EXTREMELY HALOPHILIC ARCHAEA FROM ANCIENT EVAPORITE DEPOSITS

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

by

Renia T. Gemmell BSc. (Hons.)
Department of Microbiology and Immunology
University of Leicester

March 1996

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346


LEICESTER UNIVERSITY
LIBRARY

Extremely Halophilic Archaea from Ancient Evaporite Deposits

Renia T. Gemmell

Extremely halophilic Archaea (halobacteria) were isolated from salt deposits of diverse geographical locations and ranging in geological age from Permian (250-260 million years (Ma)) to Recent. The diversity of these halobacteria was established by analyses of polar lipid patterns, protein electrophoresis patterns and ribotype bands; isolates from any one salt deposit were shown to be unique. Those halobacterial groups which predominate in present-day salt lakes and solar salterns, namely *Halobacterium*, *Halorubrum* and *Haloarcula* spp., were also shown to comprise the majority of isolates from ancient salt deposits. The inaccessibility of the deep subsurface mines, the nature of the sampling, and the 'extreme' requirements of these organisms for survival make contamination by contemporary halobacteria from surface sites highly unlikely. It is proposed that (some of) these salt mine isolates are autochthonous - presumably remnant populations derived from ancient hypersaline seas. This thesis aimed to determine whether these organisms are, in fact, 'living fossils' which have been revived. The approach used to test this hypothesis was to use the 16S rRNA gene as a molecular chronometer for the phylogenetic analysis of selected isolates representing most sites, and seek a correlation between phylogenetic position and geological age. This method proved not to be appropriate, possibly because of the natural phylogenetic diversity of the halobacteria. However, phylogenetic analysis has revealed new lineages in salt mines, distinct from known genera. That some of these halobacteria are 'living fossils', which remain locked within the salt in a state of dormancy until revived, is still an open question.

Statement.

This thesis, submitted for the degree of Doctor of Philosophy, entitled *Extremely Halophilic Archaea from Ancient Evaporite Deposits*, is based on work conducted by the author in the Department of Microbiology and Immunology, at the University of Leicester, mainly during the period between October 1991 and March 1995, under the supervision of Professor W. D. Grant. All the work recorded in this thesis is original unless otherwise acknowledged in the text or by references. None of the work has been submitted for another degree in this or any other University.

This thesis may be made available for consultation, photocopying and for use through other libraries, either directly or via the Leicester University Library.

Renia T. Gemmell
March, 1996.

Acknowledgements.

I am especially indebted to my mentor and supervisor, Professor Bill Grant, for his advice, support and encouragement.

My thanks to the Natural Environment Research Council for funding my PhD research.

Much is owed to the salt mine companies for allowing sampling, and their personnel who freely gave their time and assisted with the sampling: Dr. W. A. Longley-Cook, at Winsford Rock Salt Mine (Imperial Chemical Industries); Dr. R. Holmes and Dr. G. M. Jones, at Boulby (Cleveland Potash Ltd.); Teresa Kuc and Pawel Barmuta, at Wieliczka Salt Mine; and Director Misterski, at Inowroclaw Salt Mine. Thanks also to Grzegorz Czapowski, at the Polish Institute of Geology and Director Karnkowski of the Polish Oil and Gas Company, for advice and information. The salt and brine samples have been invaluable.

Thanks to Chris D'Lacey and Dr. C. Ockleford for confocal micrographs.

Thanks to present and past members of Lab. 105/107 and the Department of Microbiology and Immunology, for support and friendship; in particular Anne Abbott, Andy Duckworth, Rachel Greedy, Dr. D. Jones, Sheila Mackley, Terry McGenity, Gerald Owenson, Gary Pinch, Zeynep Sanal, Professor P. H. A. Sneath, John Sutherland and Maria-José Valderrama.

Many thanks to my mother, Christy, Natasha and Amelia, for their constant support and interest.

CONTENTS

List of Tables.

List of Tables ... Continued.

List of Figures.

List of Figures ... Continued.

I

GENERAL INTRODUCTION

1.1 THE EXTREMELY HALOPHILIC ARCHAEA (HALOBACTERIA).

In the first three decades of this century numerous investigations were made into the cause of a deteriorative reddening of fish and hides, a condition known variably as *'Red Heat'*, *'Red Eye'*, *'pink' etc.*; various rod and coccoid halophilic bacteria were identified as the causative agents and solar salt was given as the source. Not unusually, the classification was initially in some disarray, with the same bacteria being described under various names by different workers. These halophiles were commonly assigned to the genera *Bacillus*, *'Bacterium'*, *Sarcina*, *Micrococcus*, *Serratia* and *Pseudomonas* (Grant and Larsen, 1989; Tindall, 1992). Anderson (1954) had proposed that these red-pigmented, obligately halophilic organisms, which he described as requiring at least 15 % NaCl for growth, with optima between 25 and 30 % salt, should be classified in the genus *Halobacter* (in preference to *Halobacterium*), in the family *Pseudomonadaceae*; but this name was not adopted. By the time of publication of Bergey's 8th Manual of Determinative Bacteriology, in 1974, the rod-shaped and pleomorphic types were classified in the genus *Halobacterium*, and the coccoid types in the genus *Halococcus* (Gibbons, 1974). The names *Halococcus* and *Halobacterium* had originally been proposed by Schoop in 1935 (see Tindall, 1992).

The application of molecular analysis as a taxonomic tool enabled the revolutionary discovery, in the late 1970s, that the highest taxonomic division of Life was a tripartite one, and that the halobacteria were *bona fide* members of a newly-recognized Domain, the archaebacteria - now called the Archaea (Woese and Fox, 1977; Woese *et al.*, 1990). However, the systematics of the halobacteria remained essentially unchanged until the 1980s. Javor *et al.* (1982) then suggested a new genus, *Haloarcula*, for very pleomorphic types, but this was not formally proposed until 1986, when Torreblanca *et al.* described *Haloarcula* and *Haloferax* as two new genera of halobacteria, based on phenotypic characteristics and some chemotaxonomy, notably polar lipid composition. The haloalkaliphilic genera *Natronobacterium* and *Natronococcus*, were described by Tindall *et al.* in 1984, indicating that halobacteria may be a more diverse group than previously

realized (Grant and Ross, 1986). In 1985, Ross and Grant made a major impact in clarifying the taxonomic groupings of the halobacteria by defining nine distinct groups on the basis of 16S rRNA-DNA hybridization data. In 1989, Grant and Larsen proposed the new taxon *Halobacteriales* to include the (only) family *Halobacteriaceae*, in recognition of the ordinal status of the halobacteria. This decade therefore marked a major reworking of the systematics of the extremely halophilic Archaea (detailed in Bergey's Manual of Systematic Bacteriology; Grant and Larsen, 1989).

The genus *Halobacterium*, as redefined by Grant and Larsen (1989) included a number of species *incertae sedis*, notably '*Halobacterium trapanicum*', '*Halobacterium sodomense*', and '*Halobacterium saccharovorum*'. On the basis of a considerable accumulation of immunological (Conway de Macario *et al.*, 1986), chemotaxonomic (Ross and Grant, 1985; Lodwick *et al.*, 1986; Tindall, 1990) and molecular data (Lodwick *et al.*, 1991) it became clear that these species, together with more recent isolates such as '*Halobacterium lacusprofundi*' (Franzmann *et al.*, 1988), probably represented a distinct generic grouping, and the accompanying call for this recognition became a frequent one (Grant and Larsen, 1989). The taxonomic status of the '*Hb. saccharovorum* group' was finally clarified by McGenity and Grant in 1995, when they formally reclassified these halobacteria in a new genus, *Halorubrum*.

Also in 1995, Oren *et al.* characterized the elusive Dead Sea bloom-former and proposed *Halobaculum* as a new phylogenetic line of the family *Halobacteriaceae* for this organism. The taxonomic position of the eight currently-recognized genera of the *Halobacteriaceae* is as follows:

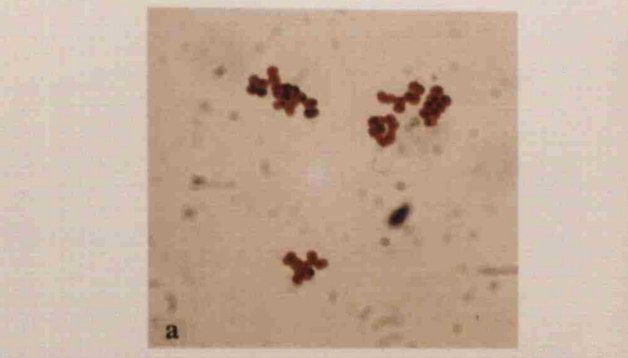
Domain Archaea
Kingdom Euryarchaeota
Division Mendosicutes
Class Archaeobacteria
Order Halobacteriales
Family Halobacteriaceae
Genus *Halobacterium*
Halorubrum
Haloarcula
Haloferax
Halococcus
Halobaculum
Natronobacterium
Natronococcus

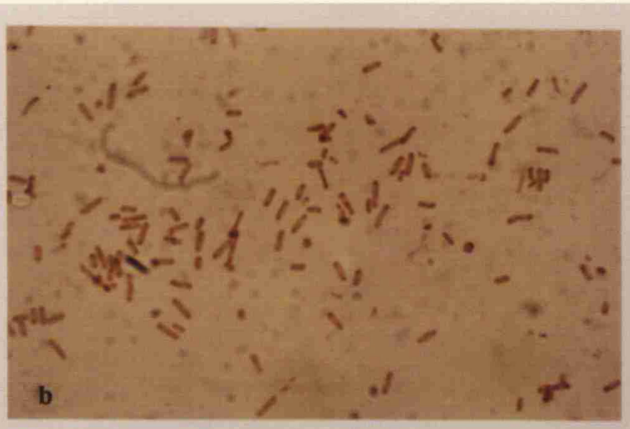
(The species belonging to these genera are given in Table 1.3.)

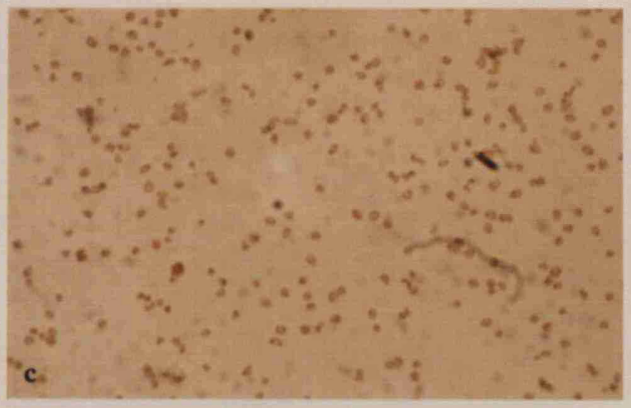
Ross and Grant (1985) had predicted seven of the above eight genera (excepting *Halobaculum*, which is a newly-recognized genus). The remaining two groups of Ross and Grant, 'Group 2' including "*Halobacterium salinarium*" (NCIMB 786), and 'Group 9' represented solely by "*Halobacterium cutirubrum*" (NCIMB 763), require to be fully described and formally given new genus status; and the respective names '*Halonema*' and '*Halothrix*' have been suggested for these genus-level groups (Grant and Ross, 1986). The confusion about the identities of certain strains of "*Halobacterium*" species, due to incomplete histories, is a current problem that is discussed by Tindall (1992).

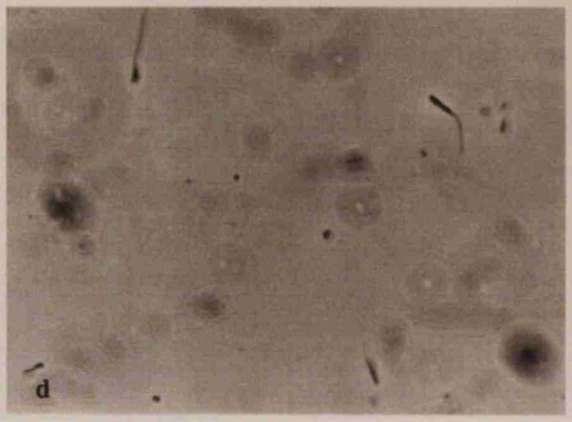
Recent 16S rRNA sequencing studies (Lodwick *et al.*, 1991; McGenity and Grant, 1995; Oren *et al.*, 1995) have confirmed the taxonomy given above, but it is additionally clear that other genera remain to be described. Kamekura and Dyall (1995) have recently proposed the new genus *Natrialba* to include two isolates which are distinct from other halobacterial genera on the basis of polar lipid analysis and phylogenetic analysis of 16S rRNA gene sequences. Despite the extremely halophilic Archaea occupying a very specific and extreme environment, the diversity of types is only now being realized.

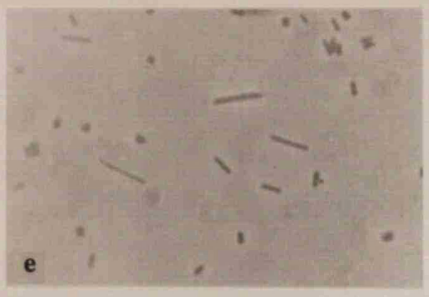
Halobacteria are variously shaped and described as cocci (Figure 1.1a), rods, ribbons (Figure 1.1b), cups, discs, triangles (Figure 1.1c), trapezoids, squares and irregular less-definable morphologies (Figures 1.1d-f). Individual species may be pleomorphic, a feature which may be due to sub-optimal growth conditions; "*Halobacterium vallismortis*" can be pear-shaped with a filamentous tail (Gonzalez *et al.*, 1978) and the normally disc-shaped "*Halobacterium volcanii*" can form hyphomicrobium-like structures (Mullakhanbhai and Larsen, 1975). The square halophilic bacterium ("*Quadra*") discovered by Walsby (1980; Parkes and Walsby, 1981) is intriguing, but has yet to be isolated in pure culture for full characterization and identification. The flat shapes of the square and triangular halobacteria are attributed to the absence of turgor pressure in these cells, which is related to the high salt concentration of the external environment (Walsby, 1980). Wais (1985) has reported the isolation from cyanobacterial mats of a halophilic archaeon which exhibits cellular morphogenesis, ranging from submicron cuboid cells within a thallus structure to elongated, motile cells that divide by budding.
The characteristic shapes of the Gram-negative species (excepting the cocci) are thought to be maintained by the glycoprotein S-layer on the cell surface. The cell envelope has been investigated in *Halobacterium halobium* and *Haloferax volcanii*. In *Halobacterium halobium* the surface glycoprotein is very acidic; over 20% of the amino acids are aspartic











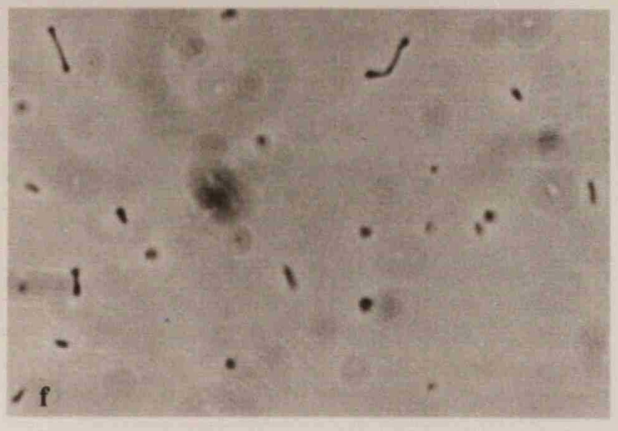


Figure 1.1 **Halobacterial Cell Morphologies.**

Gram Stains of:

a the *Halococcus*-like isolate B3br.19 (*c.* 2 μm diameter),
b the ribbon-shaped, unidentified isolate, B4br.3 (*c.* 2-3 μm x 1 μm),
c the *Haloarcula*-like isolate B3br.14 (*c.* 2 μm across).

Phase Contrast Micrographs of:

d PW4.5, a *Halorubrum*-like isolate exhibiting a tail-like morphology (*c.* 10 μm in length),
e GSL5.21, a *Halorubrum*-like isolate with a filamentous tail (*c.* 9 μm in length),
f B1E1, a *Halobacterium*-like isolate with a dumbell shape (*c.* 7 μm in length).

and glutamic acids (Kandler and Konig, 1993). It has long been recognized that Mg^{2+} ions are required to maintain the cell shape of halobacteria, and that low Mg^{2+} levels (10 mM in *'Pseudomonas salinaria'*) cause cells, excepting those of the halophilic cocci, to become spherical, which may be irreversible (Brown and Gibbons, 1955). This appears to be due to the requirement by the cell surface glycoproteins for Mg^{2+} ions. Reducing the $[Mg^{2+}]$ concentration to < 41 mM in *Haloarcula japonica* causes the cells to become spherical (Nakamura *et al.*, 1992). *Halococcus morrhuae* and *Natronococcus occultus* have no S-layer, but instead a rigid sacculus - of a sulfated heteropolysaccharide and "glycosaminoglycan", respectively (Kandler and Konig, 1993). Halobacteria, in common with other Archaea, but unlike the Bacteria, do not have peptidoglycan in their cell envelopes.

Most halobacteria require at least 1.5 M NaCl for growth (Grant and Larsen, 1989). They have an internal solute concentration of about 3.5 M KCl, 1 M NaCl and 0.1 M $MgCl_2$ (Kushner, 1986). As observed by Brown and Gibbons, as early as 1955, halobacteria store K^+ in their cells and osmoregulate with K^+, rather than the organic compatible solutes (osmolytes) found in halophilic Bacteria. Halobacteria have been shown to have glycine betaine, a compatible solute in certain halophilic Bacteria, but this is complexed with phosphatidyl glycerophosphate in the cell membrane and not free in the cytoplasm, and so is not believed to function as an osmoprotectant (Nicolaus *et al.*, 1989). The glycine betaine may, however, have some function in stabilizing the cell membrane; trehalose is thought to protect phospholipids and maintain membrane fluidity during desiccation, preventing damage on rehydration (Potts, 1994). Salt is required by halobacterial cells for stability and activity of ribosomes and enzymes, and for stability of the cell wall (see Kushner, 1986; Grant *et al.*, 1996).

Halobacteria have a chemoorganotrophic metabolism, utilizing amino acids or carbohydrates as a source of carbon for growth (Grant and Larsen, 1989). As a result of the early identification of halobacteria from salted proteinaceous products, such as fish and hides, various complex media were used for their isolation and growth (see Kushner, 1993). *Halobacterium* species, which require amino acids for growth (Grant and Larsen, 1989), were therefore frequently isolated from these media. Thai fish sauce is produced by the fermentation of fish in concentrated brine and a *Halobacterium salinarium* strain has been isolated from this product, its proteolytic activity thought to be essential to the fermentation process (Thongthai *et al.*, 1992). Similarly, *Halococcus* species may also spoil proteinaceous products and amino acids are needed for growth and also used as an energy source; glucose is not a major carbon source (Grant and Larsen, 1989). *Halorubrum* species utilize sugars as a carbon source, and some species also require

amino acids (McGenity and Grant, 1995). *Haloferax* and *Haloarcula* species also utilize carbohydrates, and other substrates, as sources of carbon and energy, but do not require amino acids for growth (Grant and Larsen, 1989). *Halobaculum* growth is stimulated by certain carbohydrates.

Halobacterial colonies are generally orange-, pink- or red-pigmented, due to the presence of C_{50} carotenoids (bacterioruberins) and, although recent studies have identified halobacteria which apparently lack this pigmentation (McGenity, 1994; Kamekura and Dyall-Smith, 1995), dense populations of these organisms impart a characteristic red coloration to hypersaline environments (Grant and Larsen, 1989).

1.2 HYPERSALINE ENVIRONMENTS: THE HALOBACTERIAL HABITAT.

'How surprising it is that any creatures should be able to exist in brine, and that they should be crawling among crystals of sulphate of soda and lime! ... Thus we have a little living world within itself, adapted to these inland lakes of brine. ... Well may we affirm, that every part of the world is habitable!' (Darwin, *The Voyage of the Beagle*; August, 1833).

Hypersaline environments are those where the total salinity is greater than that of seawater; *i.e.* >3.5 % (Grant *et al.*, in press). As discussed in Section 1.3, hypersaline environments are generally of either a marine or non-marine origin with, therefore, different modes of genesis. The terms *thalassohaline* and *athalassohaline* are used broadly to describe brines with salt compositions of a marine or non-marine nature, respectively (Table 1.1), and several examples are given below. The Great Salt Lake, though of freshwater origin, has a salts composition like that of seawater (Table 1.1) and so is frequently referred to as thalassohaline.

The hypersaline ecosystem is often highly productive, though limited in diversity, with bacteria, fungi (*Cladosporium* sp. *etc.*), algae (especially *Dunaliella* spp.; also *Nannochloris* spp. *etc.*), diatoms (*Navicula* spp., *Amphora coffeaeformis*, *Nitzschia* spp. *etc.*), protozoa (ciliates and zooflagellates), and invertebrates (rotifers, nematodes, gastropods, insects and the brine shrimp *Artemia salina, etc.*), (see Javor, 1989). These organisms, and presumably neighbouring plants, inevitably contribute to the significant levels of Dissolved Organic Carbon (DOC) in most of these brines. As examples, DOC levels are in the region of 4.2-8.3 mg l^{-1} in the Dead Sea, 43 mg l^{-1} in the Great Salt Lake, and up to 100 mg l^{-1} in Californian salterns (Javor, 1989). The compatible solutes (osmolytes) which accumulate and/or are produced in/by halophilic Bacteria (ectoine, hydroxyectoine, glycine betaine and trehalose, *etc.;* Severin *et al.*, 1992) or algae (glycerol in *Dunaliella*), as osmoprotectants in hypersaline environments, may also provide a nutrient source for other microorganisms. In certain environments, hydrocarbons, which are often associated with evaporites (Hite and Anders, 1991), may be a carbon and energy source for the growth of halophiles (Bertrand *et al.*, 1990; Gauthier *et al.*, 1992).

Table 1.1. The Ionic Composition of Seawater and Lake Brines (from Tucker, 1991).

	Seawater (ppm)	Lakewater (ppm)		
		Great Salt Lake, Utah	Dead Sea	Mono Lake, California
Cl^-	18 000	112 900	208 020	15 100
Na^+	10 770	67 500	34 940	21 400
SO_4^{2-}	2 715	13 590	540	7 530
Mg^{2+}	1 290	5 620	41 960	32
Ca^{2+}	412	330	15 800	11
K^+	380	3 380	7 560	1 120
HCO_3^- / CO_3^{2-}	140	180	240	26 430
Br^-	67			
H_3BO_3	26			
Sr^{2+}	8			
F^-	1.3			
H_4SiO_4	1			
Total Salinity	33 810.3	203 490	315 040	71 900

Table 1.2. Categories of Halophilic Microorganisms (Kushner, 1993; James *et al.*, 1990).

	Salt Concentration (M)	
Category	**Range**	**Optimum**
Non-halophile	0 - 1.0	< 0.2
Slight halophile	0.2 - 2.0	0.2 - 0.5
Moderate halophile	0.4 - 3.5	0.5 - 2.0
Borderline extreme halophile	1.0 - 4.0	2.0 - 3.0
Extreme halophile	2.0 - 5.2	> 3.0
Halotolerant	0 - 1.0	< 0.2
Haloversatile	0 - 5.2	0.2 - 0.5

Table 1.3. Halobacteria: Habitats and Salt (NaCl) Requirements for Growth.

Species	Source	Salt Range (M)	Salt Optimum (M)	Reference
Halobacterium salinarium	proteinaceous products	3-5.2	3.5-4.5	Grant and Larsen, 1989.
Halorubrum saccharovorum	saltern, San Francisco Bay	1.5-5.2	3.5-4.5	McGenity and Grant, 1995; Tomlinson and Hochstein, 1976.
Halorubrum lacusprofundi	Deep Lake, Antarctica	1.5 to sat^{n}	2.5-3.5	McGenity and Grant, 1995; Franzmann *et al.*, 1988.
Halorubrum sodomense	Dead Sea	0.5(2.5 Mg^{2+}) to 4.3	2 (0.6-1.2 M Mg^{2+}) (35 °C)	McGenity and Grant, 1995; Oren, 1983.
Halorubrum trapanicum	Trapani solar salt, Sicily	-	-	McGenity and Grant, 1995.
*"Halobacterium distributum"**	saline soils, Turkmen	> 1.7	2.6-4.3	Zvyagintseva and Tarasov, 1987.
Haloarcula vallismortis	salt pools, Death Valley	> 2.6	4.3	Torreblanca *et al.*, 1986; Gonzalez *et al.*, 1978.
Haloarcula marismortui	Dead Sea	1.7-5.2	3.4-3.9 (40 °C)	Oren *et al.*, 1990.
*"Haloarcula sinaiiensis"**	Sabkha Gavish Red Sea	2-5	3-4	Javor *et al.*, 1982.
Haloarcula japonica	salt field soil, Japan	~2.6-5.2 (42 °C)	~3.4 (42 °C)	Takashina *et al.*, 1990.
Haloarcula hispanica	solar saltern Alicante, Spain	1.9-5.2	3.2 (40 °C)	Juez *et al.*, 1986.
*"Haloarcula californiae"**	saltern, Baja California	> 2-5	3-4	Javor *et al.*, 1982.
Haloferax volcanii	bottom mud, Dead Sea	1- < 5	1.7 (30 °C) 2.5 (40 °C) (0.2M Mg Cl_2)	Torreblanca *et al.*, 1986; Mullakhanbhai and Larsen, 1975.
Haloferax mediterranei	marine saltern Alicanti, Spain	1.3-4.6	2.9 (40 °C)	Rodriguez-Valera *et al.*, 1983.
Haloferax gibbonsii	solar saltern Alicante, Spain	1.5-5.2	2.9-3.6 (40 °C)	Juez *et al.*, 1986.

Table 1.3 Continued ... 2 ...

Species	Source	Salt Range (M)	Salt Optimum (M)	Reference
Haloferax denitrificans	salterns, San Francisco	1.5-4.5+ (37 °C)	2-3 (37 °C)	Tomlinson *et al.*, 1986.
Halococcus morrhuae	Dead Sea	> 2.5	3.5-4.5	Grant and Larsen, 1989.
Halococcus saccharolyticus	saltern	~2.6-5.2	3.4	Montero *et al.*, 1989.
Halococcus salifodinae	salt mine (rock salt), Austria	> 2.6	~3.4-4.3	Denner *et al.*, 1994.
*"Halococcus turkmenicus"**	saline soil, Turkmen	> 1.7	2.6-4.3	Zvyagintseva and Tarasov, 1987.
Halobaculum gomorrense	Dead Sea	>1 (+ 0.8 M Mg Cl_2)	2.1 (0.8 M Mg Cl_2) (40 °C)	Oren *et al.*, 1995.
Natronococcus occultus	Lake Magadi, Kenya	1.4-5.2 (37 °C)	3.5-3.6 (37 °C)	Tindall *et al.*, 1984.
Natronococcus amylolyticus	Lake Magadi, Kenya	1.4-5.2	2.6-3.4	Kanai *et al.*, 1995.
Natronobacterium gregoryi	Lake Magadi, Kenya	2-5.2	3	Tindall *et al.*, 1984.
Natronobacterium magadii	Lake Magadi, Kenya	2-5.2	3.5	Tindall *et al.*, 1984.
Natronobacterium pharaonis	Wadi Natrun, Egypt	2-5.2	3.5	Tindall *et al.*, 1984.
Natronobacterium vacuolata	Lake Magadi, Kenya	2.6-5.2 (37 °C)	3.5 (37 °C)	Mwatha and Grant, 1993.

*" species" *incertae sedis.*

The halophilic microorganisms which inhabit these environments are generally categorized according to the range of salt (NaCl) concentrations within which they can grow, and the salt concentration for optimum growth (Table 1.2). The halophilic Bacteria, which belong to a number of different phylogenetic groups (see Grant *et al.*, in press), tend to predominate in salt concentrations up to 2 M NaCl. Of particular abundance in salterns are Gram-negative aerobic chemoorganotrophs, many of which are members of the family *Halomonadaceae*, which comprises the genera *Deleya* and *Halomonas*, and the species *Alcaligenes aquamarinus* and *Halovibrio variabilis* (Franzmann and Tindall, 1990).

The extremely halophilic Archaea (halobacteria) are usually the dominant populations in NaCl concentrations greater than 2.5 M and their ecology has been recently reviewed in detail (Oren, 1994a). The habitats and salt requirements of currently described species of halobacteria are given in Table 1.3. Most halobacteria require at least 1.5 M NaCl for growth and, except for *Halococcus*, lyse in distilled water. *Haloferax* spp., *Halobaculum gomorrense*, and *Halorubrum sodomense* in particular, can tolerate lower salt concentrations and have lower optima; this is probably related to their high Mg^{2+} requirements. Optimal salt concentrations for halobacterial growth are generally around 3 M NaCl, which may be less than the salt concentration of the source environment (often at NaCl saturation).

The association of halobacteria with hypersaline environments has been alluded to since ancient times, and there appear to be innumerable references to this (Baas-Becking, 1931). Solar salt manufacture is described in an ancient Chinese treatise believed to date to about 2700 B.C.: *'clear sea water ... is left for a long time until the color becomes red...'* (Baas-Becking, 1931). The Roman, Pliny, in his *Natural History* (*c.* 77 A.D.), remarks on differences in the colour of salt, as *"at Memphis it is a deep red, russett-coloured in the vicinity of the Oxus, purple at Centuripa ..."*. But millenia were to pass before it was attributed to the presence of red-pigmented bacteria; salt, used for centuries as a preservative of fish, meat and hides, was considered sterile and devoid of life. This 'red brine' phenomenon was correctly, though incompletely, explained by Darwin (1845), in his account of a visit to a salina near the Rio Negro, in Patagonia, in August 1833; he observed that *'Parts of the lake ... appeared of a reddish colour, and this perhaps was owing to some infusorial animalcula'*. This typical red colouration of halobacterial environments is due to the production of C_{50} carotenoids (bacterioruberins) by these bacteria.

Oxygen solubility (22 °C) decreases with increasing salinity, from 7 mg kg^{-1} in seawater to 2 mg kg^{-1} at the onset of halite precipitation (Javor, 1989), and so oxygen availability may be quite limited in hypersaline environments. A variety of anaerobic halophiles have

now been isolated, as reviewed in a recent paper by Ollivier *et al.* (1994), and include fermentative bacteria, phototrophs, sulphate reducers and methanogens. The fermentative, obligately anaerobic halophilic chemoorganotrophic Bacteria, of the family *Haloanaerobiaceae* (Oren, 1984), have been isolated from sediments of the Dead Sea, the Great Salt Lake, other salt lakes, saline lagoons, a saltern, an oil field brine, and subsurface brines, and eight genera are currently recognized (see Grant, *et al.*, in press): *Halobacteroides*, *Haloanaerobium*, *Haloanaerobacter*, *Haloincola*, *Halocella*, *Halothermothrix*, *Sporohalobacter* and *Acetohalobium*. Moderately and extremely halophilic sulphur-oxidizing phototrophs of the family *Ectothiorhodospiraceae*, are commonly isolated from the sediment surface (where light reaches) in marine salterns, salt lakes and soda lakes. Although generally considered to be aerobic organisms, certain of the halobacteria can grow microaerophilically in the light by utilizing bacteriorhodopsin, in those strains which produce it, or anaerobically by the fermentation of arginine (Hartmann *et al.*, 1980) or by using nitrate (Grant and Larsen, 1989); no strictly anaerobic halobacteria are known to have been isolated.

1.2.1 Great Salt Lake.

In the Great Salt Lake, thalassohaline in character, the indigenous bacterial populations are halophilic and include phototrophs (*Ectothiorhodospira* sp.), obligate anaerobes (*Haloanaerobium praevalens* and *Methanohalophilus mahii*), and members of the extremely halophilic Archaea. These latter dominate the bacterial flora of the more saline North Arm (about 33 % (w/v) salinity). Primary producers (cyanobacteria, algae and diatoms) have been isolated from both arms of the lake, but the populations of each arm differ. The alga *Dunaliella salina* is found only in the North Arm, where it outgrows *Dunaliella viridis*, the dominant primary producer in the South Arm (about 12 % (w/v) salinity). The cyanobacterium *Aphanothece halophytica*, which grows optimally in 16-23% (w/v) NaCl, is common and has been isolated from microbial mats from both arms. Several diatom species (*Amphora coffeaeformis* is the most common) have been isolated from the South Arm. The brine shrimp, *Artemia salina*, also makes a considerable contribution to nutrient recycling in this ecosystem (see Javor, 1989).

1.2.2 Dead Sea.

The Dead Sea is athalassohaline and has a salt composition quite different to that of the Great Salt Lake (Table 1.1), in particular a high [Mg^{2+}] which tends to limit growth of the alga *Dunaliella*, which is the main or only primary producer in the Dead Sea and upon which the bacteria depend for organic nutrients. Glycerol, which *Dunaliella* accumulates as a compatible solute, is readily used as a carbon and energy source by halobacteria and may be the main source, made available by leakage or released when the algae die (Oren,

1993). Halobacterial blooms are therefore closely linked with the growth of *Dunaliella*, which in addition is growth-limited by phosphate availability and a requirement that the salinity be below 12.1-12.2 % (w/v) (Oren, 1994a). The ecosystem is dominated by various halobacteria (Table 1.3) and certain of these have a characteristic requirement for a relatively high [Mg^{2+}] for growth: *i.e. Haloferax volcanii*, 0.2 M; *Halorubrum sodomense*, 0.6-1.2 M; and *Halobaculum gomorrense*, 0.8 M. These halobacteria can tolerate lower NaCl concentrations if that of $MgCl_2$ is increased: *Halorubrum sodomense* can grow fairly well in 0.5 M NaCl with 2.5 M $MgCl_2$, but requires 1.7 M NaCl if the [$MgCl_2$] is reduced to 1.8 M (Oren, 1983). These species are therefore well-adapted to their natural habitat.

1.2.3 Solar Salterns.

The thalassohaline solar saltern is designed as a series of evaporation ponds with a step-wise gradient of increasing salinity, from above that of seawater to salt saturation. The environments of these ponds differ; with increasing salinity, the oxygen concentration drops, pH decreases and temperature increases. The more saline brines have a lower specific heat capacity and, where they are coloured red by a halobacterial community, heat faster and attain higher temperatures than the clearer brines (Javor, 1989). Salt-saturated ponds have a pH one unit lower, and a temperature 10 C° greater, than ponds with less than 15 % (w/v) total salts (*c.* 10 % (w/v) NaCl); (Rodriguez-Valera *et al.*, 1985). It would be expected, therefore, that the ecology of these ponds will differ, and so this multi-pond complex has become the focus of several studies on the microbial ecology of saline environments (Ventosa *et al.*, 1982; Rodriguez-Valera *et al.*, 1981, 1985; Marquez *et al.*, 1987). The populations of both eukaryotic and prokaryotic halophilic microorganisms change with total salts concentration, with a decrease in diversity as salt concentration increases (Rodriguez-Valera *et al.*, 1985). The most abundant Bacteria in a solar saltern in S.E. Spain, studied by Ventosa *et al.* (1982) and Rodriguez-Valera *et al.* (1985), appeared to be Gram-negative aerobic chemoorganotrophic species of the genera *Pseudomonas*, *Alteromonas*, *Alcaligenes*, *Vibrio*, *Flavobacterium*, and *Acinetobacter* (and probably *Deleya* spp.), and these were dominant in the region of 10-30% salts. Of the eukaryotes *Dunaliella* numbers were highest at 20-30 % (w/v) salts, beyond which they decreased. Halobacteria were the dominant bacteria above 30 % (w/v) salts (20.5 % (w/v) NaCl), but there were differences in the halobacterial populations of ponds at different stages of the crystallization process: *Haloferax mediterranei* were dominant in ponds with 10-20 % (w/v) total salts and not isolated from brine with over 40 % (w/v) salts, whereas *'Halobacterium' saccharovorum* (now *Halorubrum saccharovorum*) numbers increased with increasing salt concentration; *Haloferax gibbonsii* was one of the dominant forms in ponds with over 20 % (w/v) salts. Marquez *et al.* (1987) found a

similar predominance of Bacteria, as described above, in a solar saltern in S.W. Spain, and additionally identified *Deleya* spp. as forming the majority.

Del Moral *et al.* (1987a,b; 1988) studied an athalassohaline saltern, near Granada, where brine was drawn from a subterranean well (La Malá) into ponds of salinities ranging from 18 % (well water; 5.7 % Na^+) to 36 % (w/v) total salts. Like the thalassohaline saltern studied (above), most of the Bacteria isolated were Gram-negative aerobic chemoorganotrophic species of the genera *Vibrio*, *Pseudomonas*, *Deleya*, *Acinetobacter*, *Flavobacterium*, *Alteromonas* and *Alcaligenes*. Similar populations (*Vibrio*, *Acinetobacter*, *Alteromonas* and *Marinomonas* spp.) were found in a geographically distinct site, in Chile (Prado *et al.*, 1991). Halophilic Gram-positive cocci (*Planococcus*, *Paracoccus*, *Micrococcus* and *Sporosarcina* spp.) were less frequently isolated. Archaeal extreme halophiles predominated in all samples on 25 % (w/v) salts media. To demonstrate the evolution of the halophile communities when brines are diluted or concentrated, enrichments of well water were subjected to decrease / increase in salinity (Del Moral *et al.*, 1987b). Halotolerant bacteria were only found in low salt concentrations, but never dominant; marine bacteria predominated in salt concentrations up to 18-25 % (w/v) (depending on the temperature), but were not isolated in a separate study of this saltern (Del Moral *et al.*, 1987a); moderate halophiles (Table 1.2) predominated in greater than 18 % (w/v) salts at 25 °C, and in 30 % (w/v) salts at 35 °C; and archaeal extreme halophiles (halobacteria) were the dominant organisms in over 35 % (w/v) salts at 35 °C.

1.2.4 Antarctica.

The freezing conditions of Antarctica, having a similar effect to evaporative concentration, have enabled the evolution of hypersaline lakes (Javor, 1989). One of these is Deep Lake, in the Vestfold Hills, which is thought to be thalassohaline (Javor, 1989), and has 21-28 % (w/v) salinity (Franzmann *et al.*, 1988). This lake can have surface temperatures as low as -20 °C, but, with a freezing point of -28 °C, is ice-free (Javor, 1989); the highest recorded temperature is +11.5 °C. Although *Dunaliella* has been isolated from this lake, primary productivity is low, and the presence of autochthonous bacteria has been arguable. However, a new species of halobacteria, *Halobacterium lacusprofundi*, has been isolated from Deep Lake. This organism is able to grow at 4 °C (optimum is 31-37 °C), albeit slowly, but as the temperature of the lake is above +1 °C for only 3 months of the year, the growing period is a short one (Franzmann *et al.*, 1988). New species of halotolerant and moderately halophilic Bacteria (*Halomonas* spp. and *Flavobacterium* spp., respectively) have been isolated from other hypersaline lakes in the Vestfold Hills (James *et al.*, 1990; Dobson *et al.*, 1993).

1.2.5 Saline Soils.

In arid regions, groundwater may be drawn upwards by capillary action, and the dissolved salts consequently precipitated subaerially or on the land surface. These saline soils are subject to fluxes in salt concentration, particularly as a result of rainfall fluctuations. In a study of hypersaline soils near Alicante, in Spain (Quesada *et al.*, 1982), most of the bacterial isolates had NaCl optima of moderate halophiles (5-10% (w/v)), but grew within a wide salt range, indicating haloversatility (James *et al.*, 1990; Table 1.2). The predominance of this type is explained by the variability of the salinity of these soils. Only 3, of 223 strains characterized, were halobacteria (*"Halobacterium"*), but this could be due to the moderate salt concentration of the soil samples (5.0-10.7% (w/v)), the moderate incubation temperature (30 °C) and insufficient incubation time (15 days). Similarly, both halophilic and halotolerant bacteria, but only one halobacterium (*Haloarcula japonica*), were isolated from the Agehama salt field, in Japan, which is subject to seasonally varying salinity (Takashina *et al.*, 1994). The validly published but probably misclassified halobacteria *"Halobacterium distributum"* and *"Halococcus turkmenicus"* were isolated from sulphate saline soils in Russia (Zvyagintseva and Tarasov, 1987).

1.2.6 Transport of Halobacteria.

The halobacterial habitat is a restricted one. If halobacteria are the predominant populations in salt-saturated brines, then the question of how they arrived in these environments inevitably arises. Marine salterns are the product of the evaporative concentration of seawater and so one might assume that at least some of the halobacteria came from the seawater. Recent investigations of the communities in the picoplankton of coastal (DeLong, 1992), pelagic (DeLong *et al.*, 1994) and open ocean subsurface (Fuhrman *et al.*, 1992,1993) environments, using 16S rRNA phylogenetic analysis, have identified two novel (uncultivated) archaeal groups; but no members of the *Halobacteriaceae* appear to have been found in this way.

It is generally accepted that, except for *Halococcus* spp., halobacteria will lyse in distilled water, but the point at which lysis occurs will depend on the medium (Grant and Larsen, 1989). *Halococcus* spp. can survive for at least 2 hours, and sometimes for as long as 24 hours, when suspended in distilled water (Denner *et al.*, 1994), so marine transport of *Halococcus* spp. is easily accepted. Rodriguez-Valera *et al.* (1979) isolated 2 to 35 *Halococcus* spp. from 5 litres of filtered seawater from a site 5 km off the coast of Spain, within 15 km of onshore salterns. The isolates were shown to survive in seawater for at least 7 days (10 % of the population were still viable after 3 months), but required at least 2 M NaCl for growth. *Halococcus* spp. are able to maintain metabolic activity in seawater (*e.g.* amino-acid transport and respiration), although growth does not occur (Rodriguez-

Valera *et al.*,1982). Ventosa *et al.* (1984) isolated 1 or 2 colonies of (presumably) *Halococcus* spp. from 10 L of filtered seawater, distant from hypersaline sites.

Whilst *Halococcus* spp. can conceivably be transported by seawater, the medium of dispersal of the remaining halobacteria is open to debate, although there have been occasional reports of these tolerating very low NaCl concentrations. Rodriguez-Valera *et al.* (1983) found that 10 % of *Haloferax mediterranei* cells survived exposure to 1 % seawater, and most of the population to 3 % seawater (equivalent to about 0.14 and 0.44 M NaCl respectively). Torreblanca *et al.* (1986) found that *Haloarcula* spp. required 0.17-0.65 M NaCl to maintain cell stability and *Haloferax* spp. only 0.51 M NaCl to prevent lysis.

It therefore seems feasible that, although halobacteria cannot grow at seawater salt concentrations, some can survive to colonize new habitats and this should be investigated more thoroughly. Those halobacteria more sensitive to low salt concentration may be protected. One may speculate that *Halobacterium* spp., for example, which are known to colonize salted fish, may survive marine transport by living commensally on fish or crustaceans. Halobacteria may exist as ultramicrobacteria in the oceans, having a survival form akin to spores, or they may form microbiofilms on sand, or be nucleated in crystals in seabed sediments. Other means of transport of halobacteria between hypersaline surface environments are within salt crystals by aeolian means, or on the feet of migrating birds.

1.2.7 Ancient Salt Deposits.

Halophilic microorganisms have been isolated from brines and rock salt from within ancient halite deposits. Namyslowski (1913) described halophilic bacteria (*e.g.* *'Bacterium halophilum'* and *'Bacterium salinum'*) in NaCl-saturated brines from underground lakes in Wieliczka salt mine in Poland which he was able to culture in liquid broth saturated with NaCl, though not on solid media. *'Bacterium salinum'*, described as imparting a *roseroter* colour to the broth, was likely a halobacterium. Smyk (1992) reported on research showing the activity of fungi and bacteria in Wieliczka Salt Mine, though with emphasis on the potential destructive effect on the relicts and sculptures. Fungi included several *Aspergillus* species. Bacteria found in the mine were *Streptomyces* spp., *Sporocytophaga* spp., *Cytophaga* spp., *Micrococcus halodenitrificans*, *Halobacterium halobium* and *Halobacterium salinarium*. These microorganisms were isolated from mining chambers, a brine tub and wooden transport equipment.

Halophilic Bacteria and halobacteria have also been isolated from Triassic and Permian salt deposits (Stan-Lotter *et al.*, 1993; Norton *et al.*, 1993; Denner *et al.*, 1994; McGenity, 1994; this study). This will be discussed in detail in the Results and Discussion section (3.1).

1.3 THE GEOLOGY AND GEOGRAPHY OF EVAPORITE DEPOSITS MODERN AND ANCIENT.

In order to appreciate salt deposits as natural environments for halophilic bacteria it is necessary to understand the nature of salt deposits - how they evolve, their physico-chemical attributes, where they are found, *etc.* This also provides the foundation required to assess the feasibility of the hypothesis that these bacteria are able to survive within salt deposits as 'living fossils' for millions of years.

