Volume 11 Number 14 1983 Nucleic Acids Research

Sequences of the 5S rRNAs of the thermo-acidophilic archaebacterium *Sulfolobus solfataricus (Caldariella acidophila)* and the thermophilic eubacteria *Bacillus acidocaldarius* and *Thermus aquaticus*

Erna Dams, Paola Londei+, Piero Cammarano+, Antoon Vandenberghe and Rupert De Wachter*

Departement Celbiologie, Universiteit Antwerpen, Universiteitsplein 1, B-2610 Wilrijk, Belgium, and +Istituto di Biologia Generale, Università di Roma, Viale Regina Elena 161, I-00100 Roma, Italy

Received 14 June 1983; Accepted 23 June 1983

ABSTRACT

We have determined the nucleotide sequences of the 5 S rRNAs of three thermophilic bacteria : the archaebacterium Sulfolobus solfataricus, also named Caldariella acidophila, and the eubacteria Bacillus acidocaldarius and Thermus aquaticus. A 5 S RNA sequence for the latter species had already been published, but it looked suspect on the basis of its alignment with other 5 S RNA sequences and its base-pairing pattern. The corrected sequence aligns much better and fits in the universal five helix secondary structure model, as do the sequences for the two other examined species. The sequence found for Sulfolobus solfataricus is identical to that determined by others for Sulfolobus acidocaldarius. The secondary structure of its 5 S RNA shows a number of exceptional features which distinguish it not only from eubacterial and eukaryotic 5 S RNAs, but also from the limited number of archaebacterial 5 S RNA structures hitherto published. The free energy change of secondary structure formation is large in the three examined 5 S RNAs.

INTRODUCTION

To our knowledge, some 163 different sequences of 5 S rRNAs are presently known, among which 105 isolated from eukaryotic species, 44 from eubacteria, 6 from archaebacteria, 6 from chloroplasts and 2 from plant mitochondria. The last collection published in this journal (1) lists 93 of these sequences. This collection is becoming one of the most valuable tools for studying evolution in a broad perspective (2, and papers cited there). The size and diversity of the collection also has made possible the derivation, on a comparative basis, of a rather detailed picture of 5 S RNA secondary structure (3-7).

In this study, we examine the structure of the 5 S RNA from three thermophilic bacteria. One of them is the archaebacterium originally described (8) as Caldariella acidophila, but in this text we follow the suggestion of Zillig et al. (9) to name it Sulfolobus solfataricus since the 5 S RNA sequencing supports their conclusion that it belongs to the

genus Sulfolobus. The second species examined is the moderately thermoacidophilic eubacterium Bacillus acidocaldarius (10). The third is the eubacterium Thermus aquaticus (11). Although a 5 S RNA sequence for the latter species has been previously published (12), it has been pointed out (4) that this sequence, as well as that originally reported (13) for Pseudomonas fluorescens, is difficult to align with other 5 S RNA sequences and shows a distorted base pairing pattern. The Pseudomonas fluorescens sequence has been recently revised (14) and the correction of the Thermus aquaticus sequence is given here. The places of isolation and growth characteristics of the three examined species are listed in Table 1.

METHODS

Cultures of each of the three bacteria were grown under the conditions described in the references in Table 1. Ribosomes of Sulfolobus solfataricus and Bacillus acidocaldarius were prepared according to Nirenberg and Matthaei (15) with minor modifications. Ribosomes of Thermus aquaticus were prepared by a simplified version (14) of Kurland's (16) method. Ribosomes were phenol-extracted and the RNA fraction subjected to gel electrophoresis to isolate 5 S RNA (17).

Of the three 5 S RNA preparations, those of S. solfataricus and T. aquaticus proved to be heterogeneous in composition, as detected by gel electro-

Table 1. Growth characteristics and origin of the examined bacteria.

Characteristics		Species		
		Caldariella acidophila or Sulfolobus solfataricus	Bacillus acidocaldarius	Thermus aquaticus
Reference		8 (see also 9)	10	11
Strain used		MT 4	--	YT-1 (ATCC 25104)
Optimal growth	temperature	87°	60°	70°
	pH	3.0-4.5	3.0	7.5-7.8
Place of isolation		Pisciarelli Solfatara, Agnano Italy		Yellowstone Park U.S.A.

