & Sjöberg, B.-M. (1988) *Nucleic Acids Res.* **16,** 4174.
27. Hatfull, G. F. & Sarkis, G. J. (1993) *Mol. Microbiol.* **7,** 395–405.
28. Yang, F., Lu, G. & Rubin, H. (1994) *J. Bacteriol.* **176,** 6738–6743.