Salt deposits, or *evaporites*, result from the evaporitic concentration of dissolved salts in water of marine, lacustrine, or mixed origin, and their consequential precipitation and crystallization as chemical sediments. Climatic factors such as temperature (air and water difference), wind speed and relative humidity of the atmosphere, determine the rate of evaporation. Water loss by evaporation should exceed input by rainfall, runoff or inflow, and loss by outflow should be limited for these salt deposits to accumulate and be maintained. In the deposition of marine evaporites a topographical barrier, such as a sill or reef, contains and isolates the salt basin (*e.g.* in a lagoon), but allows inflow of seawater to replenish the supply of salts. Salt lake survival requires a closed drainage basin, or one with restricted outflow. The salts are precipitated subaqueously, close to the air-water interface and upon the bottom sediment.

These physico-chemical processes, as well as the origin of the salt-water, determine the nature of the salts deposited and their relative proportions. Typically, the evaporative concentration of seawater precipitates salts (simply) in the following order, with the least soluble being deposited first: Ca CO_3 and other carbonates, anhydrite ($CaSO_4$) or gypsum ($CaSO_4.2H_2O$), halite (NaCl), magnesium and potassium chlorides and sulphates (bittern salts). The highly soluble bittern salts are the last to crystallize out, after prolonged desiccation. McCaffrey *et al.* (1987) showed, with brines sampled at a Bahamas solar salt production facility, that the evaporation path of seawater began with the precipitation of calcium carbonate at a brine concentration of 1.8 times that of seawater, followed by

Table 1.4. Common Evaporite Minerals and their Origins (Braitsch, 1971, and Tucker, 1991).

Mineral	Formula	Occurrence
halite	NaCl	marine and non-marine; the most important halide; primary and metamorphic
sylvite	KCl	marine; primary and from solution metamorphism (*e.g.* from carnallite)
carnallite	$KMgCl_3.6H_2O$	marine; primary and secondary (less common; cooled metam. solns.). Also non-marine, diagenetic (Casas *et al.*, 1992).
kainite	$K_4Mg_4Cl_4(SO_4)_4.11H_2O$	marine; primary; secondary from kieserite + sylvite/carnallite
anhydrite	$CaSO_4$	marine and non-marine; from gypsum
gypsum	$CaSO_4.2H_2O$	marine and non-marine; primary ppt. from seawater; secondary from anhydrite, polyhalite, *etc.*
polyhalite	$K_2MgCa_2(SO_4)_4.2H_2O$	most important complex Ca salt in marine salt deposits; met. from kainite + Ca^{2+}-rich solns.
kieserite	$MgSO_4.H_2O$	marine; from diagenetic recrystallization; metamorphic with sylvite
epsomite	$MgSO_4.7H_2O$	non-marine; possibly primary ppt. at low temperatures
glauberite	$Na_2Ca(SO_4)_2$	marine and non-marine; primary and alteration of other salts
mirabilite	$Na_2SO_4.10H_2O$	non-marine; in Na_2SO_4 - rich recent salt lakes (winter ppt.)
thenardite	Na_2SO_4	non-marine; often from mirabilite
bloedite	$Na_2Mg(SO_4)_2.4H_2O$	primary in Na_2SO_4 salt lakes. Secondary from kieserite in marine salts.
trona	$Na_2CO_3.NaHCO_3.2H_2O$	non-marine; in soda lakes
gaylussite	$Na_2CO_3.CaCO_3.5H_2O$	non-marine; in soda lakes, from calcite + alkaline brines

gypsum at 3.8 times -, halite at 10.6 times -, and (in the laboratory) magnesium sulphate at 70 times -, and potassium salts at concentrations over 90 times those of seawater. The solubility of most natural salts increases with rise in temperature. Calcium sulphate, however, has decreased solubility with increase in temperature and so is deposited in summer, alternating with halite in 'year rings' (Hatch *et al.*, 1971). Theoretically, the complete evaporation of a column of seawater 300 m deep would precipitate about 4.8 m of salts, of which 0.0192 m would be carbonates, 0.173 m gypsum, 3.7 m halite and 0.85 m potassium and magnesium salts (Handford, 1991).

In an earlier study of non-marine brine evolution, in closed basin lakes, Hardie and Eugster (1970) showed that the path to a brine type was determined very early in its evolution by the precipitation of calcite (always the first), sepiolite (a magnesium silicate) and gypsum. They determined four major naturally occurring brine groups: (i) Na-CO_3-SO_4-Cl brines (*e.g.* Lake Magadi, Kenya; Deep Springs Lake, California); (ii) NaCl-Na_2SO_4 brines (*e.g.* Saline Valley, California; Great Salt Lake); (iii) Na-Mg-Ca-Cl brines (*e.g.* Bristol Dry Lake, California) and (iv) Na-Mg-SO_4-Cl brines (*e.g.* Poison Lake, Washington), and predicted the mineral types precipitated with continued evaporation. They found considerable agreement between this model and the actual minerals precipitated by well-studied closed basins.

Although the concentration of dissolved salts in seawater (salinity) varies between sites (*e.g.* salinities of 34-36 *per mille* over much of the Pacific, Atlantic and Indian Oceans, to 38 *per mille* in the Sargasso Sea and above 40 *per mille* in the Gulf of Suez) the ionic composition is virtually constant (Hatch *et al.*, 1971). The ionic composition of salt lakes, on the other hand, is quite varied (Table 1.1). The nature of the precipitating salts is therefore quite different in marine and lacustrine environments (Table 1.4).

Modern, or *Recent*, salt deposits can be found in many arid and semi-arid regions of the world and include marine and lacustrine evaporites, and those which fall outside these categories, such as saline soils. However, the environmental conditions required for salt deposition have recurred periodically throughout geological time, and ancient salt deposits, from late Pre-Cambrian to Quaternary, which are invariably exploited as rich mineral resources, can be found world-wide.

1.3.1 Halite Crystallization (partially from Handford, 1991).

The development and nature of halite crystals is determined by the environment - the temperature, ionic composition and purity of the brine, the properties of the bedrock and sediments, *etc.* In a hypersaline environment, halite crystallization can take place at a number of sites: on the floor of a brine pool, at the brine / air interface, on detritus within the pool, displacively within sediments, within rock cavities, fractures and fissures, or as efflorescence on exposed rock surfaces. Organic matter can provide nucleation sites for

halite crystal growth. Bacteria which can survive in saturated brine could hypothetically become living nuclei to seed the growth of salt crystals.

In a brine pool, for instance, crystal nucleation usually occurs first at the brine / air interface, where, commonly, pyramidal hopper crystals then develop (plates at lower temperatures). If undisturbed, these hollow, inverted hoppers coalesce to form chains and then rafts or mats, possibly drifting across the brine pool, until these halite cumulates eventually sink to the bottom, where they aggregate. At the bottom of the brine pool, in addition to overgrowth of sunken rafts, cuboid (where uncrowded), cornet and chevron (upward-facing) crystals develop and cement the thickening halite layer. Arthurton (1973) showed experimentally that halite crystallization was initiated both on the surface and the floor of the brine pool, but was greater at the surface during the day, when it was hotter, and on the floor at night when it was cooler. He found that the deeper the brine pool, the larger and fewer the crystals.

Within brine-soaked sediments, halite crystals may displace the sediment as they grow, or, if the growth is particularly rapid, the sediment may be incorporated within the crystal as mud inclusions.

Fluid Inclusions.

Changes in environmental conditions (*e.g.* temperature) and therefore crystal growth rates often result in the formation of growth bands within the crystals, as chevrons, marked by the sediment and fluid inclusions they contain. Fluid inclusion-rich zones give a cloudy appearance to the crystals and indicate rapid crystallization of the salt. In the case of primary crystals, *i.e.* those which have not resulted from recrystallization, these primary fluid inclusions contain the mother liquor (Na^+, Mg^{2+}, K^+, Cl^- and SO_4^{2-} ions in solution) and atmospheric gases (variable amounts of, *e.g.*, vapour, CO_2, N_2, and also CH_4 under pressure) of the original brine pool. It has been proved possible to extract the brine from a single fluid inclusion, of more than 250 μm in diameter, and determine the concentrations of ions and pH of the brine (Lazar and Holland, 1988). Analysis of the chemical composition of these primary fluid inclusions is useful in evaluating the evolution of the surface brine pool and may be used to establish whether a juxtaposed evaporite is primary or diagenetic in nature (Casas *et al.*, 1992). Fluid inclusions may also contain solid organic matter and tiny droplets of petroleum are sometimes found. The latter prompted Roedder (1984) to remark: *"fossil halophilic (salt-loving) bacteria and algae may be present and should be looked for"*.

Primary fluid inclusions *"might easily survive deformation and recrystallization"* (Shlichta, 1968). They might also be incorporated within bedded halite deposits during early post-depositional diagenetic processes. Dissolution and recrystallization of the salt deposit on the floor of the basin produces clear halite which overlays truncated chevron

(primary) salt. The fluid inclusions in this clear halite may have the same composition as those in the original primary halite (Roedder, 1984), but are generally larger (several hundred microns to several millimeters) than those in primary halite (<1 to <100 μm) and are useful in determining the nature of the post-depositional diagenetic processes which occurred (Stein and Krumhansl, 1988). Secondary fluid inclusions may arise as a result of recrystallization within the deposit. Different fluid inclusions within the same salt deposit can vary in composition and may be explained by localized differences in the original evaporating salt basin (Stein and Krumhansl, 1988) or migrating fluids (Roedder, 1984).

Primary fluid inclusions of ancient salt deposits would therefore contain ancient brines, gases and organic matter. Microorganisms found within these fluid inclusions would be autochthonous with the salt deposit. If these microorganisms were capable of tolerating salt-saturation (*i.e.* halophilic bacteria and halobacteria) and somehow survived any physico-chemical changes that would have taken place since primary crystallization of the salt deposit, perhaps adopting some survival strategy, they may be cultured from the crystals, and under stringent aseptic conditions, could be shown to represent living fossils, survivors of the populations that once inhabited the ancient lake or sea.

Diagenesis.

Primary halite may be modified by syndepositional or early postdepositional diagenesis, due to physical and chemical processes resulting from solar heating and flooding by surface waters. Changes due to dissolution and reprecipitation are significant in shallow brine basins subject, for example, to periodic flooding. The halite may be reprecipitated as a clear cement, filling dissolution cavities. Crusts of halite deposition on salt flats in intensely evaporative environments such as desert areas, where the diurnal temperature ranges are also significantly high, are subject to thermal expansion and contraction, which may result in the formation of polygonal fractures; (Tucker and Tucker, 1981; Handford, 1991). These were seen on the roofs of tunnels at Winsford Rock Salt Mine (Section 3.1.7) suggesting that the Triassic climate in this region was a particularly arid one.

1.3.2 Recent Salt Deposits.

Salinas.

Salinas develop on coastland, where a sandy barrier isolates a shallow lagoon or arm of the sea, but where salts are replenished by seepage and periodic flooding, and deposited subaqueously by evaporative concentration. Coastal salinas are found in many arid or semi-arid regions of the world, such as the Mediterranean, on the Spanish (see Section 3.1.3) and Portugese coasts and Cyprus, the Sinai coast, North Africa, the Bahamas,

Central America and Australia. Lake MacLeod, in Western Australia, developed as an extensive salina after complete isolation from the Indian Ocean some 5,000 years ago. It is 140 km long, 50 km wide and 3-4 m below sea level, it is isolated from the Indian Ocean by a Pleistocene dune ridge up to 90 m high, 3-15 km wide and 180 km long. The salina brines are replenished by seepage of seawater through this barrier, allowing the accumulation of deposits of carbonate, gypsum and halite (Kendall and Warren, 1988; Handford, 1991). Salinas can also occur in non-marine settings. Salinas, in coastal or continental areas, are frequently exploited for the harvesting of salt (sodium chloride) and artificial salinas or salt pans have been constructed for this purpose.

Sabkhas.

In high intertidal or supratidal flats, where there is periodic flooding by seawater and subsequent evaporation, gypsum and anhydrite are deposited subaerially by capillary brines. For example, along Baja California, the Trucial Coast of the Arabian (Persian) Gulf, the Gulfs of Suez and Eilat, and parts of the Mediterranean coast. Of most renown is the Abu Dhabi sabkha on the Arabian Gulf, which covers an area of some 100 sq km with deposits of carbonate, gypsum and anhydrite, and a halite crust in the central region, to total about 1m in depth (Kendall and Warren, 1988). Sabkhas can also exist inland, around salt lakes.

'Cyclic Salt'.

A significant, but perhaps not obvious source of salt deposition, is 'cyclic salt'. Salt from sea spray and surface deposition can be carried inland for hundreds of miles by strong winds. In 1839, storms blew salt from the Irish Sea as far as Yorkshire, in quantities of tons per acre (Hatch *et al.*, 1971). In an analysis of an aeolian dustfall, thought to be of Saharan origin, which occurred over eastern Britain in 1984, cubes of halite were occasionally found (Wheeler, 1985). The vast quantities of salt deposited on the Desert of Rajputana in north-west India have been explained (Holland, 1911) by strong south-westerly winds carrying, in the hot dry season (April to June), salt dust from the Rann of Kutch, an extensive marginal marine sabkha system on the south-east of the Indus River delta, on the coast of the Arabian Sea. It was estimated that at least 200,000 tons of salt were carried to the desert in this way each hot dry season. During the ensuing monsoon season the salt dust is carried in solution to local hollows, which become temporary lakes. The largest of these is Sambhar Salt Lake which, at the end of the rainy season, was estimated to cover 70-80 square miles, with only about four feet in depth at its centre (Holland, 1911). These lakes eventually shrink as the brine is concentrated by evaporation in the cold dry season, depositing the salt. The annual cycle continues, allowing the salt to accumulate. This aeolian transport of salt may partially explain the

Triassic salt deposits in north-west England and may be significant in other desert regions today (Holland, 1911).

Continental Salt Lakes.

Salt lakes exist today in restricted or closed drainage basins, in arid or semi-arid regions. Examples are Great Salt Lake of Utah (see Section 3.1.3), Lake Magadi in Kenya, and the various salt lakes of California. The solute composition of the water largely depends on the salts dissolved in the inflowing rivers and mineral springs as a result of weathering reactions, and therefore the lithology of the land through which they flow, and so usually differs from that of sea-water, although that of Great Salt Lake of Utah is very similar. Despite the individual characteristics of these various lakes and their waters, and the differing proportions of sodium chloride which they contain, many of them are depositing this salt, which reaches saturation due to the influence of the other salts present.

1.3.3 Ancient Salt Deposits.

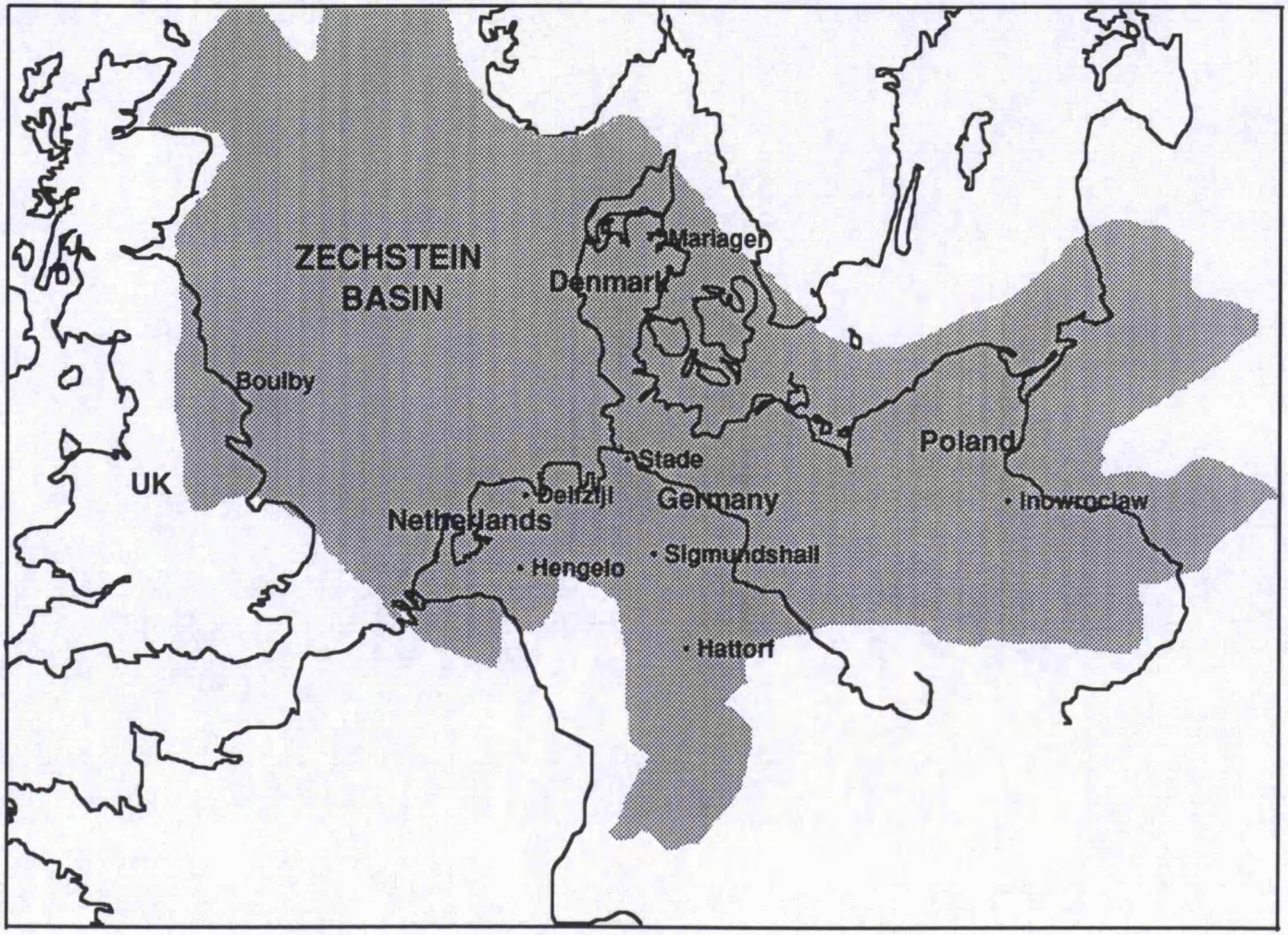
Ancient salt deposits are the results of hundreds of thousands (marine) to millions of years (nonmarine) of deposition. For example, the Messinian (Upper Miocene) evaporites of the Mediterranean took 250,000 to 300,000 years to accumulate to over 2 km thickness, whereas the continental Miocene of the Ebro Basin required 10 to 15 million years to deposit evaporites 0.6 km in thickness (Schreiber, 1988).

Both marine and non-marine ancient salt deposits can be found, though it is not always possible to distinguish between these (Smoot and Lowenstein, 1991). Perhaps the most intensively-studied ancient lacustrine evaporite deposit is the Eocene Green River Formation of Wyoming (and Colorado and Utah), which is exploited predominantly for trona, and of which Lake Magadi in Kenya is the Recent equivalent. Ancient salina and sabkha sediments occur in the Upper Permian Delaware Basin of Texas and the Devonian Elk Point Basin of western Canada. Ancient marine evaporites, deposited in shallow basins, include the Messinian deposits of the Mediterranean and the Permian Zechstein Basin (Figure 1.2).

Not all salt deposits exist as bedded salt. An overburden of sediments can cause the less dense salt bed to deform and parts of it to flow plastically upwards, intruding into the sedimentary layers, to form salt domes (or diapirs). Hundreds of such salt domes exist in the Gulf Coast region of the USA, and often have associated with them deposits of sulphur and petroleum (Martinez, 1991). The North-West German (Zechstein) Basin, is a salt dome basin, in which *halokinesis* began by the Triassic and the typical salt dome features took until Tertiary times to develop (Halbouty, 1967).

Figure 1.2. The Zechstein Basin.

Sketch map of the Zechstein Basin (After Fuchtbauer and Peryt, 1980), indicating sites sampled.



Several of the salt samples used in this study were taken from various sites in the evaporite deposit of the Zechstein Basin (Fuchtbaüer and Peryt, 1980; Harwood and Smith,1986.): from N.E. England, the Netherlands, Denmark, Germany and Poland (Figure 1.2). The Zechstein Sea existed in the later part of the Permian period of geological time, said to be the greatest period of salt deposition in Earth's history, when the climate was an arid one, with prevailing low rainfall, low humidity, and moderate to high temperatures. The Zechstein Sea is thought to have formed when a relative rise in sea level (due to eustatic changes) resulted in the flooding of a chain of inland drainage basins in the north. This inland sea was initially about 250 m deep, with a northern rim that formed a barrier to continued influx or efflux of water. It is assumed that, on the whole, water loss by evaporation exceeded input by rivers and rainfall, allowing deposition of salts. With continued shrinking of the Zechstein basin, sub-basins (*e.g.* the English, Dutch, Werra-Fulda and North-West German Basins) became isolated and subject to local influences, evolving into marginal salinas. Eventually the entire basin became a salt flat, with isolated playas of fluctuating area and depth, to be later overlain by the red beds of the Triassic. The Zechstein sequence is based on five cyclic changes of salinity, due to changes in sea level: the Z1, Z2, Z3, Z4 and a minor, incomplete Z5 Cycle. These five cycles of Zechstein sedimentation resulted in an evaporite basin 1500 m thick.

Several ancient salt deposits, of different geological ages, were sampled for this study and these are described in detail in Section 3.1. Table 1.5 lists ancient evaporite deposits and their geological ages, from around the world, also illustrated in Figure 1.3.

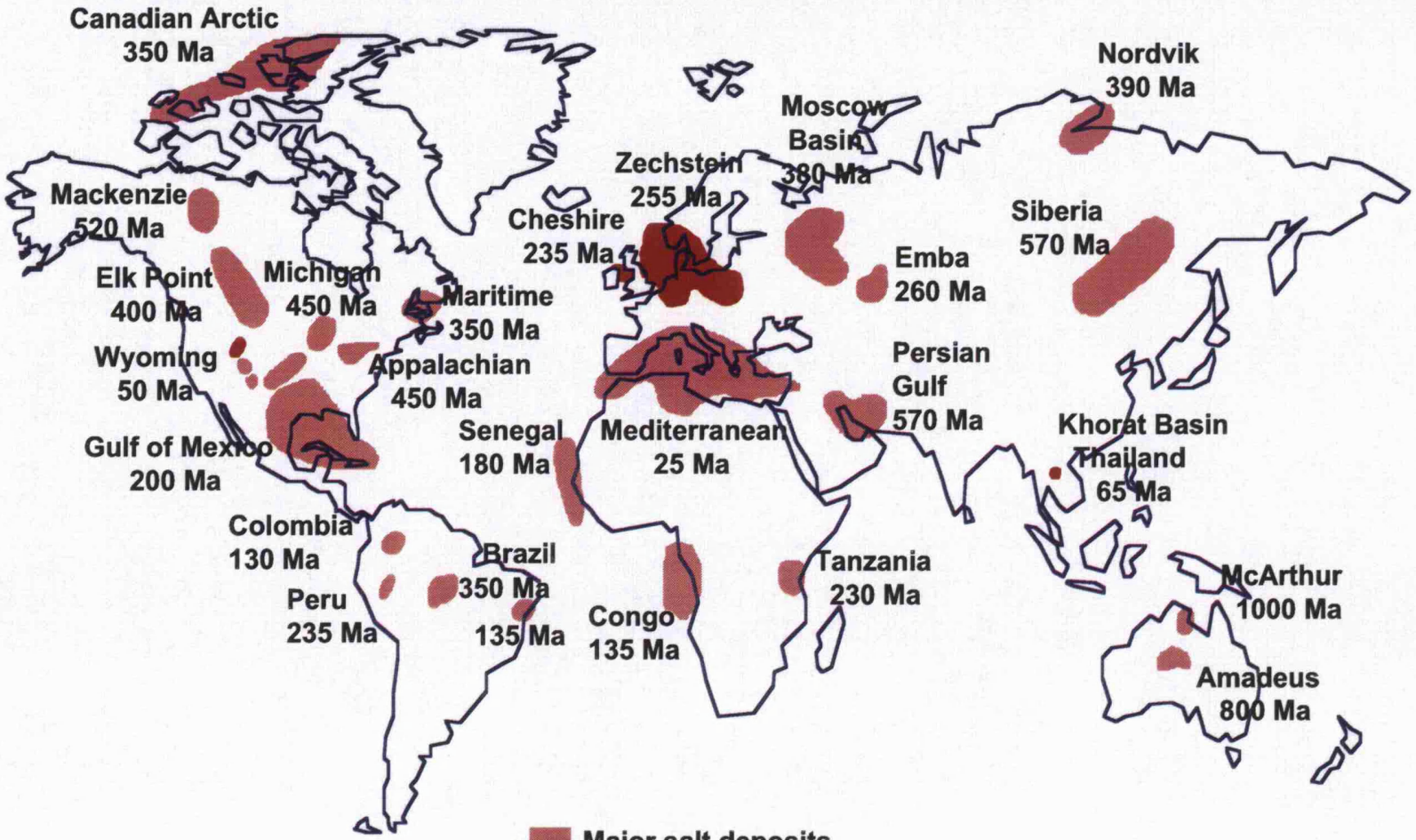
Gases.

Combustible gases are found in micro-inclusions in halite and potash deposits: mostly nitrogen, minor amounts of methane, higher alkanes, H, He, Ne and argon. Generally, halite contains less gas than potash and less hydrocarbons. The gas microinclusions frequently have high internal pressures and if the walls are weakened can be released explosively, a particular hazard in potash mines. Hydrogen is thought to be a metabolic product of fermentative bacteria; a study of the Great Salt Lake showed that with increase in salinities of the sediments there was an increase in H_2 production and a simultaneous decrease in consumption of H_2 by sulphate-reducing and methanogenic bacteria (Hite and Anders, 1991).

Table 1.5. Evaporite Deposits and their Geological Ages (Halbouty, 1967; Lefond, 1969; Cowie and Bassett, 1989). Geological ages given in Ma BP (mega annons, or 10^6 years, Before Present). Deposits studied in this thesis are in **bold**.

ERA	PERIOD / SERIES	EVAPORITE DEPOSIT
CENOZOIC	Holocene	Present-day sabkhas, **salinas, salt lakes**, etc.
	Pleistocene ~2 Ma	California; Nevada; Russia; Mexico; Israel
	Pliocene 5 Ma	Virgin Valley, Nevada; Utah; Italy; Jordan
	Miocene 23 Ma	Mediterranean; Red Sea; Trucial Coast (Arabian Gulf); Dominican Rep.; Romania; **Wieliczka, Poland**
	Oligocene 36 Ma	France; Germany; Spain; Turkey; Iran; Iraq
	Eocene 53 Ma	Green River, **Wyoming**; Potwar, Pakistan
	Palaeocene 65 Ma	None known
MESOZOIC	Cretaceous 135 Ma	Gabon, Congo & Angola; Brazil; Colombia; Russia; Florida; **Khorat Basin, Thailand**; Morocco
	Jurassic 205 Ma	Montana; Bulgaria; Black Sea; Cuba; Chile; Idaho; Gulf Coast
	Triassic 250 Ma	**Cheshire Basin;** Portugal; Spain; N. & S.W. France; Netherlands; Germany; Switzerland; N. Africa; Peru; Persian Gulf
PALAEOZOIC	Permian 290 Ma	**Zechstein Basin, N.W. Europe**; Permian Basin, Texas; Salado Formation, Mexico; Emba, Caspian Sea
	Carboniferous 355 Ma	Canadian Arctic Islands; Paradox Basin. S.W. USA
	Devonian 410 Ma	Elk Point Basin Canada; Mongolia; Moscow Basin
	Silurian 438 Ma	Salina Basin (New York, Ohio, W. Virginia, Pennsylvania, Michigan, Ontario)
	Ordovician 510 Ma	Williston Basin, USA/Canada
	Cambrian 570 Ma	Mackenzie, NW Territories; Siberia; Iran (Persian Gulf); India; Australia
PROTEROZOIC	Pre-Cambrian	Amadeus Basin, C. Australia; McArthur Group, Queensland, Australia; Ontario, Canada; Iran

Figure 1.3 Major Salt Deposits of the World



Organics in Halite and Potash.

The association between evaporites and petroleum deposits is an established one. It can be assumed, from knowledge of modern evaporitic environments, that the organic matter anoxically sealed within halites is algal and/or bacterial, though there have been few studies of this. Halite deposits may also contain spores, pollen and woody plant material. Samples of coalified plants have also been collected from potash and halite, for example in the Cretaceous halite of northeast Thailand (Hite and Anders, 1991).

In the Eocene rocks of the Green River Basin of Wyoming large deposits of trona occur associated with kerogenous marlstones (oil shales) which formed as a result of bacterial and algal burial. The richest rocks in this Formation have TOC values as high as 35%. A sodium carbonate brine, associated with the sequence, contains as much as 10 wt. % organic acids which, when precipitated out, are shown to be very similar in composition to the kerogen in the oil shales. It is suggested that this 'black trona brine of Wyoming' is fossil brine from the original lake (Hite and Anders, 1991).

1.4 THE SURVIVAL, LONGEVITY AND ISOLATION OF ANCIENT MICROORGANISMS.

"... if a spore or some other phase of a microörganism could persist alive for forty years there is no real reason why it might not do so for much longer periods even to millions of years." (Lipman, 1931).

1.4.1 Ancient DNA and Life in Amber.

In recent years there has been a surge of interest by scientists in the extraction and sequencing of autochthonous DNA from ancient preserved biota (see Johnson, 1991), although some claims of success have been met with considerable scepticism (Lindahl, 1993a). Current research seems to be focused on amber as the preservative material. The age of recoverable DNA is being put increasingly further back, with the amplification and sequencing of DNA from an extinct weevil entombed in 120-135 million year old amber (Cano *et al.*, 1993). In *'The Quest for Life in Amber'*, Poinar and Poinar (1994) have, for over two decades, examined many lifeforms preserved in amber, including bacterial cells inside nematodes, as well as various insects, some with intact tissues and cells complete with organelles. In 225 million-year-old amber, the Poinars identified intact bacteria, algae, fungi, and protozoa, in a *'flashphoto of life in a drop of Triassic water, preserved in three-dimensional form'*. In 1984 they succeeded in isolating *Bacillus* sp., presumably from spores, from a stingless bee embedded in 25-40 million-year-old Dominican amber, but were unable to prove this conclusively. Cano continued this particular study, and in 1995 (Cano and Borucki) announced the revival and culture of 25-40 million-year-old bacillus spores from this amber under stringent aseptic and controlled conditions.

With claims that DNA, and organisms, have survived intact for millions of years the possibility of contamination by contemporary organisms has to be considered and controls need to be stringent. In addition there is the problem of chemical decay, particularly of nucleic acids, to be reconciled with longevity. DNA is subject to endogenous hydrolytic depurination and deamination, oxidative damage, and nonenzymatic methylation by *S*-adenosylmethionine, which are remedied by excision-repair processes in metabolically active cells. In fossil tissue, hydrated DNA, deprived of these repair mechanisms, spontaneously degrades to short fragments within several thousand years (reviewed by Lindahl, 1993b). Just as DNA in bacterial spores is protected from damage by special mechanisms (Lindahl, 1993b) certain properties of amber may preserve encapsulated lifeforms. Amber has preservative and antibiotic

properties, which may be attributed to components of the resin: the sugars withdraw moisture and dehydrate the tissues; terpenes and alcohols inhibit microbial activity, and fix and preserve the tissues. The amber would also protect the embedded organisms by excluding atmospheric moisture and oxygen (Poinar and Poinar, 1994).

1.4.2 Bacteria from Ancient Rocks.

Possibly the earliest claim that living microorganisms can be isolated from rocks was made by Lipman in 1928, who followed this with a claim of having isolated bacteria from coal, suggesting that they were survivors since the coal was formed, remaining *in a state of suspended animation* as a *spore or some similarly resistant resting stage* until conditions became favourable (Lipman, 1931). Despite rigorous precautions against contamination, Lipman's claims that viable bacteria could be recovered after millions of years were later refuted. Strong (1956) described contemporaneous 'iron bacteria' within primary crystallized halite in ferrugenous salt samples from Permian salt cores from a deep borehole at Eskdale in N. E. Yorkshire (Boulby area). He observed *colonies of microorganic matter* and *thread bacteria several millimetres in length* and up to 1 μm thick, and associated with *cyst like bodies*. These were absent or fragmentary in salt from disturbed beds. Salt and organic fragments were placed in distilled water for microscopic observation, so any halobacteria present (except *Halococcus* spp.) would not have survived this treatment. No suggestions of viability were made.

Reiser and Tasch (1960) also looked at Permian salt from Kansas and observed diplococcus-like bacteria inside salt crystals. Taking extreme care to prevent contamination, they claimed to have cultured these halotolerant bacteria from a small proportion of the salt samples.

Dombrowski made several reports from the late 1950s, on the isolation of bacteria from salt deposits. Following the isolation of *Pseudomonas halocrenaea* (1960), from thermal saline (Zechstein) carbon dioxide springs at Bad Nauheim and also from salts mined from the Zechstein deposit, Dombrowski made further bacteriological investigations of Zechstein salt (1963). Using elaborate modes of sterilization of equipment and surroundings, he claimed to have isolated bacteria from surface-sterilized pieces of primary, undisturbed Zechstein salt. Microscopic observation of thin sections of salt showed the bacteria to be embedded in the crystalline matrix. Additionally, bacteria were reported to have been isolated from Middle-Devonian, Silurian and Precambrian salt. Dombrowski suggested that a combination of dehydration and salification conserved these bacteria for millions of years.

Schwartz was involved in several investigations in the early 1960s, into the halophilic bacterial inhabitants of various saline environments, and whilst he was able to isolate

halophiles (*e.g. 'Sarcina litoralis'* and *'Halobacterium halobium'*) from salt lakes and saltworks brine (Nehrkorn and Schwartz, 1961), was unable to isolate living bacteria from rock salt samples (Bien and Schwartz, 1965), although variable amounts of diverse forms of conserved dead bacteria were observed in rock salt. Consequently, Bien and Schwartz questioned the results of Tasch and Dombrowski, suggesting that their isolates had been introduced into the rock salt from outside, via fissures. This view was also taken by De Ley *et al.* (1966), who showed, by several molecular, biochemical and physiological tests, that the *Pseudomonas halocrenaea* isolated by Dombrowski was identical to the pathogen *Pseudomonas aeruginosa*, and suggested that the organism may have infiltrated the salt deposit via groundwater or air. This is not unreasonable, as bacteria have been shown to be transported in underground waters through permeable rocks via pores (see Lappin-Scott and Costerton, 1990), but ancient salt deposits are generally impermeable (Sonnenfeld, 1984). Dombrowski's experiments were reproduced with more success by Bibo *et al* (1983), who used stringent precautions against contamination, surface sterilization of the salt samples, and conducted control experiments, to show that viable microorganisms could be isolated from salt samples from primary salt deposits and cultivated.

More recent reports have been made of halobacteria isolated from Triassic and Permian salt deposits (Norton *et al.*, 1993; Stan-Lotter *et al.*, 1993; Denner *et al.*, 1994; McGenity, 1994). In one of these deposits, the presence of well-preserved pollen suggests that they either did not experience or were able to tolerate metamorphic events; this may hold true for embedded microorganisms (Stan-Lotter, 1993). Norton *et al.* (1993) isolated approximately one halobacterium per 500 g of freshly-blasted, surface-sterilized rock salt and suggested that viable halobacteria are found, though with rarity, within such deposits.

The isolation of bacteria from deep subsurface environments has received much attention in recent years, particularly in relation to repositories for petroleum (Bock *et al.*, 1994) or radioactive waste (Pedersen and Ekendahl, 1990). Viable bacteria have been isolated from a rock matrix at 450 m depth where, without nutrient flux, they may have existed in a state of starvation for several million years (Amy *et al.*, 1993). L'Haridon *et al.* (1995) isolated thermophilic Bacteria and hyperthermophilic Archaea from a continental oil reservoir, 1,670 metres below the surface, and suggested that they were autochthonous, and *'they may have been deposited with the original sediment and survived over geological time'*.

1.4.3 Longevity.

Questions need to be asked (and have been) about the feasibility of microorganisms surviving in a viable state for millions of years - *i.e.* what is the physiological basis of their longevity? If they are in a state of suspended animation, with a basal metabolic rate, then spontaneous decomposition of metabolites, in particular proteins and nucleic acids, needs to be considered, and over very long time periods (10^9 years) ionizing radiations, such as that of ^{40}K, present in the environment and inside cells, may be deleterious (Sneath, 1962). Gest and Mandelstam (1987) determined that natural radiation does not have a significant effect on mutation rate and spore longevity, although the time period under investigation was only 16 years. Lotus seeds have survived in a peat bog for over 1,000 years, and *Bacillus* can be cultured from dry soil from the roots of plants stored for over 300 years (Sneath, 1962). There is evidence, apart from the recent work on amber, that bacterial spores can survive for at least 6,000 years in suitable conditions, and estimations based on an exponential rate of death due to radiation damage indicated that a proportion of the spore population could survive for at least 200,000 years (Gest and Mandelstam, 1987) and possibly millions of years (Sneath, 1962).

Kennedy *et al.* (1994) have compiled a database of reports of over 5000 preserved and/or revived microorganisms older than 50 years. These ancient bacteria are often spore-formers, such as *Bacillus* spp. Sources of these ancient microorganisms include 166-year-old Porter beer, mastodon intestines, soils, coal, deep sediment cores, and rock salt. Supportive evidence for a truly ancient origin of such a microorganism includes inaccessibility of the source environment, strict avoidance of contamination during sampling and culture, unique characteristics of the organism, its repeated culture from related samples, and its phylogenetic distinctness.

1.4.4 Strategies for Survival.

Certain bacteria are known to respond to stress, such as starvation, by entering dormancy, in the form of cysts or spores. Dormancy can be defined as *'a reversible state of low metabolic activity, in which cells can persist for extended periods without division* ... until ... *conditions are more favourable'*, when dormant cells may be resuscitated (Kaprelyants *et al.*, 1993). Sporulation is a well-known survival strategy. Bacterial spores can survive exposure to extremes of heat and cold, including the vacuum of space (see Young, 1985). It is believed that some non-sporulating bacteria possess resistant forms which they can adopt in unfavourable conditions (Roszak and Colwell, 1987).

When bacteria are subjected to starvation they may respond by cell division without growth, and consequential miniaturization, to become *ultramicrobacteria* (Torrella and Morita, 1981), which appear to be common in oligotrophic environments such as the deep subsurface (Amy *et al.*, 1993) and aqueous environments (Torrella and Morita, 1981).

This miniaturization is accompanied by several other responses (reduced RNA levels, protein degradation, synthesis of new proteins, *etc.*) which may allow long-term survival in a low nutrient environment (see Morita, 1988; Lappin-Scott and Costerton, 1990). Morita (1988) defines this physiological state, due to insufficient availability of nutrients for growth and reproduction (starvation), as *starvation-survival.*

Another possible mechanism for surviving environmental stress (*e.g.* a drop in temperature) is to enter the '*viable but non-culturable*' (VBNC) state, whereby the bacterial cells are metabolically active, but no longer able to undergo cell division in/on a medium which would normally support growth (Oliver, 1993). VBNC cells are generally reduced in size and rounded. Resuscitation from this state can be a slow process, sometimes taking days.

Many bacteria have the capacity, in varying degrees, to withstand drying (Potts, 1994). One mechanism is the accumulation, by transport or *de novo* synthesis, of osmolytes (see Section 1.2), which allow a balance in water activity between the cell and its external environment. These osmolytes, or compatible solutes, are also thought to stabilize proteins by 'preferential exclusion' (Potts, 1994), which allows a hydration shell around the proteins. Stabilizing salts, such as KCl, however, bind to charged regions on the protein. According to the 'water replacement hypothesis' (see Potts, 1994), during extreme drying (desiccation) polyhydroxyl compounds protect macromolecules by replacing the shell of water; but this is only known for trehalose, in particular, and sucrose. Trehalose also stabilizes the cell membranes during desiccation (Potts, 1994). Potts (1994) speculates on the entry into a 'glassy' state by bacteria subjected to desiccation; in this state cytoplasmic water, and therefore cell reactions, would be immobilized.

Sporulation is not a known feature of halobacteria, and ultramicro- forms and VBNC cells are not known, but other forms of cellular morphogenesis, which may be a response to sub-optimal environmental conditions, are known. The unidentified halophilic archaeon studied by Wais (1985), exhibited cellular morphogenesis, one stage of which was existence within a thallus which, unlike the individual unicells, was resistant to lysis in distilled water. This thallus may well be a structure enabling long-term survival of these halophilic Archaea in an environment (cyanobacterial mats) reported to be subject to extreme variations in salinity. Similarly, Kostrikina *et al.* (1991) described pleomorphic halobacteria (*'Halobacterium distributum' 'Halococcus turkmenicus', Natronobacterium pharaonis* and *Natronococcus occultus*) from saline soils, also subject to fluctuations in salinity. *'H. distributum'* had four morphological forms; in one form cells were packaged within a common capsule and in another, single cells were rounded with a thick, multilayered cell wall, and thought to be cyst-like resting cells (*halocysts*).

One strategy for long-term survival which may be taken by microorganisms in evaporite deposits is incorporation within crystals (see Section 1.3.1). It has been shown experimentally (Tew, 1980) that *Ectothiorhodospira* species can survive for at least 100 days within crystallized mirabilite ($Na_2SO_4.10H_2O$) and Del Moral *et al.* (1987c) showed that moderately halophilic bacteria, in common with other bacteria (see Del Moral *et al.*, 1987c), participate in the precipitation of calcium carbonate crystals in laboratory culture. Bacteria have been observed within fluid inclusions in natural salt crystals from a solar saltern and a soda lake, and halobacteria have been shown to survive within fluid inclusions in laboratory-grown salt crystals for at least six months (Norton and Grant, 1988). Recent microscopic examination of the original halite crystals containing isolate BH1 (Norton and Grant, 1988) showed intact rod-shaped cells within fluid inclusions (*i.e.* after eight years). The cells appeared to be motile, but Brownian motion could not be ruled out. Baas-Becking (1931) described *'crystals of NaCl ... from Searles' Lake, California, which were inside still a dark purple, because of a large Thiospirillum!'*, but the outsides of which were bleached by the light. The bacteria (viable or not) promote crystallization, apparently by acting as nuclei for crystal-formation (Norton and Grant, 1988). Lopez-Cortez *et al.* (1994) found that a greater number of cubic halite crystals, with a greater range of sizes, were formed in the presence of halobacteria, than in their absence, and that halobacteria also caused the formation of dendritic halite crystals, which may be attributable to the *S*-layer protein. The crystals may form a protective shell, enabling bacterial survival.

Few investigations have, apparently, been made of the responses of halobacteria to starvation, except with respect to the stringent response to amino acid starvation. In *Hf. volcanii* and *Hc. morrhuae* stable RNA synthesis is under stringent control, whereas *H. halobium* is partially relaxed and *H. salinarium* is fully relaxed (Cimmino *et al.*,1993).

1.5 PHYLOGENETICS AND EVOLUTION.

1.5.1 What is Phylogenetics?

Phylogenetics is the study of hierarchical relationships between organisms and their evolution through time from common ancestors, inferred by comparing characters (*e.g.* molecular sequence data) between the extant organisms. When represented on a phylogenetic tree the genealogy of, say, two extant species (terminal nodes) can be traced back along their branches to the point (internal node) where they diverged from a common ancestor. The distance from this internal node to the terminus of the branch (*i.e.* the branch length) is therefore a measure of evolutionary time, during which the species accumulated random changes (mutations) in its genes. A tree may be unrooted, and therefore the direction of evolution is not specified, or rooted to indicate the evolutionary pathway from a common ancestor. In the Universal Tree of Life, if all branches are followed back in this way, from node to node, to the root of the tree, one must inevitably contemplate the origins of life and the nature of the first organism, the Universal Ancestor, which may (Woese and Fox, 1977) or may not (Linkkila and Gogarten, 1991) have been a simple progenote.

1.5.2 Early Approaches to Phylogenetic Tree Reconstruction.

Classically, the evolutionary relatedness between organisms has been based on morphological characteristics and the fossil record. Higher plants and animals, in their complexity, comprise a wealth of morphological information, with which it is possible to look at constant characters using the fossil record. Such information can be both subjective and imprecise, however. With microorganisms, such as bacteria, there is little or nothing to distinguish them morphologically and the fossil record is rarely applicable and often inconclusive. But there is a kind of fossil record in the molecular sequences of living organisms. An accurate phylogenetic picture can be drawn with data derived from some universal property of organisms, a feature which has been passed unfailingly down the generations since the first organism replicated - the replication machinery, for example. With the development of molecular techniques evolutionary relationships between organisms from all branches of the Universal Tree of Life could be analysed, and algorithms were devised whereby these relationships could be defined in mathematically precise terms.