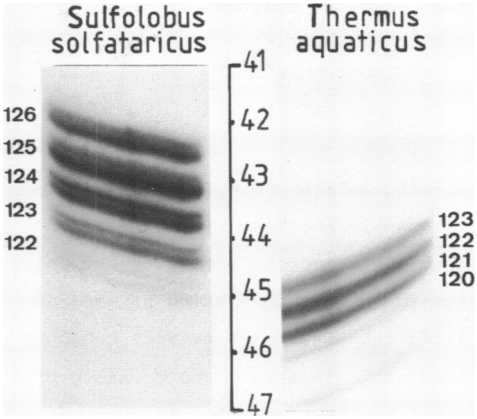


Fig. 1. Length heterogeneity in 5 S RNA preparations.

10 μg 5 S RNA, labeled at the 3'-terminus with about 60.10^6 dpm ^{32}P, was loaded on 0.4 mm thick and 80 cm long gels containing 8% polyacrylamide in 0.05 M Tris-borate pH 8.3, 7 M urea, and subjected to electrophoresis at 1700 V for 16 h. Only the relevant part of the autoradiographs is shown, the scale in the middle indicating the distance from the start in cm. The chain length of each component, as derived by sequence analysis, is marked alongside each band.

phoresis of the material after ligation with [5'-^{32}P]pCp at the 3'-terminus (18), and as demonstrated in Fig. 1. Each band was extracted and sequenced separately, which proved that the heterogeneity was due to the presence of components different in length but not in sequence.

The major part of each sequence was determined by Peattie's (18) partial chemical degradation method applied to 3'-terminally labeled RNA. Gels containing 8%, 12%, and 20% polyacrylamide were used to resolve sequence areas near the 5'-terminus, in the middle, and near the 3'-terminus of the molecules respectively. For each of the three 5 S RNAs the sequencing presented more problems than usually encountered, a fact that we attribute to the exceptionally stable secondary structures (see below) characteristic of these thermophilic bacteria. In each case, two areas of band compression were found on 8% gels, although not systematically in corresponding parts of the sequences. The compression areas are indicated in Fig. 3. The areas closest to the 3'-termini could be resolved on the 20% gels, which give superior resolution but can only be used up to a certain distance from the 3'-end. Those in the middle of the sequence or near the 5'-terminus were resolved by running 8% gels at 65°C (19). In the cases of Bacillus acidocaldarius and Thermus aquaticus, the 5'-terminal end-group was identified and the adjacent sequence confirmed by 5'-terminal ligation of $(Ap)_4A$, followed by kinase labeling with [γ-^{32}P]ATP and partial nuclease degradation (14). In the case of Sulfolobus solfataricus 5 S RNA, we were unable to ligate $(Ap)_4A$ to the 5'-terminus, a fact which may be attributable to the protruding oligo-U-tail present at the 3'-terminus (Fig.3). The 5'-end group was therefore identified

by high pressure liquid chromatography of an alkaline hydrolysate of the 5 S RNA (20).

RESULTS AND DISCUSSION

Primary structure

The three examined 5 S RNA primary structures are represented in Fig.2a. The alignment is consistent with the one chosen in the last review of known 5 S RNA sequences (1). The latter alignment contained 150 positions but only 130 of these were needed in Fig.2a, the remaining ones consisting of gaps in the three examined sequences. Slightly different alignments between eubacterial and archaebacterial sequences are conceivable in area D and possibly D', and future data on additional archaebacterial 5 S RNAs may confirm or invalidate our choice.

The sequence found for Sulfolobus solfataricus 5 S RNA is completely

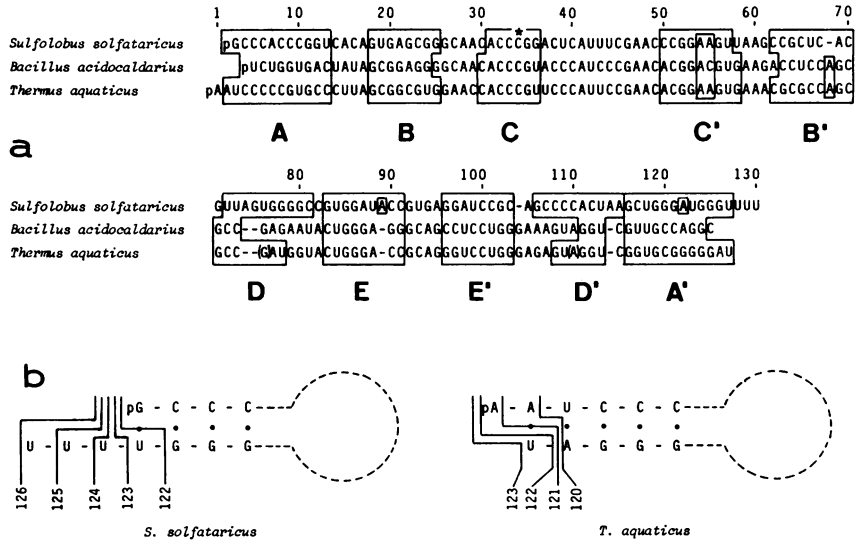


Fig. 2. Primary structure of the 5 S RNAs.