The potential of polypeptides in defining phylogenetic relationships was recognized by Zuckerkandl and Pauling (1965a; 1965b) in the 1960s, when they suggested that

molecules carrying genetic information, *i.e.* genes, RNA and most polypeptides (primary, secondary and tertiary semantides, respectively), carry the evolutionary history of the organism. Thus, comparison of a number of homologous polypeptide chains (*e.g.* of haemoglobins), with a consideration of the genetic code (in view of its degeneracy) and of the constraints imposed by the higher order structure of the molecule, would provide information about the evolutionary relationships between organisms. They tentatively proposed that nucleic acid sequences, which contain stretches of non-coding DNA in addition to code for the potential transcription / translation to proteins, would therefore be more informative of evolutionary history.

At the time that Zuckerkandl and Pauling were considering the evolution of the globin chains (1965b), Margoliash and Smith (1965) were similarly looking at cytochrome *c* amino acid sequences from a wide variety of organisms. They observed that the more distantly-related the species, the more dissimilar (number of residues different) the cytochrome *c* sequences, and that the dissimilarity was relatively constant between individual members of two different classes of organisms (*e.g.* between a mammal and a bird). This quantification was seen to allow an evolutionary timescale for the molecule and led to the first phylogenetic tree using a distance method (Fitch and Margoliash, 1967; see below). This period marked the dawn of molecular phylogenetics.

In the early 1970s, before sequencing entire 16S rRNA genes on a large scale became realistic, Woese and colleagues developed 16S rRNA oligonucleotide cataloguing as a means for examining organism interrelationships. By this method, the 16S rRNA for each organism was digested with T_1 ribonuclease, which cleaves at the 3' end of guanine residues, and the resulting fragments separated by 2-D electrophoresis and then sequenced. Oligonucleotides of hexamers and above were catalogued and these data were used to compile association coefficients for pairs of species. The association coefficient was calculated by dividing twice the number of nucleotides in common (in hexamers and above) to both species A and species B, by the total number of nucleotides (in hexamers and above) in A and B; *i.e.*

$$S_{AB} = \frac{2N_{AB}}{(N_A + N_B)}$$

A matrix of S_{AB} values could then be used to construct a tree (Woese & Fox; 1977; Woese, 1981). It was the use of this method which allowed comparisons at a range of taxonomic levels and which elucidated the path to the 'archaebacteria' (see below). However, it required the analysis of full 16S rRNA sequences to resolve branching orders amongst the phyla and their subdivisions (Woese,1987). For example, 16S cataloguing showed there to be one major group of extreme halophiles, equivalent to a

(eu)bacterial phylum; *H. halobium*, *H. volcanii* and *H. morrhuae* clustering together between two groups of methanogens (Fox et al, 1980). 16S rRNA cataloguing did not therefore have the resolving power required for finely-tuned analysis, and with technological advances full 16S rRNA sequences became the choice for phylogenetic tree reconstruction.

1.5.3 The Three Domains.

The current, popular view of a tripartite phylogenetic structure to the tree of life (Eukarya, Bacteria and Archaea) is a development of that first propounded by Woese and Fox in 1977. They coined the name 'archaebacteria' for the third primary kingdom, which was thought initially to be represented solely by the methanogens. By 1978 (Woese *et al.*) the extreme halophiles and certain thermoacidophiles were also recognized as members of the 'archaebacteria', a therefore phenotypically diverse group, though with particular properties in common. However, although this tripartite phylogeny based on 16S rRNA sequence analysis was strongly supported by certain distinct phenotypical characters (cell wall chemistry, lipid linkage, characteristic tRNA and rRNA base modifications, etc.) common to each group, the dogmatic acceptance that life was fundamentally dichotomous, being either prokaryotic or eukaryotic in form (first proposed by E. Chatton in 1937), persisted. In 1990 Woese *et al.*, arguing against the phylogenetic validity of the negative usage of *Prokaryotae* as a taxonomic rank, formally proposed that Eukarya, Bacteria and Archaea should each head a domain, as the highest taxonomic unit, above the level of kingdom, to comprise the three domains of life. Organisms could then be classified in a more natural way and assigned a taxonomic rank to reflect their evolutionary history.

1.5.4 The Archaea.

The question of where and how the Archaea branch on the phylogenetic tree is the cause of some contention, but evidence at present appears to be in favour of a divergence from the eukaryotic line of descent, rather than a simultaneous divergence of the three lineages from the 'progenote', over 3 billion years ago, as originally proposed by Fox *et al.* (1980). This was recognized as early as 1979, by Hori and Osawa, who placed *Halobacterium cutirubrum* at a point early on the eukaryotic line of descent, in a distance matrix tree derived from nucleotide substitution rates (*Knuc*) for 54 5S rRNA sequences. This placement was recognized as being in support of (then) recent work by Matheson *et. al.* (1978) who showed that the 'A' protein (in the 70S ribosome) of *H. cutirubrum* had properties in common with the eukaryotic 'A' protein and that the 5S rRNA secondary structure "*is in a unique class, showing properties somewhat closer to those of eukaryotic 5S RNA*". More recent studies on the organization, structures and sequences of archaeal

r-proteins support the monophyly of the Archaea and a common line of descent with the Eukarya (Matheson, 1992).

Earlier work by Iwabe *et al.* (1989), with the protein sequences of the duplicate genes of ATPase (α and β) and elongation factors Tu and G, showed Archaea to be phylogenetically more closely related to Eukarya than to Bacteria. The recent discovery of a transcription factor (TATA-Binding Protein) in an archaean organism (*Pyrococcus woesei*) that is homologous to the TBP in eukaryotes (Rowlands *et al.*, 1994), and which is apparently absent in Bacteria, also supports the divergence of the Archaea from a common lineage with the eukaryotes in the phylogenetic tree.
Analysis of RNA polymerases (Iwabe *et al.*, 1991) has given variable results according to the subunits compared. Whereas a comparison of the largest subunits supports a likely monophyletic origin for the Archaea, when the second-largest subunits were compared this was not confirmed. However, when both data sets were combined the monophyletic origin of the Archaea was again favoured, supported by more recent RNA polymerase sequence analysis (Klenk and Zillig,1994).

This monophyletic origin of the Archaea, where they cluster as one unit, along a single branch which diverged from the Eukarya line, is not universally agreed. It has been alternately proposed that separate branchings have given rise to different lines of 'archaeobacterial' evolution (the polyphyletic hypothesis, *e.g.* the eocyte tree of Lake (1988); see below). The latter precludes the three domain theory. This monophyly / polyphyly controversy is the subject of continual dispute.

Opposing the monophyly hypothesis is Lake, who argues for a polyphyletic branching of the Archaea. In Lake's tree, there are two phylogenetic superkingdoms: the 'karyotes', which groups the 'eocytes' (extreme thermophiles) and eukaryotes, and the 'parkaryotes', which groups the eubacteria, halobacteria and methanogens (with *Thermoplasma* and *Archaeoglobus*); (1988).
Phylogenetic analysis of glutamate dehydrogenase sequences has provided a different interpretation of the universal tree of life, not only supporting the non-monophyly of the Archaea, but placing the root of the tree outside the Bacterial branch - either in the archaebacterial kingdom or in the eukaryotic branch according to the treeing method used (Benachenhou-Lahfa *et al.*, 1993). This is based on there being two families of paralogous *gdh* genes, distributed universally (Bacteria and Eukarya in Family I; representatives of all three domains in Family II), presumably resulting from their duplication and coexistence in the universal ancestor prior to the divergence of the three domains. However, inconsistencies, such as the eukaryotes *Giardia lamblia* and *Chlorella sorokiniana* clustering with Bacteria (Family I), a *Clostridium sp.* in each *gdh* Family (a

more recent divergence of the *gdh* genes here?), the placement of the bovine gene with the chicken and not other mammals, and the changing positions of *Halobacterium salinarium* and the root according to the treeing method, gives this little credence.

In accordance with the monophyly of the Archaea (the viewpoint taken in this study), the archaeal domain (see review by Kandler,1993) is divided into two kingdoms: the *Euryarchaeota*, which contains the methanogens, extreme halophiles, the *Thermococcus - Pyrococcus* group and *Thermoplasma*, and the *Crenarchaeota*, which comprises the thermoacidophiles (Woese *et al.*, 1990). The phylogenetic distance of the methanogen and halophile lineages from *Thermococcus* suggests they underwent a rapid burst of evolution (macroevolution), presumably in response to environmental changes. A further saltatory event led to the *Methanomicrobiales*, some of which developed a halophilic phenotype (Woese, 1987). The extreme halophile lineage, which seems to cluster closest to the *Methanomicrobiales* (Yang *et al.*, 1985; Woese,1987; Rouvière *et al.*, 1992), is thought to have arisen from a branch of the methanogens at about the time aerobicity appeared in the Bacteria, about 1.5 billion years ago (Fox et al 1980; Woese,1987).

1.5.5 The Use of the 16S rRNA Gene in Phylogenetic Analysis.

16S rRNA has become the molecule of choice for phylogenetic analysis, fulfilling 'Woese's Postulates' (Woese & Fox, 1977; Fox et al,1980; Woese, 1987) in that:

i) it is readily isolated;
ii) it is large enough to provide adequate and statistically significant information;
iii) it occurs in all organisms;
iv) it exhibits constancy of function over great evolutionary distances, for direct comparison;
v) its sequence changes randomly and different positions change at different rates, allowing comparisons between a great range of phylogenetic relationships, from domain (highly conserved regions) to species (greatly variable regions) level;
vi) it consists of a large number of independent domains such that nonrandom changes in one region do not affect the rest of the molecule - so the clock runs smoothly.
vii) it has a large and expanding database.

1.5.5.1 Sequence Alignment.

The first and most crucial step in phylogenetic analysis is the correct alignment of the homologous sequences so that homologous 'character states' (nucleotides or amino acids) for a given position are arranged in a columnar fashion, for each position in the sequence

('positional homology'; Swofford and Olsen, 1990). Errors in alignment will produce 'wrong' trees. This is not a problem in conserved regions of the molecule, but difficulties can arise in highly variable regions, especially if, in addition to substitutions, there are insertions or deletions. It may then be necessary to introduce a limited number of gaps to facilitate alignment. To maximize matching, mismatches and gaps should be kept to a minimum and in computer-based alignment algorithms this is achieved by an *a priori* assignment of costs. However, optimal matching may produce an erroneous alignment, masking true homology between sequences (Mindell,1991). Computer-derived alignments are therefore often adjusted manually, with phylogenetic relationships borne in mind (the most closely related are aligned first). A consideration of the secondary structure (base pairings) of the molecule (see below), can also aid in improving an alignment satisfactorally obtained by computer. (See Appendix 1, for sequence alignments in this thesis.)

1.5.5.2 Phylogenetic Inference and Methods for Tree Reconstruction.
There are many methods for inferring trees from phylogenetic data, comprehensively reviewed recently by Swofford and Olsen (1990). Only the main methods will be outlined here.

There are two main categories of methods for inferring phylogeny from aligned sequences:

(1) ***Distance Methods*** - pairwise comparisons of aligned sequences give similarity (or dissimilarity) values, which can then be converted to phylogenetic distances (*e.g.* **Jukes-Cantor** and **Kimura** methods). The 'distance matrix' produced is used (*e.g.* in the **Fitch and Margoliash** method) to infer a phylogenetic tree.

(2) ***Discrete Character Methods*** (a 'character' is a position in a sequence) - where each sequence is given a value (*e.g.* **maximum-likelihood** and **parsimony** methods). An advantage of this method is that different kinds of data can be amalgamated into the same analysis (*e.g.* Iwabe, *et al.*, 1991).

*1. **Distance Methods*** (reviewed by Olsen *et al.*, 1986; Swofford and Olsen, 1990).
Dissimilarity values are calculated from pairs of sequences in the alignment, with no consideration of individual characters, and these values are converted to evolutionary distances in a matrix. The most popular models for this are:

(i) the **Jukes-Cantor** one-parameter model (1969), shown below, which assumes independent change at all sites with equal probability, that a base change is independent of its identity and all types of substitution are equally likely;
(ii) the **Kimura** two-parameter model (1980) which assumes that all four nucleotides are equally frequent, but rates for transitions and transversions are independent of each other and can be treated separately.
The generalized two-parameter model is based on (i) and (ii), and allows individual nucleotides to have different frequencies, though substitution rates must be balanced. These models therefore predict the evolutionary change, as nucleotide substitutions, of DNA sequences in pairwise comparisons (Nei, 1991).

As only homologous positions are considered in a pairwise analysis, if sequences vary in length the ends should be truncated accordingly. In addition, treatment of gaps in the analysis needs to be considered carefully. Gap positions can be ignored (weighted 0), effectively increasing similarity of already similar sequences, treated as substitutions (weighted 1) or weighted between 0 and 1 (*e.g.* 0.5, as a compromise). Regions with large numbers of alignment gaps should be discounted as positional homology here would be uncertain. Other positions in the sequence can also be weighted according to phylogenetic usefulness. For example, bases in helices of the 16S rRNA secondary structure can be given a lower weighting (see below). Ambiguities are best counted as completely ambiguous (N), rather than as a purine (R) / pyrimidine (Y), *etc.* (Swofford and Olsen, 1990.)

Phylogenetic distances are calculated as follows (Swofford and Olsen, 1990):

To calculate similarity from an aligned pair of nucleotide sequences:

$$\text{(i)}\quad S = M / L$$
$$\text{(ii)}\quad L = M + U + w_G G$$

S = similarity
M = number of positions with identical bases
L = sequence length
U = number of positions with non-identical bases
w_G = weight given to gaps

To calculate evolutionary distance:

(i) $D = 1 - S$

(ii) $d = -b \ln (1 - D / b)$

D = dissimilarity

S = similarity

d = estimated evolutionary pairwise distance

b = variable, according to the model; *e.g. if all four bases are considered equally likely:*

$b = 3 / 4$ (Jukes and Cantor, 1969)

This distance matrix data is then used to infer a phylogenetic tree by one of a number of methods, such as the Fitch-Margoliash method or Neighbour-Joining. As the original sequence data is not used directly, these methods suffer from a loss of information, particularly as the number of taxa increases (Penny *et al.*, 1992).

The best-fit tree (usually unrooted) is one which minimizes the differences between observed distances and expected distances. It is generally assumed that the phylogenetic distance between two species is represented by the sum of the branch lengths separating them on the tree. One disadvantage of this method is that multiple changes at one site are obscured and usually underestimated, so that branch lengths will tend to be too short (*e.g.* when using the Fitch-Margoliash method with the Jukes-Cantor model); but this can be corrected for (Felsenstein, 1988). This is particularly a problem with distant species, where observed distances underestimate actual distances. The assumption of a molecular clock is optional with these methods. Two examples of methods which produce phylogenetic trees from distance matrix data are:

a) **Least-squares** - *e.g.* the **Fitch - Margoliash** (weighted least squares) method, which was the first distance matrix method for estimating branch lengths for phylogenetic tree inference (Fitch and Margoliash, 1967), producing an additive tree (therefore unrooted). Each pairwise distance is assumed to be independent, which is not the case with related molecules, so trees are less well resolved (Swofford and Olsen, 1990).

b) **Neighbour-Joining** - (Saitou and Nei, 1987).

Neighbour-joining uses the principle of minimum evolution, as intrinsic to the algorithm, to produce a unique final tree. In this method, the sum of branch lengths is computed, as a distance matrix, for all pairs of taxa (terminal nodes). These distances are then corrected to take the average clock rate of each taxon into account. This method does not, therefore, assume that all lineages have diverged equally and allows for unequal rates.

Starting with a star-like tree, this modified distance matrix is used to infer a final tree: the pair of taxa with the smallest sum of branch lengths (neighbours) are joined and their

common ancestral node becomes the new terminal node to pair with its closest neighbour, and so on; pairs of closest neighbours are successively joined in this way, minimizing the sums of branch lengths at each stage, until a final unrooted tree, with both topology and branch length estimates, is produced. Saitou and Nei (1987) claim that computer simulations show this method to be efficient in obtaining the correct tree topology.

2. Discrete Character Methods.

a) **Maximum parsimony** (reviewed in Felsenstein, 1982 and 1983; Stewart, 1993).

The most parsimonious tree is one with a topology that requires the minimum number of changes for the observed characters to have evolved from ancestral species. Maximum-parsimony uses sequence alignments (character data) directly, rather than distances, and considers change at each individual site. It is assumed that characters evolve independently. Characters can be weighted according to probabilities of rate of change, with less weight being given where the probability is higher (*e.g.* for transitions; see below), enhancing the power of parsimony (Hillis *et al.*, 1994).

Compositional bias, *e.g.* G + C variation, can be removed by using transversion distance or transversion parsimony methods (Woese *et al.*, 1991) which ignore transitions (*i.e.* sites are R or Y), but have the disadvantage of reduced information content.

Parsimony would give the correct tree if parallelisms and reversals did not occur in evolution. Also, if the rates of evolution vary between branches in the tree, the results may be misleading (Sidow and Wilson, 1991). Rapidly-evolving lineages,for example, tend to branch too deeply with this method. Assuming the taxa are monophyletic, an outgroup sequence can be included to root the tree.

b) **Maximum-likelihood** (Felsenstein, 1981).

Sequence data is used directly with this method and input order can affect the result. The model does not take deletions and insertions into account.

This is a statistical inference approach, where the maximum-likelihood tree is the one which maximizes the probability of evolving the observed data (constant) given the tree (variable) and a specified model (a constant; b in the evolutionary distance equation, above). The model may itself may be partially derived from the sequence data and the Jukes-Cantor and Kimura models are the most popular.

Assumptions made by the current (DNAML version 3.5c) maximum-likelihood model (similar to those described in the 1981 paper) are:

i) all sites in the sequence are included;

ii) each site in the sequence evolves independently;

iii) each site undergoes substitution at an expected rate; different rates of evolution at different sites are allowed for in the present model;

iv) two types of events leading to substitutions are recognized:

a) a Y (pyrimidine) is replaced by a Y, or a R (purine) by a R, leading to no change or a transition, and

b) a base is replaced by any other base, at random, from a pool of bases, leading to no change, a transition or a transversion.

The ratios of A and G in the purine pool and of T and C in the pyrimidine pool are the same as those in the entire pool. Different rates of transitions and transversions are allowed for.

v) after speciation two lineages evolve independently.

Maximum-likelihood is a more powerful (*i.e.* the program settles on a single tree, or 'converges', with short sequences) method than parsimony, and these both are more powerful than distance methods, with large data sets, because more sequence information is used (Penny *et al.*, 1992). However, Penny *et al.* estimate that six taxa are an approximate maximum to calculate an exact (considering all trees and finding the best for the data, *i.e.* the 'global optimum') tree.

A significant disadvantage of the maximum-likelihood method is that it is computationally-intensive, therefore slow and expensive to run.

Nei (1991) has reviewed the relative efficiencies of different phylogenetic inference methods in obtaining the correct tree, given the tree models for up to 8 species, under various conditions, such as sequence length, use of d (total number of nucleotide substitutions) or p (number of nucleotide differences per site) for distance methods, expected number of nucleotide substitutions per site (U), a constant or varied substitution rate, and whether or not the tree is rooted. For example, in a comparison of Maximum Likelihood with Maximum Parsimony, Fitch and Margoliash, Minimum Evolution and Neighbour-Joining, where the rate of evolution is taken to be constant, the probability of obtaining the correct tree was generally greatest for the ME and NJ methods, ML was usually better than MP, but FM was consistently the poorest method. The optimal conditions in this analysis were the ME or NJ method, with 600 nucleotides (compared with 300), $U = 0.5$ and using p distances. When the substitution rate varied, the ML method was the most superior in obtaining the correct tree, closely followed by ME and NJ; d distances were always better than p distances, most markedly so for more distant sequences, and the FM method usually had the lowest probabilities. Considering only the main methods of tree reconstruction discussed earlier, Nei appears to advocate NJ (with d) > ML > MP > FM. The main message given by these examples is the larger the

number of nucleotides used, the greater is the probability of obtaining the correct tree topology.

The abundance of available methods for phylogenetic tree inference presents a problem of choice. It is difficult to determine, from published papers, which treeing method, or model of evolutionary change, has the most advantages and least disadvantages for DNA sequence analysis. All methods involve information loss in some way (Penny *et al.*, 1992). Authors appear to be biased in favour of a particular method, and perhaps obviously so if it is a method they were involved in developing. It is important to be aware of the assumptions made (about sequence evolution) by the method being used, and therefore of its limitations, when interpreting tree topologies.

1.5.5.3 Reliability of Tree Topology.

The ideal phylogenetic tree for a given (large) set of species is one which shows the same tree topology (i) with various phylogenetic tree reconstruction methods and (ii) when different marker molecules are used. However, a look at the multitude of phylogenetic trees published in the last 15-20 years reveals that there is considerable disagreement on tree topology according to the data set and methods used for its analysis.

Klenk and Zillig (1994) addressed these problems using both amino acid sequences and nucleic acid sequences (nucleotides of second codon positions) of RNA polymerase B (large subunit) as the molecular marker, and three phylogenetic methods (distance matrix, maximum parsimony and maximum likelihood) to investigate the optimal topology of the archaeal tree. They constructed a decision matrix combining the results of these methods for ten possible branching topologies, including 'optimal' topologies, 'worst' topologies and topologies which are optimal when other marker molecules are used. On this basis they were able to determine an optimal tree topology, which was further confirmed by other methods which were unsuitable for inclusion in the decision matrix, and by resampling methods (bootstrapping and jackknifing) to determine the statistical confidence of the best tree.

Certain anomalies in tree topology may be explained by horizontal gene transfer, which places species within the 'wrong' group. This explanation is supported if different treeing methods produce the same 'wrong' tree and if trees using a different molecule, but with the same organisms, produce the conventional topology. For example, a phylogenetic tree showing the relationships of glutamine synthetases I and II has separated two archaebacteria, with the extreme thermophile *S. solfataricus* clustering with the Gram-negative bacteria and the methanogen *M. voltae* clustering with the Gram-positive bacteria, suggesting horizontal transfers of the *GS* gene could have taken place (Smith *et al.*, 1992). The displacement of the two *Clostridium* species in the *gdh* tree

(Benachenhou-Lahfa *et al.*, 1993; see above) could possibly be explained by lateral gene transfer.

There can also be problems in tree interpretation with a rapidly (tachytelically) evolving lineage, e.g. the *Planctomycetaceae* - 16S rRNA cataloguing shows very low S_{AB} values and so a great phylogenetic distance from other Bacteria, perhaps suggesting particularly deep (early) branching in the Bacterial line of descent; but full 16S rRNA sequencing analysis does not support this. Variations in dendrogram topology according to the planctomycete sequence(s) used and the large phylogenetic distance between these and other phyla suggested that the planctomycetes are, in fact, a rapidly-evolving group (Woese,1987; Liesack *et al.*, 1992). Rapidly-evolving lineages can also give problems with full 16S rRNA sequence comparisons if different treeing methods are used. *Sulfolobus solfataricus*, for example, branches more deeply into the tree than the other extreme thermophiles if the distance matrix method and outgroup sequences are used (Woese,1987). Variation in evolution rates ('unequal rate effects'), not only between organisms in a tree, but between different molecular markers and also of positions within a molecule, can change tree topology, producing artifactual trees. Lake (1991) recommends his evolutionary parsimony algorithm as one which is less sensitive to unequal rate effects.

When DNA sequences used in phylogenetic tree construction differ widely in nucleotide composition branching order artifacts can arise. A high G+C content in a sequence, in particular, can give that species an artificially deep branching in the tree. This problem can be avoided, if the purine (and therefore pyrimidine) content of the sequences is relatively constant, by using transversion distances and transversion parsimony analyses (Woese *et al.*, 1991). Variability in sequence is most marked in the first and third nucleotide positions of a codon and so a tree derived from the second codon positions only can circumvent this risk (Klenk and Zillig, 1994). Alternatively, the amino acid sequences can be used (Iwabe *et al.*, 1991; Benachenhou-Lahfa *et al.*,1993; Klenk and Zillig, 1994).

Most methods for inferring phylogenetic trees admit data to the algorithm in the input order (PHYLIP; Felsenstein). A bias may therefore be introduced in the branching order of the tree, and can be removed by randomizing the input order.

Trees can be rooted by including a closely-related outgroup which is known to have diverged earlier from the ingroup being analysed. If no outgroup is used, the root can be placed at the mid-point of the longest lineage between two taxa. However, the latter assumes that the rate of evolution is relatively equal along all lineages of the tree.

It must be remembered that a phylogenetic tree inferred with gene sequence data is a gene tree, though often such trees are, by extrapolation, considered as species trees. If other genes of the same species produce congruent trees, then one can feel confident that the tree topology is a good one. A gene producing a different topology, however, may have a different evolutionary history, such as different substitution rates or gene transfer events.

If the topology is supported by other treeing methods and if it remains consistent with additional data, the tree is probably the true one.
Felsenstein (1983; 1988) advises the use of simulation tests to check phylogenetic methods. A method can be tested for robustness by using the 'best' tree from a real data set to simulate new data sets, with a time scale and rates that approximate to the original conditions, inferring phylogenies from each of these, then using the variability amongst these trees to assess the reliability of the 'best' tree. Simulations can also be used to compare the ability of different methods to estimate a true tree from a data set (Saitou and Nei, 1987; Nei, 1991; Hillis *et al.*, 1994).
Penny *et al.* (1992) discuss five criteria that should be met by methods for phylogenetic tree construction. These are: efficient (fast); powerful (converges to a single tree with short sequences); consistent (converges to the correct tree); robust (consistent when there are deviations from the model) and falsifiable (the data must be able to reject the model, as not all models would be suitable for a particular set of data, *e.g.* if it was not generated by a tree-like process).

1.5.5.4 Resampling (Testing) the Data.

The reliability of the tree topology can be tested by repeatedly resampling the data, using methods such as **bootstrapping** and **jackknifing**. These methods assume that data points (*e.g.* sites in a sequence) evolve independently and are identically distributed.
The usual application of jackknifing involves dropping data points one at a time, each time reestimating a branch length, to test its reliability.
With bootstrapping, the better method when the data set is large, data points are randomly removed and/or replaced within the original data set, whilst keeping nucleotides associated with the same species, until a new data set of the same size as the original is obtained. In this way, data points may remain, be duplicated, or removed, in each bootstrapped data set. By repeatedly resampling in this way, to produce, say, 100 bootstrapped trees, a consensus tree can be drawn with the % confidence intervals for the internal nodes. These methods, and their limitations, are reviewed in Felsenstein (1988).

1.5.5.5 Signature Analysis.

A correct alignment of 16S (or 23S) rRNA sequences can be used to determine sequence signatures. Sequence signatures are bases or short oligonucleotides, located at specific positions in the sequence, which are constant for all members within the defined group, but different for any organisms outside that group. They can be used to assign species to a group at any taxonomic level, differentiating between the three domains, between kingdoms, orders, or be so finely-tuned to be species-specific. (Woese, 1987).
Winker and Woese (1991) have determined 16S rRNA sequence signatures to differentiate the three domains and between the *Euryarchaeota* and *Crenarchaeota*

kingdoms of the *Archaea*. Their signatures were of two types: (i) homologous positions; which define a set of positions which are constant in each group and where each position has a base composition characteristic of that group, and (ii) nonhomologous structures; where positions in the sequence possess features (individual bases, base-pairs or a series of bases, absent or present) which are reflected as secondary structure conformational idiosyncrasies characteristic for a given group (see below).

Rouvière *et al* (1992) used this procedure to distinguish the *Methanomicrobiales* from the other two main groups of methanogens (and other Archaea) and, within this order, the 'methanosarcina' and 'methanogenium' sub-groups could be similarly differentiated.

Unidentified sequences, for example from a non-culturable environmental population of bacteria, can be assigned to a group on the basis of signature sequences. Signature oligonucleotides can be used to design group-specific rRNA probes.

1.5.5.6 16S rRNA Secondary Structure.

Alignments of 16S rRNA sequences from diverse sources make it apparent that the primary structure consists of both constant and variable regions and varies between species. 16S rRNA secondary structure, on the other hand, has a universal constancy of shape that reflects its functional importance. However certain features of the secondary structure, in particular those shown by a comparative analysis of the individual helical regions, can be used to distinguish between Bacteria, Archaea and Eukarya (Woese *et al.*, 1983).

Whilst the core of this secondary structure is universally conserved, the remainder is variable on a universal level, but with a degree of constancy that shows intra- or inter-group relationships (kingdom/ phylum/ *etc.*). For example, the region between 179 and 220 (*E.coli* numbering) differentiates the Eukarya (three helices) from the chloroplasts, Bacteria, and Archaea (two helices). The Archaeal structures can be distinguished from the others and, for finer tuning, the 198-220 helix is much shorter (two to four base pairs) in the halophiles and methanogen studied than in *Sulfolobus* (nine to ten base pairs). (Gutell *et al.*, 1985; Winker and Woese, 1991). Gutell *et al.* (1985) have provided a comprehensive comparative analysis of 16S rRNA secondary structures and their phylogenetic value.

The frequency and pattern of 16S rRNA sequence variation correlates strongly with the secondary structure of the molecule. *i.e.* base pairs at a particular position on a particular helix may change frequently, whilst others on the same helix change exceedingly rarely (Woese,1987). The unpaired regions of the molecule (loops) tend to be more highly conserved; 59% of the universally conserved nucleotides occur in such single-stranded regions, which contain only 39% (in *E. coli*) of the total number of bases (Gutell *et al.*, 1985).

Phylogenetic treeing methods which assume that all bases are evolutionarily independent (*e.g.* Maximum Likelihood) do not therefore take into account the fact that bases in helices are not completely independent, that a change in one of these bases would require a compensatory change in its pair, to maintain secondary structure. This problem can be alleviated by giving bases in helices a lower weighting in the analysis. Wheeler and Honeycutt (1988), in their analysis of 5S rRNA sequences with three groups of organisms, observed that differences in base-paired positions (helices) exceeded those expected by chance, whereas differences in unpaired (loop) regions were very rare. They explained this by positive selection pressure on base changes that restore complementarity and so maintain secondary structure. They suggested that phylogenies based on the unpaired regions produced more accurate trees, but if the paired regions were used the phylogenies were misleading. Based on these findings, Wheeler and Honeycutt suggested "*nucleotide positions that pair should be downweighted, perhaps by one-half, or even excluded entirely*".

Wheeler and Honeycutt's conclusions are contradicted by Smith (1989) who analysed the 18S rRNA sequences of echinoderms and compared these results with established phylogenies based on morphology, to establish to what evolutionary distance phylogenies based on 18S RNA sequences could be considered reliable. Smith concluded that molecular phylogenies were less reliable at greater evolutionary distances (150-250 Ma). He found that phylogenies based on nucleotides in paired regions provided the most reliable results for divergences up to 200 Ma and usually found the expected tree as one of the most parsimonious solutions.

Dixon and Hillis (1993), in an analysis of the 28S rRNA genes of vertebrates, also contradicted Wheeler and Honeycutt, finding that phylogenies based on stem regions supported conventional trees, but those of the loop regions generally did not, although the best results were obtained when both data types were combined. Based on observed substitutions in stem and loop regions, they concluded that bases in helices are not entirely inter-dependent, despite secondary structure constraints, and proposed a weighting between 0.5 (complete dependence of a base on its partner so that all base changes are compensatory) and 1.0 (complete independence of base changes), of $\approx$0.8.

It seems reasonable to suggest that weighting, in phylogenetic analysis, should be considered separately according to the molecular data used and the nature of the secondary structure of the molecule, as properties of, say, the smaller 5S (5.8S) rRNA molecules will not be the same as those of 16S (18S) or 23S (28S) rRNA molecules. The substitution rate in 5S rRNA is twice as fast as that in 16S rRNA (Ochman and Wilson, 1987).

A database of 16S rRNA (and -like) secondary structures, representative of the three Domains, is available from R.Gutell (1994) and can be accessed via the WWW and the *Ribosomal Database Project* (Larsen *et al.*, 1993).

1.5.6 23S rRNA Sequence Analysis.

The large subunit (23S) rRNA sequence is beginning to receive attention as a molecular tool for phylogenetic analysis (Ree *et al.*,1993; Ludwig and Schleifer, 1994). 23S rRNAs are twice the length (*ca.* 3000 nucleotides) of 16S rRNAs (*ca.* 1500 nucleotides), with twice the number of secondary structure elements (*ca.* 100; *ca.* 50 in 16S rRNA; Ludwig and Schleifer, 1994) and so as historical documents are phylogenetically much more informative. Comparison of 16S and 23S trees shows that they generally correlate well (Martínez-Murcia *et al.*, 1993; Ludwig and Schleifer, 1994). Martínez-Murcia *et al.* describe the two evolutionary chronometers as 'synchronous'. Lodwick *et al.* (1994) have shown that halobacterial phylogeny based on 23S rRNA is in good agreement with that based on 16S rRNA.

1.5.7 Calibrating the Molecular Clock.

The concept of a *'molecular evolutionary clock'*, first put forward by Zuckerkandl and Pauling (Zuckerkandl and Pauling, 1965), is based on the idea that divergence of homologous molecules, such as haemoglobins from different species, can be related to evolutionary time. In addition to protein sequencing, early attempts to calibrate the molecular clock involved DNA:DNA hybridization, immunological responses and protein electrophoresis (discussed in Thorpe, 1982; with references.).

Sequence changes do not occur at regular time intervals, but the probability of a substitution is predictable and approximately constant, and so the molecular clock is of stochastic regularity. Problems arise in estimating evolution rate from greatly divergent molecules (greater than 10% differences), as the probability of multiple substitutions at one site increases, so that observed differences will be less than actual differences (Thorpe, 1982). Base substitutions can be either transitions (purine to purine, or pyrimidine to pyrimidine) or transversions (purine to pyrimidine, or *vice versa*). Transition type substitutions tend to be more frequent (Kimura, 1980).

Different molecules evolve at different rates, and homologous molecules may evolve at different rates in different organisms, so rate heterogeneity needs to be considered when attempting to calibrate the molecular clock. This can be evaluated by the 'relative-rate test', whereby the evolution rate is the same in each lineage if the genetic distance between each species and an outgroup is approximately the same (discussed in Avise, 1994, pp.101-109).

The fossil record has been used as a reference for estimating substitution rates in sequence data for the Metazoa, but is not generally a useful tool for prokaryotes. Ochman and Wilson (1987) used the fossil record and specific palaeoecological events to date divergence times on the 16S rRNA phylogenetic tree, with upper and lower limits to the

age of each lineage. Examples are the appearance of legumes and their symbiotic bacteria, such as *Rhizobium*, and the commensal association of *Escherichia coli* with mammals. They related these divergence times to observed divergences between protein-coding genes, of bacteria on this tree (*e.g.* between *Escherichia coli* and *Salmonella typhimurium*), to determine substitution rates. As a result, they gave an evolution rate for 16S rRNA of 1% (0.01 nucleotide substitutions per site) per 50 Ma. This rate was similar for both eukaryotes and prokaryotes, implying a universal substitution rate associated with cellular, rather than organismal, generations. Estimations for the 5S rRNA were for 1% substitution in 25 Ma, but Ochman and Wilson attached uncertainty to the use of this molecule, due to its small size (120 nucleotides in length). Earlier work by Kimura and Ohta (1973) used distance data for 5S rRNA and cytochrome *c* sequences, with palaeontological dating for the mammal-fish divergence, to date the divergence of the eukaryotes and prokaryotes at about 1.8 billion years ago. Based on their eukaryote-prokaryote 5S rRNA distance value, K = 0.817 +/- 0.158 (*i.e.* average number of nucleotide substitutions per site), this would give a substitution rate Knuc (Hori and Osawa, 1979; Kimura, 1980) of approximately 1% per 22 Ma. Though this appears to be supported by Ochman and Wilson's calculations, their estimates for the eukaryotic-prokaryotic divergence are given as a more recent 800 Ma. With this date in mind, Kimura and Ohta's distance data would give much faster rate estimates, of approximately 1% per 10 Ma.

Recently, Moran *et al.* (1993) estimated the evolution rate of the 16S rDNA sequences of endosymbiont bacteria of aphids. Based on their concordant phylogenies, the fossil record of the aphid hosts (and, in the case of one group, host plant fossils and biogeographic evidence) was used to date divergences of the bacterial lineages. Nucleotide substitution rates were estimated by two methods: (i) numbers of base substitutions separating pairs of sequences from their common ancestor at the dated node, and (ii) the minimal number of substitutions between the ancestral sequence at the dated node and its descendants, determined for each branch, then averaged. The evolution rate was found to be roughly constant among lineages (based on relative rate comparisons; see above), at 0.01-0.02 substitutions per site (Knuc) per 50 Ma, endorsing the use of 16S rRNA as a molecular clock. However, due to possible rate heterogeneity, caution should be exercised in extrapolating these results.

1.6 AIMS OF THIS THESIS.

This study had two main aims:

1. **To investigate the diversity and distribution of halobacteria in ancient salt deposits and compare these with modern hypersaline environments.**
Halobacteria were isolated from salt deposits of diverse geographical locations, ranging in geological age from 250-260 Ma to Recent. A polyphasic approach was taken in assessing the diversity of these halobacteria: they were grouped to genus level by analysis of polar lipid patterns; within-group diversity was investigated by analysis of SDS-PAGE protein profiles and restriction length polymorphisms of the 16S rRNA gene. Geographical and geological patterns of distribution were sought.

2. **To test the hypothesis that the halobacteria isolated from ancient salt deposits are autochthonous and 'living fossils'.**
Selected halobacterial isolates, representing most sites, were chosen for 16S rRNA gene sequencing and phylogenetic analysis, to seek a correlation between phylogenetic distance and geological age, which may provide evidence that the halobacteria from the salt mines are 'living fossils'. For example, the phylogenetic distance between an isolate from a 200 Ma salt deposit and a contemporary isolate should be twice that between an isolate from a 100 Ma deposit and a contemporary isolate, if they are on the same lineage. One might also suppose that these ancient isolates would be represented by points (rather than branches) on an extant lineage. Sequencing should also test the reliability of polar lipid analysis in grouping halobacteria.

II

MATERIALS AND METHODS

2.1 HALOPHILE ISOLATION MEDIA.

The nature of the salt deposit samples, and the mode of sampling, are described for each site in Section 3.1.
The media used for the isolation of neutrophilic halobacteria were based on that of Norton and Grant (1988), which was a modification of the halophile medium of Payne *et al.* (1960). The main changes to the Norton and Grant medium were: (i) part I was in a final volume of 300 ml and part II in a final volume of 700 ml and (ii) the pH of part I was adjusted to 7.6 with 4M NaOH. This was the CHM medium (Classical Halophile Medium).

Initially, for the samples from Great Salt Lake, Winsford and Boulby (see Section 3.1), in order to optimize the number and variety of halophilic bacteria isolated, cultures were incubated at 30 °C, 40 °C and 50 °C, and modifications were made to CHM, as follows:

(i) CHMst.
After adjusting the pH, 5.0 g l^{-1} soluble starch (stirred in) and *Difco* agar (for solid media), decreased to 15.0 g l^{-1}, were added to part I.

(ii) CHMbr.
Half of the 2.0 g l^{-1} KCl was replaced with KBr.

(iii) CHMtween.
10.0 g l^{-1} Tween-80 (polyethylene sorbitan monooleate) was added to part I.

(iv) CHMmilk.
All the CHM ingredients were in a final volume of 500 ml and autoclaved separately from 500 ml pasteurized milk; these components were then mixed.

(v) 10 % KCl and 18 % KCl - Halophile Medium.

As for CHM, but the NaCl was partially replaced with 100.0 g l^{-1} and 180.0 g l^{-1} KCl, respectively (used for the Boulby and Inowroclaw samples; Boulby is a potash mine and some samples were expected to be KCl-rich).

In addition, the Rodriguez-Valera minimal medium (RVM) was used (Rodriguez-Valera *et al.*, 1980), as the richer complex media may have precluded the growth of some organisms.

To standardize the viable counts, incubations were for one month. A representative selection of halophilic bacteria was sub-cultured from plates for each sample. Colonies with orange, pink or red pigmentation were presumed to be halobacteria. Only halobacteria were investigated further in this study.

For the reasons given in Section 3.1.2, the CHMst medium was subsequently adopted as the standard medium for the isolation of neutrophilic halophiles; and cultures were incubated at 30 °C and 40 °C only.

For the alkaline-saline samples, the standard alkaliphilic halophile medium was prepared (HALK), in which $MgSO_4.7H_2O$ was decreased to 1.0 g l^{-1} and added to the organics moiety, and 18.52 g l^{-1} Na_2CO_3 were mixed with the NaCl (Tindall *et al.*, 1980).

Halobacterial isolates required for polar lipid or DNA analyses were cultured in 10-15 ml of liquid CHMst medium (HALK, for the haloalkaliphiles), incubated at 40 °C in a shaking incubator. When greater cell mass was required, as for polar lipid analysis, a 10-15 ml culture was used to inoculate a further 250-300 ml medium. Broths were usually streaked onto media plates to check purity.

Cells were prepared for Gram-staining (Figure 1.1a-c) by fixing in 2 % acetic acid (Dussault, 1955), which was allowed to evaporate before applying the stains.

2.2 POLAR LIPID ANALYSES.

2.2.1 Polar Lipid Extraction.

Polar lipids were extracted from freeze-dried cells of 167 isolates of halobacteria (those appearing to be duplicates on the basis of the preliminary protein profile results were not included) and the following reference strains: the neutrophilic halobacteria *Halobacterium salinarium* CCM 2084, *Halorubrum saccharovorum* NCIMB 2081, *Haloarcula vallismortis* ATCC 29715, *Haloferax volcanii* NCIMB 2012, *Halococcus morrhuae* NCIMB 787, '*Halobacterium* sp.' NCIMB 777, *'Halobacterium cutirubrum'* NCIMB 763 and *'Halobacterium trapanicum'* NCIMB 767, and the alkaliphilic halobacteria *Natronobacterium gregoryi* NCIMB 2189 and SP1 (*Nb. pharaonis*).

The polar lipid extraction method used here was based on that of Collins *et al.* (1980), which was a modification of the procedure of Bligh and Dyer (1959), who used a chloroform-methanol-water solvent system to efficiently extract and purify total lipids from cells. The method used here does not involve water, the lipids being extracted from dried cells by heating with chloroform-methanol.

Halobacterial cell pellets were each obtained from 250-300 ml of stationary phase cultures in standard media. These were centrifuged in a cool (0-10 °C) 'Sorval' at 9K rpm for 20 minutes, the cell pellets washed by resuspension in sterile isotonic salts solution, then re-centrifuged to retrieve the pellets. Each cell pellet was transferred to a Universal, the top covered with 'clingfilm' pierced with holes to allow the escape of water vapour. Pellets were frozen at -70 °C for about 1hour, then freeze-dried until dry. The freeze-dried cells were ground until finely-divided. 5-7 ml (depending on the amount of cell material) of 1:1 chloroform:methanol was added to the freeze-dried cells (c. 100-150 mg), two closed staples added as magnetic stirrers and the Universal bottle tightly sealed with a Teflon-lined metal cap. Universals were heated on a magnetic heated stirring stand at 55-60 °C for 12 - 20 hours. After cooling, extracts were centrifuged at 2,500 rpm. for 10 minutes, at about 10 °C. The supernatant was transferred to a clean, dry Universal and the lipid extract was then dried completely under a stream of oxygen-free nitrogen gas. When dry, the lipid extract was resuspended in approximately 500 µl 1:1 chloroform:methanol and stored in a bead vial at 4 °C. The polar lipid extract was then ready for applying to TLC plates. If, after storage, there appeared to be particulate matter suspended in the extract, this was re-centrifuged and the clear supernatant transferred to a clean bead-vial prior to applying to TLC plates.

2.2.2 Thin Layer Chromatography.

The polar lipids were separated in two dimensions, using two different solvent systems, by thin-layer chromatography on silica gel plates (Collins *et al.*, 1980).
Chromatography solvent mixtures were as follows:

1st Dimension:-
chloroform / methanol / water (65 : 25 :4, by vol.)

2nd Dimension:-
chloroform / methanol / glacial acetic acid / water (80 : 12 : 15 : 4, by vol.)

The polar lipid extract was spotted (1.5-2.0 mm) onto the bottom left-hand corner (~1 cm in) of the plate using a capillary tube. The total amount of extract applied depended upon its concentration, but there was sufficient when there was a 'starring' effect.
After chromatography, to visualize the polar lipids, the TLC plates were sprayed with a solution of 0.1 % Ce $(SO_4)_2$ in 1M H_2 SO_4, then heated at 150 °C for about 3 minutes, until the charred grey spots of lipids appeared. (Glycolipids were initially pink/purple, before darkening.)
Two-dimensional thin-layer chromatography is a simple and rapid, yet sensitive technique, with high resolving power (Kates, 1986).
The resulting polar lipid profiles were compared with reference strains and, on this basis, isolates were placed into recognized groups of halobacteria.

2.3 SDS-PAGE

The method used for the separation of water-soluble whole-cell halobacterial proteins by SDS polyacrylamide gel electrophoresis was based on the discontinuous gel method of Laemmli (1970), in which an upper stacking gel concentrates the samples before separation of the proteins in the resolving (separating) gel. Gels and buffers contain the anionic detergent sodium dodecyl sulphate (SDS), a protein denaturant which binds to the proteins, inducing conformational changes, and producing linear complexes with a constant negative charge per unit mass, so that protein electrophoretic mobility is related only to molecular weight. In one-dimensional electrophoresis, the proteins are separated as bands, to form a protein profile, or fingerprint. These profiles can be used to distinguish one bacterial strain from another, indicate relationships, or be used to identify unknown isolates. Kersters and De Ley (1975) demonstrated that a computer-based numerical analysis of electrophoretic protein patterns required a highly standardized procedure for reproducible results, as patterns could change with variations in such as growth medium and centrifugation times. It was stressed that the electrophoresis conditions were also required to be standardized for reproducibility. The numerical analysis of one-dimensional electrophoretic protein profiles is a high resolution classification method, comparable to DNA-DNA hybridization (see Vauterin *et al.*,1993) and other typing methods, such as serotyping and phage typing (see Costas, 1992). The technique is becoming increasingly popular for the classification and identification of bacteria within genera, species, or groups (see Costas, 1992). Protein profile analysis has recently been used to compare species of halobacteria (Denner *et al.*, 1994; Hesselberg and Vreeland, 1995), but these studies involved only a limited number of halobacteria.

When first using this technique, samples were electrophoresed on small gels, using the Bio-Rad Mini-PROTEAN II slab cell kit. The protocol was basically as described below, but without attempting full standardization. The protein profiles were compared visually and sorted into groups to determine those samples to be co-electrophoresed on the same gel. This screening also served to reduce the number of isolates for polar lipid analysis, by eliminating duplicates.

Colonies were taken directly from a pure culture on a 'fresh' plate of solid medium with a sterile plastic loop, and mixed into an equal volume of PAGE loading buffer (usually *c.* 100 µl) in a screw-capped microfuge tube. Cells were thoroughly mixed by vortexing. Microfuge tubes were then placed in a preheated heating block, set at 100 °C, for 15 minutes, to break the cells, destroy proteolytic activity and extract the denatured whole-cell proteins by solubilization. The effect of SDS is as described earlier, and β–mercaptoethanol in the boiling loading buffer breaks disulphide bridges between

polypeptide chains. After cooling, the extracts were centrifuged at 13 K rpm for 15 minutes, and the supernatants, containing the extracted proteins, transferred to clean, labelled eppendorfs. These extracts were stored at -70 °C.

20 cm, 12 % gels were prepared and assembled for vertical electrophoresis using the Bio-Rad PROTEAN™ II slab cell kit and protocol. In the standardization of this protocol, 1.5 mm spacers were used between the glass plates and separating gels were poured to a depth 1.0 cm below the base of the wells, this also determining the depth of the stacking gel. The separating gels were prepared the day before electrophoresis, covered with a layer of water (to prevent evaporation and also to produce a level, horizontal surface) and 'clingfilm', and allowed to set overnight. The stacking gel was poured about an hour prior to assembly on the electrode core unit. A 15-toothed comb was used. A line, 10 cm from the top of the separating gel, was drawn on the outer glass plate of each gel to mark the length of the gel run.

Prior to loading onto the gel the extracts were again boiled, for 5 minutes, cooled, then centrifuged at 13 K rpm for 5 minutes to precipitate any remaining debris. Generally, 10 µl of each protein extract was loaded into the designated well in the stacking gel, using a Hamilton syringe which was rinsed 3 times with distilled water between each application. '*Pasteurella* 27', provided by the NCTC laboratory at Colindale and recommended for its well-spaced protein banding pattern, was used as the reference strain, and loaded into a central well, adjacent to the molecular weight markers (BDH 'Electran', molecular weight range 12,300-78,000) and also into the two pairs of end wells (subject to 'smiling' during electrophoresis). The remaining 9 wells were loaded with halobacterial protein extracts.
To maintain a relatively constant running temperature, the inner cooling unit was connected to a continually-flowing cold water supply. Proteins were electrophoresed using a Pharmacia Electrophoresis Constant Power Supply ECPS 3000/150 power pack, at a constant current of 30 mAmps per gel (*i.e* . 60 mAmps if 2 gels were being run), until the protein profiles were 10 cm in length. This gave an initial voltage of 70-80 volts and, on termination, a final voltage of 220 volts. The running time was in the region of 4 hours 40 minutes for 2 gels, or 6 hours if a single gel was run.

After electrophoresis, gels were immersed in Coomassie Brilliant Blue Stain, in a lidded shallow plastic box, and rocked gently, on a DENLEY A600 Rocker, overnight (or at least 2 hours), to simultaneously stain and fix the proteins. The stain was then drawn off with a 50 ml syringe, to avoid damaging the gel, and replaced with destain solution. The destain solution was periodically replaced with a fresh solution, until the gel background

was colourless and only the protein bands were stained. This process generally took 6-7 hours.
(If continued destaining was thought to be making the protein bands too faint whilst removing stain from the gel background, then the destain solution could be replaced with distilled water. Leaving the gel in water for at least a day (8 hours was sufficient if the gel was rocked) resulted in an enhancement of the bands. This may have been due to preferential adsorption of the stain from the gel background by the proteins, with only some of the stain going into solution.)

Polyacrylamide gels were dried using the NOVEX system (*R&D Systems Europe Ltd.*), whereby the gel was washed in water, soaked in a drying solution, then dried between two pieces of cellophane, clamped within a plastic drying frame, at room temperature. In place of the commercially-available NOVEX Gel-Dry™ drying solution, the solution used in this protocol was a mixture of 20 % (v/v) absolute ethanol, 10 % (v/v) glycerol and 70 % (v/v) distilled water. Gels were dry in less than 48 hours and kept flat by storing between the pages of a thick book. The dried gels were then suitable for scanning and analysis.

2.3.1 Analysis of Protein Electrophoresis Patterns.

The SDS-PAGE gels were scanned on a flatbed scanner and the scanned images loaded into the GelCompar™ (version 3.1b; Applied Maths BVBA, Kortrijk, Belgium) software package for analysis. A *Pasteurella* reference track was used as the database standard, with which the various gels could be calibrated. The 'converted' gels were normalized with background subtracted, by automatic association of the reference tracks by pattern recognition and subsequent alignment. In the analysis program, gel tracks were selected for cluster analysis. Track (densitometric curve) similarities were calculated using the Pearson product-moment correlation coefficient, optimized by the fine alignment option; the dendrogram was produced from the similarity matrix by the Unweighted Pair Group Method with Arithmetic averaging (UPGMA).

2.4 DNA EXTRACTION.

The DNA extraction method was based on that of Pitcher *et al.* (1989), in which guanidinium thiocyanate, a strong protein denaturant, was used to lyse the cells, and ammonium acetate to salt-out proteins and stabilize the DNA, which was precipitated by propan-2-ol. The halobacterial cells (except *Halococcus* spp.) could have been lysed in distilled water alone, but it was thought that the protein-destabilizing properties of guanidinium thiocyanate would probably produce a cleaner preparation. The procedure differed from that of Pitcher *et al.* in that sarkosyl was not used, and the extracted DNA was kept in propan-2-ol at -20 °C for 1 hour, or at 4 °C overnight, to allow the DNA to precipitate. Centrifugation at 8 K rpm for at least 2 minutes was usually required to precipitate the DNA pellet. After the wash steps (two were sufficient), the DNA was allowed to dry at room temperature. 1 µl 10 mg ml^{-1} RNAse was added to break-up any RNA which may be present. The DNA was stored at -20 °C, or 4 °C if in use.

The quality of the DNA was assessed by electrophoresis in a 1 % agarose gel at 50 v, for approximately 2.5 hours. The gel was stained in 1 x TAE buffer containing 0.5 µg ml^{-1} ethidium bromide, for 1.5 hours, then washed briefly in distilled water. The DNA was viewed on a UV transilluminator and the gel photographed.

Concentration and purity of the DNA were estimated by measuring the Optical Density of the solution at 260 nm and at 280 nm, using a Philips PU8720 UV/VIS Scanning Spectrophotometer. Optically matched quartz cuvettes were used to contain the samples for making the spectrophotometric readings at ultra-violet wavelengths. Nanopure water was used as a blank.

The DNA concentration was estimated from the reading at 260 nm.
At 260 nm, OD 1.0 = 50 µg ml^{-1} double-stranded DNA (Maniatis *et al.*,1982; p. 46).

2.5 ANALYSIS OF 16S rRNA GENE RESTRICTION PATTERNS IN HALOBACTERIA (RIBOTYPING).

The use of restriction fragment length polymorphisms (RFLPs) of rRNA genes as a means of identifying bacteria was originally proposed by Grimont and Grimont (1986). The procedure, now familiarly termed *ribotyping*, combines the techniques of restriction analysis, PCR amplification, Southern transfer and hybridization, in the following steps:

1. The DNA is digested with a restriction endonuclease.
2. The DNA fragments are separated by electrophoresis in an agarose gel.
3. The DNA fragments are transferred to a membrane (Southern Blotting).
4. A labelled rRNA gene probe (*e.g.* digoxygenin-labelled 16S rDNA) is hybridized to the rRNA (*e.g.* 16S) gene fragments on the membrane.
5. The bound probe is detected, in this method in a colorimetric reaction, thereby identifying the 16S rRNA gene fragments in a **ribotype**.

When whole-cell DNA is digested by a restriction endonuclease a complete range of sizes of DNA fragments is produced. These separate out along the agarose gel during electrophoresis and appear, after ethidium bromide staining, as a smeared track consisting of innumerable bands (Figure 3.12d, p. 141). The number of visible bands can be considerably reduced, to a countable number, by the ribotyping procedure (Grimont and Grimont, 1986). Following Southern transfer to a membrane, the DNA fragments are probed with labelled 16S rRNA gene PCR product (16S rDNA) to enable detection of DNA fragments containing all, or part of, the 16S rRNA gene (see Figure 3.12, p. 141). (This differs from the method of Grimont and Grimont, who used a 16+23S rRNA probe to detect 16S and 23S rDNA.) These fragments will be visualized as distinct bands on the membrane, at distances proportional to size.

The result will depend on the restriction enzyme chosen. With an enzyme cutting only outside the 16S rRNA gene, each band (and there may only be one) represents one or more (if of equivalent size) 16S rRNA gene copies (Sanz *et al.*, 1988). An enzyme with cleavage sites inside the 16S rRNA gene will produce a few, or several, fragments of DNA containing parts of this gene, which will be detected as multiple bands on the membrane. The result is a 'bar-code', or ribotype profile. Ribotyping is potentially highly sensitive as a taxonomic tool, with a resolving power equivalent to that of DNA/DNA hybridization in differentiating between strains of a species (Grimont and Grimont, 1986). It provides an effective method for the discrimination and typing of isolates on a

large scale, such as in epidemiological (Brown and Ison, 1993; Owen *et al.*, 1993) or environmental studies (Guillot and Leclerc, 1993).
To avoid using radioactive substances in this procedure, the 16S rRNA gene probe was labelled with digoxygenin (Liesack *et al.*, 1990) and visualized colorimetrically using the Boehringer Mannheim DIG detection system. This system has been used successfully by Brown and Ison (1993), to identify different ribotypes of *H. ducreyi* from diverse geographical locations, and by Priest *et al.* (1994) to show extensive genetic heterogeneity in *Bacillus thuringiensis* strains, with results which were congruent with serotyping.

The main aim in applying ribotyping to the halobacterial isolates was to confirm groupings determined by polar lipid analysis and protein profiling, differentiate between like-strains and, if possible, determine the minimal number of 16S rRNA gene copies in each isolate. With these aims in mind, the *map* program in the GCG software package (see Section 2.6.4) was used to determine the number of restriction sites in published halobacterial 16S rRNA sequences, for all the restriction enzymes in the *map* program, to choose the appropriate enzymes for this analysis, *i.e.* (i) an 'external cutter', only cutting outside the halobacterial 16S rRNA gene, and (ii) an 'internal cutter', frequently cutting within the halobacterial 16S rRNA gene.

2.5.1 Preparation of the 16S rRNA Gene Hybridization Probe.

The Boehringer Mannheim PCR DIG labeling mix (Cat. No. 1585550) was used to PCR-amplify and directly label with digoxygenin-11-dUTP the 16S rRNA gene of isolate GSL5.48, in a procedure similar to that of Liesack *et al.* (1990). This product was then used as a probe for ribotyping halobacterial isolates.
The PCR DIG labeling mix is a solution of 2 mM dATP, 2 mM dCTP, 2 mM dGTP, 1.9 mM dTTP, and 0.1 mM digoxygenin-11-dUTP, as their lithium salts, in water, at pH 7.0. The PCR protocol was as used in the preparation of amplified DNA for sequencing (see Section 2.6.1), except that 6.25 µl of the Boehringer Mannheim PCR DIG labeling mix was used (≡ 125 µM dNTP) and the volume of water added was reduced to 77.75 µl. A HYBAID Thermal Cycler was used for the PCR reactions (see Section 2.6.1 for the program), which took 3 h 40 min to complete. As this DIG-labelled PCR product was to be used as a 16S rRNA probe it was not necessary to purify it, as any unincorporated DIG-11-dUTP would not bind to the target DNA and remaining unlabelled DNA would not be detected (Boehringer Mannheim; personal communication).

The success of the PCR reaction was determined by electrophoresis of the DNA alongside a 1 Kb DNA marker and previously successful 16S rRNA PCR products, on a 1 %

agarose gel, and detection under UV light after staining with ethidium bromide (0.5 μg ml^{-1} in 1 x TAE buffer).

To confirm that DIG-11-dUTP was successfully incorporated into the halobacterial 16S rRNA gene product during the PCR reaction it was necessary to show that the 16S probe could be detected by the DIG system. Dilutions of the probe (10^0-10^{-7}) were prepared in pre-hybridization solution (*i.e.* 100 ng μl^{-1} to 10 fg μl^{-1}). Each probe preparation (1 μl) was spotted onto HYBOND™-N membrane, then fixed by U.V. cross-linking (face down on a U.V. transilluminator) for 3 minutes. (A sample of Boehringer Mannheim positively-charged nylon membrane was also used, for comparison.) The hybridization and detection reactions were followed as for the ribotyping protocol, below.

To confirm that the DIG probe could be successfully detected after hybridization to halobacterial DNA (and therefore that the DIG-11-dUTP label had been incorporated during the PCR reaction), to determine the level of detection of DNA and also to confirm the probe concentration required for detection, a 10^0-10^{-6} dilution series of isolate T3.1 16S rDNA (from undiluted 100 ng μl^{-1} to 100 fg μl^{-1}, diluted in nanopure water) was prepared. Each dilution (1μl) was dot-blotted on to each of three labelled pieces (*c.* 2 cm x 7 cm) of HYBOND™-N hybridization membrane. The DNA was fixed to the membrane by UV cross-linking for 3 minutes. The membranes were pre-hybridized for 2 hours, then hybridized for 20 hours (see below for protocol) using three different concentrations (100 ng, 300 ng, and 500 ng per 3 ml prehybridization buffer, respectively) of freshly heat-denatured probe (heated in a boiling water-bath for 10 minutes, then chilled immediately on ice). Following hybridization, the membranes were washed and the detection reactions were followed as for the DIG-detection protocol, below.

2.5.2 Restriction Endonuclease Digestion.

The restriction endonucleases *Eco*R I (an 'external cutter') and *Taq* I (an 'internal cutter') were principally used (independently) for DNA digestion; *Sal* I, *Xho* I, *Bam*H I and *Kpn* I were also separately used to confirm or supplement results. These enzymes have the following DNA cleavage sites:

Taq I	5' ... T↓CGA ... 3'
*Eco*R I	5' ... G↓AATTC ... 3'
Sal I	5' ... G↓TCGAC ... 3'
Xho I	5' ... C↓TCGAG ... 3'

*Bam*H I 5' ... G↓GATCC ... 3'
Kpn I 5' ... GGTAC↓C ... 3'

The reagents were mixed in the following order, in a 0.5 ml microfuge tube: sterile nanopure water (to bring the final volume to 10 µl), 1 µl buffer, 1 µg DNA and 1 µl (10 units) restriction enzyme. λ DNA was included as a control. When possible, to reduce handling time and pipetting error, volumes were scaled-up to make a 'mastermix' of water, buffer and enzyme, from which 9 µl aliquots were removed for the addition of DNA. This was briefly vortexed to mix, then pulse-centrifuged, before incubating at 37 °C (65 °C for *Taq* I) in the water-bath, for about 20 hours.

2.5.3 Agarose Gel Electrophoresis.

After incubation, the reaction tubes were pulse-centrifuged, then 2 µl 6 x DNA loading buffer added to each. The digested DNA fragments were separated by electrophoresis in a horizontal 1 % agarose gel at 50 v, for approximately 2.5 hours. The gel was stained in 1 x TAE buffer with 0.5 µg ml^{-1} ethidium bromide, for 1.5 hours, then washed briefly in distilled water. The DNA was viewed on a UV transilluminator and the gel was photographed alongside a fluorescent ruler, so that sizes of ribotype bands can later be estimated if required. Streaks of DNA on the agarose gel (Figure 3.12d) indicated that the DNA had been cut by the enzyme.

2.5.4 Southern (Vacuum) Blotting.

The gel was washed in a denaturation (alkaline) solution for 2 x 15 minutes, to denature and linearize the DNA, so that it could be easily transferred to the membrane. This was followed by neutralization washes for 2 x 15 minutes, to lower the gel pH to less than 9.0. With the transfer buffers 2 x SSC beneath the gel, and 20 x SSC on top of the gel, a Hybaid vacuum blotter was used to transfer the DNA from the gel to the HYBOND™-N membrane, using the protocol recommended in the Hybaid Vacu-Aid instruction manual (Hybaid Limited, 1992).

After Southern blotting (1.5 hours), the positions of the gel slots and of any bubbles were marked with a pencil. If required, the gel was restained with ethidium bromide to check that the DNA had been successfully transferred (1.5 hours). The membrane was washed in 2 x SSC buffer for 2 minutes at room temperature, allowed to air-dry on 3MM filter paper, then wrapped in Saran Wrap and placed face down on a UV transilluminator for 3 minutes, to cross-link the DNA to the membrane.

2.5.5 Hybridization.

The membrane was placed in a hybridization bag with 40 ml pre-hybridization buffer, heat-sealed and incubated at 68 °C (in a water-bath) for 1.5 hours. This step was to block non-specific DNA binding sites on the membrane.

The pre-hybridization buffer was then removed from the hybridization bag and replaced with hybridization buffer (*i.e.* 6 ml pre-hybridization buffer with 6 µl probe added), the bag re-sealed and incubated at 68 °C for about 20 hours.

After hybridization the membrane was washed, to remove unbound probe, as recommended (Boehringer Mannheim, 1993).

2.5.6 Detection of the DIG-Labelled 16S rRNA Gene Restriction Fragments.

The 16S rRNA gene restriction fragments, labelled with a DIG-11-dUTP-16S rRNA gene probe during the hybridization reactions, were detected colorimetrically following a reaction between anti-digoxigenin[Fab]-alkaline phoshatase conjugate and the colour substrate solution NBT with X-phosphate, using the DIG Nucleic Acid Detection Kit for colorimetric detection (Cat. No. 1175041) and following the recommended protocol (Boehringer Mannheim, 1993). Membranes were kept wet in TE buffer, for photographing.

2.5.7 Confirmation of Ribotype Patterns.

The number of 16S rRNA gene cleavage sites for the halobacteria sequenced for this thesis, in addition to those for published halobacterial 16S rRNA sequences, were determined for the enzymes *Taq* I, *Eco*R I, *Sal* I, *Xho* I, *Bam*H I and *Kpn* I, using the *map* program in the GCG software package (see Section 2.6.4). The expected number of bands in the ribotype could be determined from this information and compared with the observed results, to aid in their interpretation.

2.6 PHYLOGENETIC ANALYSIS.

The 21 salt deposit isolates and 3 culture collection strains, chosen for phylogenetic analysis as part of this study, are described in Section 3.5. The 16S rRNA gene was amplified by PCR, sequenced using the DyeDeoxy™ Terminator Cycle Sequencing method, and sequences determined by the Applied Biosystems automated 373A DNA Sequencer, as described below.

2.6.1 16S rRNA Gene Amplification by the Polymerase Chain Reaction (PCR).

Theoretical Basis and Methodology of PCR.

The concept of the polymerase chain reaction was evolved in 1983 by Kary B. Mullis, then working for the Cetus Corporation (Mullis, 1990), who received the Nobel Prize for its invention in 1993. Patents for the PCR process are owned by Hoffmann-La Roche AG (covered by U.S. patents 4,683,195 and 4,683,202; 1987).

The PCR reaction is an *in vitro* method for DNA synthesis involving three linked stages of DNA replication which are cyclically repeated to produce multiple copies of the original template DNA. DNA synthesis is specified by two oligonucleotide primers, a 'forward' and 'reverse' primer, which hybridize to opposite strands of the DNA and bracket the region to be copied. The three stages of the process are (i) thermal denaturation, to melt and separate the duplex DNA to single-stranded templates, (ii) cooling, to allow annealing of the forward and reverse primers to opposite DNA strands, and (iii) extension of the primers in a 5'→3' direction by a DNA polymerase, drawing on a pool of dATP, dCTP, dGTP and dTTP, to produce complementary strands of DNA. At the end of each cycle this results in a doubling of the target duplex DNA. These reactions are controlled by temperatures and times that are optimal for each stage. After 25-40 cycles, in an exponential increase of DNA copies, the original template DNA can theoretically be amplified around a billion-fold (x 10^9 for 30 cycles), but in practice limiting factors (*e.g.* thermal denaturation of the enzyme and reannealing of target strands competing with primer annealing) reduce the efficiency of the process and tend to restrict replication to about a million copies (Newton and Graham, 1994).

In the early days of PCR, before the discovery of the thermal-stable enzyme from *Thermus aquaticus*, *Taq* DNA polymerase, the DNA polymerase had to be added after the denaturation step in each cycle. The thermal cycling reactions were achieved by the use of three heated waterbaths and the timed transfer of the reaction tubes from one water-bath to the next. The use of thermal cyclers, which are essentially microprocessor-controlled

heating blocks, allow the automation of the polymerase chain reaction, under more controlled conditions.

Primers are typically 18-28 nucleotides in length, designed to hybridize specifically to flanking regions of the DNA to be copied, in this case the 16S rRNA gene. It is recommended that the G+C composition should be 50-60%, the T_m 55-80 °C and equal or similar for each primer. Complementarity of the 3' ends of the primer pairs (which can result in primer-dimers), palindromic sequences within the primers, self-complementarity (secondary structure) and stretches of one base type, should all be avoided. If the template DNA has significant secondary structure this can hinder primer annealing or extension. This can be circumvented by replacing dGTP with the helix-destabilizing analogue 7-deaza-2'-deoxyguanosine, c^{7}dGTP (Innis, 1990). The reaction temperature for primer annealing is taken to be 5 °C below the T_m of the primers, which can be calculated using the rule of Suggs *et al.* (1981):

$$T_m = 2\ °C\ (\text{number of A} + \text{T}) + 4\ °C\ (\text{number of G} + \text{C})$$

The primers used for 16S rRNA PCR amplification of halobacterial DNA in this study were synthesized on an ABI 380B DNA Synthesizer (Department of Biochemistry, University of Leicester) and designed by T.J. McGenity (PhD Thesis, 1994), and were as follows:

Forward:

5'- TCCGG TTGAT CCTGC CGGAG -3' (65% G+C)

(position 3 to 22; sequence alignment numbering, Appendix 1)

Reverse (complementary): 5'- AAGGA GGTGA TCCAG CC -3' (59% G+C)

(position 1482 to 1466; sequence alignment numbering, Appendix 1)

The DNA polymerase used in this study was AmpliTaq® (Perkin-Elmer Cetus), a modified, recombinant form of *Taq* polymerase, the thermostable DNA polymerase from *Thermus aquaticus*. AmpliTaq has an optimal temperature of about 75 °C and retains activity at denaturation temperatures (a half-life of 40 minutes at 95 °C). AmpliTaq still has over 70% activity remaining after exposure to 96 °C for 30 cycles (AmpliTaq® data sheet). The high temperatures allowed by this enzyme also minimize secondary structure problems.

The amount of DNA template added to the 100 µl reaction mix was usually about 100 ng.

PCR Conditions.

The PCR temperature cycling reactions were automated, using either a HYBAID (IHB 2024) 'Intelligent Heating Block' or a PERKIN ELMER Cetus DNA Thermal Cycler. The method was based on that of McGenity (1994), with some modification of the thermal cycling program, partially based on recommendations in Innis and Gelfand (1990), when the PERKIN ELMER Cetus DNA Thermal Cycler was used.

The reactions began with an initial denaturing step of 96 °C for 5 minutes, after which the *Taq* polymerase was added, immediately followed by the PCR amplification program, which was as follows:

	HYBAID (IHB 2024)	PERKIN ELMER Cetus
Denaturation	95 °C 15 seconds	95 °C 30 seconds
Annealing*	55 °C 40 seconds	55 °C 40 seconds
Extension	72 °C 2 minutes	72 °C 2 minutes
Number of cycles	30	30
Final Extension	72 °C 3 minutes	72 °C 4 minutes

*The T_m for the primers was calculated using the rule of Suggs *et al.* (1981). The forward primer T_m = 66 °C; the reverse primer T_m = 54 °C; the mean T_m = 60 °C. The primer annealing temperature for this reaction was taken to be 55 °C.

The success of the PCR reaction was assessed by visualizing the products on a 1 % agarose gel. The 16S rRNA PCR products were stored at 4 °C prior to purification by the QIAGEN PCR purification kit.

2.6.2 Direct 16S rRNA Gene Sequencing.

A potential problem with gene amplification by the PCR reaction can be erroneous incorporation of bases during the amplification process. Direct sequencing of double-stranded amplified DNA, as an alternative to cloning, can alleviate this as errors introduced in early cycles will result in ambiguities, whereas those appearing later in the process will be greatly outnumbered by correctly copied products and so not be detected (Sogin, 1990).

Double stranded DNA can be sequenced by the cycle sequencing method, which utilizes the chain termination reaction of Sanger *et al.* (1977) in a linear PCR amplification of the double-stranded PCR product (Murray, 1989; Embley, 1991). Only one primer is used, therefore extension is unidirectional, resulting in a linear increase in DNA copies (not

exponential, as in PCR gene amplification, when forward and reverse primers are employed). Cycle sequencing involves the denaturation, primer annealing and extension reactions of a conventional PCR protocol, though the details of the programs are different.

The PRISM™ Ready Reaction DyeDeoxy™ Terminator Cycle Sequencing Kit and Protocol (no. 401388; Rev. A) used in this study was developed by Applied Biosystems, specifically for analysis on the Applied Biosystems Automated Sequencing system. This method uses the 16S rRNA PCR product directly. In this kit, each dideoxynucleoside triphosphate (ddNTP) is labelled with a different fluorescent dye, so that all reactions can be performed in a single tube (instead of one per dNTP). The reaction product is then run on a single lane of a sequencing gel and the coloured bands identify the terminating ddNTP bases of the DNA copy fragments.

'Dye Terminator' sequencing was chosen in preference to 'Dye Primer' sequencing (primers are dye-labelled rather than the ddNTPs) as the primers used for sequencing halobacterial 16S rRNA DNA are custom-made, non-standard primers, which would otherwise have to be specifically dye-labelled. Dye terminator sequencing also has the advantage that prematurely terminated fragments are not detected, whereas with the dye primer method, as the terminal base of a prematurely terminated fragment does not necessarily correspond to that indicated by the primer label, this results in higher background 'noise'. Secondary structure impedence can be seen in dye primer sequence data as 'false stops' and 'compressions', although the use of a thermally stable polymerase (*Taq*) and/or c^{7}dGTP (Innis, 1990) should minimize this problem.

Three forward primers (AF, BF and CF) and four reverse, complementary primers (ER, FR, GR and HR) were used, so that the overlapping sequencing products would enable the entire 16S rRNA gene to be sequenced. The primers used for cycle sequencing of the halobacterial 16S rRNA genes were designed by T.J. McGenity (PhD Thesis, 1994) and synthesized on an ABI 380B DNA Synthesizer by the Protein and Nucleic Acid Chemistry (PNAC) laboratory of the University of Leicester. The primers were as follows:

Primer	Sequence (5' to 3')	Position ϕ	G+C%	Annealing Temp. °C (Suggs *et al.*, 1981)
ER*	CTGCT GCGCC CCGTA G	337-322	75	51
AF	CTACG GGGCG CAGCA G	322-337	75	51
FR*	TACCG CGGCG GCTGG C	472-457	81	53
BF	GCCAG CCGCC GCGGT A	457-472	81	53
GR*	CGGGT CTCGC TCGTT G	1059-1044	69	49
CF	ACGCG GGCTA CAATG G	1180-1195	62	47#
HR*	ACGGG CGGTG TGTGC	1359-1345	73	47

* reverse and complementary sequence.

ϕ base positions based on the 16S rRNA sequence alignment in this study (Appendix 1).

due to a calculation error, an annealing temperature of 51 °C was used.

The sequencing reactions were performed in a PERKIN ELMER Cetus DNA Thermal Cycler (as for the PCR amplification reactions). Ramping times approximated 1 C° /second.

The tubes were placed in the pre-heated thermal cycler, with a drop of mineral oil in the wells (for efficient thermal contact and to prevent sticking), and the cycle sequencing program was immediately begun. The programs were based on one recommended for the Perkin Elmer Cetus Model 480 thermal cycler, in the Applied Biosystems, Inc. protocol (Part Number 901497; Rev. E). DNA was denatured at 96 °C. As annealing temperatures varied with the primer, so did timing, so that the total run time remained similar. Extension was at 60 °C for 4 minutes.

The programs (based on that for the PERKIN ELMER Cetus Model 480) were as follows:

	Primer →	HR	GR	ER AF CF	FR BF
Denat-uration	1 second ramp to	96 °C	96 °C	96 °C	96 °C
	Hold at 96 °C for	30 secs.	30 secs.	30 secs.	28 secs.
Anneal-ing	1 second ramp to	47 °C	49 °C	51 °C	53 °C
	Hold for	15 secs.	15 secs.	15 secs.	13 secs.
Extens-ion	1 second ramp to	60 °C	60 °C	60 °C	60 °C
	Hold at 60 °C for	4 mins.	4 mins.	4 mins.	4 mins.
Number of cycles		25	25	25	25
Run Time (approx.)		2h.57min.	2h.59min.	2h.52min.	2h.48min.

The ramp rate of the thermal cycler was approximately 1 C° per second. Run times approximated 3 hours (Applied Biosystems, Inc. recommend 2 hours 47 minutes ± 5 minutes for 25 cycles).

On completion of the cycle sequencing program the reactions were rapidly cooled to 4 °C (achieved by linking the thermal cycling program to a 4 °C soak program), then transferred to a refrigerator for storage.

The sequencing reaction products were purified by phenol / chloroform extraction (Applied Biosystems).

2.6.3 Automated Sequencing of the 16S rRNA Gene.

Sequences were determined by the Applied Biosystems automated 373A DNA Sequencer. This is a service provided by the Protein and Nucleic Acid Chemistry (PNAC) laboratory of the University of Leicester, who receive the cleaned sequencing products, process them, and later return the results as sequence chromatograms, readable with the *ABI* software installed on Macintosh computers in customer laboratories.

Each sample, containing all four fluorescent dye-labelled ddNTP terminator sequencing products, is loaded into a single well of a vertical acrylamide gel, so that one template sequence can be read from one lane. The Applied Biosystems automated 373A DNA Sequencer incorporates a laser which scans the width of the gel as the samples run down it, during an electrophoresis run which takes about 12 hours. The emission from each of the fluorescent dyes is collected by a lens, directed through a filter (one for each dye

alternately moving into place), focussed onto a photomultiplier tube, with an output to a computer. The ABI software kit processes this raw data from the electrophoresis run and uses it to call the sequence, with each base represented by a different coloured peak on a chromatogram. The ABI SeqEd™ software kit allows sequence editing and alignment, *etc.* The sequence data can be reformatted for entry into alternative sequence editing programs, such as GCG (see below).

2.6.4 From 16S rRNA Sequences to Phylogenetic Trees.

The procedures undertaken to derive phylogenetic trees from the 16S rRNA sequences are outlined, in order, below:

ABI SeqEd™.

The ABI SeqEd™ software (version 1.0.3) provides the facilities to edit sequences. These include reverse-complementation of sequences derived from reverse primers and sequence alignment (multiple and overlap). Chromatograms can be displayed (Figures 3.14 and 3.15).

Successful sequences were exported from SeqEd, in GCG format, to a temporary folder on the Macintosh hard-disk and transferred from there, by ftp, to the user directory on IRIX (mainframe computer). The data was organized so that all primer sequences for a single isolate were in one directory. Programs for sequence analysis, CLUSTAL V, GCG and ReadSeq, were accessed from the Molecular Biology Packages, version 6.05, at Leicester.

GCG

The GCG (1991; version 7.3 - UNIX, June 1993) molecular biology software package was used to edit, assemble and align the sequence data:

GelStart this program had to be run before *GelEnter* and the other sequence assembly programs. New fragment assembly projects were created in *GelStart.*

GelEnter allowed the primer sequence data to be entered, one at a time. Preliminary editing was done, *e.g.* removing the stretch of Ns usual at the 3' end.

GelOverlap searched for sites of overlap between sequence fragments in the fragment assembly project, producing a file for input into *GelAssemble.*

GelAssemble is a multiple sequence editor, allowing sequence assembly in 'contigs'. Sites of overlap, determined by *GelOverlap*, were adjusted if necessary. Changes based on analyses of the chromatograms, such as identification of 'Ns' or correction of miscalled bases, could be made. The consensus sequence was then saved.

All 16S rRNA sequences for multiple alignment and analysis were saved in a single file. These included, for comparison, published sequences made available to the public from EMBL, GenBank or the Ribosomal Database Project (RDP), though in some instances these were added at a later stage, in ***lineup***. (See Table 3.17 for sequenced strains, and Table 3.18 for published strains, used in phylogenetic analysis).

topir this organized the sequences, contained in a single file, in a sequential format (x.pir) acceptable to the multiple alignment program, CLUSTAL V, or the later version, CLUSTAL W.

CLUSTAL V(W).

This program was used to align the sequences, in a multiple sequence alignment, producing a multiple sequence file (x.msf).

GCG.

lineup the x.msf file was edited. The alignment was optimized by manual adjustment, particularly in gap regions, and with reference to the secondary structures. Sequences could be removed. Additional sequences were included either by manually entering them or, if they were in GCG format and in the same directory, by using the command 'get'. Input to *lineup* was limited to 30 sequences, however. Apparent sequence anomalies were checked against the original chromatograms.

ReadSeq.

This program (version 30 Dec 92) reads sequences and changes their format (*e.g.* to PIR, MSF, GCG, EMBL, PHYLIP, *etc.*).

The x.msf file was renamed x.infile and all sequences within this file were reformatted for PHYLIP3.2.

Eve.

As PHYLIP does not recognize '.' as gaps, all spaces (.) in x.infile were replaced with a '-' in the 'eve' editor (on IRIX).

PHYLIP 3.4

Once the 16S rRNA gene sequence data had been optimally aligned and in the correct, PHYLIP 3.2 format (suitable for PHYLIP 3.4), it was ready for use in deriving phylogenetic trees. PHYLIP, the PHYLogenetic Inference Package of Joseph Felsenstein and the University of Washington (copyright 1986-1993), comprises a whole suite of programs for inferring and drawing phylogenetic trees from molecular sequence data. The SEQNET computing facilities of DRAL Daresbury Laboratory (e-mail: seqnet@dl.ac.uk) were used for this. Each x.infile was transferred to Daresbury by ftp (file transfer protocol) and the trees derived using PHYLIP 3.4 in SEQNET.

The treeing methods and options used were as follows:

1. Distance Tree.

DNADIST (1986-1993, by Joseph Felsenstein and the University of Washington).

A matrix of distances between each pair of sequences was computed from the sequence data (x.infile) under the **Jukes and Cantor** (1969) model of nucleotide substitution (see Section 1.5.5.2). Settings used in this program were the default ones, of a transition/transversion ratio of 2.0 and one category of substitution rates. The distance matrix outfile was input into the FITCH program for phylogenetic tree inference.

FITCH (Fitch-Margoliash method version 3.5c) estimates total branch lengths from pairwise distances in the distance matrix. The default settings were used, except that the outgroup root was requested (*Methanospirillum hungatei*, member of the *Methanomicrobiales*, a sister group of the extreme halophiles (Burggraf *et al*, 1991)), as were global rearrangements (reconsiders the position of each species on the tree) and randomized input order of the species (random number seed 1315; three times). These latter two options were chosen to optimize the trueness of the tree.

The output from FITCH is a 'treefile' and an 'outfile'. The outfile contains a tree with all nodes named (terminal nodes) or numbered (internal nodes), and all branch lengths are listed. It also gives statistical information: sum of squares, average percent standard deviation and number of trees examined. The treefile is used to plot a tree.

Bootstrapping.

Distance trees were resampled by bootstrapping the data, to generate multiple (100) data sets, from which a consensus tree was derived. The procedure was as follows:

SEQBOOT (1991-93, by the University of Washington and J. Felsenstein) was run (random number seed 1315) on the x.infile data set to produce 100 bootstrapped data sets, each a resampled version of x.infile. The outfile from SEQBOOT was then used as infile to DNADIST (Jukes-Cantor method), with the multiple data sets option set to 100. The outfile was input to FITCH, again with the multiple data sets option set to 100. An outfile and treefile, each with data for 100 trees, were produced.

CONSENSE was run on the treefile from FITCH, to derive a majority rule consensus tree from the 100 trees (a treefile and outfile are produced). At each internal node on the consensus tree, the number of occurrences of the descendent group is enumerated; *e.g.* 98 indicates a particular group descending from that node in 98% of trees. The tree could be plotted from the treefile.

2. Maximum Likelihood Tree.

DNAML (version 3.4c in PHYLIP 3.4).

The sequence data (x.infile) was input directly into DNAML. The default settings, a transition/transversion ratio of 2.0, using empirical base frequencies (observed in the input sequences) and one category of substitution rates, *etc.*, were used. *Methanospirillum hungatei* was included as an outgroup (Burggraf *et al*, 1991), to root the tree. The Maximum Likelihood trees were not bootstrapped as the operation would be excessively computer-intensive.

An 'outfile' and 'treefile' were produced from DNAML. The outfile provided information on the number of species, number of sites and the base frequencies for A, C, G and T. The base frequencies were used to compute the actual transition/transversion ratio (the default or chosen transition/transversion ratio was also given). A tree was given, with its log likelihood and the number of trees evaluated; also a table showing the lengths of each tree segment (units of expected nucleotide substitutions per site) and the associated confidence limits (rough). The tree was plotted from the treefile.

Plotting the Trees.

The files were transferred from Daresbury to IRIX, and from here to the user temporary file on the laboratory Macintosh computer. The trees were plotted from the treefiles using the PHYLIP programs **DRAWTREE** and **DRAWGRAM** (Phylip 3.54c). The files

were then transferred to **PowerPoint** (MS-PowerPoint 3.0), where the trees could be improved for presentation (labels, legends, *etc.*). Trees were given a scale indicating a phylogenetic distance of 0.015 Knuc (average number of nucleotide substitutions per site), this being equivalent to 50 million years, based on the calibration of Moran *et al.* (1993).

Secondary Structures.

The 16S rRNA secondary structure of *Haloferax volcanii*, obtained from the Ribosomal Database Project via the Internet, was used as a template to draw the 16S rRNA secondary structures of the sequences (e.g. Figure 3.16). This was useful in many ways, for example in looking for putative structure-related anomalies in the sequences, such as the 'N' sites or positions of base heterogeneity (see Section 3.5.1). They were also used as a guide in determining sequence alignments and afforded some test of the reliability of the primary structures, providing a check against chimeric sequences.

III

RESULTS AND DISCUSSION.

3.1 SAMPLING SITES.

Salt samples for microbiological analysis were taken from several geographically distinct sites of a spectrum of geological ages. Some of these samples were collected by the author, some by other members of the laboratory and the remainder were posted by mine personnel. These are summarized in Table 3.1.

Table 3.1. Sampling Sites for Microbiological Analysis.

Site	Geological Period and Age (Ma BP)
Great Salt Lake	Recent (< 30 000 years)
Spanish salterns	Recent
Wieliczka Salt Mine, S. Poland	Lower Miocene (*c.* 20)
Green River Valley, Wyoming, USA	Lower / Middle Eocene (*c.* 50)
Khorat Basin, Thailand	Upper Cretaceous / Lower Tertiary (*c.* 65)
Winsford Salt Mine, Cheshire, N.W. England	Middle Triassic (Anisian stage; 235-240)
Boulby Potash Mine, Cleveland, N.E. England	Upper Permian (EZ3; 250-260)
Solno and Gora Solution Mines, Inowroclaw, Poland	Upper Permian (PZ3; 250-260)
Mariager, Denmark (AKZO)	Upper Permian (250-260)
Delfzijl, Netherlands (AKZO)	Upper Permian (250-260)
Hengelo, Netherlands (AKZO)	Upper Permian (250-260)
Stade, Germany (AKZO)	Upper Permian (250-260)
Werra and Hannover, Germany (Kali und Salz)	Upper Permian (Z1 and Z3; 250-260)

3.1.1 Choice of Growth Media for the Isolation of Halobacteria.

Various growth media had been used for the Great Salt Lake, Winsford and Boulby samples in order to optimize the isolation of halobacteria, in numbers and variety (Section 2.1). One significant observation arising from this was that viable counts were increased when CHM + starch medium was used (Oren 1990), by at least 10% and sometimes as much as x2 or x3. For example, at 40 °C, the viable count for WM2S was estimated at 1.3×10^5 g^{-1} on CHM, whereas it was 2.7×10^5 g^{-1} on CHMst. The variety of bacteria on the CHMst medium appeared to be equivalent to that on the CHM medium.

The CHM + milk medium sometimes gave the highest viable counts and also appeared to be selective for halobacteria, although Bacteria were not excluded. For example, at 30 °C, isolates from the WB4br sample, from Winsford, were predominantly orange-pigmented (presumed halobacteria) on CHMmilk medium, whereas on CHMst there were few orange colonies and Bacteria predominated; and B3br from Boulby, at 30 °C, produced 92% 'reds' on CHMmilk, but only 53% 'reds' on CHM.

The CHMbr medium usually gave results which were similar to the CHM medium. Counts on CHMtween were usually comparatively low. They tended to be least on the RVM minimal medium and colonies were often 'tiny', presumably because the nutrients were inadequate.

CHMst was thenceforth used as the standard isolation medium for halophiles and viable counts on CHMst are given in the results below.

Over 860 isolates of halophilic bacteria and halobacteria were initially sub-cultured from the hundreds of culture plates of the sites described below. This number was reduced by eliminating those with poor growth or in mixed culture, and many which appeared identical and were from the same sample; unusual forms, and those representing the major colony form isolated, were retained. A second screening procedure used analysis of protein electrophoresis patterns (small gels; Section 2.3) to eliminate many of the identical (same sample) isolates. Only the halobacteria (about 170, presumed from colony pigmentation) were to be investigated further.

3.1.2 Great Salt Lake, USA.

The Great Salt Lake in Utah, North America, is a 30,000 year old remnant of the ancient Lake Bonneville, which existed in Pleistocene times as a freshwater lake. During its history, the Great Salt Lake has been subjected to fluctuations in depth, area and salinity

in response to changes in climate. Today, the lake covers about 3,900 km^2, averages 10 m in depth and has a salinity up to 33.0 %. Prior to 1959, when the construction of a railroad causeway divided the lake into a north and south arm, the lake had a salinity of about 20 %. With the causeway, the South Arm, which receives the inflow, was diluted to about 12 %, whereas the isolated North Arm had an increased salinity of about 33 %. (Javor, 1989). The causeway was breached in 1984, allowing some equilibration between these extremes.