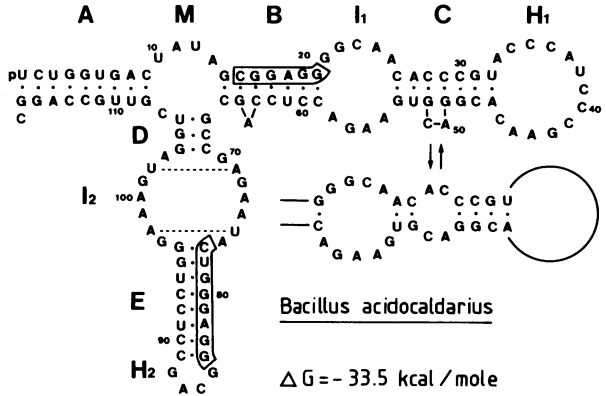
a) Alignment of the 3 sequences. The numbering of positions does not necessarily indicate the distance from the 5'-terminus, nor is it identical with the numbering system employed for the 5 S RNA sequence collection (1). Boxed areas are double-stranded in the secondary structure models (Fig.3). The asterisk indicates the presence of a modification, probably a 2'-O-methylribose residue, in the Sulfolobus sequence.

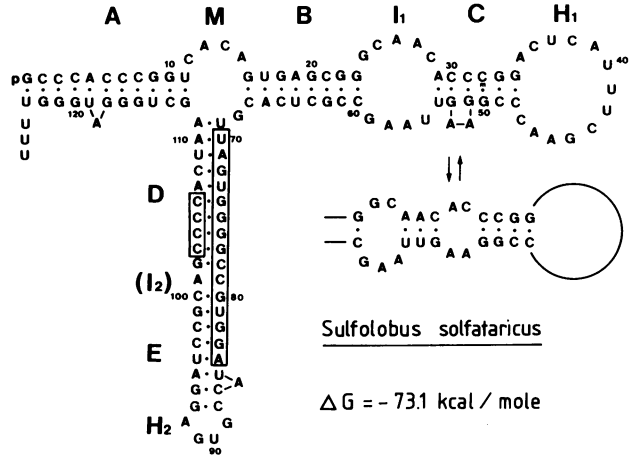
b) Length heterogeneity in S. solfataricus and T. aquaticus 5 S RNA. The termini and the chain lengths of the different components are indicated on a drawing of the base-pairing extremities of each molecule.

identical to the one reported (21) for Sulfolobus acidocaldarius. The identity includes the presence, at position 32 from the 5'-terminus, of a ribose-modified cytidine, probably 2'-O-methylribosylcytosine. This was inferred from the resistance of the adjacent 3'-linked phosphodiester bond to hydrolysis, but its susceptibility to ß-elimination after chemical modification by the C-reaction (18). On the sequencing gels, this is detected as a gap in the acid hydrolysis ladder corresponding with a band in the C-reaction lane.

The application of gel sequencing methods to the structural analysis of 5 S RNAs has greatly facilitated the detection of length heterogeneity, in other words the presence of one or more terminal residues in submolar amounts. Such heterogeneity seems to be the rule rather than the exception. Of the three species examined here, only B. acidocaldarius possesses a homogeneous 5 S RNA. The heterogeneity of the preparations isolated from S. solfataricus and T. aquaticus is demonstrated in Fig.1 and the structural forms present are indicated in Fig.2b. In the case of S. solfataricus, each of the 5 bands visible in Fig.1 corresponds with a difference in length of one U at the 3'-terminus. Each of these bands, however, seems to be split further into two components. We have not been able to discover the structural difference between the members of each pair. One possibility is that another residue, beside C_{32}, is partly modified and that this results in the splitting of each band. In alkaline hydrolysates of S. solfataricus 5 S RNA separated by HPLC (20), we detected large peaks of nucleoside 3'(2')-phosphates, a small pGp peak resulting from the 5'-end, and a small peak with intermediate mobility which we attributed to the presence of CmpGp, although we could not rigorously identify it due to the lack of methylated dinucleotide markers. If another, incompletely methylated ribose is present, it would have to be in a sequence yielding an alkali-resistant dinucleotide identical to CmpGp or indistinguishable from it in the chromatographic system used.