The ionic composition of the lake is given in Table 1.1. The principal minerals, produced on evaporation of the lakewater, are halite (86 %) and sodium sulphate. The latter is deposited as mirabilite on the lake shores in winter (Javor, 1989).

Brine samples were collected by W.D. Grant in Summer, 1991, from the South Arm of the lake (GSL2) and, adjacent to the lake, from a solar evaporation pond (GSL5) which was coloured red, presumedly due to a population of the red-pigmented halobacteria. The brine was collected in sterile plastic tubes.

GSL2

Viable counts were highest on plates which had been incubated at 18 °C for 1 week, then at room temperature (25-27 °C) for 3 weeks, with an optimal count of 2.9 x 10^3 ml^{-1} on CHMbr medium. Counts at 30 °C were greater than at 40 °C, and growth was absent or negligible at 50 °C. There was no growth on RVM minimal medium at the incubation temperatures used. On all media where growth occurred, colonies were predominantly pale dusky pink and presumed to be Bacteria. Presumed Archaea were only isolated on CHMcasein medium and were 8/12 isolates at 40 °C.

GSL5

At 18 °C / room temperature, the colonies were predominantly transparent or translucent white forms (10^4-10^5 ml^{-1}), but orange colonies also grew. At 30 °C, 40 °C and 50 °C, on all media where growth occurred, isolates were predominantly orange-pigmented, presumed Archaea. Optimal viable counts were on CHMst, with *c.* 2 x 10^6 ml^{-1} at 30 °C and 40 °C, and 3.8 x 10^4 ml^{-1} at 50 °C. Counts on the CHMcasein plates were 'too numerous to count', but probably exceeded those on CHMst. Again, there was no growth on RVM minimal medium.

3.1.3 Spanish Salterns.

Spain has ancient salt deposits of Upper Triassic-Lower Jurassic, Upper Eocene-Oligocene and Miocene ages. The Triassic salt is diapirized (see Section 1.3.3) and mined conventionally, or by solution-mining. Some 90% of the salt springs of Spain are associated with the Triassic salt, which is dissolved by circulating subterranean waters, and wells may be dug to collect the brine. However, it is the solar salterns, which evaporate salt from seawater, which are the prime source of salt in Spain, and the most important of these are located on the Mediterranean coast around Alicante. (Lefond, 1969).

Samples were collected from three sites in the South of Spain, at the end of December, 1991, by M.-J. Valderrama. Although October to January is usually the wettest time of the year, rainfall had not been noticeably greater than usual prior to sampling; the low salt concentration in some of these salt ponds (Table 3.2) is not, therefore, thought to be due to dilution by recent rainfall (M.-J. Valderrama; personal communication). Properties of these brines are given in Table 3.2.

1. **La Malá (Granada)**: an inland saltern.
In this inland saltern, subterranean saline groundwater is drawn via a well into a series of shallow evaporation ponds, which are operational in the warmer months of the year (Ramos-Cormenzana; 1993).
M1 - well water.
M2 - salt pond 1.
M3 - salt pond 2.
M4 - salt pond 3.

2. **Cabo de Gata (Almeria)**: a solar saltern.
C1 - salt pond.

3. **Roquetas de Mar (Almeria)**: a solar saltern (no longer in use).
R1 - salt pond 1.
R2 - salt pond 2.
R3 - water from a hole in a dried salt pond.
R4 - salt pond 3; salt crystallization pond.

Table 3.2. Properties of the Spanish Brines. In some cases there was insufficient sample to measure ion concentrations. (See also Del Moral *et al.*, 1987a). Viable counts were after 1 month's incubation on CHMst medium. Proportions of 'reds' (presumed halobacteria) are given in **bold**.

	SM1	SM2	SM3	SM4	SC1	SR1	SR2	SR3	SR4
Sample	well water	salt pond 1	salt pond 2	salt pond 3	salt pond	salt pond 1	salt pond 2	salt pond	final pond
pH	7.5-8.0	-	-	7.0	-	7.5	-	7.5	8.0
Total Salts*	17%	18%	11%	11%	16%	1%	1%	9%	40%
M[Na^+]	2.85	-	-	1.8	-	0.195	-	1.1	5.0
M[Cl^-]	2.8	-	-	1.75	-	0.23	-	1.4	6.6
Viable Count ml^{-1} @ 30 °C	10 **0**	7×10^5 **<<1%**	8×10^4 **0**	5×10^4 **0**	5×10^3 **minor**	230 **minor**	700 **>3%**	8×10^4 **0**	2×10^6 **major**
Viable Count ml^{-1} @ 40 °C	0	2×10^4 **major**	1×10^3 **0**	3×10^3 **<1%**	1×10^3 **major**	430 **major**	260 **major**	0	1×10^6 **major**
Viable Count ml^{-1} @ 50 °C	0	4×10^3 **major**	0	0	7×10^3 **>75%**	570 **major**	40 3/4	0	7×10^5 **maj?**

* M.-J. Valderrama (pers. comm.).

Halophilic or halotolerant bacteria were isolated from all the brines (Table 3.2). Halobacteria usually require at least 2M NaCl (11.7 %) for growth and 'reds' (presumed halobacteria) were the majority isolates from the more saline SC1, SM2 and SR4 brines at 40 °C and 50 °C. At 30 °C, Bacteria predominated in all brines except SR4, in which 'reds' again formed the majority; this would be expected in a brine with NaCl at saturation. Unexpectedly, there was only one isolate from a 100 µl inoculum of the SM1 well water, which had an estimated 2.8M NaCl. However, this is in accord with the 1700 colonies l^{-1} found by Del Moral *et al.* (1987b).

SM3, SM4 and SR3 brines have salt concentrations below those normally required for halobacterial growth, and only SM4 produced red colonies (2 out of about 500 colonies at 40 °C), later shown by lipid analysis to be halobacteria (SM4.2). Viable counts for brines SR1 and SR2 were low, which was expected, considering the total salts concentration was estimated at 1 % (Table 3.2). What was highly surprising, however, was the isolation of red-pigmented bacteria, presumably halobacteria, from these brines. This was confirmed by later lipid analyses (isolates SR1.4, SR1.5 and SR2.2) and 16S rRNA gene sequencing (SR1.5).

3.1.4 Wieliczka Salt Mine, Poland.

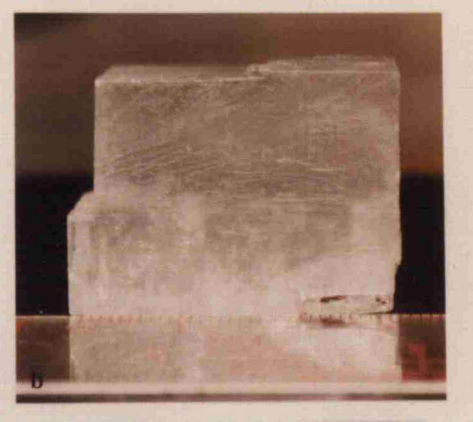
The salt beds at Wieliczka, in southern Poland, were deposited from a sea which lay in a depression between a desert and the Carpathian Mountains, to the South of Poland, about 20 million years ago, in the Miocene period. Within the folded halite strata there are large blocks of salt (the largest about 100 m) which fell into the evaporating sea after removal from the adjacent desert deposits by erosion.

There is archaeological evidence that salt production at Wieliczka goes back to Neolithic times (c. 3500 B.C.), when salt was evaporated from brine ground springs. Rock salt was discovered in the second half of the 13th Century and has been continuously mined since then. Today, there are over 200 km of corridors and 2,040 chambers, in 9 mining levels, from level I at 64 m depth, to level IX at 327 m depth. Wieliczka Salt Mine has been a popular tourist attraction since the 15th Century. The tourist route, on levels I, II and III, takes at least 600,000 visitors a year through galleries and chambers which were cut between the 17th and 20th Centuries, and exhibits underground lakes, and statues and chapels carved out of the rock salt by the miners. The museum on level III (135 m. depth) documents the history of the mining craft and houses maps, art-work and Polish mineral specimens. Of particular interest are specimens of fossilized Miocene wood, pine cones, seeds, magnolia leaves, shells and gastropods, *etc.*, from the salt deposit. The

remaining 6 mining levels are not open to the public, although on the 6th (or 5th) level, at 211 m deep, a sanitorium treats patients with asthma and other lung complaints.

Until about 1890 the salt was mined manually. Then, until 1966, drills and cutters excavated the mine mechanically. Now, the salt is also removed by leaching. In the western part of the Wieliczka deposit, at Barycz, the salt is solution-mined only.
(Information derived from a visit to the mine and museum, with geologist guide Teresa Kuc, 16-6-92; and from Grzesiowski *et al.*, 1992.)

Samples were collected on 18-06-92 by R.T. Gemmell, accompanied by Pawel Barmuta, the mine geologist. From 4 to 5 hours were spent down the worked part of the mine (not open to tourists), collecting samples. Both the older parts of the mine, where galleries were rectangular, of variable widths and supported by wood, and the newer parts, where the galleries resembled rail underground tunnels and were supported by metal, were visited. The different modes of excavation - cut with pick-axes or machine-cut - could be seen. The plasticity of rock salt was much in evidence, with wooden roof supports bent and cracked. Some of this timber was from a Wieliczka forest which had collapsed into the mine chambers. In Leopold Channel, on level III, were 100-year-old wooden channels (wood preserves well in salt) which had been used to transport brine from saturation towers, where rock salt was piled and dissolved with water. Efflorescence was sampled (**PWA**) from the wall of this corridor, near to an inflow (W (Wyciek) III 37).
Several samples, of brine (**PW1**) and crystals, were taken from the lake in the 'Crystal Grotto' (Figure 3.1a) on the upper 2nd level. In the 1960s, the surface collapsed into this chamber, about 70 m. below, resulting in an inflow of Pleistocene water which had to be dammed. The shallow lake which was formed became saturated with salt and, because it was virtually undisturbed, a rich mass of salt crystals formed on the bottom. Pieces of wood were encrusted with salt crystals and some crystals contained inclusions of mud and wood fragments. This lake was contaminated with detritus from the surface and with human urine. The walls of this chamber were clear, **primary Miocene salt crystals** (Pawel Barmuta, pers. comm.) and samples of these (Figure 3.1b) were broken off the wall, adjacent to a deep ledge about 3 feet off the ground. Some samples, which had been cut by miners for souvenirs, were picked from the floor of this ledge. These primary crystals clearly contained fluid inclusions and released a distinctive paraffin-like odour when broken.



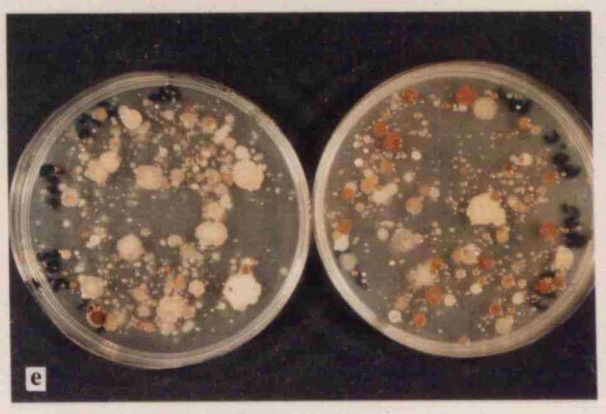


Figure 3.1 **Wieliczka Salt Mine.**

a PW1 - the 'Crystal Grotto';

b primary salt crystal containing fluid inclusions (from the walls of the 'Crystal Grotto' chamber);

c PW4 - inflow brine;

d PW6 - salt stalactites and drips of brine, due to seepage from a leaching chamber above;

e halophile colonies derived from a spread of PW2 brine (10^{-1} dilution) on CHM plates, after incubation at 30 °C (left) and 40 °C (right); note the greater proportion of red-pigmented colonies (presumed halobacteria) after incubation at 40 °C.

First seen about 15 years previously, inflow water from Quaternary rocks above was leaching into the salt deposit and through into the mining levels, from where it had to be pumped out into sumps. One such inflow was sampled (PW2) from its sump (1m x 1m x ?depth) on the 3rd level (W III 25). This inflow brine is utilized in the solution mining process. A second efflorescence sample (PWB), was taken from high on the gallery walls, near an old shaft (Górsko Shaft). This recrystallized salt was a result of quaternary sand waters leaching through.

On level VII (180 m deep?), at the northern border of the salt deposit, is the pumping station, which controls the removal of inflow waters. Three inflows were sampled here, directly at the exit from the pipes, before entry into the tanks. All were from the northern border of the salt deposits, emanating from the surrounding Miocene rocks (contain a lot of water) and had breached the clay layers which coat the salt deposits, but all had different sources. At the time of sampling, the largest inflow (PW3), which was the result of a 1970 inflow from surrounding rocks into the 7th level (W VII 16), was flowing at a rate of 10 cm^3 hr^{-1} and had a salt concentration of only 70 g l^{-1}. The other two samples (PW4 and PW5) were of different inflows, from the 6th level (Z-25 and Z-32, respectively). Figure 3.1c shows the PW4 inflow.

These inflows are utilized for solution mining and account for 50 % of the water used in solution mining here; the remaining 50 % is from the surface. The water is pumped to the leaching chambers to produce fully saturated brine, which is then pumped to the surface for drying. The final sample (PW6) was drips of brine from a salt stalactite (Figure 3.1d), of which there were many, growing from the roof of level VII, and which were due to seepage from a leaching chamber above, between levels VI and VII (the leaching chambers are situated between levels).

Inflows are a perpetual problem at Wieliczka. In the October following this sampling trip, *New Scientist* (3 October 1992; p.7) reported *"threatening to destroy a remarkable feat of engineering ... last month, some 5000 litres of water a minute began pouring into the mine in an area 200 metres below the surface ...* then *... the flow ... had stabilised at between 200 and 300 litres a minute ... contingency plans for evacuation have been posted around the town"*.

Brine was collected in sterile plastic 'universal' tubes. Efflorescences were scraped off the salt walls, with sterile wooden tongue depressors, into sterile 'WhirlPak' bags, which were then sealed. Temperature and pH of the brine samples were measured on-site; conductivity and ionic concentrations were measured in the laboratory (Table 3.3).

(Much of the information about the mine was given en route, by Pawel Barmuta.)

Table 3.3. Properties of the Wieliczka Brines. Viable counts were after 1 month's incubation on CHMst medium. Proportions of 'reds' (presumed halobacteria) are given in bold.

	PW1	PW2	PW3	PW4	PW5	PW6
Site	crystal lake	sump W III 25	inflow W VII 16	inflow Z-25	inflow Z-32	leachate level VI/VII
Temperature	12.0 °C	13.0 °C	16.0 °C	17.0 °C	16.5 °C	-
pH	6.5-7.0	7.0	7.0	6.5-7.0	7.0	-
Conductivity (ms.)	520	500	119	480	290	530
M[Na^+]	5.6	5.4	1.1	5.2	3.1	5.6
M[Cl^-]	6.9	6.1	1.25	5.5	3.35	6.0
M[K^+]	0.0115	0.0135	0.0045	0.00465	0.0085	0.0375
Viable Count ml^{-1} @ 30 °C	3×10^5 **c. 40%**	2.7×10^5 **c. 30%**	20 **0**	190 **2+/19**	150 **1/15**	5×10^5 **c. 30%**
Viable Count ml^{-1} @ 40 °C	6.9×10^4 **>70%**	1.4×10^5 **c. 80%**	0	70 **6/7**	230 **1/23**	2.7×10^5 **c. 70%**

Viable counts (Table 3.3) were in the region of 10^5 ml^{-1} for the crystal lake (PW1), the brine sump (PW2) and the leaching chamber brine (PW6). These were brines which had been in prolonged contact with the salt deposit and a variety of both Bacteria and Archaea were isolated, the latter predominating at 40 °C, as illustrated in Figure 3.1e. In contrast, the inflow brines, PW3, PW4 and PW5, had very low viable counts, which might have been expected as they originated from outside the salt deposit. PW3, the largest inflow, had only about 1M NaCl, and just two bacteria were isolated from a 100 µl sample - it could be presumed that this inflow had had minimal contact with the salt. However, PW4 and PW5 had NaCl concentrations within the optimal range for halobacterial growth and the scarcity of halophilic bacteria is unusual. It is possible that the salt in these brines was solutioned from a more metamorphosed part of the deposit. Alternatively, more halophiles were in the brine, but they were not cultured because they had not had time to recover since the inflow excavated the salt. The 30 °C culture of PW4 brine produced just 19 colonies (some 'tiny') from 100 µl after 1 month, but another look at the plate after a further 2 months (at room temperature) revealed several more (~6) small red colonies (PW4.4), later identified by polar lipid analysis as halobacteria. The efflorescence sample PWA had halophile counts of 5 x 10^5 g^{-1} at 30 °C and approximately 10^3 g^{-1} at 40 °C; the proportion of halobacteria was unclear at 30 °C, but they predominated at 40 °C. Sample PWB had counts in the region of 10^5 g^{-1} at 30 °C and 40 °C, with 63 % and 96 % halobacteria, respectively.

3.1.5 Green River Basin, Wyoming.

The Green River evaporite deposit is the product of Gosiute Lake, a freshwater sea that existed for about 4 million years in the Eocene period and extended to a maximum of about 15,500 square miles. Evaporative concentration of the lakewaters during arid cycles allowed the deposition of extensive beds of trona and mixed trona-halite (Bradley and Eugster, 1969). This 50 million year old deposit is mined for trona by continuous mining in a room and pillar method, at some 500 m depth (*General Chemical* advertisement). Salt is usually mixed with trona, although there are some areas of near pure salt a few inches thick (Lefond, 1969).

One sample of brine, collected in October, 1994, was received (26-10-94) from the Tg trona mine at Granger, Wyoming. Brine from this mine is sold to nearby coal-fired power plants for use in emissions control (information provided with sample). This brine had a distinctive oily odour and, as would be expected from a trona ($Na_2CO_3.NaHCO_3.2\ H_2O$) mine, was alkaline, with a pH of 10.5. Ionic measurements were as follows: 2.4M Na^+, 0.58M Cl^- and 0.0057M K^+. These results imply a maximum [NaCl] of 0.58M.

Direct observation of this brine, using phase-contrast microscopy with x500 magnification, revealed a mixed bacterial community of spirilla (motile), rods (different forms, some motile), and triangular forms.

Viable counts on HALK medium, after 2 weeks, were 5.27 x 10^3 at 30 °C and 8.7 x 10^3 at 40 °C, with the proportion of red-pigmented colonies (presumptive halobacteria) approximately 3% and 20%, respectively. Sub-cultures of these red-pigmented colonies showed them to be rods or pleomorphic bacteria. The comparatively low counts suggests that a significant number were not isolated.

3.1.6 Khorat Basin, Thailand.

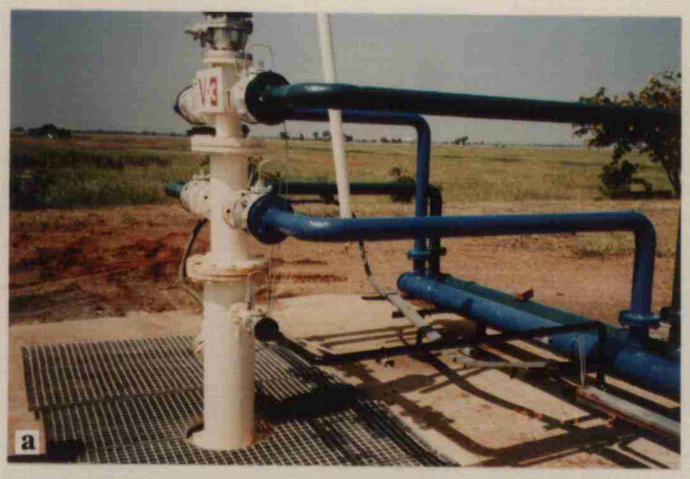
Samples are from the Khorat Basin, in N.E. Thailand. The salt deposit here is of Upper Cretaceous-Lower Tertiary age and originated as a shallow, closed marine basin. Local uplifting resulted in flow of the plastic salt beds into smooth folds, but there is no evidence of faulting as a result. With continued uplifting, the salt anticlines developed into domes. (Suwanich; *c.*1992).

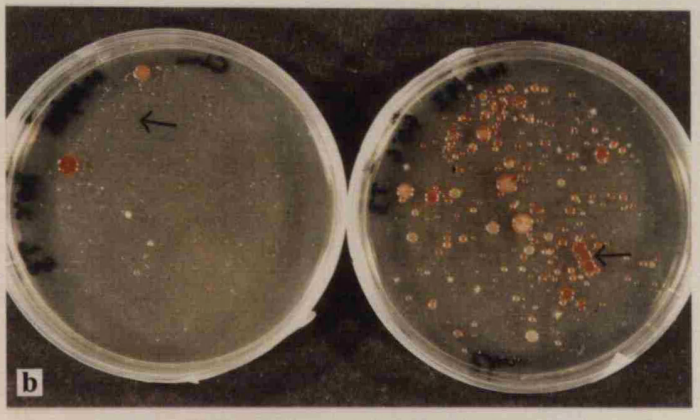
The V3 well, from which brine was sampled (Figure 3.2a), mines bedded salt from about 270 m depth. The plant was opened in 1987 and salt production began in 1989. In the solution-mining process, water is heated to 105 °C (steam) and chlorine is injected into the water before it enters the well head (V3). The water entering the well head is therefore sterile, and by this stage it has cooled to 60 °C (therefore pasteurised); the water temperature equilibrates within the cavern. An oil blanket on the water surface contains it, and prevents solution of the overlying salt. There is 150,000 m^3 brine underground at any one time and it is pumped out at 50 m^3 per hour, giving a residence time of about 125 days. The saturated brine exits at a temperature of about 40 °C.

All samples (see Table 3.4 for properties) were collected on 2-11-92 by W.D. Grant.

T1 and T2 are brine samples taken from the sedimentation tank, about one mile from the well head. The brine was piped from the well head to the sedimentation tank, where the oil was allowed to settle out. The tank was uncovered.

T3 and T4 were brines sampled directly from a tap at the well head, at the top of the V3 rig (Figure 3.2a).





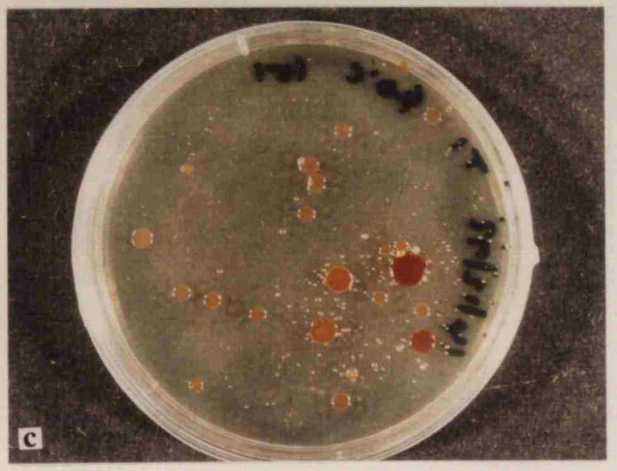


Figure 3.2 **Khorat Basin, Thailand.**

a V3 rig well head - T3 and T4 brine samples were taken from here;

b halophile colonies from a spread of T3 brine on CHM plates, after incubation at 30 °C (left) and 40 °C (right); tiny white colonies indicated on the 30 °C plate were probably the unidentified isolate T3.1; orange colonies indicated on the 40 °C plate were the *Haloarcula*-like isolate T3.5;

c halophile colonies from a spread of T1 brine (10^{-1} dilution) on CHM plates, after incubation at 40 °C.

(Red-pigmented colonies are presumed halobacteria).

Table 3.4. Properties of the Thai Brines. Except for the temperature, measurements were made in the laboratory. Viable counts were after 1 month's incubation on CHMst medium. Proportions of 'reds' (presumed halobacteria) are given in bold. ND = Not Determined.

	T1	T2	T3	T4	T5	T_{208}
Source on Site	sediment-ation tank	sediment-ation tank	well head	well head	disused saltern	disused saltern
Temperature (on-site)	(~30 °C)	(~30 °C)	40 °C	40 °C	28 °C	30 °C
pH	7.0	7.0	7.0	7.0	7.0	7.5
Conductivity (ms.)	500	493	490	500	71.8	211
M[Na^+]	5.7	5.5	5 .5	5.5	0.7	2.4
M[Cl^-]	6.3	6.3	6.1	6.0	0.76	2.4
M[K^+]	0.0043	0.0028	0.0028	0.0033	0.0020	0.0047
Viable Count ml^{-1} @ 30 °C	8.9×10^3 **ND**	9.1×10^3 **ND**	2.5×10^4 **ND**	1.7×10^3 **ND**	2.0×10^4 **~12%**	773 **~15%**
Viable Count ml^{-1} @ 40 °C	1.3×10^5 **ND**	1.4×10^5 **ND**	4.2×10^5 **ND**	$7.7. \times 10^4$ **ND**	1.1×10^4 **10-11%**	450 **>70%**

T5 and **T208** were brine samples, each from a different disused saltern (disused for ~10 years). Salterns **T5** and **T208** were about 2 miles from the well head and only ~10 m. apart, with a walkway between. A soil sample, **Ts**, was taken from beside saltern **T5**. These salterns have never gone red, as the brine does not contain sufficient nutrients to sustain prolific halobacterial growth.

Halophilic bacteria were isolated from all the brines. Viable counts were 10^3-10^4 ml^{-1} at 30 °C and 10^4-10^5 ml^{-1} at 40 °C, except for the T_{208} brine, which had counts only in the hundreds. The saltern brines, T5 and T_{208}, were undersaturated with NaCl (0.7M and 2.4M respectively) and the lower viable counts for T_{208} were therefore expected. However, T_{208} did harbour halobacteria (T_{208}.9 and T_{208}.10 were about 5 % of the isolates), with the usual higher counts at 40 °C. The comparatively high counts for T5 could be due to a high proportion of halotolerant species in this brine, although halobacteria were not precluded, as some 12 % of these isolates were thought to be 'reds', and later confirmed as such by polar lipid analysis (T5.3, T5.5, T5.6 and T5.7) and 16S rRNA gene sequencing (T5.7). The proportion of putative halobacteria in the sedimentation tank and well head brines was difficult to establish as the majority of the colonies were very tiny and of uncertain colour (Figures 3.2b,c). Some of these tiny colonies, not originally thought to be 'reds', were later identified as halobacteria (T2.1 and T3.1). The soil sample, which gave viable counts of about 2 x 10^5 g^{-1} at 30 °C, and 2.6 x 10^5 g^{-1} at 40 °C, produced about 7 % and 12 % 'reds', respectively.

3.1.7 Winsford Rock Salt Mine, Cheshire.

The 235-240 Ma Triassic (Anisian stage) salt deposits of the Cheshire Basin are saliferous beds within the Keuper Marl Group. The Lower Keuper Saliferous Beds, known as the Northwich Halite, are 290 m. thick; the Upper Keuper Saliferous Beds, known as the Wilksley Halite, are up to 400 m thick. The Northwich Halite at Winsford has nine zones, of salt interspersed with marl, and only two of these (zones B and F) have so far been mined; (Longley-Cook, 1989). By volume, 24 % is clay and silt; gypsum occurs sporadically; calcium and magnesium carbonates are in trace amounts (Earp and Taylor, 1986). Evans (1970) interprets these Triassic salt beds as products not of a desiccating desert environment, but of sub-aqueous deposition from supersaturated bottom brine in basins, in a low-lying land flooded by a shallow extension of the ocean. There is evidence (Arthurton, 1973) that these layered Northwich halite beds were produced by the upward and lateral growth of crystals which nucleated on the floor of the shallow saline pools. Polygons on the bedding surfaces of the halite, seen on the roof of the mine, are the result

of thermal contraction and expansion, and so indicate long periods of exposure of the precipitated salt (Tucker and Tucker, 1981).
The Northwich Halite is overlain unconformably by Pleistocene glacial drift, which is porous and has allowed the percolation of groundwater. This has dissolved the surface of the salt bed sub-outcrop to produce subterranean brine lakes, which lay on the impervious zones of rocksalt and marl. This saturated brine protects the salt from further dissolution, except where large-scale artificial brine removal (*i.e.* by pumping) has led to replacement by fresh water and therefore increased solution of the salt (Earp and Taylor, 1986). For centuries, salt has been produced from this natural brine ('wild brine'), originally from springs and later from wells in the Cheshire Basin. Artificial brine is produced by solution-mining at Holford, east of Northwich (Longley-Cook, 1989).

Rock salt has been mined at Winsford Rock Salt Mine, in Cheshire, N.W. England, since 1844 (except for a short period of closure from 1892 to 1928). Salt is extracted by cutting and blasting in the 'room and pillar' method, with the salt pillars (24 m x 24 m) left uncut to support the mine. The rock fragments are crushed underground before the loads are hoisted to the surface; (Longley-Cook, 1989). The rock salt is piled on the surface and covered with potassium ferrocyanide, as an anti-caking agent. It is sold mainly as 'grit' for deicing roads.

Winsford Mine is ventilated forcedly by fans which draw air from the surface at 160.5 m^3 sec^{-1}.; these are situated near the base of No. 4 shaft, 183 m deep; (G.Hall, 1987; information sheet). In the summer, when this air is particularly moist, it is directed into dewatering tunnels, where the moisture condenses on the cool walls and the dried air is then directed to the mining faces (Longley-Cook, 1989). This condensate forms brine pools on the floor. Such brine pools, and salt efflorescences produced by recrystallization of salt dissolved from the tunnel walls by the condensing moisture, were sampled at Winsford.

Brine and salt samples were collected from three sites in Zone B of Winsford salt mine, on 20-11-91, by R.T. Gemmell. Also present were W.D. Grant, T.J. McGenity, M.-J. Valderrama, and a Winsford geologist. The sites were located in the dewatering tunnels at B4 (thought to be Bostock 4 panel; **WB4** samples), Moulton Cross (Moulton 2 panel; **WM** samples) and B3 (thought to be Bostock 3 panel; **WB3** samples).

Figure 3.3 **Winsford Rock Salt Mine.**

a sampling salt efflorescence beside the B4 brine pool;

b Moulton Cross brine pool, with mats of halite crystals on the surface;

c sampling brine from a B3 brine pool;

d piles of mined salt on the surface - samples were taken from the salt piles and the nearby puddles.

The shallow brine pools at B4, Moulton Cross and B3 were a consequence of the accumulation of condensed moisture from fan-driven air. The B4 brine pool (~ 4 m x 2 m) was sampled for brine (**WB4br**) and 'sediment' (bottom crystalline sludge, sampled anaerobically; **WB4S**), and efflorescence was collected from the wall adjacent to the pool (**WB4E**; Figure 3.3a). The Moulton Cross brine pool was of greater area than the other two pools (~ 20 m x 7 m), with localized deeper parts. Salt crystals had aggregated as typical mats (Arthurton, 1973) on the surface of the pool (Figure 3.3b), and these were sampled (**WM2C**), as well as brine from different parts of the pool (**WM2br1** and **WM2br2**), 'sediment' (**WM2S**, collected anaerobically) and efflorescence from beside the pool (**WM2E**). The smallest brine pools, at B3 (~2 m x 1 m), were more recent (~ 2 months), formed when the ventillation fans were moved. Efflorescence (**WB3E**) was sampled from beside the pools and brine (**WB3br**) taken from one of them (Figure 3.3c). A sample of crushed salt was taken from the salt piles on the surface (**WSP**) and brine from a nearby puddle (**WSPbr**); (Figure 3.3d).

Brine was collected in sterile plastic 'universal' tubes. Efflorescences were scraped off the salt walls with sterile wooden tongue depressors, into sterile 'WhirlPak' bags, which were then sealed. Temperature and pH of the brine samples were measured on-site; conductivity and ionic concentrations were measured in the laboratory (Table 3.5).

Both halophilic Bacteria and halobacteria were isolated from all brine and salt samples. The viable counts on CHMst medium at 30 °C, 40 °C and 50 °C, for the mine brines, and the proportion of presumed halobacteria, are shown in Table 3.5. The viable counts for Bacteria exceeded those of halobacteria, but fell with increase in temperature, whereas halobacterial numbers increased with increase in temperature. The brine sample taken from a puddle near the salt pile (WSPbr) had much lower viable counts, of less than 10^3 ml^{-1}, and most of the isolates at 30 °C, and all at 40 °C and 50 °C, appeared to be halobacteria. Presumably this puddle had had minimal contact with the salt. A comparison with the counts for the salt pile sample (Table 3.6) suggests that halobacteria (and halophilic Bacteria) are not prolific on the surface.

The viable counts for the solid salt samples (Table 3.6) showed the same trend for decreased numbers with increase in temperature, but tended to be higher than the corresponding counts for the brines from the same sites. As the efflorescence, 'sediment' and crystalline sludge samples are the result of salt recrystallization, it may be that the bacteria 'co-precipitated' with the salt and so become 'concentrated'. The proportions of halobacteria again increased at 40 °C, compared to 30 °C, but the trend does not appear to continue at 50 °C with most samples.

Table 3.5. Properties of the Winsford Brines. Viable counts were after 1 month's incubation on CHMst medium. Proportions of 'reds' (presumed halobacteria) are given in **bold**. ND = Not Determined.

	WB4br	WM2br1	WM2br2	WB3br	WSPbr
Temperature	13.5 °C	12.0 °C	12.0 °C	14.0 °C	-
pH	7.2	7.0	7.0	7.1	-
Conductivity (ms.)	-	315.0	315.0	316.2	-
M[Na^+]	5.0	5.0	5.0	5.0	-
M[Cl]	6.0	6.2	6.2	7.3	-
M[K^+]	0.051	0.0305	0.029	0.047	-
Viable Count	2.5×10^5	1.5×10^5	1.3×10^5	1.4×10^5	960
ml^{-1} @ 30 °C	**<10%**	**0**	**<3%**	ND	**?majority**
Viable Count	1.6×10^5	2.7×10^4	1.6×10^4	1.3×10^4	760
ml^{-1} @ 40 °C	**majority**	**<30%**	**~44%**	**~15%**	**100%**
Viable Count	1.7×10^4	6.4×10^3	1.3×10^4	70	480
ml^{-1} @ 50 °C	**100%**	**>95%**	**majority**	**>80%**	**100%**

Table 3.6. Viable Counts for Winsford Salt Samples. Viable counts were after 1 month's incubation on CHMst medium. Proportions of 'reds' (presumed halobacteria) are given in **bold**. ND = Not Determined.

	WB4E	WM2S	WM2E	WM2C	WB3E	WSP
Viable Count g^{-1} @ 30 °C	5.1×10^5 **'many'**	4.2×10^5 ND	3.2×10^5 **'few'**	1.9×10^5 **'few'**	4.6×10^5 **many**	3.5×10^5 **<3%**
Viable Count g^{-1} @ 40 °C	2.4×10^5 **majority**	2.7×10^5 **majority**	1.9×10^5 **?majority**	9.5×10^4 **majority**	3.4×10^5 **majority**	2.3×10^4 **majority**
Viable Count g^{-1} @ 50 °C	9.7×10^4 **?majority**	4.9×10^4 **~6%**	6.4×10^4 **~19%**	1.5×10^4 ND	1.8×10^4 **~2%**	6.0×10^3 **majority**

3.1.8 Boulby Potash Mine, Cleveland.

Evaporite beds of the third and fourth cycles (EZ3 and EZ4) of English Zechstein deposition in the Upper Permian (250-260 million years ago) occur in the Boulby area, in Cleveland, N.E. England. Boulby Halite and Boulby Potash of EZ3 are mined at Boulby. Evidence suggests (Smith, 1973) that these deposits formed not within a barred basin, the model often applied to Zechstein deposition, but on extensive salt flats on the margins of the diminishing Zechstein Sea. The Boulby Potash sylvinite (sylvite + halite) bed has been subjected to two stages of diagenesis (Woods and Powell, 1979): (i) upward brine migration, probably provided by the conversion of gypsum to anhydrite, causing almost complete recrystallization, probably of carnallite to produce sylvite, and (ii) significant plastic flow of the sylvite when the evaporites were at their maximum depth, undulating the top of the bed. Magnesium minerals are virtually absent in the Boulby Potash. The Boulby Halite bed has suffered very little mineralogical change, and is comparatively undisturbed, except over fault zones. A 'pink halite band' near the top of the halite bed contains a calcium and magnesium chloride brine which seeps from the rock when it is breached. It also contains occasional lenses, up to 2 m in diameter, of pure translucent white halite overlying a band of anhydrite which are presumably due to the solution activity of calcium sulphate brines (Woods and Powell, 1979).

Boulby Potash Mine (Cleveland Potash Limited) has been operational since 1974, following earlier exploratory boreholes. Boulby Potash is situated at about 1,050 m depth and is mined by cutting and blasting roadways, leaving pillars for roof support. The underlying Boulby Halite is also cut where the potash seam is thin, and in addition has been mined to carry roadways for transport and ventilation. The potash is hoisted to the surface where it is crushed and the KCl concentrated by flotation in brine. Halite is kept separate from the potash, for sale as road salt. Processing waste (mostly salt, with insoluble minerals) is discharged into the sea. (From information about Boulby Mine, provided by C.W. Mackenzie of Cleveland Potash Limited, to T. Morse, 15-11-84.)

1) Samples were taken from eight sites within Boulby Mine, by R.T. Gemmell on 3-12-91. Also present were W.D. Grant, T.J. McGenity and Boulby geologist G. M. Jones.

Boulby Mine has brine inflows of various sources. The first sampling site (**B1**) was a sump collecting piped-in inflow brine from 120 panel; (Figure 3.4a). The inflow (about 10 gal. $min.^{-1}$), which was believed to be residual brine, had flowed along the salt floor, in contact also with potash and carnallite. Another residual brine inflow may also have mixed with this. Brine samples were taken from the pipe exit (**B1bra**) and from within the sump (**B1brb**, anaerobically, and **B1brc**). Efflorescence (**B1E**) from the side of the

rock wall behind the sump, and a stalactite (**B1st**) at the drop to the sump, were also sampled here. Fresh water was trickle-fed to this sump, to prevent brine crystallization. About 600 metres away, on the 202 panel (**B2**), was a brine pool fed primarily by the 120 residual inflow and also by an inflow of low flow rate (about 1 gal. $min.^{-1}$) and very good carnallite contact from a junction of the 202 panel. The brine was periodically pumped out of this sump. Samples of brine (**B2br**) and 'sediment' (**B2S**; anaerobically) were collected. **B1** and **B2** brines were expected to have a high $[Mg^{2+}]$.

There are two shafts into the Boulby Potash deposits: No.1 is the rock shaft, for hoisting rock and exhausting stale air; No.2 is the Man shaft, for transporting personnel and for fresh air intake to the mine. The shafts are sunk through beds of boulder clay, Jurassic shales, Triassic Keuper marl, Bunter sandstone and Bunter mudstone, and Permian mudstone, to reach the evaporites. As the Bunter sandstone contains brine at hydrostatic pressure, the shafts needed to be sealed (No.1 with steel and No.2 with concrete) to prevent inflow (C.W. Mackenzie to T. Morse; see earlier). The Bunter brine is not of seawater composition, and is 5 % NaCl at the top, about 13 % NaCl in the middle and 17-18 % NaCl near the bottom. As the shaft seal is imperfect the Bunter brine weeps into the mine from above the evaporites, with, therefore, little evaporite contact. This brine (sample **B3br**) was drained from the rock shaft, at West Link. On the adjacent rock wall was a curtain of stalactites (sample **B3st**), the 'Rock Shaft Garland' (Figure 3.4b). This location also housed a dump hole for the D lines (by D30).

Three brine samples were taken from tanks at the Wirth Pump Lodge: **B4br**, **B5br** and **B6br**. These tanks were constantly emptying and filling, and contamination from other sources should not be discounted. Sample **B4br** was from a tap in a tank collecting brine pumped from the 113 panel sump. The 113 inflow had the highest flow rate, of about 270 gal./min. and its source was thought to be the Bunter sandstone. It was believed that the Bunter brine had entered the EZ4 Upper Halite through faults in the Bunter and Permian mudstones, and solution-mined a natural vertical shaft through, and into the EZ3 Boulby Potash and Boulby Halite. As a result, the brine was now in free-fall from the Bunter sandstone to the halite, with little halite or potash contact. 19 % NaCl and <1 % KCl (undersaturation) were expected in this sample.
Sample **B5br** was from a tank receiving sump water from the 'D-lines' (D13). An unsaturated inflow of about 50 gal./min. had flowed into and pooled in the 114 panel. Contact was predominantly with halite and NaCl saturation was expected; there was some contact with potash and 4-5 % KCl was expected. The inflow was carried from the panel to the tanks in the pump lodge via two drilled boreholes (*i.e.* the **D lines**; thought to be D13 and D11), which were lined to prevent damage to the strata.

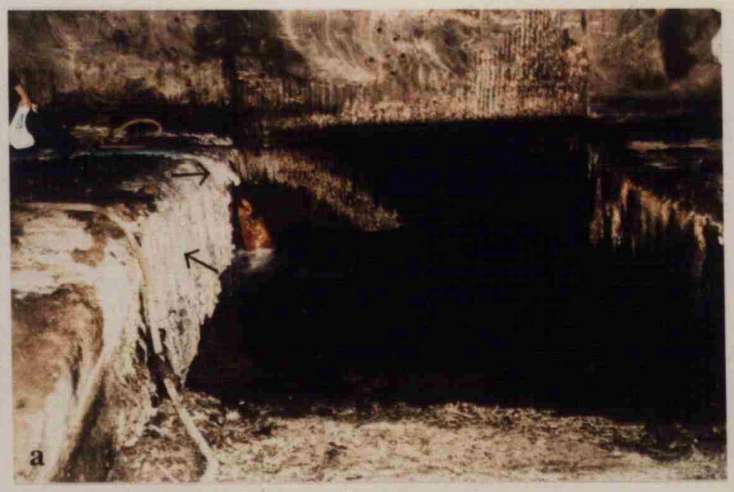


Figure 3.4 **Boulby Potash Mine.**

a B1 - the 120 panel sump - top arrow indicates inlet pipe, bottom arrow indicates efflorescence on the sump wall;

b B3 - 'Rock Shaft Garland' salt stalactites; arrow indicates brine pipeline;

c B8 - sampling brine from a tap on the South inflow brine pipeline;

d 325 panel - brine inflow seeping along the bedding plane; efflorescence was collected.

B6br brine was from inside a tank collecting sump water from the flooded 302 panel, in the S.E. of the mine. The total panel inflow here was 60 gal.min.$^{-1}$ and had good halite contact, including dissolution of the Upper Halite rock. The brine was thought to be fully saturated with respect to both NaCl and KCl (12-13 % expected, but may be lowered in the cooler pump lodge because of precipitation), with low Ca^{2+} and Mg^{2+}.

To the south of the Wirth Pump Lodge, at the South Sump, brine samples were taken from taps on the pipelines draining the West inflow brine (**B7br**) and the South inflow brine (**B8br**; Figure 3.4c). The West brine was expected to be saturated with NaCl (~ 24 %) and have a fairly low [KCl]. In the West workings brine collects in the potash panels and eventually overflows onto the salt roadways, where it is collected by the first drainhole, D28. The second drainhole, D29, collects brine which has tipped into the potash panel. Shaft sump water, used for washing, is also pumped to the West workings, and sometimes water from 113 sump and the D lines. This brine is therefore a mixture of residual brine, Bunter brine from Rock Shaft Garland and washing water. The South brine was from a large flooded area and, because it was piped out from the top, least dense layers, was expected to be undersaturated, with about 9 % NaCl and 2 % KCl.

Brine was collected in sterile plastic 'universal' tubes. Efflorescences were scraped off the salt walls with sterile wooden tongue depressors, and stalactites broken off; these were put into sterile 'WhirlPak' bags, which were then sealed.

2) On 23-12-92, brine samples were collected by R.T. Gemmell and W.D. Grant, accompanied by Boulby geologist R. Holmes. We had been informed by R. Holmes that the mine had a new brine inflow (350 panel) in the S.E. of the mine, and there was a possibility that it was residual. Before our arrival, a second inflow had occurred in a new panel (325). Both brine flows were sampled.

The **350** panel had been mined the 1st week of June, at about 1200 m depth, in the Boulby Potash, and the brine inflow had started 19-10-92, close to its source, trickling in in several places. Individual flows were less than 0.25 gal. min.$^{-1}$; the total was 1-2 gal. min.$^{-1}$ The chemistry of the brine had changed little, suggesting it could be residual (Bunter brine chemistry changes). Brine drips, with a flow of more than 1 gal. min.$^{-1}$, were collected (**B350A**) from the wall of panel 350. A second sample was taken from ~50 m further S.E. - drips of brine from a salt stalactite (**B350B**).