The sequence previously published (12) for Thermus aquaticus 5 S RNA contains a number of errors, the most conspicuous one being that the UG residue at positions 10-11 in area A (Fig.2a) was originally assigned to area D'. This accounts for most of the difficulties in properly aligning the sequence with other 5 S RNAs and obtaining the universal secondary structure model (4). There are 6 other errors that were probably made during the sequencing of oligonucleotides obtained by T_1 RNAse and pancreatic RNAse digestion. Although the corrections have made the sequence 3 nucleotides longer than it was originally estimated (12) they make it fit much better in





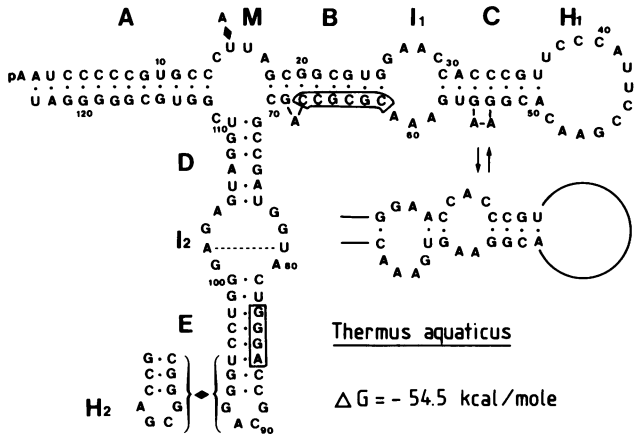


Fig. 3. Secondary structure models
The models are drawn and the ΔG values calculated, according to the conventions described earlier (4), for the longest component of each 5 S RNA. Additional base pairs that may exist at the expense of loop I_2 are indicated by dotted lines. The secondary structure equilibrium in area I_1-C is universal. Boxed segments correspond with areas of band compression on the sequencing gels. On the Thermus aquaticus sequence, the changes required to obtain the Thermus thermophilus sequence (22) are indicated by losenges.

the complete alignment (1) of 5 S RNAs, from which one position (No 134) can now be deleted. The 5 S RNA sequence of the related species Thermus thermophilus has recently been revised (22). It has the same length as T. aquaticus 5 S RNA and differs from it by 9 substitutions, as indicated in Fig. 3.

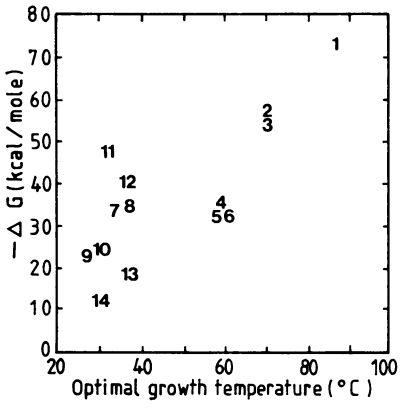
Secondary structure

Fig.3 represents secondary structure models for the three examined 5 S RNAs, drawn according to a set of topological conventions stated explicitly elsewhere (4,1). There are a number of features that distinguish the S. solfataricus base-pairing pattern from those of B. acidocaldarius and T. aquaticus, and in fact from those of nearly all 5 S RNAs. These are the presence of a bulge in helix A, the absence of such a bulge in helix B, and the fact that helix D is continuous with helix E except for the presence of an A•C pair. Another exceptional feature is that helix D is separated by a non-paired base - a G - from helix B, but is adjacent to helix A. In eubacteria helix D is usually separated from A, but adjacent to B. In eukaryotes and in archaebacteria other than Sulfolobus, helix D is adjacent to both helices A and B, or is separated from both by an equal number of unpaired bases. Finally, the Sulfolobus structure is distinct from the eubacterial structures by the presence of a bulge on helix E, a property that it has in common with eukaryotic and with most archaebacterial 5 S RNAs.

Arguments have been put forward elsewhere (4,14) in favour of the hypothesis that non-standard base pairs such as the A•G in helix D of T. aquaticus and the A•C between D and E in S. solfataricus do not disrupt the helical structure but form part of it. Stahl et al. (21) have gone even further by postulating that helices D and E may be continuous not only in Sulfolobus but also in other 5 S RNAs. In eubacteria, such a continuous helix is easiest to conceive because internal loop I_2 contains equal numbers of bases in both strands. These bases might then form a series of non-standard base pairs, a structure which might be stable within the ribosome if not in the free 5 S RNA. In eukaryotes and in archaebacteria other than Sulfolobus, loop I_2 is asymmetrical and stacking of helices D and E would require the formation of a bulge in the 5'-proximal strand. At any rate, a weak spot is always present between helices D and E. This spot could serve as a hinge allowing the D-E helix to alternate between a stacked state and some other conformation. In Sulfolobus, the weak spot may be reduced to a single A•C pair because this already provides sufficient flexibility at the high growth temperature characteristic of this bacterium. The universal existence of two alternative base pairing schemes in area I_1-C (Fig.3) has

Fig. 4. Stability of 5 S RNA secondary structure and optimal growth temperature of bacteria.

Species are numbered as in the list below. Optimal growth temperature was taken from the reference mentioned after each species name. ΔG was calculated as previously explained (4) on the basis of secondary structure models defined in Fig.3 for species 1,2,3, and 6, and elsewhere for Pseudomonas fluorescens (14) and the remaining species (1).



1 Sulfolobus solfataricus (8)
2 Thermus thermophilus (28)
3 Thermus aquaticus (28)
4 Thermoplasma acidophilum (29)
5 Bacillus stearothermophilus (29)
6 Bacillus acidocaldarius (10)
7 Bacillus subtilis (30)
8 Streptomyces griseus (30)
9 Pseudomonas fluorescens (29)
10 Paracoccus denitrificans (29)
11 Rhodospirillum rubrum (29)
12 Escherichia coli (29)
13 Clostridium pasteurianum (29)
14 Lactobacillus viridescens (29)

also been taken as evidence (4,17) for a conformational switch in this area of the molecule.

The sequence segments that give rise to band compression on the gels are boxed in Fig.3. Kramer and Mills (23) have explained the origin of this phenomenon. It is observed at the 3'-side of double-stranded areas when the fragments separated by electrophoresis have a common 5'-terminus, and on the 5'-side of double strands when the fragments have a common 3'-terminus, as is the case on the Peattie gels used here. Exceptions to this rule are residues 103-106 of S. solfataricus 5 S RNA and 62-67 in T. aquaticus 5 S RNA, which are at the 3'-side of helices. This can be explained by the presence, in the fragments subjected to gel electrophoresis, of hairpins that do not exist in the intact 5 S RNA. In S. solfataricus, the sequence $CCCA_{107}$ can pair with $UGGG_{117}$ to form the 5'-side of a hairpin. In T. aquaticus, $CGCGCC_{67}$ can similarly pair with $GGUGCG_{117}$. Thus, although there is a relation between band compression and secondary structure, the phenomenon should be interpreted cautiously since it does not necessarily reflect base-pairing as it occurs in the native structure.

As expected, the ΔG values corresponding with the secondary structures in Fig.3 are comparatively large since these structures have to withstand denaturation at elevated temperatures. A correlation between optimal growth temperature for several bacteria and ΔG of the 5 S RNA secondary structure

model can be seen in Fig.4. Although the ΔG values show a rather large dispersion among mesophilic bacteria, the minimal ΔG value observed in each temperature range clearly increases as a function of optimal growth temperature.

Identity of Caldariella acidophila and Sulfolobus solfataricus

Zillig et al. (9) distinguish three species belonging to the Sulfolobus genus : Sulfolobus acidocaldarius (24), Sulfolobus brierleyi (25) and Sulfolobus solfataricus (9). Whereas the former two species were isolated from American sources, Sulfolobus solfataricus isolates DSM 1616 and DSM 1617 originate from the same volcanic hot acidic springs at the Pisciarelli Solfatara (Agnano, Italy) from which the Caldariella isolates MT 3 and MT 4 of De Rosa et al. (8) were obtained. It was therefore suggested (9) that Caldariella acidophila and Sulfolobus solfataricus are the same species. The assignment of Caldariella acidophila to the genus Sulfolobus was supported by the close similarity between its ribosomal subunit protein patterns (26) and those of Sulfolobus species obtained from American, Italian, and Japanese sources (27). The results of the 5 S RNA sequence analysis reported here are also consistent with this assignment, which is why we have followed the suggestion of Zillig et al. (9) to use the species name Sulfolobus solfataricus.