Table 3.7. Properties of the Boulby Brines. Viable counts (VC) were after 1 month's incubation; proportions of 'reds' (presumed halobacteria) are given in bold. 10KCl = CHM with 10% (w/v) KCl + 10% (w/v) NaCl. ND = Not Determined. gpm = gallons per minute.

	B1(a)	B2	B3	B4	B5	B6	B7	B8	B350A
Flow Rate (gpm) *	10	1	-	270	50	60	75	175	1-2
T °C	32.5	34.0	30.4	31.7	32.7	33.6	35	33.9	28.3
pH	6.6	7.6	6.2	6.03	6.01	7.0	6.0	6.25	5.1
% DO_2	48	19	33	12	5	2	11	5	-
Density g/cm^3 *	1.2278	1.2206	1.1210	1.1708	1.1742	1.2194	1.2091	1.2160	1.2299
% w/w KCl *	9.9 (1.3M)	9.6 (1.3M)	0.1 (0.01M)	0.2 (0.03M)	0.5 (0.07M)	9.3 (1.2M)	6.0 (0.8M)	8.1 (1.1M)	10.7 (1.4M)
% w/w NaCl *	14.0 (2.4M)	14.0 (2.4M)	17.2 (2.9M)	23.4 (4.0M)	23.8 (4.1M)	20.7 (3.5M)	22.4 (3.8M)	21.2 (3.6M)	17.4 (3.0M)
% Ca^{2+} *	0.19	0.19	0.18	0.14	0.14	0.13	0.15	0.15	0.55
% Mg^{2+} *	1.64	1.56	0.08	0.07	0.07	0.17	0.17	0.19	0.60
Conductivity (ms.)	327	339	221	295	306	363	367	375	530
M[Na^+]	4.65	4.50	3.55	4.75	4.80	4.75	4.80	4.75	4.60
M[Cl^-]	7.8	8.0	3.4	7.8	7.8	6.6	6.6	6.6	10.5
M[K^+]	1.450	1.350	0.017	0.056	0.100	1.500	1.100	1.450	1.800
VC ml^{-1} 30°C CHMst	2.9 x 10^5 **major**	8.0 x 10^5 **~100%**	990 **> 50%**	1.2 x 10^4 **~100%**	7.8 x 10^3 **~100%**	7.7 x 10^3 **major**	140 **100%**	40 **100%**	0
VC ml^{-1} 40°C CHMst	4.6 x 10^5 **major**	6.8 x 10^5 **~100%**	1.2 x 10^3 **> 64%**	9.5 x 10^3 **major**	7.3 x 10^3 **~100%**	8.2 x 10^3 **~100 %**	890 **100%**	660 **~100 %**	0
VC ml^{-1} 50°C CHMst	3.9 x 10^5 **~100%**	4.5 x 10^5 **~100%**	670 **> 94%**	1.9 x 10^3 **~100%**	1.8 x 10^3 **100%**	1.3 x 10^3 **100%**	120 **100%**	120 **100%**	0
VC ml^{-1} 30°C 10KCl	6.6 x 10^4 **major**	1.3 x 10^5 **major**	140 3/14	1.0 x 10^3 **~100%**	1.3 x 10^3 **~100%**	780 **~99%**	20 2/2	40 4/4	-
VC ml^{-1} 40°C 10KCl	1.5 x 10^5 ND	2.8 x 10^5 **major**	260 2/26	1.8 x 10^3 **major**	1.1 x 10^3 **>82%**	790 **major**	0	50 4/5	-

* Information provided by Boulby Mine. B350A - analysis of sample of 21-12-92.

Table 3.8. Viable Counts for B1brc Brine and Boulby Salt Samples. Viable counts (VC) were after 1 month's incubation; proportions of 'reds' (presumed halobacteria) are given in bold. ND = Not Determined.

	B1brc	B1E	B1st	B2S	B3st
VC ml^{-1}/g^{-1}	8.4×10^6	2×10^5	7.4×10^5	2.4×10^6	3.2×10^3
@	**100%**	**100%**	**100%**	**100%**	ND
30°C;CHMst					
VC ml^{-1}/g^{-1}	6.5×10^6	7.4×10^4	1.5×10^6	1.9×10^6	7.2×10^3
@	**100%**	**100%**	**~100%**	**majority**	**>39%**
40°C;CHMst					
VC ml^{-1}/g^{-1}	4.9×10^6	1.7×10^4	1.1×10^6	2.2×10^6	1.6×10^3
@	**100%**	**majority**	**100%**	ND	**75%**
50°C;CHMst					

The brine inflow into **325** panel, on the salt access road, was directly from the salt (below the potash), and thought to be residual. Brine was seeping along the bedding planes (Figure 3.4d) and had appeared almost immediately following mining. Sample **B325A** was damp efflorescence on the walls, which had been mined only the previous weekend. A second efflorescence sample (**B325B**) was taken on the same roadway where it had been mined a little earlier, in the first week of December. Sample **B350A** was collected in a sterile Durans bottle; the remainder in sterile plastic 'universal' tubes.

Temperatures, DO_2 and pH values of the brines were measured on-site; conductivities and ionic concentrations were measured in the laboratory; additional analysis results were provided by Boulby Mine* (Table 3.7).

Halobacteria were isolated from all samples from Boulby, except the B350 and B325 samples, which appeared to be devoid of bacteria. Viable counts (Tables 3.7 and 3.8) were highest (generally 10^5 to 10^6 ml^{-1} / g^{-1}) for samples from the B1 and B2 sites, intermediate (10^3 to 10^4 ml^{-1}) for sites B4, B5 and B6, and low (10^3 and less) for B3, B7 and B8. There was no significant change in count with change in incubation temperature, probably because the isolates were predominantly, if not exclusively, halobacteria, which can grow at 30 °C and 50 °C. The counts do not bear any relationship to the % w/w NaCl values given by Boulby after this visit, although the low counts and low [NaCl] for B3 brine were expected for a Bunter brine with little halite contact. However, B4 brine was also believed to derive from the Bunter sandstone, with little halite contact, yet it had about 4M NaCl and intermediate bacterial counts. B7 brine had a high [NaCl] and was from a number of sources, yet the bacterial count was particularly low. The efficieny of halophile isolation may bear more relation to the precise source of the salt in the brines; regions which had been exposed to metamorphic events would be less likely to harbour many bacteria.

High K^+ CHM media were also used for isolation, as the brines from B1, B2, B6, B7 and B8 had a high K^+ content. (B350A had a high K^+ content, but as no bacteria were isolated on CHM, high K^+ medium was not tried.) The KCl partially replaced NaCl in the CHM medium, so that in a 10% KCl (1.34M) medium there was 10 % NaCl (1.71M) and in a 18 % KCl (2.4M) medium there was 2 % NaCl (0.34M).
Viable counts on the 10 % KCl medium were lower than, but remained proportional to, those on CHMst (Table 3.7). Halobacteria were usually the predominant isolates on 10 % KCl medium. The isolate B1bra.5, which comprised about 10 % of the isolates on such medium, was included in phylogenetic analysis.

Only B1brc brine produced colonies on 18 % KCl, with 650 ml^{-1} at 30 °C and 400 ml^{-1} at 40 °C. Selected isolates were studied in more detail, for growth on KCl. Isolates B7br.21 (provided by T.J. McGenity), B1brc.4, B1brc.40 and B1brc.42 were able to grow on 0-25 % KCl, in the absence of NaCl, but also on 0-20 % NaCl, in the absence of KCl. It was not clear what the optimal KCl and NaCl concentrations were, but these isolates, which were identified as Bacteria, could be categorized as both haloversatile (James *et al.*, 1990) and 'kaloversatile'. When several culture collection strains of halophilic Bacteria were compared, only *Micrococcus halobius* could grow in greater than 10 % KCl. The salt ranges for growth of this organism were 5-25 % KCl in the absence of NaCl and, equally, 5-25 % NaCl in the absence of KCl. Venkataraman and Sreenivasan (1954) described an isolate (H_5), identified as a *Micrococcus* sp., which could grow in a medium where the NaCl has been replaced with KCl.

3.1.9 Inowroclaw, Poland.

There are 3 salt mines at Inowroclaw, towards the N.W. of Poland, each exploiting domes of the Zechstein basin: Solno, Góra and Mogilno. Brines were obtained from Solno, above which the town of Inowroclaw lies, and from Góra.

The Solno salt dome diapirized from a Zechstein 3 (PZ3) salt bed 5,000 metres thick, to lie 150 metres below the surface, 2,500 x 1,200 x 5,000 metres in dimension. Mining began at Solno in 1870, with the extraction of salt from the upper layers of the dome, at 160-180 metres in depth. At the beginning of the 20th Century, this mine was flooded by accident, by water from cavities in the gypsum cap of the dome, causing the land above to sink. The second mining operation was to a depth of 470 metres, 300 metres below the old mine, when the first shaft (of three) was built, in 1924-1929, to this first (of ten) mining level. The tenth level was at 637 metres. These mining levels were excavated into galleries and chambers, solution-mined from below and the brine pumped out. However, brine from the upper (old) mining level and the gypsum cavities was threatening to empty into the middle mining levels at about 470 metres, which would have caused collapse from above and put the town in danger of sinking. To prevent this happening, there was controlled flooding of the mine in 1986 and 1991, with brine from the Góra mine, and Solno is no longer in operation.

The salt dome at Góra, 10 kilometres from Inowroclaw, is diapirized from the same layer of salt as the Solno dome, but is smaller, with a circular cross-section and diameter of about 900 metres. Exploitation of the Góra dome began at the end of the 1960's. This mine is entirely solution-mined from the surface, with water being pumped through a well (18 5/8" diameter?) into the dome, and saturated brine (310 g l^{-1}) forced out via a central

tubing. Brines are transported from the well head by pipeline, at a rate of about 50 m^3 h^{-1}, to Inowroclaw and to industries in surrounding towns. Góra brine, with its low [Mg^{2+}], is favoured by industry (*e.g.* for soda and for Cl_2; a local PVC industry utilizes Cl_2 from these brines). (Information provided in a meeting with Dyrektor Misterski of Inowroclawska Kopalnia Soli; 23-6-92.)

Samples were taken of Solno and Góra brines, directly from the pipeline, by mine personnel, on 23-06-92. Information about the concentrations of ions in these brines was provided by Dyrektor Misterski; Na^+, K^+ and Cl^- concentrations were also measured in the laboratory (Table 3.9).

Viable counts were in the region of 10^5 ml^{-1} for both the Solno and Góra brines, but whereas the proportion of 'reds' in the less saline Solno brine (*c.* 2M NaCl) was negligible, presumed halobacteria were of significant numbers in the salt-saturated Góra brine and, as expected, greater at 40 °C (Table 3.9).

As the Solno brine had a high K^+ content, as for the Boulby samples, CHM medium with a high KCl concentration was also used for bacterial isolation. With the Solno brine, viable counts on CHM with 10 % (w/v) KCl were virtually identical to those on standard CHM and were only a little lower on 18 % (w/v) KCl. The colony forms also appeared to be the same: predominantly cream-coloured, and no halobacteria were isolated. This suggested that the Solno Bacteria could grow in a wide range of both NaCl and KCl concentrations. Individual isolates were able to grow on CHM media where the 20 % (w/v) salts composition ranged from 0 % KCl + 20 % NaCl to 18 % KCl + 2 % NaCl. Three isolates (PIS11, 12, 13) were studied in more detail and shown to grow in a KCl range (no NaCl) of 0-10 % (optimally 5 % KCl), and in a NaCl range (no KCl) of 0-25 % (optimally 5-10 % NaCl). These Bacteria do not fall neatly into a category of halophilicity (Kushner, 1978), but as they could grow in the absence of NaCl, and had optima of 0.85-1.71M, can best be termed haloversatile (James *et al.*, 1990) and, considering their optimal growth in 5 % KCl, 'kaloversatile'.

Table 3.9. Properties of the Inowroclaw Brines. Ion concentrations provided by Inowroclawskie Kopalnie Soli are given (*). Na^+, K^+ and Cl^- concentrations measured in the laboratory are shown in **bold**, beneath the equivalent in the mine analyses. Viable counts were after 1 month's incubation; proportions of 'reds' (presumed halobacteria) are given in bold.

	SOLNO	GORA
Density (g / cm^3)*	1240-1250	1196
$[Na^+]$ (g l^{-1})* M $[Na^+]$	- **2.1**	106.3 (= 4.6 M) **5.4**
$[K^+]$ (g l^{-1})* M $[K^+]$	15.4-20.4 (= 0.39-0.52 M) **0.900**	0.0833 (= 0.002 M) **0.011**
$[Mg^{2+}]$ (g l^{-1})*	22.1-25.2	0.1933
$[Ca^{2+}]$ (g l^{-1})*	0.0	0.3078
$[Cl^-]$ (g l^{-1})* M $[Cl^-]$	180.0-192.0 (= 5.07-5.40 M) **5.7**	107.0 (= 3.01 M) **7.3**
$[SO_4^{2-}]$ (g l^{-1})*	10.2-15.1	0.3822
$[HCO_3^-]$ (g l^{-1})*	0.45-0.60	-
$[CO_3^{2-}]$ (g l^{-1})*	-	0.0789
pH	7.0	7.5
Conductivity (ms.)	440	510
Viable Count ml^{-1} @ 30 °C (CHMst)	2.4×10^5 **<<1%**	2.1×10^5 **>12%**
Viable Count ml^{-1} @ 40 °C (CHMst)	8.5×10^4 **~1%**	1.5×10^5 **67%**
Viable Count ml^{-1} @ 30 °C (10% KCl)	2.4×10^5 **0**	1.6×10^4 **>7%**
Viable Count ml^{-1} @ 40 °C (10% KCl)	8.4×10^4 **0**	1.8×10^4 **>58%**
Viable Count ml^{-1} @ 30 °C (18% KCl)	9.6×10^4 **0**	13 **0**
Viable Count ml^{-1} @ 40 °C (18% KCl)	2.3×10^4 **0**	7 **0**

With the Góra brine, on CHM with 10 % KCl, viable counts were reduced by about a factor of ten, and the number of halobacteria isolated was also reduced; counts were negligible with 18 % KCl (Table 3.9). This could be because the Góra isolates were not from a high [K^+] environment and so were not adapted to tolerate these K^+ levels, or alternatively they were not able to grow in a [NaCl] as low as 0.34M. As expected, archaeal isolates would not grow on less than 10 % NaCl, although the [KCl] was correspondingly increased (PIG.8 showed a little growth on 5 % NaCl + 15 % KCl). Halobacteria PIG.5 and PIG.8 were included in phylogenetic analysis.
Bacterial isolates from the 18 % KCl medium (PIG.9, 10, 13, 14, 15 and 16) showed preferential growth on 15-20 % KCl (5-0 % NaCl respectively). These had a KCl growth range (no NaCl) of 0-25 % (to <15 % for PIG.14, and 20 % for PIG.15 and 16) and a NaCl growth range of 0-15 % (20 % for PIG.14). Optima were not clear from the results, but, like the Solno Bacteria, these are best described as haloversatile and kaloversatile.

3.1.10 Kali und Salz AG Mines, Germany.

Kali und Salz AG mine potash and rock salt, by the conventional drill and blast and cutting method (Dr. Käding and Dr. Beer, pers. comm.), from the Zechstein Basin in Germany. Dr. Käding and Dr. Beer, of the Kali und Salz Geological Department, sent two brine samples from their mine workings, together with geological and chemical details, and some of the properties given in Table 3.10. These were received in June, 1992.

1. *HA*

A natural brine from a potash seam within the Werra Formation (Zechstein 1), from Hattorf potash and rock salt mine in the Werra region of Germany. This deposit has been partly influenced by the tectonic and thermal effects of Miocene vulcanism, and from its chemical composition, the brine appears to have resulted from diagenesis or metamorphism of carnallite ($KMgCl_3.6H_2O$). The source of this brine had a flow rate of 4.5 l min^{-1}. (Dr. Beer, pers. comm.).

2. *SI*

From Sigmundshall potash mine in the Hannover region. This was not a natural brine, but from a brine pool with no detectable flow rate, derived from backfilled waste salt near an underground brine pond. The raw ore was 'produced from a highly deformed potash seam within a salt diapir' of the Leine Formation (Zechstein 3), originally some 2,500 to 3,500 m deep. (Dr. Beer, pers. comm.).

No bacteria were isolated from these brines, on plates of, or as enrichments in, CHMst medium. The brines appeared to be sterile. This was highly unusual, considering the ubiquity of halobacteria in hypersaline environments. Either these brines had never been populated by halophilic bacteria (*e.g.* an ionic composition incompatible with growth), or the populations had since been killed by metamorphic events in the salt deposits, or killed more recently by industrial treatments (heating or toxic chemicals). The *SI* brine may have been contaminated by organic products of mining activity (Dr. Beer, pers. comm.). Polyacrylamide, stearylamine, fatty acid polyether, coal-tar oil and pine oil, in concentrations less than tenths of a ppm, were suggested.

The ionic measurements (Table 3.10) indicated a low NaCl concentration: < 0.61 M in *HA* and < 2.05 M in *SI*, which may have been too low to support the growth of halobacteria, although moderately halophilic Bacteria could have grown. With ~0.6M NaCl and ~0.5M KCl as the dissolved salts in *HA* brine, the Cl^- ion excess was probably partially balanced by Mg^{2+} ions; ~4.9 M $MgCl_2$ would account for this and this brine is known to have been derived from carnallite. In the *SI* brine, with ~2 M NaCl and ~1 M KCl, the Cl^- ion excess could be accounted for by ~3 M $MgCl_2$. But this is probably an over-simplification; it was attempted to produce the above salt solutions artificially, but salts remained in solution. The solubility of $MgCl_2$ (542.5 g l^{-1}, or 5.69 M, at 20 °C; Weast, 1964-5) would have been reduced by the KCl and NaCl.

To test whether the *HA* and *SI* brines were each capable of supporting the growth of halobacteria, they were used as the salts moiety in a preparation of the Classical Halophile Medium. *i.e.* filter-sterilized brine added to the autoclaved organics. (2 % filter-sterilized $MgSO_4.7H_2O$ was also added. It was not known at this stage that the *HA* brine was derived from carnallite and so an already high [Mg^{2+}] was not considered.) Liquid media prepared in this way were inoculated with culture collection strains of neutraphilic halobacteria and the results, after 4-13 days, were as shown in Table 3.11. Non-sterilized brine was also used to make up the medium as, effectively, a large-scale inoculum; but no growth ensued.

Table 3.10. Properties of the Kali und Salz Brines.

	HA	*SI*
Sampled*	29 - 5 - 92	6 -5 - 92
Source*	possibly 2° from carnallite	from waste processed salt
Flow Rate (l /min)	4.5	none definable
Temperature °C*	27	27
Density (g /cm^3)*	1,279	1,296
pH	6.5	7.0
Conductivity (ms.)	695	619
M[Na^+]	0.61	2.05
M[Cl^-]	11.0	9.0
M[K^+]	0.54	1.1

* Information provided by Kali und Salz.

Table 3.11. Growth Capability of Halobacteria in Kali und Salz Brines.

	Growth in CHM + *HA*	Growth in CHM + *SI*
Haloferax volcanii 2012	+	+
Halobacterium sp. 777	-	+
Halobacterium cutirubrum 763	-	+
Halorubrum saccharovorum 2081	-	+
Halorubrum sodomense 33755	-	+
Haloarcula vallismortis 29715	-	+
Halobacterium salinarium 2084	-	-
Halococcus morrhuae 787	-	-
Negative Control (no inoculum)	-	-

It can be concluded from these results that the Kali und Salz brines have an ionic composition which is not incompatible with the growth of halobacteria, although the low NaCl concentrations probably limited the types which grew: whereas *SI* (2 M NaCl) supported the growth of a range of halobacteria, only *Haloferax volcanii* could grow in *HA* brine (0.6 M NaCl). The Mg^{2+}, which was probably in greater concentrations in the *HA* brine, may have had some effect in restricting growth. *Haloferax volcanii* was originally isolated from the Dead Sea, which has 1.73 M Mg^{2+} (Table 1.1), and the medium commonly used to support the growth of *Haloferax volcanii* has 2.14 M NaCl and 0.25 M Mg^{2+}. Also from these results, the brines therefore contained no substances, or none of sufficient concentration, that were toxic to halobacteria.
Why were no halobacteria isolated from these brines when brines from other locations in the Zechstein Basin yielded growth? The remaining explanation could be that the halobacteria which originally populated these deposits were killed by the excessive heating and tectonic effects of vulcanism, or diapirization. If this was the case, these sterile brines provide a natural negative control, obviating the argument that halobacteria may be introduced in to the evaporite deposit at a later stage.

3.1.11 'AKZO' Solution Mines, Western Europe.

AKZO has four locations in Western Europe where salt is produced from solution-mining of the Zechstein deposit: Mariager in Denmark, Stade in Germany, and Delfzijl and Hengelo in the Netherlands. Brine samples from each mine were provided by Dr. Meijer of AKZO. These were received, with information (Table 3.12), on 11-9-92.

The Netherlands has both Triassic and Permian (Zechstein) deposits, the latter lying at greater depths. At Hengelo, the Triassic Röt salt lies at a depth of 290-360 m and the Zechstein at about 590 m deeper (Lefond, 1969). The 'Boring Z1' well is the only well in Hengelo which draws from Zechstein salt (Dr. Meijer, pers. comm.). In Germany, at Stade, the Zechstein salt is only 250 m thick, but can reach thicknesses of over 3,000 m in the north of Germany, due to tectonic forces since its deposition (Lefond, 1969).

Both Bacteria and presumed halobacteria were isolated from all of the AKZO brines but, similar to the Thai brine isolates, a large proportion of the colonies in these brines were tiny and transparent, and of indeterminate colour, so the given percentages of 'reds' may have been underestimated. This was particularly true for the German brine, from which all the colonies were of this nature. After a further 2 months (at room temperature) more colonies had grown (*cf.* PW4 brine) from the German brine, but none of the isolates survived sub-culturing, so identification by polar lipid analysis was not possible. Red-

pigmented colonies from the Dutch and Danish brines were later confirmed as halobacteria.

Table 3.12. Properties of the AKZO Brines. Viable counts were after 1 month's incubation on CHMst medium. Proportions of 'reds' (presumed halobacteria) are given in **bold**. ND = Not Determined.

	Nh	Nd	D	Gst
Sampled*	27-8-1992	14-5-1992	12-5-1992	20-5-1992
Location*	Hengelo, Netherlands	Delfzijl, Netherlands	Mariager, Denmark	Stade, Germany
Source*	raw brine from 'Boring Z1' well	raw brine from 'Pekelbassin'	raw brine	raw brine from storage tank
pH	7.0	7.0	7.0	7.0
Conductivity (ms.)	490	490	500	500
M[Na^+]	5.4	5.3	5.3	5.3
M[Cl^-]	5.8	6.0	6.0	5.9
M[K^+]	0.0021	0.045	0.0035	0.0021
Viable Count ml^{-1}	837	3.9×10^5	9.7×10^4	5.3×10^3
@ 30 °C	**>48%**	**>1%**	**>50%**	ND
Viable Count ml^{-1}	833	4.0×10^5	1.0×10^5	593
@ 40 °C	**>66%**	**~100%**	**~40%**	ND

* Information provided by AKZO.

3.2 POLAR LIPID ANALYSIS.

3.2.1 Membrane Lipids.

Core Lipids.

The 'core' lipids of the halobacterial cell membrane are derivatives of the saturated, isopranoid glycerol diether, *sn*-2,3-diphytanylglycerol (C_{20} - C_{20}), commonly termed 'archaeol', which is unique to the Archaea and therefore important in distinguishing halobacteria from halophilic Bacteria (Ross *et al.*, 1981). The variants of archaeol found in halobacterial species are C_{20} - C_{20} (in all halobacteria), C_{20} - C_{25}, and C_{25} - C_{25} (*Natronococcus* and *Natronobacterium*) diethers. A *Natronobacterium* from India has a novel 3-hydroxydiether core lipid. An individual halobacterium may have a mixture of core lipid types in its membrane, *e.g.* the haloalkaliphilic isolate SP8 (De Rosa *et al.*, 1983) has core lipids that are 9 % C_{20} - C_{20}, 85 % C_{20} - C_{25} and 6 % C_{25} - C_{25} diethers.

Polar Lipids.

The high concentration of acidic lipids in the membranes of halobacteria maintain a high negative charge surface density, which is required for halobacterial survival in high salt concentration. The negative charges are shielded by the high [Na+], preventing disruption of the lipid bilayers due to charge-charge repulsion (Kates, 1993).

Phosphate, sugar and sulphated sugar groups linked to the core lipids constitute the membrane phospholipids, glycolipids and sulphated glycolipids, respectively. These polar lipids are important as taxonomic markers in halobacteria. All halobacteria examined have archaeol analogues of phosphatidylglycerolmethylphosphate (PGP-Me; Kates, 1993) as the major phospholipid, and the minor phospholipid phosphatidylglycerol (PG); phosphatidic acid (PA) is often present in small or trace amounts. Phosphatidylglycerolsulphate (PGS), also a minor lipid component, is not present in all halobacteria. The glycolipids vary in their sugar residues, the sugar linkages, and the position of the sulphate group, if present. The presence or absence of PGS and the nature of the glycolipids appear to be genus-specific, and so polar lipid profiles are a useful taxonomic tool in assigning halobacteria to their generic groups. There is a good correlation between groupings based on polar lipid analysis and those based on 16S rRNA-DNA hybridization data (Ross and Grant, 1985) and, more recently, on 16S rRNA gene sequence analysis (McGenity and Grant, 1995). The polar lipid profiles of the halobacterial groups are as follows (Kates, 1993):

Halobacterium.
C_{20} - C_{20} core lipids.
PGP-Me; PG; PA; PGS.
Major glycolipid: S-TGD-1; minor glycolipids: TGD-1, S-TeGD and TeGD.

Halorubrum. (Lanzotti *et al.*, 1989).
C_{20} - C_{20} core lipids.
PGP-Me; PG; PGS.
S-DGD-1, but no DGD-1. *Hr. sodomense* has S-DGD-3.
Hr. trapanicum - S-DGD-5.

Haloarcula.
C_{20} - C_{20} core lipids.
PGP-Me; PG; PGS.
No PGS in strain (A4-1) from Granada, Spain (Monteoliva-Sanchez, Rodriguez and Kates - unpublished data; see Kates, 1993).
Major glycolipid: TGD-2; minor amounts of DGD-2.

Haloferax.
C_{20} - C_{20} core lipids.
PGP-Me; PG.
No PGS.
Major glycolipid: S-DGD-1; minor glycolipid: DGD-1.

Halococcus.
C_{20} - C_{20} and C_{20} - C_{25} core lipids.
PGP-Me; PG.
No PGS.
Some have S-TGD-1, TGD-1, S-TeGD and S-DGD-1. *Hc. saccharolyticus* has only S-DGD-1 and an unidentified phosphoglycolipid.

Halobaculum.
C_{20} - C_{20} core lipids.
PGP-Me; PG.
No PGS.
Glycolipid: S-DGD-1.

Natronobacterium.
C_{20} - C_{20}, C_{20} - C_{25} and C_{25} - C_{25} core lipids.
PGP-Me; PG; PA .
No PGS.
Thought to be distinctive in having no glycolipids, but DGD-4 detected in *Nb.* strain SSL1 (and others) from India, and on overloaded TLC plates of *Nb. gregoryi* (Upasani *et al.*, 1994).

Natronococcus.
C_{20} - C_{20}, C_{20} - C_{25} and C_{25} - C_{25} core lipids.
PGP-Me; a cyclic form of PGP in *Nc. occultus* (Lanzotti *et al.*, 1989); PG.
No PGS.
Distinctive in having no glycolipids, but DGD-1 detected on overloaded TLC plates.

Species Unassigned to Recognized Generic Groups.
"Halobacterium cutirubrum" (NCIMB 763): C_{20} - C_{20} derivatives of PG, PGP-Me, PGS, TGD-1 and S-DGD.

"Halobacterium halobium" (NCIMB 777), *"Halobacterium trapanicum"* (NCIMB 784) and *"Halobacterium salinarium"* (NCIMB 786): C_{20} - C_{20} and C_{20} - C_{25} derivatives of PG, PGP-Me, PGS, and uncharacterized glycolipids.

Abbreviations: S-, sulphated; DGD, diglycosyldiether; TGD, triglycosyldiether; TeGD, tetraglycosyldiether. D, diether, is interchangeable with A, archaeol; thus, DGD is synonymous with DGA, *etc.*, of Kates, 1993.

3.2.2 Identification of Halobacterial Isolates.

Polar lipid extracts from 167 halobacterial isolates, representing all sites, were analysed by Two-Dimensional Thin-Layer Chromatography, as described in Section 2.2. The isolates were then sorted into polar lipid profile groups (Ross *et al.*, 1985; Kates, 1993), which have been shown to correlate closely with generic groupings determined by 16S rRNA-DNA hybridization studies (Tindall *et al.*, 1984; Ross *et al.*, 1985; Ross and Grant, 1985). Table 3.13 summarizes the distribution of these halobacterial groups within and between the different sites studied. Those isolates which could not be assigned to a known generic group were categorized as 'unclassified' and may represent new taxa.

Isolates of the haloalkaliphilic genus *Natronobacterium* were isolated only from the alkaline brine from Wyoming. With respect to the sites with salts of neutral pH, *Halorubrum* spp. and *Haloarcula* spp. were isolated from all salt deposits, and *Halobacterium* spp. were isolated predominantly from Winsford and Boulby. Most identified isolates belonged to one of these groups and based on observations recorded during plate counts, these were the predominant isolates from several of the sites: *Halorubrum* spp. predominated in cultures of samples from Spain (SR4), Wieliczka (PW1 and PWB), Winsford (WB4 and WM2) and probably Denmark; *Halobacterium* spp. predominated in cultures from Winsford (WSP) and Boulby (B1 and B2); and *Haloarcula* spp. predominated in the Góra brine culture. The remaining groups were poorly represented. The rarity of isolation of *Haloferax* spp. could be due to the unsuitability of the media and the high NaCl concentration used. Within the unclassified category, apart from those isolates with unique polar lipid patterns, there were four clear groups, which will be called (i) the T4.2 group, (ii) the SR1.5 group, (iii) *'Halobacterium trapanicum'* NCIMB 767-like, and (iv) '004.1'-like. The 'T4.2 group', based on polar lipid analysis, appears to be well-represented, including nine isolates from Thailand (T2, T3 and T4), Boulby (B4 and B8) and Góra. Isolate T4.2 has been sequenced. The 'SR1.5' group includes five isolates, two from Spain (SR1) and three from Boulby (B1 and B2); each of these isolates represented the dominant organism on their respective plate cultures. SR1.5 has been sequenced. The '767'-like organisms were from Wieliczka (PW6), Thailand (T4, T5 and T_{208}) and Boulby (B1 and B3), and had polar lipid profiles like *Halococcus* spp., but, from microscopic observation, these isolates were not cocci and so constituted a separate group, similar to *'Halobacterium trapanicum'* NCIMB 767. The '004.1'-like profiles (see McGenity, 1994) belonged to isolates from Wieliczka (PW2), Winsford (WB3), Boulby (B1), Góra and Delfzijl. Over half of the isolates from Thailand were unidentified. One of these was T3.1 (Figure 3.5ii), which formed the majority isolated from the well head brine and had an unusual cone-like cell morphology; this has been sequenced.

The profiles have been identified by comparing them with published profiles (Ross *et al.*, 1985; Kates, 1993) and the polar lipid profiles of reference strains of halobacteria. The main purpose for the polar lipid analysis was to screen the isolates and determine their generic affiliation, to provide a picture of their distribution and to select organisms for phylogenetic analysis.

Table 3.13. Site Distribution of Halobacterial Groups According to Polar Lipid Analysis. Values given indicate the number of isolates.

Sampling Site	Viable Count ml^{-1} @ 40 °C	% Reds (Halobacteria) @ 40 °C	*Halobac-terium*	*Halo-rubrum*	*Halo-arcula*	*Halo-ferax*	*Halo-coccus*	*'Halob-acterium* sp.' NCIMB 777	*'H. cuti-rubrum'* NCIMB 763	*Natrono-bacterium*	Unclass-ified
Great Salt Lake (GSL5)	2×10^6	majority	0	2	3	0	0	0	0	0	1
Spain											
SM2	2×10^4	majority	0	2	0	0	0	0	0	0	0
SM4	3×10^3	< 1 %	0	1	0	0	0	0	0	0	0
SC1	1×10^3	majority	0	3	1	0	0	0	0	0	0
SR1	430	majority	0	0	0	0	0	0	0	0	2
SR2	260	majority	0	0	1	0	0	0	0	0	0
SR4	1×10^6	majority	0	1	1	0	0	0	0	0	0
Wieliczka											
PW1	6.9×10^4	>70%	0	3	1	0	0	0	1	0	0
PW2	1.4×10^5	*c.* 80%	1	0	1	0	0	0	0	0	1
PW4	70	6/7	0	2	1	0	0	0	0	0	1
PW5	230	1/23	1	0	0	0	1	0	0	0	0
PW6	2.7×10^5	*c.* 70%	0	2	0	0	0	0	0	0	1
PWA	*c.* 10^3 (g^{-1})	majority	0	0	1	0	2	0	0	0	0
PWB	2.2×10^5 (g^{-1})	*c.* 96%	0	4	1	0	1	0	0	0	0

Table 3.13 Continued ... 2 ...

Sampling Site	Viable Count ml^{-1} @ 40 °C	% Reds (Halobacteria) @ 40 °C	*Halobacterium*	*Halorubrum*	*Haloarcula*	*Haloferax*	*Halococcus*	*'Halobacterium* sp.' NCIMB 777	*'H. cutirubrum'* NCIMB 763	*Natronobacterium*	Unclassified
Wyoming											
	8.7×10^3	*c.* 20 %	0	0	0	0	0	0	0	6	0
Thailand											
T1/2	1.4×10^5	ND	0	0	0	0	0	0	0	0	7
T3/4	*c.* 2.5×10^5	ND	0	0	2	0	0	0	0	0	6
T5	1.1×10^4	10-11%	0	0	0	0	0	1	0	0	3
Ts	2.6×10^5 (g^{-1})	12%	0	0	0	1	1	4	0	0	0
T_{203}	450	> 70%	0	1	2	1	0	0	0	0	1
Winsford											
WB4br	1.6×10^5	majority	0	1	0	0	0	0	0	0	0
WB4E	2.4×10^5 (g^{-1})	majority	0	1	1	0	3	0	0	0	0
WM2br1/2	*c.* 2.0×10^4	< 44%	1	1	0	0	0	0	0	0	0
WM2S/C	*c.* 10^5 (g^{-1})	majority	2	2	1	0	0	0	0	0	0
WB3br	1.3×10^4	*c.* 15 %	0	0	0	0	0	0	0	0	1
WB3E	3.4×10^5 (g^{-1})	majority	2	1	0	0	0	0	0	0	0
WSPbr	760	100%	1	0	1	0	0	0	0	0	0
WSP	2.3×10^4 (g^{-1})	majority	1	0	1	0	0	0	0	0	0

Table 3.13 Continued ... 3 ...

Sampling Site	Viable Count ml^{-1} @ 40 °C	% Reds (Halobacteria) @ 40 °C	*Halobacterium*	*Halorubrum*	*Haloarcula*	*Haloferax*	*Halococcus*	*'Halobacterium* sp.' NCIMB 777	*'H. cutirubrum'* NCIMB 763	*Natronobacterium*	Unclassified
Boulby											
B1bra	4.6×10^5	majority	1	1	1	1	0	0	0	0	1
B1brc	6.5×10^6	100 %	1	0	0	0	0	0	0	0	1
B1E	7.4×10^4 (g^{-1})	100 %	3	0	0	0	0	0	0	0	1
B1st	1.5×10^6 (g^{-1})	*c.* 100 %	0	0	0	0	0	0	0	0	1
B2br	6.8×10^5	*c.* 100%	2	0	1	0	0	0	0	0	1
B2S	1.9×10^6 (g^{-1})	majority	3	1	0	0	0	0	0	0	1
B3	1.2×10^3	> 64 %	0	1	1	0	1	0	0	0	0
B3st	7.2×10^3 (g^{-1})	> 39 %	3	0	0	0	0	0	0	0	1
B4	9.5×10^3	majority	2	0	0	1	0	0	0	0	0
B5	7.3×10^3	*c.* 100%	0	0	1	0	2	0	0	0	0
B6	8.2×10^3	*c.* 100%	0	1	0	0	0	0	0	0	0
B7	890	100 %	0	0	0	0	0	0	0	0	1
B8	660	*c.* 100%	1	0	1	0	0	0	0	0	0

Table 3.13 Continued ... 4 ...

Sampling Site	Viable Count ml^{-1} @ 40 °C	% Reds (Halobacteria) @ 40 °C	*Halobacterium*	*Halorubrum*	*Haloarcula*	*Haloferax*	*Halococcus*	*'Halobacterium* sp.' NCIMB 777	*'H. cutirubrum'* NCIMB 763	*Natronobacterium*	Unclassified
Inowroclaw	8.5×10^4	*c.* 1 %	0	0	2	0	0	0	0	0	0
Solno Gora	1.5×10^5	67 %	0	1	4	1	0	0	3	0	4
Netherlands	833	> 66 %	0	2	0	0	0	0	0	0	2
Hengelo											
Delfzijl	4.0×10^5	*c.* 100 %	0	0	1	0	0	0	0	0	2
Denmark	1.0×10^5	*c.* 40 %	0	5	0	0	0	0	0	0	1
		Total:	25	39	31	5	11	5	4	6	41

***Halobacterium*-like (Figure 3.5a).**

Isolates identified by a *Halobacterium*-like polar lipid profile were mostly from Winsford and Boulby. As typified by isolate WM2C.13 (Figure 3.5b), this group possessed C_{20} - C_{20} derivatives of PG, PGP-Me and PGS, with S-TGD-1 as the major glycolipid and an unknown minor glycolipid (GL-1: probably TGD-1). Some, such as B1bra.5 (Figure 3.5c) had trace amounts of (probably) S-TeGD. Most of these isolates had minimal amounts of PG. Interestingly, B1bra.5 (also B1E.24, but not shown) appeared to have C_{20} - C_{20} and C_{20} - C_{25} derivatives of PG, PGP-Me and PGS, a core lipid structure not previously known for *Halobacterium* spp.; yet phylogenetic analysis indicated that WM2C.13 and B1bra.5 are closely related. In a *Halobacterium* subgroup initially thought to be *'Halobacterium cutirubrum'* NCIMB 763-like, the glycolipids, like the PG, are negligible; in B1E.5 (and B1E.1) they appear to be absent (Figure 3.5d). These isolates (B1E.5 and B1E.1) were shown to be similar to other *Halobacterium* species on the basis of protein profile and ribotyping results.

***Halorubrum*-like (Figure 3.5e).**

All isolates grouped as *Halorubrum*-like had C_{20} - C_{20} derivatives of PG, PGP-Me and PGS, and S-DGD as the glycolipid. This typical *Halorubrum* polar lipid profile is shown by Nh.5 (Figure 3.5f); this isolate has been sequenced. Some strains, such as WB4br.8, had minor amounts of other phospholipids (Figure 3.5g). PW6.1 had a *Halorubrum*-like profile, with two additional unidentified glycolipids (Figure 3.5h). Morphologically, PW6.1 was a rod, with some 'bullrush'-like forms, as exhibited by other *Halorubrum* species (Figure 1.1d,e). B2S.26 also has a typical *Halorubrum* polar lipid profile (Figure 3.5i), but later phylogenetic analysis showed that this organism should not be considered part of this genus.

***Haloarcula*-like (Figure 3.5j).**

All isolates grouped as *Haloarcula*-like had C_{20} - C_{20} derivatives of PG, PGP-Me and PGS, with TGD-2 as the major glycolipid, and isolate B2br.21 (Figure 3.5k) shows clearly this typical profile. Some isolates also had minor amounts of other polar lipids, possibly including the glycolipid DGD (GL-2). One of these was T3.2 (Figure 3.5l), confirmed as a *Haloarcula* sp. by phylogenetic analysis.

***Haloferax*-like (Figure 3.5m).**

The few isolates grouped as *Haloferax*-like had C_{20} - C_{20} derivatives of PG and PGP-Me, but no PGS, with S-DGD-1 as the major glycolipid. B1bra.15 (Figure 3.5n) possessed, in addition, minor amounts of a glycolipid (GL-3) which could be DGD-1, based on chromatographic mobility, whilst B4br.30 (Figure 3.5o) had minor amounts of unidentified glycolipids.

***Halococcus*-like (Figure 3.5p).**

The isolates grouped as *Halococcus*-like had the characteristic C_{20} - C_{20} and C_{20} - C_{25} derivatives of PG and PGP-Me, but no PGS, with S-DGD as the major glycolipid and minor amounts of an unidentified glycolipid, as typified by isolate WB4E.17 (Figure 3.5q). Other isolates with this polar lipid profile were initially grouped as *Halococcus*-like, but their rod-like cell morphology excluded them from this genus; these are described in the unclassified category.

***'Halobacterium* sp.' NCIMB 777-like (Figure 3.5r).**

NCIMB 777-like halobacteria were only identified from Thai samples. These isolates had the characteristic C_{20} - C_{20} and C_{20} - C_{25} derivatives of PG, PGP-Me and PGS, and three C_{20} - C_{20} and C_{20} - C_{25} glycolipids, as shown by isolate Ts.9 (Figure 3.5s). Two of the glycolipids have R_f values which identify them tentatively as TGD-2 and S-DGD-1; the third has been labelled GL-4. The identity of this group has been confirmed by phylogenetic analysis of isolate T5.7.

***'Halobacterium cutirubrum'* NCIMB 763-like (Figure 3.5t).**

The NCIMB 763-like isolates had C_{20} - C_{20} derivatives of PG, PGP-Me and PGS, and the glycolipids were probably S-DGD and S-TGD-1. Isolate PIG.5, which has been sequenced, showed this typical profile (Figure 3.5u), with additional minor polar lipids.

***Natronobacterium*-like (Figure 3.5v).**

The *Natronobacterium*-like halobacteria were only isolated from the alkaline-saline brine from Wyoming. These had one of two types of core lipids: (i) C_{20} - C_{20} and C_{20} - C_{25} derivatives of PG and PGP-Me, as shown by WTg5 (Figure 3.5w), or (ii) C_{20} - C_{20}, C_{20} - C_{25} and C_{25} - C_{25} derivatives of PG and PGP-Me (DeRosa *et al*, 1983), as shown by the triple PG spot on the WTg6 TLC plate (Figure 3.5x). The WTg6 profile resembled that of *Nb. pharaonis* (SP2, Figure 3.5v), an organism known to have C_{20} - C_{20}, C_{20} - C_{25} and C_{25} - C_{25} core lipids (Tindall, 1985). PGS and glycolipids were characteristically absent.