ACKNOWLEDGEMENTS

We want to thank Drs. M. De Rosa and A. Gambacorta for supplying Caldariella acidophila cells, and Dr. R.J. Roberts for providing strain YT-1 of Thermus aquaticus. E. Dams holds an I.W.O.N.L. scholarship. Our research was supported in part by F.K.F.O. grant nr. 2.0017.80.

*To whom correspondence should be addressed

REFERENCES

1. Erdmann, V.A., Huysmans, E., Vandenberghe, A., De Wachter, R. (1983) Nucl. Acids Res. 11, r105-r133.
2. Huysmans, E., Dams, E., Vandenberghe, A., De Wachter, R. (1983) Nucl. Acids Res. 11, 2871-2880.
3. Studnicka, G.M., Eiserling, F.A., Lake, J.A. (1981) Nucl. Acids Res. 9, 1885-1904.
4. De Wachter, R., Chen, M.-W., Vandenberghe, A. (1982) Biochimie 64, 311-329.
5. Böhm, S., Fabian, H., Welfle, H. (1982) Acta Biol. Med. Germ. 41, 1-16.
6. Delihas, N., Andersen, J. (1982) Nucl. Acids Res. 10, 7323-7344.
7. Küntzel, H., Piechulla, B., Hahn, U. (1983) Nucl. Acids Res. 11, 893-900.
8. De Rosa, M., Gambacorta, A., Bu'Lock, J.D. (1975) J. Gen. Microbiol. 86, 156-164.

9. Zillig, W., Stetter, K.O., Wunderl, S., Schulz, W., Priess, H., Scholz, I. (1980) Arch. Microbiol. 125, 259-269.
10. De Rosa, M., Gambacorta, A., Bu'Lock, J.D. (1974) J. Bacteriol. 117, 212-214.
11. Brock, T.D., Freeze, H. (1969) J. Bacteriol. 98, 289-297.
12. Nazar, R.N., Matheson, A.T. (1977) J. Biol. Chem. 252, 4256-4261.
13. Du Buy, B., Weissman, S.M. (1971) J. Biol. Chem. 246, 747-761.
14. Dams, E., Vandenberghe, A., De Wachter, R. (1983) Nucl. Acids Res. 11, 1245-1252.
15. Nirenberg, M.W., Matthaei, J.H. (1961) Proc. Nat. Acad. Sci. U.S. 47, 1588-1602.
16. Kurland, C.G. (1966) J. Mol. Biol. 18, 90-108.
17. Fang, B.-L., De Baere, R., Vandenberghe, A., De Wachter, R. (1982) Nucl. Acids Res. 10, 4679-4685.
18. Peattie, D.A. (1979) Proc. Nat. Acad. Sci. U.S. 76, 1760-1764.
19. Nazar, R.N., Wildeman, A.G. (1981) Nucl. Acids Res. 9, 5345-5358.
20. Vandenberghe, A., De Wachter, R. (1982) J. Liq. Chromatogr. 5, 2079-2084.
21. Stahl, D.A., Luehrsen, K.R., Woese, C.R., Pace, N.R. (1981) Nucl. Acids Res. 9, 6129-6137.
22. Komiya, H., Kawakami, M., Takemura, S., Kumagai, I., Erdmann, V.A. (1983) Nucl. Acids Res. 11, 913-916.
23. Kramer, F.R., Mills, D.R. (1978) Proc. Nat. Acad. Sci. U.S. 75, 5334-5338.
24. Brock, T.D., Brock, K.M., Belley, R.T., Weiss, R.L. (1972) Arch. Microbiol. 84, 54-68.
25. Brierley, C.L., Brierley, J.A. (1973) Can. J. Microbiol. 19, 183-188.
26. Londei, P., Teichner, A., Cammarano, P., De Rosa, M., Gambacorta, A. (1983) Biochem. J. 209, 461-470.
27. Schmid, G., Böck, A. (1982) Mol. Gen. Genet. 185, 498-501.
28. Brock, T.D. (1978) in Thermophilic Microorganisms and Life at High Temperatures, Starr, M.P. Ed., pp. 52-55, Springer Verlag, New York.
29. Bergey's Manual of Determinative Bacteriology (1974), Buchanan, R.E. and Gibbons, N.E. Eds., 8th edn., The Williams and Wilkins Cy, Baltimore.
30. Bergey's Manual of Determinative Bacteriology (1957), Breed, R.S., Murray, E.G.D. and Smith, N.R. Eds., 7th edn., The Williams and Wilkins Cy, Baltimore.