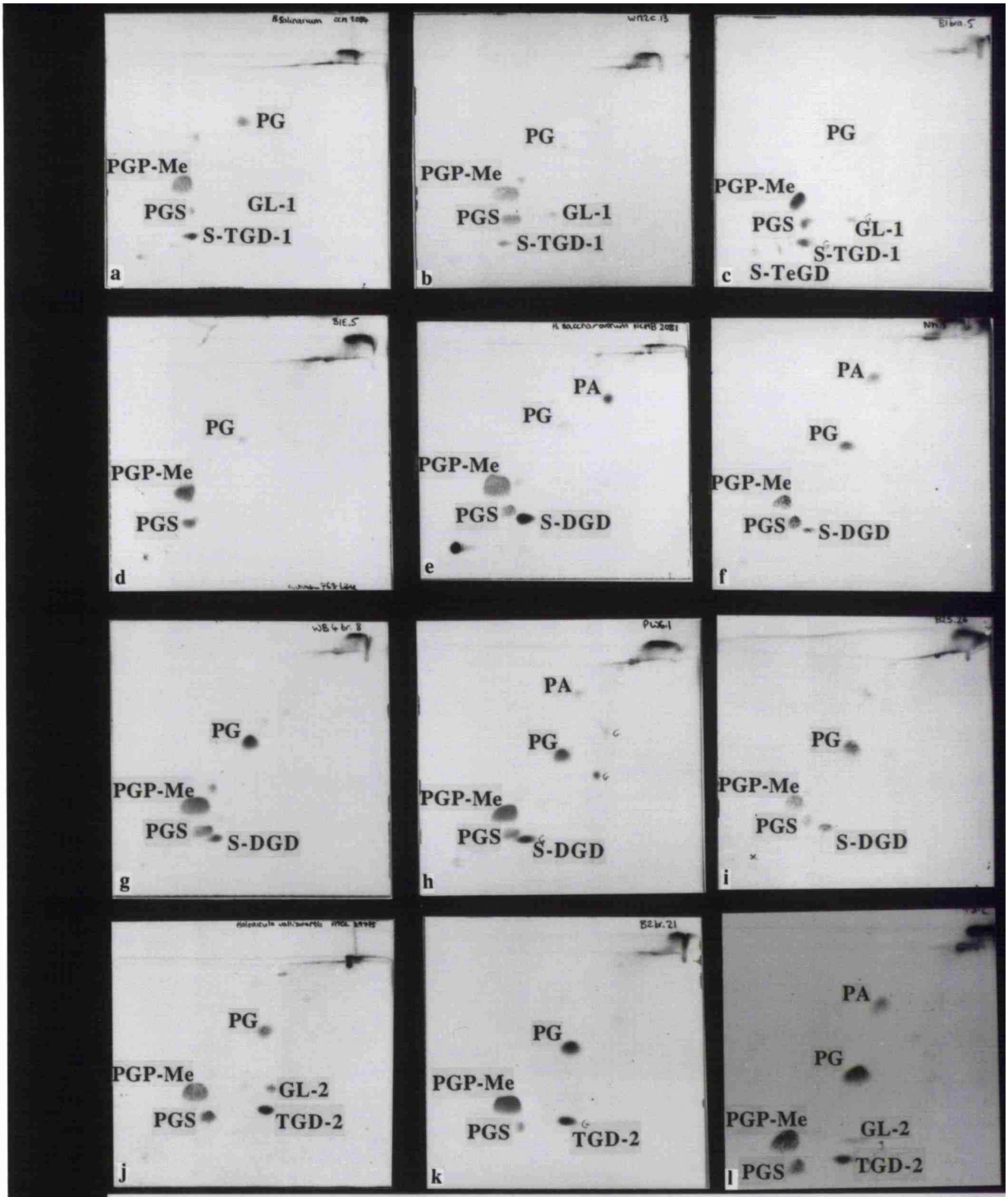


Figure 3.5 Polar Lipid Profiles.

Determination of phospholipids and glycolipids was based on comparisons with published profiles (Ross *et al.*, 1985; Kates, 1993) and the polar lipid profiles of reference strains of halobacteria. Structures are given in Kates (1993).

Abbreviations:
PG, phosphatidylglycerol;
PGP-Me, phosphatidylglycerolmethylphosphate;
PGS, phosphatidylglycerosulphate;
S-, sulphated;
DGD, diglycosyldiether;
TGD, triglycosyldiether;
TeGD, tetraglycosyldiether;
GL-1, unknown glycolipid;
GL-2, unknown glycolipid;
GL-3, unknown glycolipid;
GL-4, unknown glycolipid.

Figure 3.5a-l
a-d *Halobacterium*-like isolates;
e-i *Halorubrum*-like isolates;
j-l *Haloarcula*-like isolates.

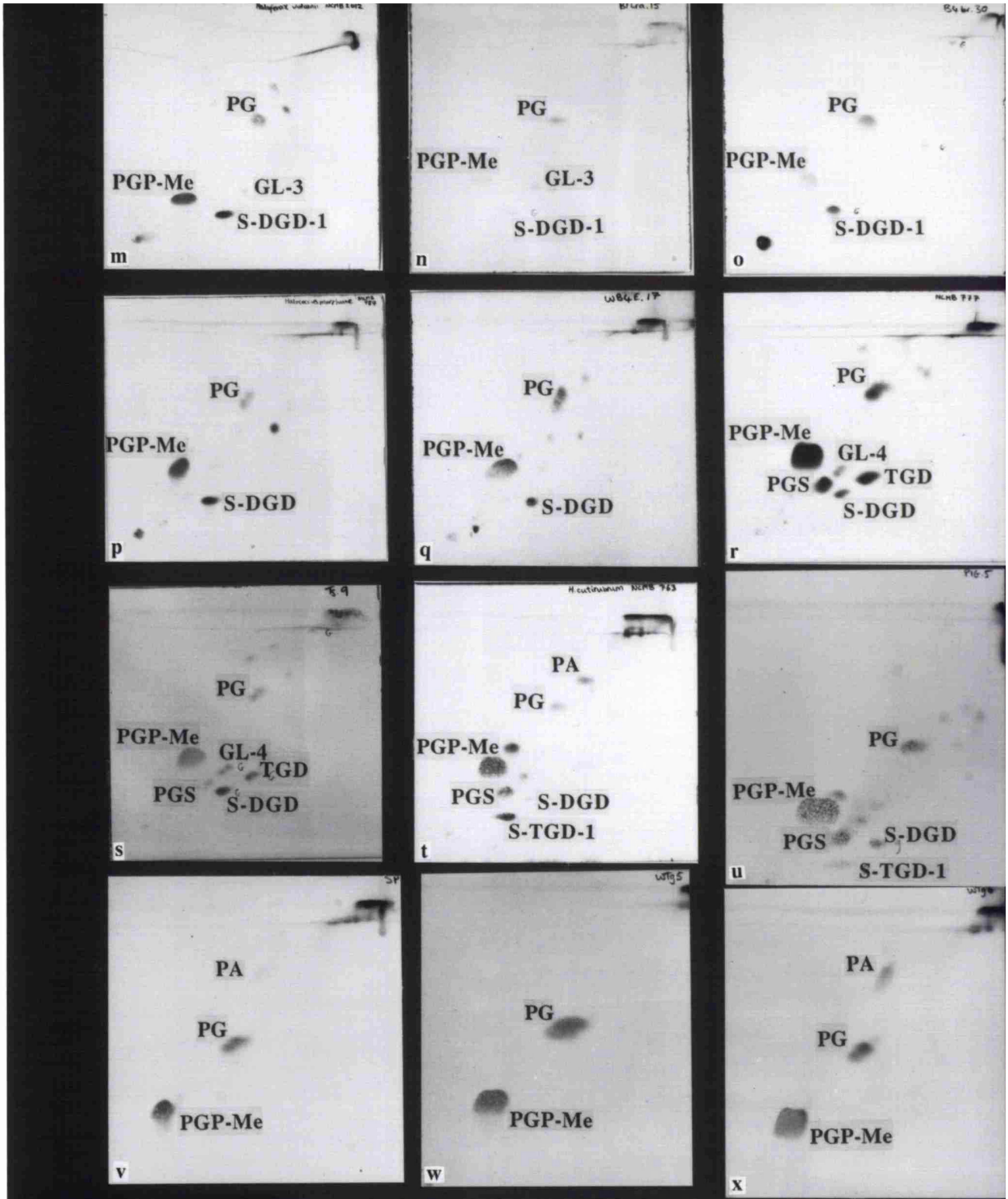


Figure 3.5m-x **Polar Lipid Profiles ... Continued.**

m-o	*Haloferax*-like isolates;
p-q	*Halococcus*-like isolates;
r-s	'*Halobacterium* sp.' NCIMB 777-like;
t-u	'*Halobacterium cutirubrum*' NCIMB 763-like;
v-x	*Natronobacterium*-like.

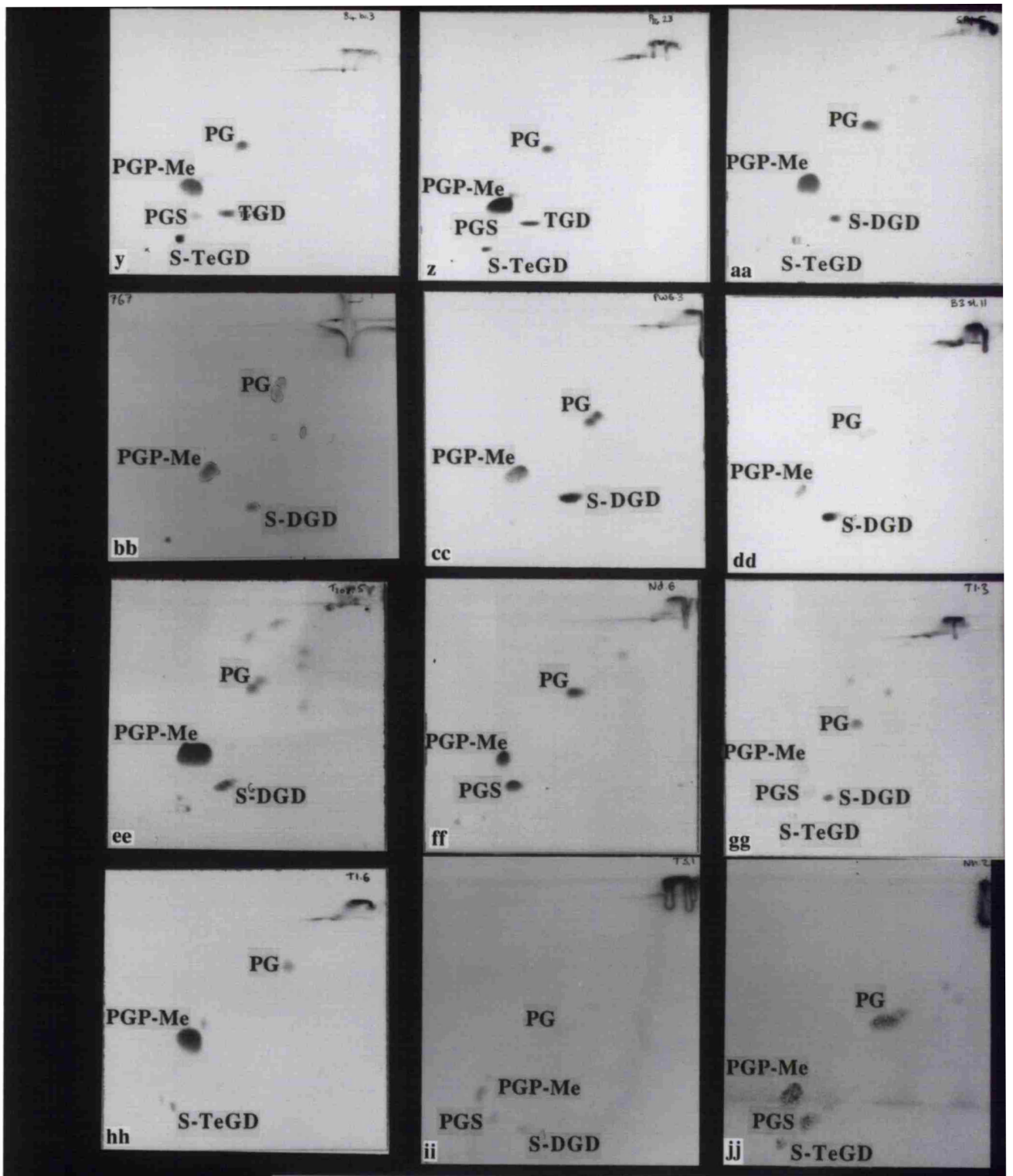


Figure 3.5y-jj Polar Lipid Profiles ... Continued.

Unclassified Halobacteria.

y-z 'T4.2 group';
aa 'SR1.5 group';
bb-ee NCIMB767-like;
ff '004.1'-like;
gg-jj 'Others'.

(Glycolipids are provisionally identified on the basis of chromatographic mobilities.)

Unclassified.

The glycolipids in the profiles of the unclassified or unidentified halobacteria (Figure 3.5 y-jj), were provisionally identified on the basis of their chromatographic mobilities.

(i) The T4.2 Group.

The 'T4.2' group isolates (B4br.3, B8br.3, T2.2, T3.3, T3.4, T4.2, PIG.3, PIG.23 and PIG.29) had a rod-shaped morphology which sometimes elongated to a 'bullrush' form. They had a very similar polar lipid profile to *Halobacterium* spp., but did not appear to have S-TGD-1, which is the major glycolipid for this genus. The major glycolipids were probably S-TeGD and possibly TGD-2. The core lipids were C_{20} - C_{20}, with possibly small quantities of C_{20} - C_{25} PGP-Me; PGS was minimal. Isolates B4br.3 (Figure 3.5y) and PIG.23 (Figure 3.5z) showed this profile clearly.

(ii) The SR1.5 Group.

The 'SR1.5' group isolates (SR1.4, SR1.5, B1st.6, B1brc.14 and B2br.13) had a rod-shaped morphology. They had C_{20} - C_{20} derivatives of PG and PGP-Me, but no PGS, with unknown glycolipids, probably S-DGD and S-TeGD (Figure 3.5aa).

(iii) *'Halobacterium trapanicum'* NCIMB 767-like (Figure 3.5bb).

Those isolates with NCIMB 767-like polar lipid profiles were characterized by C_{20} - C_{20} and C_{20} - C_{25} derivatives of PG and PGP-Me, an absence of PGS, and a glycolipid which may be S-DGD, a pattern characteristic of *Halococcus* spp., but these organisms did not have a coccoid cell morphology. There were three subgroups:

(a) PW6.3, with 'cup-shaped' cells, had C_{20} - C_{20} and C_{20} - C_{25} PG and PGP-Me, and possibly S-DGD (Figure 3.5cc);

(b) B1E.9 and B3st.11 (Figure 3.5dd), with large and irregular-shaped cells, had C_{20} - C_{20} and C_{20} - C_{25} PG and PGP-Me, and (probably) S-DGD;

(c) T2.7, T4.3, T5.5 and T_{208}.5 (Figure 3.5ee), with pleomorphic, flat, angular cells, had C_{20} - C_{20} and C_{20} - C_{25} PG and PGP-Me, (probably) S-DGD, and minor quantities of an unidentified glycolipid and other polar lipids.

(iv) '004.1'-like.

The homogeneity of this group (PW2.2, WB3br.7, B1bra.33, PIG.31 and Nd.6) was uncertain, as each member had quite different cell morphologies, but each had C_{20} - C_{20} derivatives of PG, PGP-Me and PGS, with no glycolipids (Figure 3.5 ff), a profile like that of isolate 004.1 of McGenity (1994).

Others.

The remaining profiles are of indeterminate affiliation. Some of these have been sequenced:

T1.3 (Figure 3.5gg) had C_{20} - C_{20} derivatives of PG, PGP-Me and PGS. S-DGD appeared to be the major glycolipid, and (probably) S-TeGD the minor glycolipid, with additional unidentified polar lipids.

T1.6 (Figure 3.5hh) had C_{20} - C_{20} derivatives of PG and PGP-Me, but no PGS, with (probably) S-TeGD as the major glycolipid, and an unidentified minor glycolipid.

T3.1 (Figure 3.5ii) and T3.6 had a *Halorubrum*-like profile, with PG, PGP-Me and PGS, and S-DGD as the probable glycolipid, except that the core lipids were C_{20} - C_{20} and C_{20} - C_{25} derivatives.

Nh.2 (Figure 3.5jj) had C_{20} - C_{20} and C_{20} - C_{25} derivatives of PG, PGP-Me and PGS, and a glycolipid, probably S-TeGD.

Based on the results of this analysis, the 21 isolates selected for sequencing and phylogenetic analysis were as follows:

***Halorubrum* spp.: SC1.2; GSL5.48; PW1.1; WB4E.8; B2S.26, PIG.8 and Nh.5.**

***Halobacterium* spp.: PW5.4, WM2C.13 and B1bra.5.**

***Haloarcula* spp.: T3.2 and T3.5.**

***'Halobacterium'* NCIMB 777-like: T5.7.**

***'Halobacterium cutirubrum'* NCIMB 763-like: PIG.5.**

Unclassified: SR1.5, T1.3, T1.6, T3.1, T3.6, T4.2 and Nh.2.

3.3 ANALYSIS OF PROTEIN ELECTROPHORESIS PATTERNS IN HALOBACTERIA.

Although poorly-defined tracks were omitted from the analysis, a lack of uniformity in the gels, probably due to variations in concentration of protein extracts, or to differences in staining, produced considerable problems in making accurate between-gel comparisons. For example, minor bands may not have been detected in the more dilute samples, and with overstaining or high concentrations bands may merge, or be overemphasized, introducing a bias. It has been observed that the Pearson product-moment correlation coefficient can weight similarities towards major bands, despite differences in minor bands, but this bias can be removed by excluding zones with prominent bands and analysing partial profiles (see Costas, 1992). The results may have been improved by increasing the resolution of the gel scans, and therefore the resolution of the analyses. However, Vauterin *et al.* (1993) point out that increasing the resolution of the gels (by using high electrophoresis power settings to give faster running times) to produce sharper bands results in reduced correlation between gels. In these analyses, using the option to normalize gels based on pattern recognition, rather than on peak alignment, should have reduced the effect of poor resolution. Gel runs were standardized to 10 cm, but judgement of the position of the dye 'front' was error-prone, resulting in some variation in track lengths.

Analysis of selected reference tracks from the same and from different gels showed similarities ranging from about 59 % to nearly 98 %, indicating considerable variation within and between gels. This has probably led to an underestimation of the similarities between isolates. The electrophoresis protocol therefore needs to be more stringently standardized, although in the highly standardized and controlled procedure adopted by Kersters and De Ley (1975) independently grown cultures of the same organism, or even different electrophoresis runs of the same extract, produced profiles with similarity coefficients which were little more than 82 %; this did, however, appear to depend on the organism. It has been recommended that only those gels with reference strain profiles at least 93 % (mean, 95 %) similar to the reference standard should be used in the analysis (Verissimo *et al.*,1996); in an analysis of *Helicobacter* and related species, Costas *et al.* (1993) used gels with 96.0 ± 1.0 % reference similarity. Vauterin *et al.* (1993) find 90 % and above acceptable for reproducibility.

Notwithstanding the inherent difficulties in complete standardization and reproducibility of the results, the analysis of protein electrophoresis patterns has great potential in a large-scale study of the relatedness of bacterial strains, and is inexpensive and easy to use. Its application is most suited to comparisons at the species, strain, and variant levels of relatedness.

Figure 3.6 is a dendrogram of halobacteria isolated in this study, based on the *GelCompar* computer analysis of whole-cell protein electrophoresis patterns (electrophoregrams). The generic affiliations of the isolates, given in the dendrogram, were based on the results of the polar lipid analyses. The isolates were generally in small, poorly-defined clusters, with syngeneric groups placed separately in the dendrogram. Exceptions were the deeply-branching and isolated *Halococcus* group, which clustered at the 35.9 % (± 9.0 %) level of similarity, the distinct NCIMB 777 group, which clustered (although without NCIMB 786) at 67.0 % (± 5.3 %), and the *Halobacterium* group, which was relatively cohesive at 65.5 % similarity (± 7.8 %), although the *Halobacterium salinarium* reference strains appeared separately. *Haloarcula* strains comprised two main clusters, with similarity levels of 74.1 % (± 7.0 %) and 84.9 % (± 1.7 %), respectively, and *Halorubrum*-like isolates were placed in several small, ill-delineated clusters. Most of the unclassified organisms could not be reliably identified on the basis of these results. Kersters and De Ley (1975) observed that when large numbers of protein patterns were compared and these included strains of different genera, these latter sometimes intercalated other clusters erroneously.

***Halobacterium* spp.** (Figure 3.7) clustered overall at 56.6 % (± 11.0 %), with two major sub-clusters at the 65-67 % similarity level. Boulby isolates appeared in two clusters, and within one of these, in which all samples were run on the same gel, the high similarity (95.6 %) between B7br.1 (unclassified based on the polar lipid profile) and B2S.17, was notable. Sampling sites B2 and B7 were quite separate within Boulby Mine (Section 3.1.8) and the brines from these sites, which were of mixed sources, were quite different in character (Table 3.7), but they had in common a residual brine component. The other strains in this cluster, which diverged from other *Halobacterium* spp. at the 73 % similarity level, were isolated from the residual brine sites B1 and B2. It is also of interest that Winsford surface isolates (WSP) did not cluster with isolates from within the mine (WM2 and WB3), despite their presumed same source.

Figures 3.6-3.10. Dendrograms of Halobacterial Isolates based on Protein Electrophoresis Patterns.

Scanned images of SDS-PAGE gels were analysed using the GelCompar™ software package. A *Pasteurella* reference track was included. Similarities were calculated using the Pearson product-moment correlation coefficient, and the dendrogram drawn with the UPGMA clustering algorithm. Scanned gel strips are included in the dendrograms.

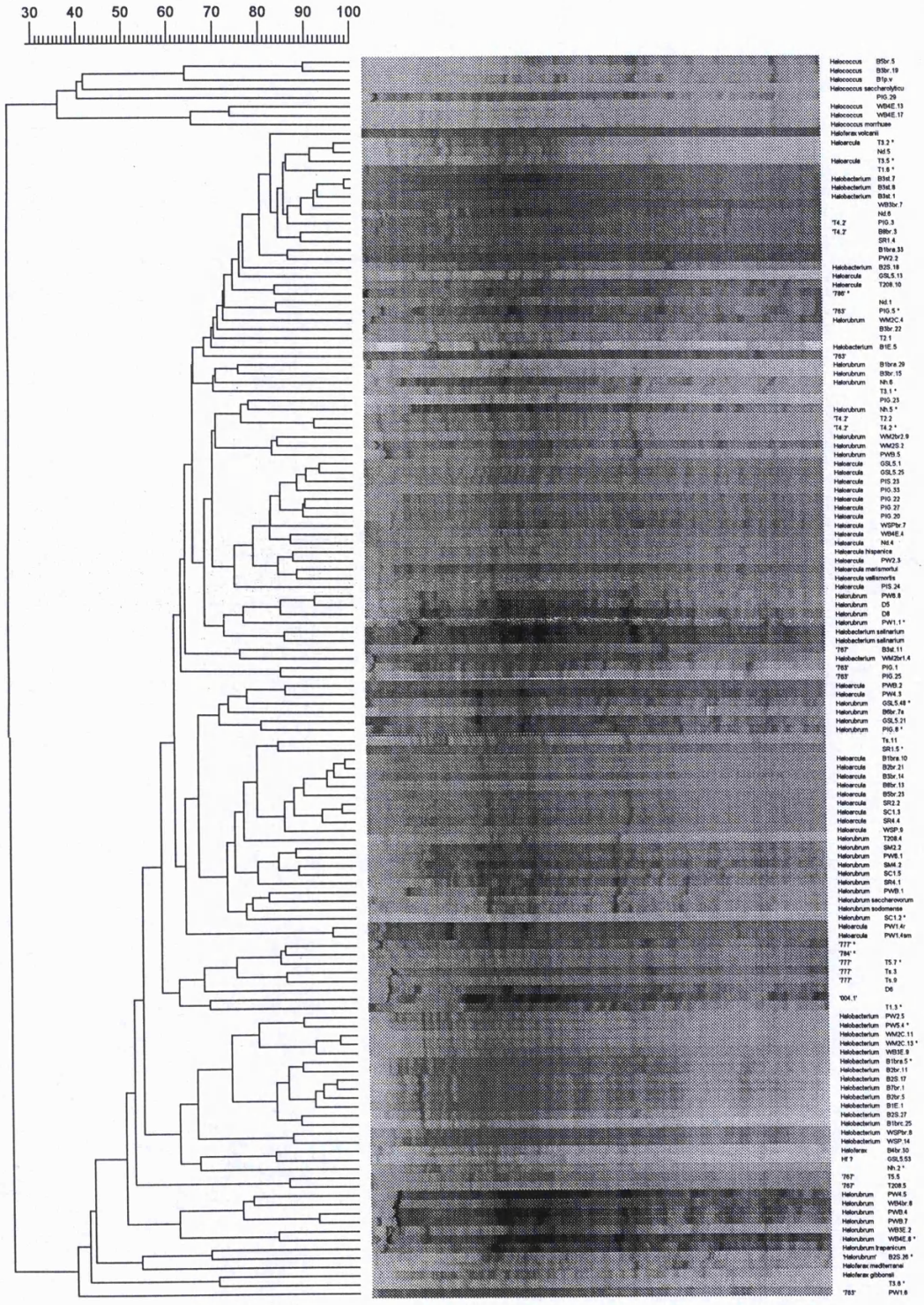


Figure 3.6. Dendrogram of all Isolates, based on Protein Electrophoresis Patterns.

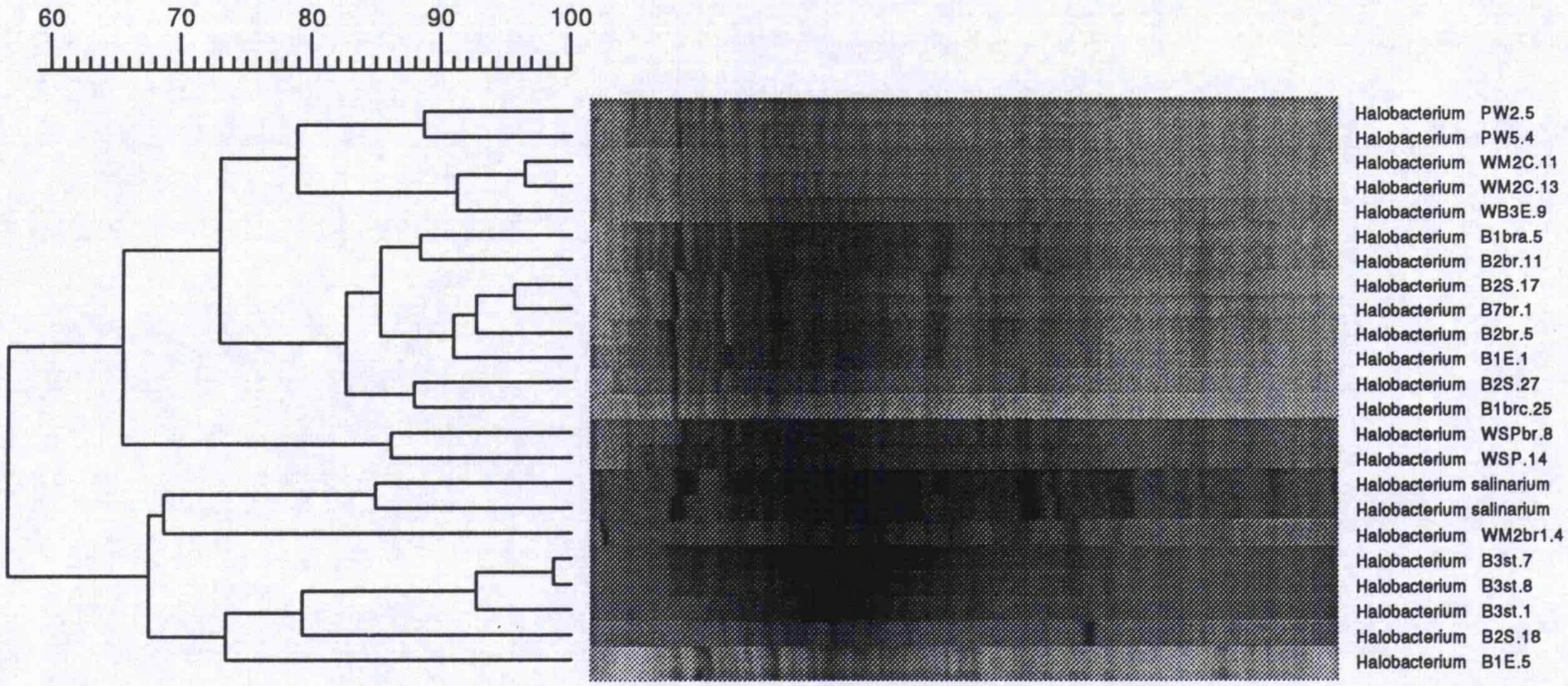


Figure 3.7. Dendrogram of *Halobacterium* -like Isolates, based on Protein Electrophoresis Patterns.

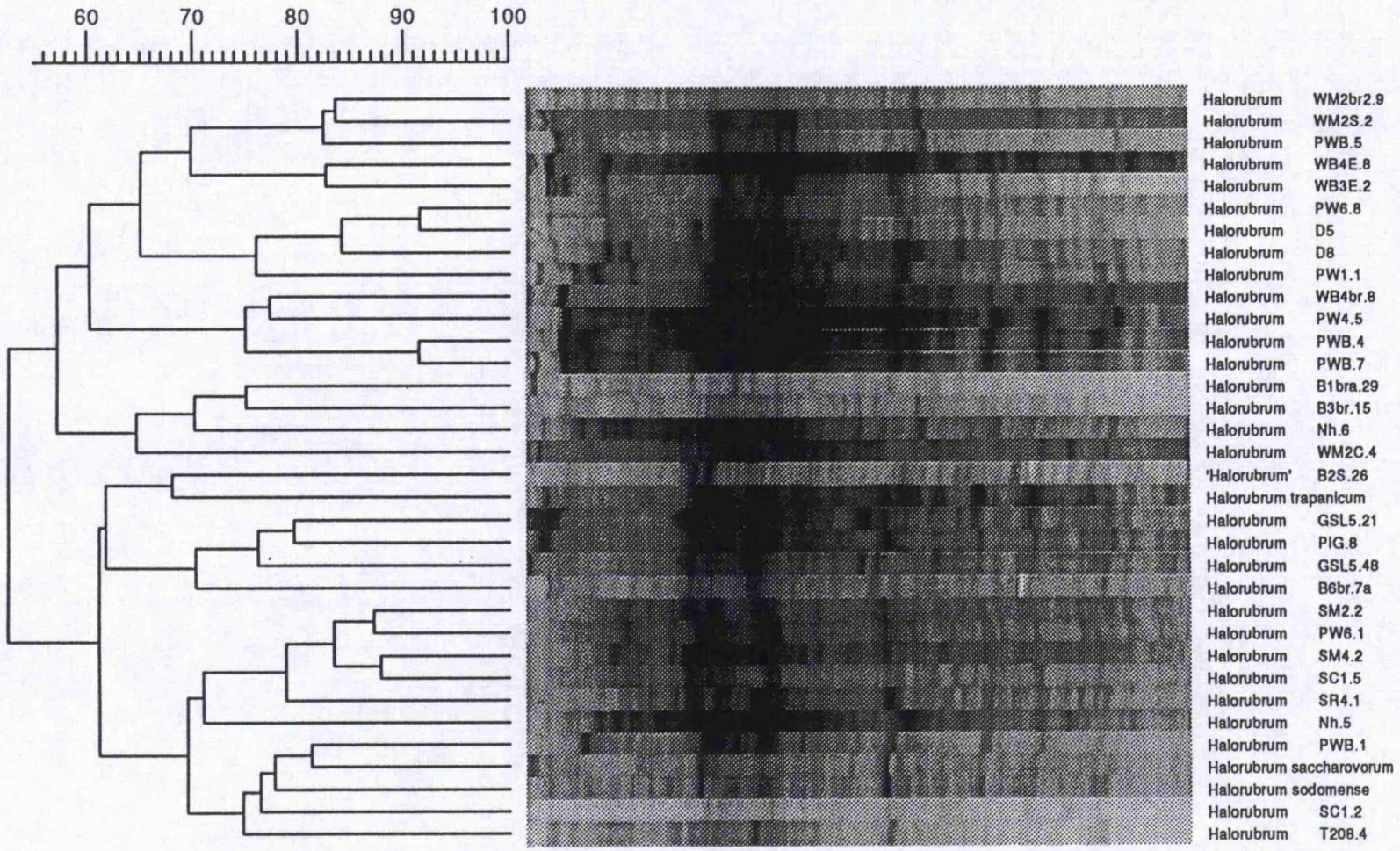


Figure 3.8. Dendrogram of *Halorubrum* -like Isolates, based on Protein Electrophoresis Patterns.

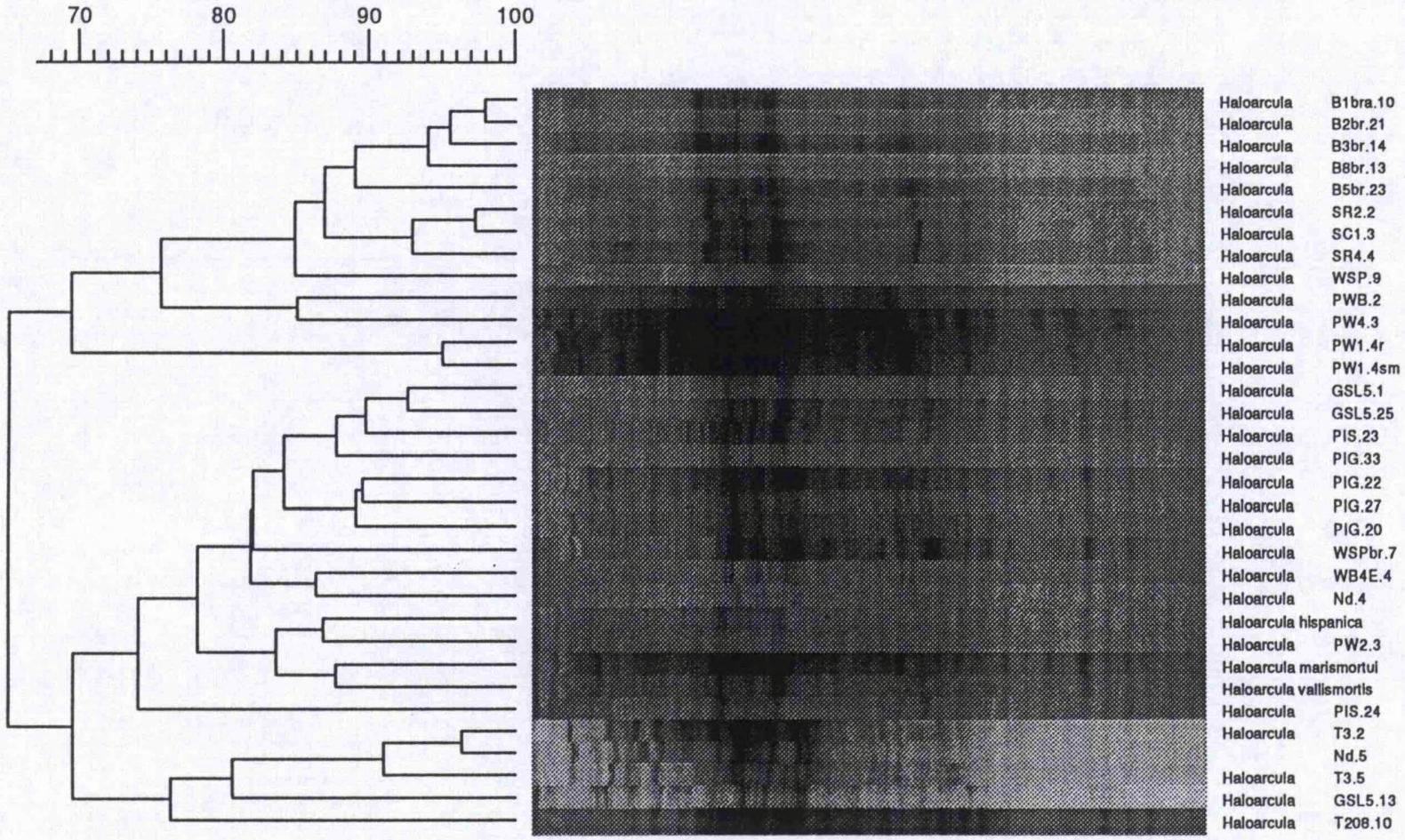


Figure 3.9. **Dendrogram of *Haloarcula* - like Isolates, based on Protein Electrophoresis Patterns.**

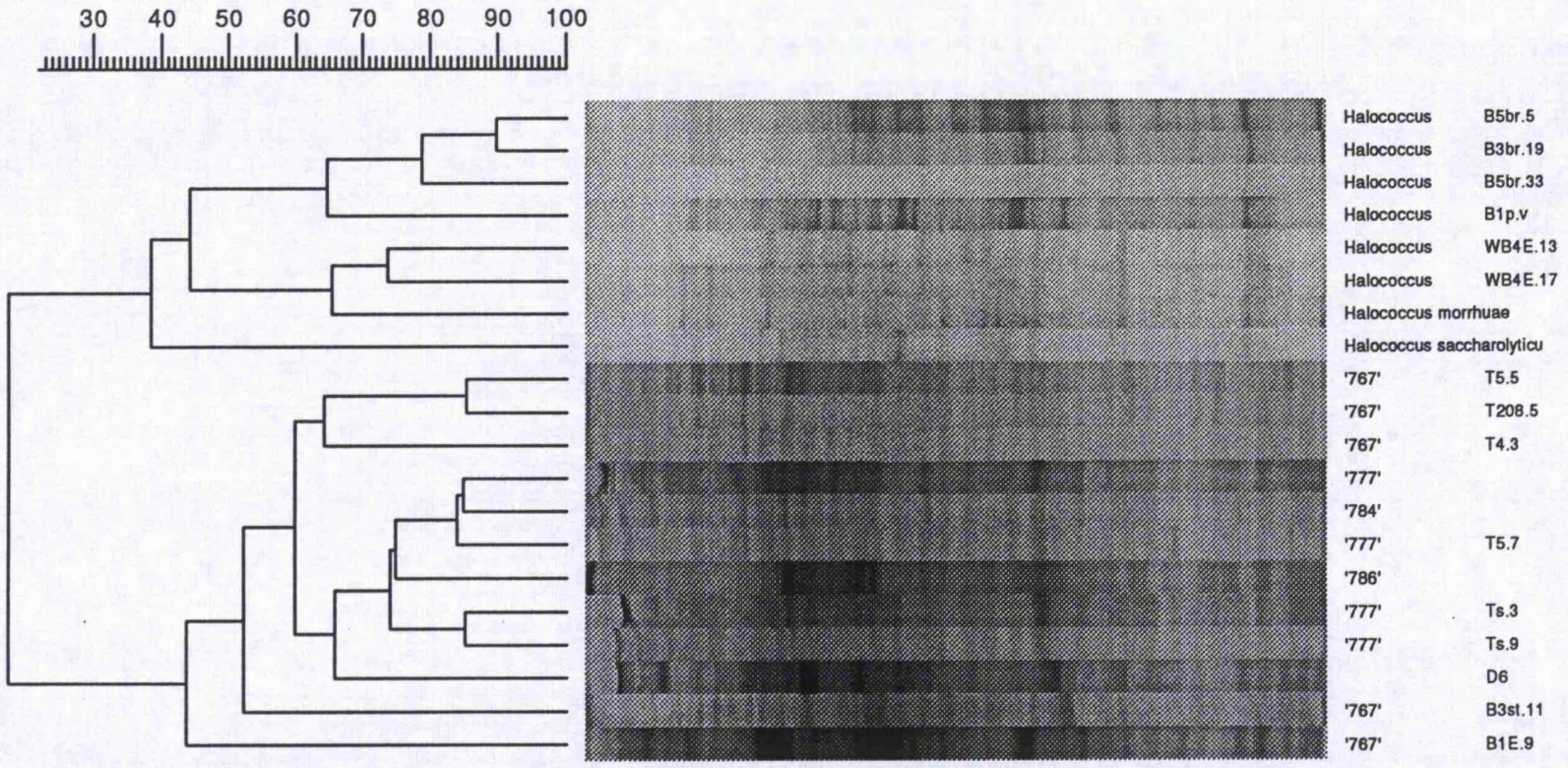


Figure 3.10. Dendrogram of *Halococcus* -like, NCIMB 777-like, and NCIMB 767-like Isolates, based on Protein Electrophoresis Patterns.

***Halorubrum* spp.** (Figure 3.8) clustered overall at 52.6 % (± 12.0 %), with two major sub-clusters at about the 57 % and 61 % similarity levels and several smaller clusters at the 68-76 % level. Irrespective of whether the proteins were run on the same gel, there appeared to be considerable heterogeneity amongst *Halorubrum* isolates not only from within the same salt deposit (*e.g.* Winsford), but from the same sample (PWB). Most strains exhibited less than 90 % similarity to any other.

***Haloarcula* spp.** (Figure 3.9) clustered overall at 65.1 % (± 11.7 %), with two major sub-clusters at about the 69 % similarity level. Isolates from five different sites in Boulby Mine clustered together at 89.0 % (± 0.9 %) similarity. PW1.4r and PW1.4sm were rough- and smooth-colony strains, respectively, of an isolate from the 'Crystal Lake', in Wieliczka (Section 3.1.4); these were similar at the 95.0 % level and together formed a deep branch which diverged from other *Haloarcula* spp. at about the 69 % level, suggesting that PW1.4 is a new species of *Haloarcula. Haloarcula marismortui* and *Haloarcula vallismortis* were 87.7 % similar, and had 83.6 % similarity to *Haloarcula hispanica.*

***Halococcus* spp.** (Figure 3.10) were the most deeply-branching group, at 38.4 % (± 10.2 %), with isolates from Boulby clustering with the Austrian (Permian salt) isolate B1p.v (*Halococcus salifodinae*; Denner *et al.*, 1994) and the Winsford isolates clustering separately, with *Halococcus morrhuae*; *Halococcus saccharolyticus* was the most deeply-branching *Halococcus* sp.

The NCIMB 777-like strains clustered closely together at about the 74 % level (Figure 3.10).
Isolates described as NCIMB 767-like on the basis of polar lipid analysis formed one clear cluster which comprised the Thai isolates; the Boulby isolates branched deeply and separately (Figure 3.10).

Although not very clear on the scanned gel tracks shown with the dendrograms, there appeared to be distinct protein bands associated with the generic groups. These marker bands, which are indicated in Figure 3.11, for the *Halobacterium*, *Halorubrum*, *Haloarcula* and *Halococcus* groups, could be useful in preliminary screening of isolates and assignment to groups.

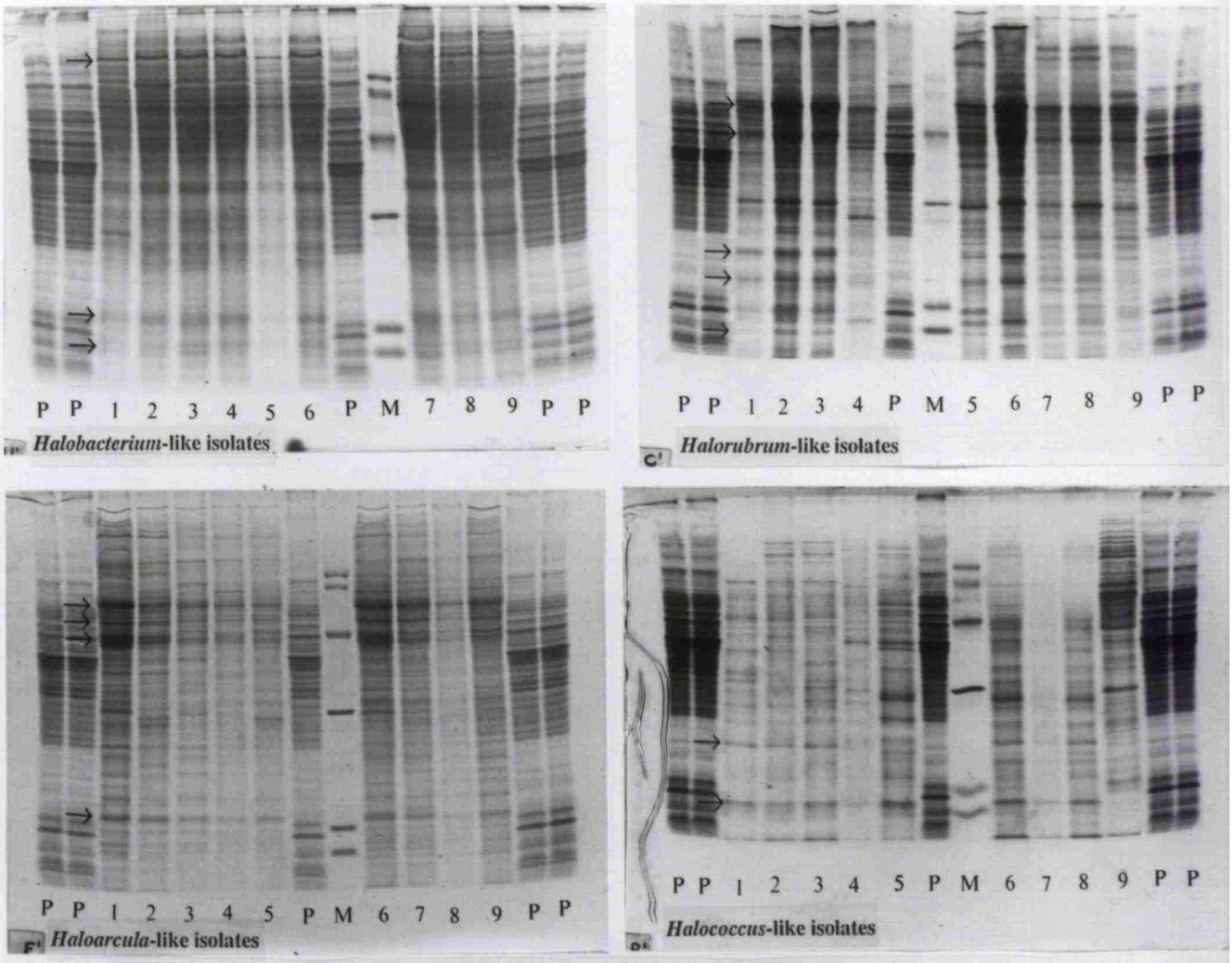


Figure 3.11 **Protein Electrophoresis Patterns of Halobacterial Groups.**

	1	2	3	4	5	6	7	8	9
Halobacterium-like isolates:	B2S.27	B1E.1	B2br.5	B2S.17	B1brc .25	B7br. 1	B4br. 13	B1bra .5	B2br. 11
Halorubrum-like isolates:	PWB.1	PWB.4	PWB.7	PWB.5	PW1. 1	PW4. 5	PW6. 8	D8	D5
Haloarcula-like isolates:	*Ha.mar-ismortui*	*Ha.vall-ismortis*	*Ha.his-panica*	PW2.3	B5br. 23	PIG. 22	PIG. 27	PIG. 33	PIG. 20
Halococcus-like isolates:	WB4E. 13	WB4E. 17	*Hc.mor-rhuae*	*Hc.sacc-harolyt-icus*	B1pV	B5br. 5	B5br. 33	B3br. 19	Ts.11

P = *Pasteurella* sp. reference tracks.
M = molecular weight markers.

← arrows indicate bands which may be group-specific.

On the basis of the entire tree, the results of this analysis of halobacterial isolates were inconclusive. Possibly because the level of similarity defining a species is imprecise within the halobacteria as a family, each generic group sub-divided at different levels, resulting in an intercalation of groups, which therefore had no clear boundaries. When groups were analysed individually, the level of similarity for the deepest node varied from about 38 % (*Halococcus*) to 65 % (*Haloarcula*) and, as mentioned, levels of sub-clustering varied; so generic or species groupings could not be determined empirically by this method. This apparent variability may to some extent be exaggerated by a lack of uniformity in the gels, as discussed earlier, and may also reflect the high sensitivity of this technique, but nevertheless emphasises the great diversity within the halobacteria. Apart from the *Halobacterium* strains showing some tendency to sub-cluster according to the salt deposits from which they were isolated, the halobacteria did not, as a whole, form distinct geographical groups - there did not, for example, appear to be a *Haloarcula* sp. var. Wieliczka as an electrophoretype distinct from a *Haloarcula* sp. var. Winsford. Nor was there any real indication that the same organism was appearing in different sites (and this was confirmed by a visual comparison of gels), with the possible exception of the *Haloarcula* species T3.2, from Thailand, and the hitherto unidentified Nd.5, from the Netherlands, which showed 96.3 % similarity; and SR2.2 and SC1.3, also *Haloarcula* spp. and with 97.3 % similarity, although both from Almeria, in Spain, were from two different sites. With these exceptions, inter-site isolate similarity coefficients were less than 93 %.

(The *GelCompar* software package was not available until the latter stages of completing this thesis, and so full utilization of its facilities has not been possible, for optimal refinement of the protein profile analysis.)

3.4 ANALYSIS OF 16S rRNA GENE RESTRICTION PATTERNS IN HALOBACTERIA (RIBOTYPING).

3.4.1 Preparation of the Digoxygenin-Labelled 16S rRNA Gene Hybridization Probe.

The DIG-DNA migrated the same distance along the gel as the control samples, to the 1.5 Kb marker level. The PCR reaction therefore appeared to have been a success, although the DIG-labelled DNA was expected to have migrated more slowly. The brightness of the DIG-DNA band appeared to be of equivalent brightness to the control samples, suggesting a similar concentration (approximately 100 ng μl^{-1}).

DIG-labelled probe subjected to the detection protocol gave a good colour reaction after 2-3 hours, with no background staining and the result appeared to be better on the HYBAID-N membrane than on the Boehringer-Mannheim membrane. There was a strong colour reaction for 100 ng to 1 ng probe; 10 pg was just detectable. After remaining in the colour substrate solution overnight there was no change in the level of detection, but background colour appeared.

This showed that the detection procedure was successful and determined the optimal probe concentration for detection. On the basis of this assay, a final probe concentration of 100 ng ml^{-1} (3 μl probe in 3 ml hybridization buffer) was the minimum concentration for a satisfactory colour reaction. However, it could not be concluded from these results that the detection was of DIG-labelled 16S rRNA PCR product. It could be that unincorporated DIG-11-dUTP had been detected. To show that the DIG-11-dUTP was successfully incorporated into the halobacterial 16S rDNA during the PCR reaction it was then necessary to show detection after hybridization of the probe to halobacterial DNA, in dot blots.

16S rDNA was detectable to a concentration of 100 pg with 1 μl (but weakly), 3 μl or 5 μl probe in 3 ml buffer. The background was clear of stain. A probe concentration of 100 ng ml^{-1} was chosen (3 μl probe in 3 ml buffer) for the hybridization reactions.

(In retrospect, the DNA should have been heat-denatured prior to dot-blotting. It may be that detection would have been more sensitive, due to more efficient hybridization, if this had been done.)

3.4.2 Halobacterial Ribotypes.

As a result of the *map* survey of restriction sites in the halobacteria, with 'all' restriction enzymes, *Eco*R I was identified as an 'external cutter' for all halobacteria examined and *Taq* I was found to cut frequently within the 16S rRNA gene of most halobacterial sequences.

The number of 16S rRNA gene cleavage sites for known halobacterial sequences, for the restriction enzymes used in this study are given in Table 3.14. For these organisms, therefore, with *Eco*R I and (except for T5.7) *Sal* I digestions, one band represents at least one gene. DNA fragments the same size could appear as one band; Liesack *et al.* found that with *Bam*H I digested genomic DNA of *E. coli*, the seven rDNA operons resolved as three bands (also see Sanz *et al.*, 1988).

Whereas multiple copies of rRNA operons are common in the genomes of Eukarya and Bacteria, this is not the case in the Archaea. In virtually all organisms, however, these copies have identical, or nearly identical sequences (Myvalganam and Dennis, 1992). Studies of gene organization in the halobacteria have shown that they have at least one or two copies (unlinked) of the rRNA operon, which is organized (as in Bacteria) as 5'-16S-23S-5S-3' (Garrett *et al.*, 1991; Dennis, 1993). *Halobacterium cutirubrum* and *Halococcus morrhuae* have one rRNA operon, *Haloferax volcanii* has two identical (or nearly) rRNA operons, but *Haloarcula marismortui* and *'Haloarcula sinaiiensis'* are unusual in each having two heterogeneous 16S rRNA genes (Myvalganam and Dennis, 1992; Kamekura and Seno, 1993). Sanz *et al.* (1988) report findings, in pulsed-field gel electrophoresis (PFGE) experiments, of four copies (minimum) of the 16S rRNA gene in *Haloarcula californiae* and *Haloferax gibbonsii*, and three copies (minimum) each in *'Halobacterium halobium'* NCIMB 777 and *'Halobacterium marismortui'*, but the presumption, based on 16S rRNA sequences of *H. halobium*, *H. volcanii* and *H. morrhuae*, that the restriction enzymes used did not have cut sites, in the six halobacteria investigated, was ill-founded. *Haloferax volcanii* (and *Haloferax gibbonsii)* and *Halobacterium salinarium* (R1) have restriction sites within their 16S rRNA genes for *Bam*H I (Table 3.14).

Table 3.14. Number of 16S rRNA Gene Cleavage Sites for Known Halobacterial Sequences with the Restriction Endonucleases *Taq* I, *Eco*R I, *Sal* I, *Xho* I, *Kpn* I and *Bam*H I.

	Taq I	*Eco*R I	*Sal* I	*Xho* I	*Kpn* I	*Bam*H I
Halobacterium salinarium R1	3	0	0	0	1	1
Halorubrum saccharovorum	6	0	0	0	0	1
Halorubrum lacusprofundi	6	0	1	0	0	1
Halorubrum sodomense	5	0	0	0	0	3
Halorubrum trapanicum	5	0	0	0	2	3
Haloarcula vallismortis	2	0	0	0	0	0
Haloarcula marismortui rrnA	4	0	0	0	0	0
Haloarcula marismortui rrnB	4	0	0	0	0	0
Haloarcula sinaiiensis Major	2	0	0	0	0	0
Haloarcula sinaiiensis Minor	3	0	0	0	0	0
Halococcus morrhuae	3	0	0	1	0	0
Haloferax volcanii	10	0	0	2	0	1
'Halobacterium halobium' 777	2	0	0	0	0	2
'Halobacterium trapanicum' 784	2	0	0	0	0	2
'Halobacterium salinarium' 786	2	0	0	0	0	2
Natronobacterium magadii	1	0	0	0	0	2
T5.7	3	0	1	0	0	2
SC1.2	5	0	0	0	2	3
GSL5.48	5	0	0	0	2	3
PW1.1	6	0	0	0	0	1
WB4E.8	6	0	0	0	0	1
Nh.5	6	0	0	0	0	1
PIG.8	5	0	0	0	0	1
B2S.26	6	0	0	1	0	1
PW5.4	5	0	0	0	0	1
WM2C.13	5	0	0	0	0	1
B1bra.5	5	0	0	0	0	1
T1.6	3	0	0	1	0	3
PIG.5	3	0	0	1	0	1
T3.2	3	0	0	0	0	0
T3.5	4	0	0	0	0	0
T4.2	2	0	0	0	0	0
T3.1	2	0	0	0	0	0
T3.6	2	0	0	0	0	0
SR1.5	3	0	0	0	0	2
T1.3	6	0	0	2	0	1
Nh.2	5	0	0	0	0	1

The results of the ribotyping are given below for the halobacterial groups recognized by polar lipid analysis.

Halobacterium-like.

All *Halobacterium*-like strains produced one band with *Eco*R I and (except B7br.1, which had 2 bands) with *Sal* I (Figure 3.12a,b). The sequenced strains, PW5.4, WM2C.13 and B1bra.5, also produced a single band with *Xho* I (Figure 3.12a). **This suggested that *Halobacterium*-like strains have one copy only of the 16S rRNA gene,** and supported earlier findings that the '*Halobacterium halobium*' and '*Halobacterium cutirubrum*' (now regarded as synonymous with *Halobacterium salinarium*) genomes each had single-copy rRNA genes in an operon (Hofman *et al.*, 1979; Hui and Dennis, 1985).

The *Taq* I digested bands were not distinct, possibly because the fragment sizes were small. At least four bands were expected (six for PW5.4, WM2C.13 and B1bra.5), whereas one, two or three were observed. This could be explained by small fragments being lost from the gel; the smallest fragment in *H. salinarium* and the sequenced *Halobacterium*-like strains would have been about 50 bases (from the *map* results). Also, if the fragment sizes were similar, they would appear as one band.

In comparing the ribotype patterns it could be concluded that:

1. **Boulby isolates B1E.1 and B1E.24 belonged to the same ribotype, as did B2br.5 and B2br.11; but this was not unexpected for isolates from the same sample.**
2. **Winsford isolates WM2C.11, WB3E.9, WSPbr.8 and WSP.14 also appeared to have highly similar, if not identical banding patterns.**
3. **With the few exceptions given above, all the *Halobacterium*-like isolates were unique; in no case was the same type found in different geographical sites. (The sequenced isolates PW5.4 and WM2C.13 had the same-sized bands with *Eco*R I, *Sal* I and *Xho* I, but patterns differed with *Taq* I.).**

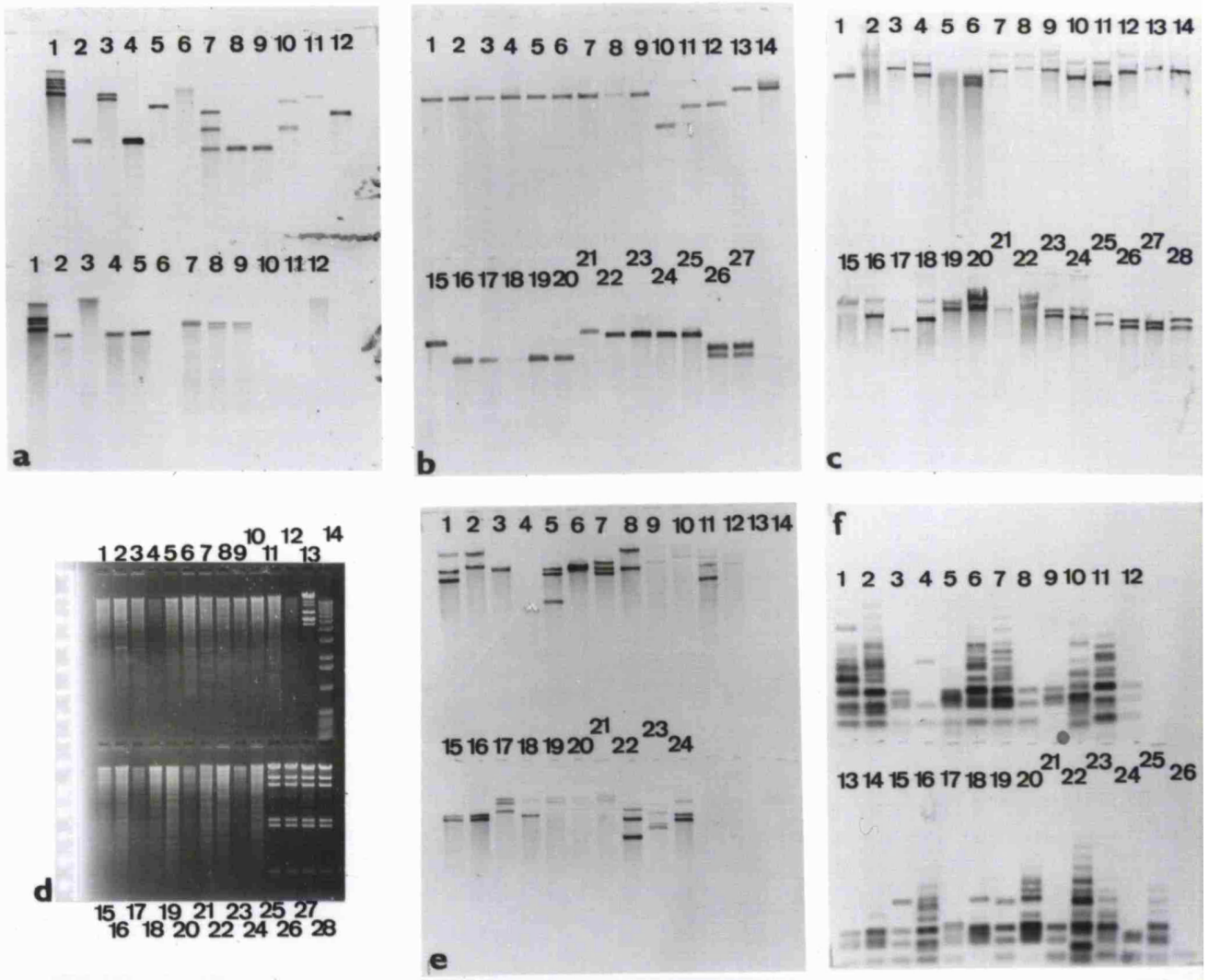


Figure 3.12 Ribotypes of Halobacterial Groups.

a Ribotypes of some of the sequenced isolates after digestion by *Sal* I (top) and *Xho* I (bottom). Isolates 1-12 are as follows: SR1.5, PW5.4, T1.6, WM2C.13, B1bra.5, PIG.5, T5.7, NCIMB 777, NCIMB 784, NCIMB 786, Nh.2 and Nh.5, respectively.

b *Halobacterium*-like isolates (1-25: B1bra.5, B1brc.25, B1E.1, B1E.5, B1E.24, B2br.5, B2br.11, B2S.17, B2S.27, B3st.1, B3st.7, B3st.8, B4br.13, B7br.1, WM2br1.4, WM2C.11, WM2C.13, WB3E.9, WSPbr.8, WSP.14, PIG.31, D1, D8, PW6.1 and WB4E.8, respectively), with one exception (B7br.1), had a single riboband after *Sal* I digestion, indicating a single-copy 16S rRNA gene. *Haloarcula* -like isolates B6br.18 and T208.5 (26 and 27) had two bands with this enzyme.

c *Eco*RI digestion. Isolates 1-28: *Hr. saccharovorum* NCIMB 2081, *Hr. sodomense* ATCC 33755, *Hr.trapanicum* NRC 34021, WB4br.8, WM2S.2, WM2br2.9, WM2C.4, WB3E.2, WB4E.8, B2S.26, SM4.2, D4, SC1.2, PW4.5, PW1.1, PW6.1, D5, Nh.5, D1, T1.3*, T3.1*, PW6.8, SC1.4, SC1.5, SR4.1, GSL5.21, PWB.5 and PIG.8, respectively. *Halorubrum* -like isolates had one or two bands with *Eco*R I, indicating one or two 16S rRNA genes, respectively.

d Successful restriction endonuclease (*Eco*R I) digestion of whole-cell (*Haloarcula*) DNA resulted in smeared tracks of innumerable DNA bands after electrophoresis - apparent after ethidium bromide staining. Tracks 1-24: B1bra.10, B2br.21, B3br.14, B5br.23, B8br.13, WB4E.4, WM2C.3, WSPbr.7, WSP.9, GSL5.1, GSL5.13, GSL5.25, λ DNA, 1 Kb DNA ladder, PW1.4sm, PW1.4r, PW2.3, PW4.3, PIG.20, PIG.27, PIG.33, PIS.24, PIS.27 and SR4.4; (tracks 25-28 were tests of *HIND* III -digested λ DNA).

e 16S rDNA detection of *Eco*R I -digested *Haloarcula* DNA (Figure d) indicated that members of this group have at least two copies of the 16S rRNA gene.

f *Taq* I ribotypes of *Haloarcula*-like isolates. Tracks 1-26: B1bra.10, B2br.21, B3br.14, B5br.23, B8br.13, WB4E.4, WM2C.3, WSPbr.7, WSP.9, GSL5.1, GSL5.13, GSL5.25, PW1.4sm, PW1.4r, PW2.3, PW4.3, PWA.1, PIG.20, PIG.27, PIG.33, PIS.23, PIS.24, PIS.27, SR2.2, SR4.4 and T5.6*.

* Unidentified isolates.

Halorubrum-like.

With the enzymes *Eco*R I (Figure 3.12c) and *Sal* I, different *Halorubrum*-like isolates from within any site produced either one or two bands. As sequenced *Halorubrum* 16S rRNA genes have been shown not to have internal cut sites for these enzymes this suggested that **some of these isolates had two copies (at least) of the 16S rRNA gene.** *Halorubrum saccharovorum* (2081) had two bands with *Eco*R I; *Halorubrum trapanicum* (34021) had one. The two 16S rRNA genes shown in some *Halorubrum* species in this study are probably identical, or virtually identical. Nuttall and Dyall-Smith (1993) reported three separate, non-heterogeneous 16S rRNA genes in the *Halorubrum* sp. Ch2; and Kamekura and Dyall-Smith (1995) analysed ssDNA sequences for *Hr. saccharovorum*, *Hr. lacusprofundi* and *Hr. distributum*, from eight to ten clones containing PCR-amplified 16S rRNA genes, and did not report any heterogeneity in the16S rRNA genes.

At least six bands were expected in the ribotype patterns resulting from *Taq* I digestion, but again there were less than expected, and, as for the *Halobacterium* ribopatterns, this was probably due to small fragments being lost from the gel. *Halorubrum* spp., unlike other halobacteria, have *dam* methylated DNA (Lodwick *et al.*, 1986) and this may have been resistant to digestion at some sites by *Taq* I, which is blocked by overlapping *dam* methylation (*i.e.* GATCGA), but an analysis of the sequenced *Halorubrum* genes showed there were no GATCGA sequences within the 16S rRNA gene.

In comparing the ribotype patterns it could be concluded that:

1. **D1 and D8, from Denmark, appeared to be the same ribotype.**
2. **All *Halorubrum* isolates seemed to be unique, with no instances of the same type appearing in different sites.**

Haloarcula-like.

Most *Haloarcula*-like isolates had ribotype patterns with at least two bands with the 'external cutters' *Eco*R I (Figure 3.12e) and *Sal* I, indicating that **halobacteria in this group generally have at least two copies of the 16S rRNA gene.** It has been shown by others that *Haloarcula marismortui* (Mylvaganam and Dennis, 1992) and '*Haloarcula sinaiiensis*' (Kamekura and Seno, 1993) have two heterogeneous 16S rRNA-encoding genes. These are non-adjacent in *Haloarcula marismortui*.

The *Taq* I digest ribotype patterns had well-defined banding (Figure 3.12f), which was useful in differentiating strains. The fewer than expected bands could be explained by the reasons given earlier for *Halobacterium* and *Halorubrum Taq* I ribotypes.

In comparing the ribotype patterns it could be concluded that:

1. **PW1.4r and PW1.4sm, rough- and smooth-colony strains, respectively, of an isolate from the 'Crystal Lake', in Wieliczka, appeared to be identical.**
2. **T_{208}.9 and T_{208}.10, isolated from a disused saltern in Thailand, were the same ribotype.**
3. ***Haloarcula* isolates were unique; the same type did not appear in other sites.**

Halococcus.

Sufficient DNA was extracted from only three of these isolates, one of which clearly had two bands with *Eco*R I, suggesting that ***Halococcus* species may have more than one copy of the 16S rRNA gene.** These results were insufficient to draw any further conclusions.

'Halobacterium sp.' NCIMB 777-like.

The 777-like strains produced one or two bands with *Eco*R I, and at least one band with *Sal* I and *Xho* I (Figure 3.12a), suggesting that **some members of this group have more than one copy of the 16S rRNA gene.** Sanz *et al.* (1988) reported three copies (minimum) in *'Halobacterium halobium'* NCIMB 777 (see above).

In comparing the ribotype patterns it could be concluded that:

1. ***'Halobacterium* sp.' NCIMB 777 and *'Halobacterium* sp.' NCIMB 784 belonged to the same ribotype.**
2. **Ts.7 and Ts.9, from a Thai saltern soil sample, appeared to have identical ribotype patterns; the results were not clear with the other isolates, all from Thailand.**

Natronobacterium-like.

The results were not very clear, but the six haloalkaliphile isolates each had different banding patterns with *Eco*R I. WTg1 had four bands with this enzyme, suggesting that, unless it has internal *Eco*R I restriction sites, it has multiple copies of the 16S rRNA gene. *Natronobacterium magadii* has a single rRNA operon (Lodwick *et al.*, 1991).

One band in the ribotype may represent more than one equal-sized DNA fragment and this should be considered when interpreting results. With this proviso, *Eco*R I is a suitable 'external cutter' for the determination of the minimum number of 16S rRNA gene copies in the halobacteria isolated from these salt deposits. Most of the halobacteria studied here had at least two copies of the RNA gene, which may or may not be identical.

Taq I digests did not produce very successful ribotypes; the bands were poorly-defined and inconsistent when repeated. These results may have been improved by optimizing the digest times; however, incubations were long to make partial digestion unlikely.

The ribopatterns produced here did not have sufficient bands for satisfactory typing of strains. Ideally, six to eleven bands give a suitable ribopattern for typing (Brown and Ison, 1993; Owen *et al.*, 1993; Dolzani *et al.*, 1994). The number of detectable fragments could have been increased by using a longer probe, such as a 16S+23S rDNA probe. Dolzani *et al.* (1994) based their ribotyping on the detection of fragments of the highly polymorphic 16S-23S rRNA intergenic spacer regions, which allowed differentiation at the genus, species and sub-species level. Other genes could be probed, but they must be in multiple copies in the genome to give complex banding patterns; and to be taxonomically valuable the gene must be conserved. If necessary, two or three restriction enzymes can be combined in the digest to increase the resolution of the profile. A cocktail of enzymes is a possibility if the same buffer can be used, for example with *Eco*R I and *Bam*H I.

Analysis of the 16S rRNA gene restriction patterns has shown that halobacteria exhibit great genetic diversity, not only between genera, but also within each group. Only a few instances were found of isolates sharing identical ribotype patterns, and these had originated from the same sample, although four *Halobacterium* isolates from different samples, but the same salt deposit (Winsford), were highly similar. This latter may be clarified with other enzymes, as there were instances of isolates from different sites (*e.g.* PW5.4 and WM2C.13) having the same banding patterns with most, but not all, of the enzymes used.

3.5 PHYLOGENETIC ANALYSIS.

The halobacteria isolated for the purposes of this study were taken from salt deposits of diverse geographical locations and a spectrum of geological ages. In addition to three as yet unsequenced culture-collection strains, a selection of twenty-one halobacteria, representing each of the sites studied, were chosen for phylogenetic analysis, with two aims in mind. Firstly, supported by the polar lipid analyses (Section 3.2) and protein profiling by PAGE (Section 3.3), a greater appreciation of the diversity of halobacteria in relation to geography and geology, and possible identification of new taxa. Secondly, to test the hypothesis that (some of) the halobacteria isolated from the salt mines may be autochthonous - phylogenetic analysis may be one means of demonstrating this. Working under the hypothesis that if a molecular chronometer is used to derive a phylogenetic tree of halobacteria from halites of a spectrum of geological ages, the degree of relatedness between these halobacteria should be proportional to geological age. Thus, an isolate from a 235 Ma deposit should branch closer to the root of the tree than an isolate from a 20 Ma deposit, and branches of 'living fossils' would be mere 'blips' on an extant lineage (assuming it is of the same lineage), frozen in evolutionary time. The molecular chronometer used in this study was the 16S rRNA gene, for the reasons given earlier (Section 1.5.5).

Following extraction of DNA from the organisms (listed in Table 3.17), PCR amplification of these DNA extracts using 16S rRNA gene primers, and purification of the PCR product using the *QIAquick-spin PCR Purification Kit*, the products were electrophoresed in 1 % agarose gels to check for purity and successful amplification of a (approximately) 1.5 Kb product. The result of one such run can be seen in Figure 3.13. The 16S rRNA gene was successfully amplified from DNA extracts from all strains. These cleaned PCR products were then of the quality required for gene sequencing.
The 16S rRNA genes were sequenced using the PRISM™ Ready Reaction DyeDeoxy™ Terminator Cycle Sequencing Kit and Protocol developed by *Applied Biosystems ABI*, with internal primers, as described in *Materials and Methods*. The cleaned sequencing products were processed by the PNAC laboratory and the results returned as sequence chromatograms, as previously described. Typical *ABI SeqEd*™ chromatogram traces are shown in Figures 3.14 and 3.15.

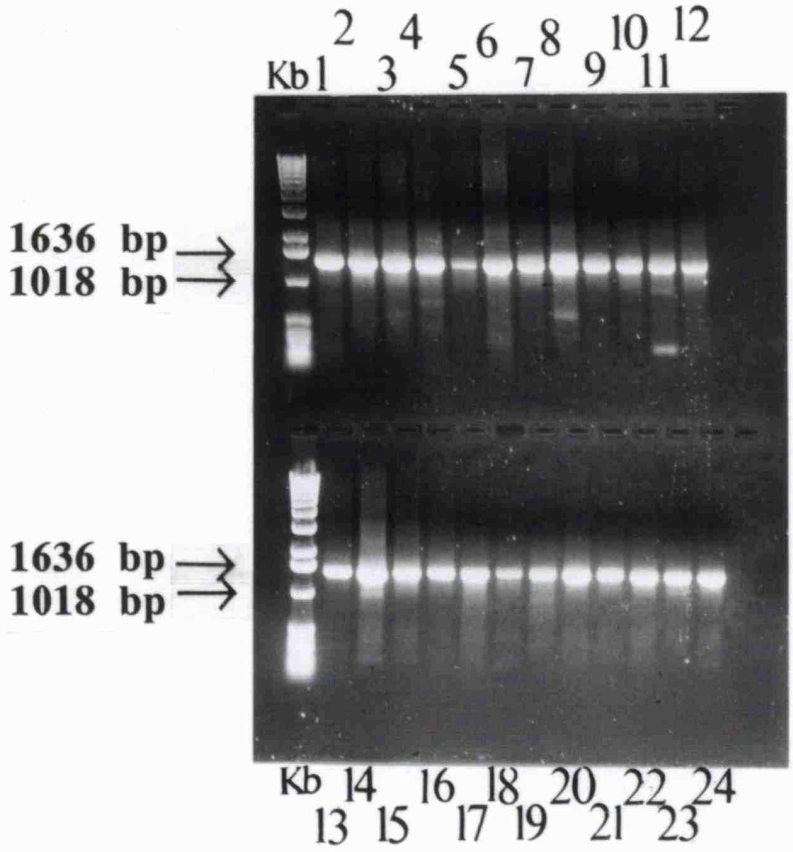


Figure 3.13 16S rRNA PCR Products.

1 Kb DNA ladder; 1- SC1.2; 2- PW1.1; 3- T1.3; 4- WB4E.8; 5- B2S.26; 6- PIG.8; 7- SR1.5; 8- PW5.4; 9- T1.6; 10- WM2C.13; 11- B1bra.5; 12- PIG.5; 1 Kb DNA ladder; 13- T3.1; 14- T3.2; 15- T3.5; 16- T3.6; 17- T4.2; 18- T5.7; 19- Nh.2; 20- Nh.5; 21- GSL5.48; 22- NCIMB 777; 23- NCIMB 784; 24- NCIMB 786.

Products of 16S rRNA gene amplification by the PCR reaction, cleaned using the *QIAquick-spin PCR Purification Kit*. Electrophoresis alongside a 1 Kb DNA ladder indicated that the primary product in each case was a double-stranded DNA fragment of approximately 1.5 Kbp.

Fig. 3·14

T4·2 GR

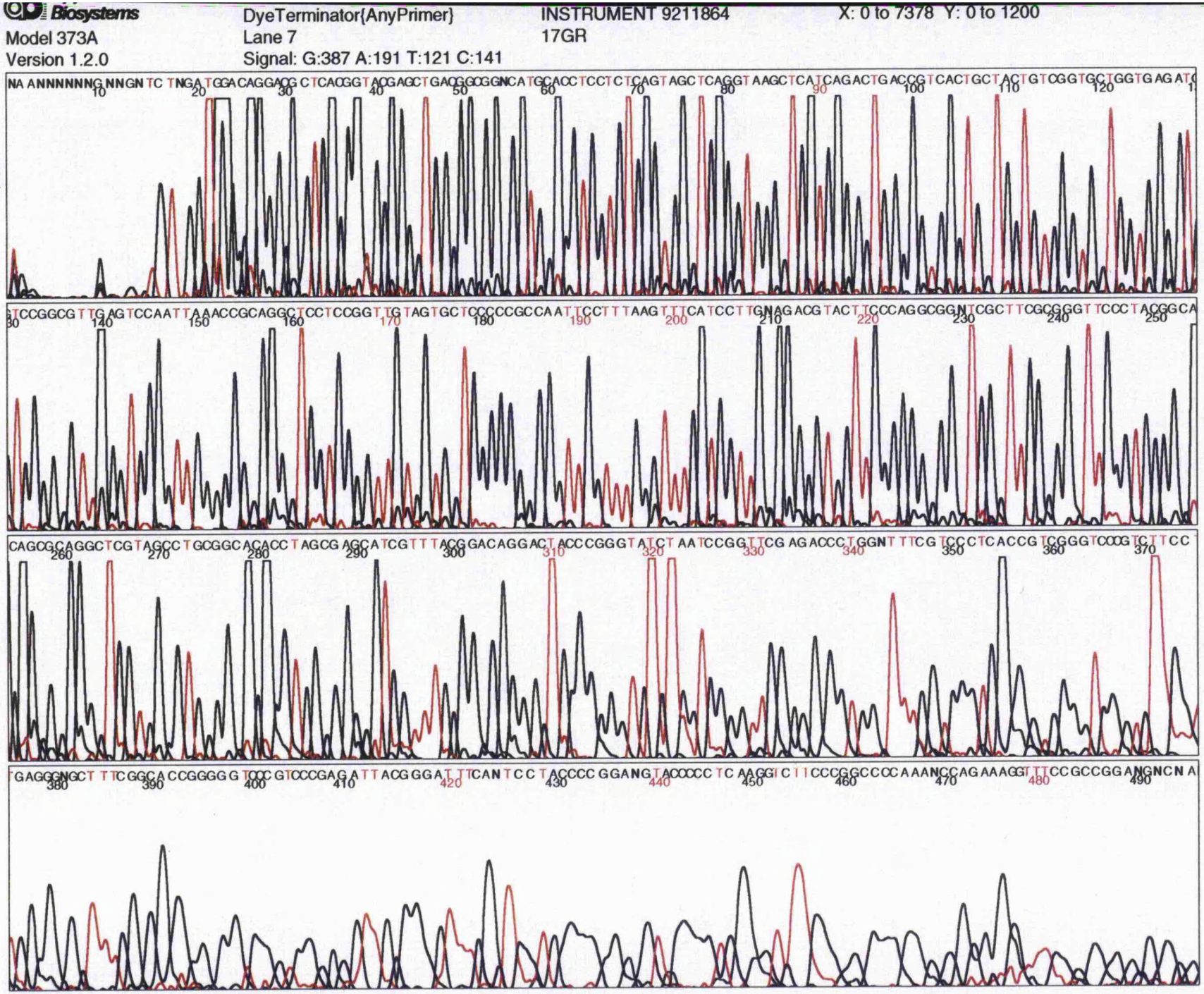
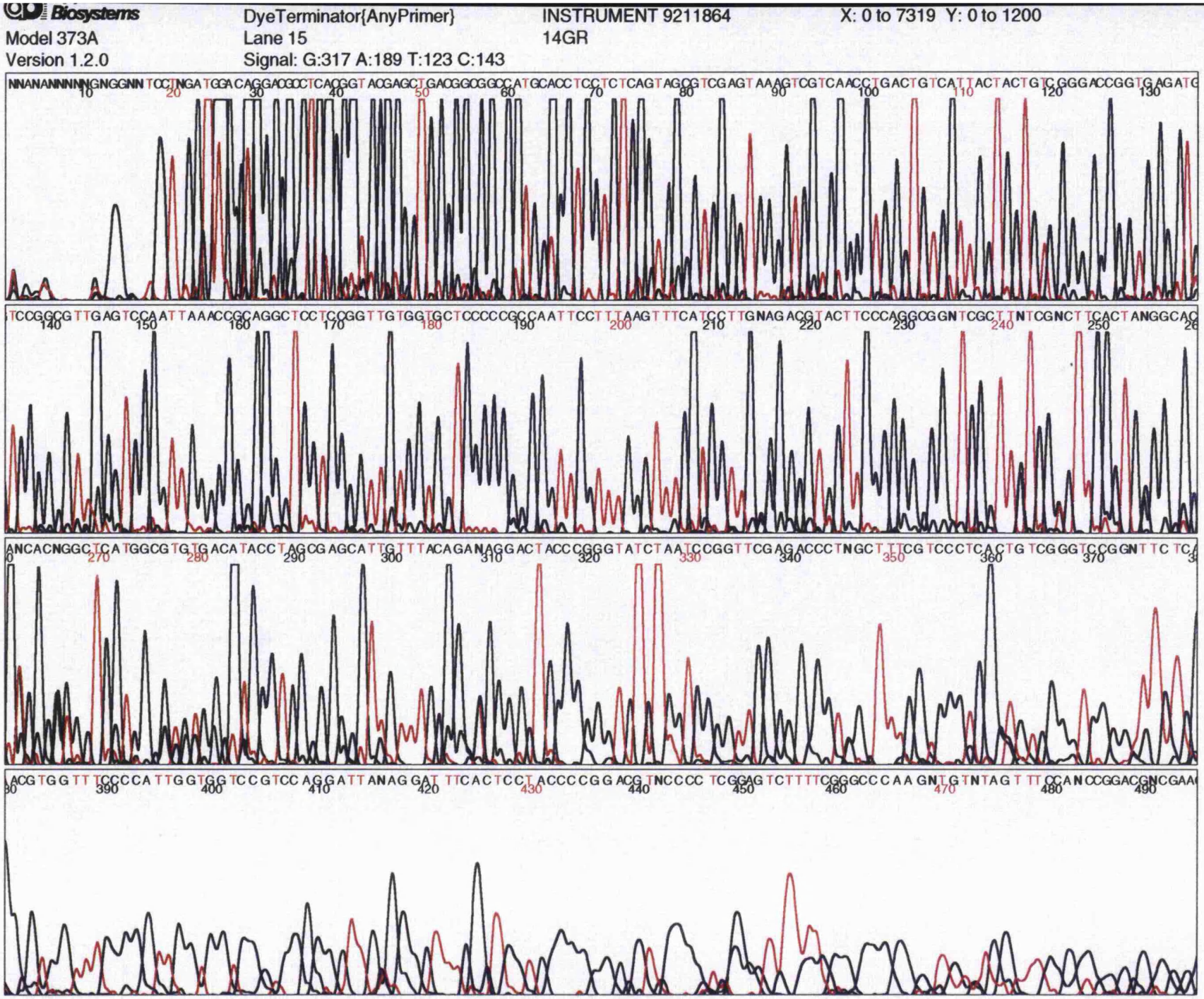


Fig. 3·15

T3·2 GR



3.5.1 Base Ambiguities.

When the aligned sequences (x.msf) were compared to the chromatograms it could be seen that certain regions of the 16S rRNA gene were consistently poor for base-reading. Positions which were 'N' for most sequences were: 48, 101, 105, 213, 216, 219, 458, 468, 480, 1002, 1043, 1046, 1072, 1092, 1113, 1119, 1120, 1123, 1137, 1141, 1144 and 1159 (sequence alignment numbering; Appendix 1). When these positions were ambiguous or unreadable and the remainder of the sequence consisted of clear, readable peaks, it suggested several possibilities: (i) a feature of the sequencing method (DyeDeoxy™ Terminator Cycle Sequencing and determination by automation), (ii) that base modification (*e.g.* methylation) impeded sequencing, (iii) that there was some interference due to the secondary structure of the molecule, or (iv) that perhaps two heterogeneous genes were simultaneously being sequenced.

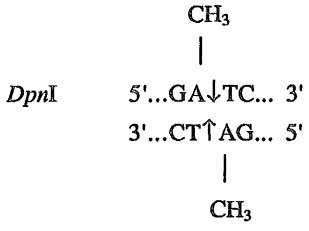
(i) The Sequencing Method.

Taq DyeDeoxy™ Terminator Cycle Sequencing is known to produce low C or T peaks after a G peak and a low third G in a string of Gs, so it is quite possible that other sequence patterns may cause problems. Careful analysis of those parts of the sequence containing Ns reveals the frequent occurrence of the trinucleotide sequences 'CCG', 'CCA' or 'GTG' flanking the N; however these occur elsewhere in the sequence without causing problems. Although it is possible certain juxtaposed bases give a poor read, it is not clear here if sequence context is responsible for the Ns.

(ii) *Dam* Methylation.

Organisms which are Dam^+ produce the enzyme DNA adenine methylase, which methylates the adenine residue in the DNA sequence GATC. This property is characterized by inhibition of the restriction endonuclease *Mbo*I, which cleaves 5' to the G of a non-*dam* methylated sequence GATC, and by the specificity of *Dpn*I, which cleaves 3' to the A in GATC only when the A is methylated:

*Mbo*I 5'...↓GATC ... 3'
3'... CTAG↑... 5'



Within the halobacteria, only members of the genus *Halorubrum* have been found to have *dam* methylated DNA (Lodwick *et al.*, 1986).

All of the halobacterial sequences analysed here have (when the sequence was readable) a GATC sequence at both the 5' (10-13) and 3' (1471-1474) ends of the gene. Representatives of the *Halorubrum* group have, in addition, up to 6 internal GATC sites, whereas other genera have 0, 1 or 2 internal sites. In no cases, for *Halorubrum* species or any other, do GATC sites appear associated with regions of base ambiguity. That is not to say, however, that there might not be some other form of base modification.

(iii) Secondary Structure.

When the 'N' bases are located on a halobacterial 16S rRNA secondary structure template (Figure 3.16) there is some pattern to their position. There are three clusters of Ns: (i) where Ns at 101 and 105 oppose those at 212, 215 and 218, on a helix; (ii) between bases 450 and 500, in the region involved in tertiary structure of the molecule; and (iii) in the region between bases 1000 and 1154. This localization suggests DNA folding may be interfering with sequencing, and although this can be a problem with such GC-rich sequences, it should be minimized by the use of a thermally stable polymerase, such as *Taq*.

(iv) Heterogeneous Genes.

Haloarcula marismortui has two non-adjacent ribosomal RNA operons in its genome, *rrnA* and *rrnB*, which are heterogeneous, differing at 74 of 1472 positions. These differences are (except one) restricted to 3 domains of the gene and in one of these domains ('508-823', *Ha. marismortui* numbering), where 51 of the heterogeneities occur, it appears that the *rrnB* gene is diverging 1.5-3 times more rapidly than the *rrnA* gene (Mylvaganam and Dennis, 1992). Similarly, *Haloarcula sinaiiensis* has two heterogeneous 16S rRNA genes, Major and Minor, which differ at 41 of 1468 positions, 25 of which are concentrated in one region ('708-883'; Kamekura and Seno, 1993). The *Haloarcula sinaiiensis* region of pooled heterogeneity therefore overlaps considerably with that found in *Haloarcula marismortui*.

In the two organisms affiliated to the *Haloarcula* group (*Ha. marismortui rrnB*) which were sequenced for this study, T3.2 and T3.5, the majority of the Ns were due to base ambiguities, identified by two superimposed, but readable, chromatogram peaks (Figure 3.15). As shown in Table 3.15, the T3.2 gene has a total of 39 ambiguous base positions: 9 are found in T3.2 alone and 29 are positions where the same ambiguity is found in T3.5. The T3.5 gene has a total of 67 ambiguous base positions, 37 of which are found in T3.5 alone. An additional position exists (*) where both T3.2 and T3.5 show ambiguity, but its nature differs. It seems that these ambiguities are due to co-sequencing of (at least) two heterogeneous 16S rRNA genes in each case.

Table 3.15. T3.2 and T3.5 Base Ambiguities, and the Number of Positions Where These Occur.

	Y (T/C)	R (A/G)	K (T/G)	S (C/G)	M (A/C)	W (T/A)	Total
T3.2 only	2	1+1*	2	3	1	0	9+1
T3.5 only	15	13	4	3	0	2+1*	37+1
T3.2 & T3.5	8	7	3	3	7	1	29

These ambiguous base positions in T3.2 and T3.5 were concentrated in the region 526-847 (Appendix 1), where they numbered 16 in T3.2 and 40 in T3.5. This corresponds to the sites in the *Haloarcula marismortui* ('508-823') and *Haloarcula sinaiiensis* ('708-883') genes in which intra-species 16S rRNA gene heterogeneity is concentrated. The other 23 ambiguous base positions in T3.2 and 27 positions in T3.5 are broadly associated with the remaining clusters of heterogeneity in the *Haloarcula marismortui* sequences ('56-321' and '986-1158'), supporting the observation of Mylvaganam and Dennis (1992) that these differences are not random, but localized. Mylvaganam and Dennis also observed that most of the differences were in helical regions of the secondary structure and were compensatory. Likewise, in the T3.5 secondary structure (not shown), 43 of the 67 ambiguous bases occured in helices; 8 of these were Y - R pairs and 1 was a K - M pair, suggesting some compensatory changes in the heterogeneous genes.

It is evident from these results that the T3.5 16S rRNA genes are more heterogeneous than the T3.2 16S rRNA genes, but it is not possible to guage how the differences are distributed between the genes within each organism.

These observations make it clear that direct sequencing of the PCR product is not a suitable method in cases where the organism has two or more heterogeneous 16S rRNA genes, as all genes are simultaneously sequenced and cannot be individually determined. In such cases, the 16S rRNA genes need to be separated prior to sequence determination. This can be done by digesting the DNA with restriction endonucleases which have no cleavage sites within the 16S rRNA gene (Section 3.4), then either cloning into plasmid vectors, to isolate the individual genes, or separating the digested DNA by electrophoresis and removing the bands identified by a 16S rRNA gene probe.
Of all the organisms sequenced, only the two *Haloarcula* isolates were found to have a significant problem of base ambiguity and this was due to gene heterogeneity. Occasionally, this problem was found in the sequences of other isolates, probably as a result of folding.

3.5.2 Secondary Structures.

When superimposed on the *Haloferax volcanii* template (Figure 3.16), the sequences produced the expected secondary structure conformation and virtually all changes in helices maintained base complementarity (the few which did not were usually adjacent to loops or 'hairpin bends'), confirming the reliability of the sequence data. Changes sometimes lengthened or shortened helices, with the simultaneous effect of narrowing or widening loops. This is illustrated below; changes affecting helix length, *etc.* (relative to *Haloferax volcanii*), are marked in bold.

Lengthened helix / narrowed loop (T4.2; *c.* 157 - 185):

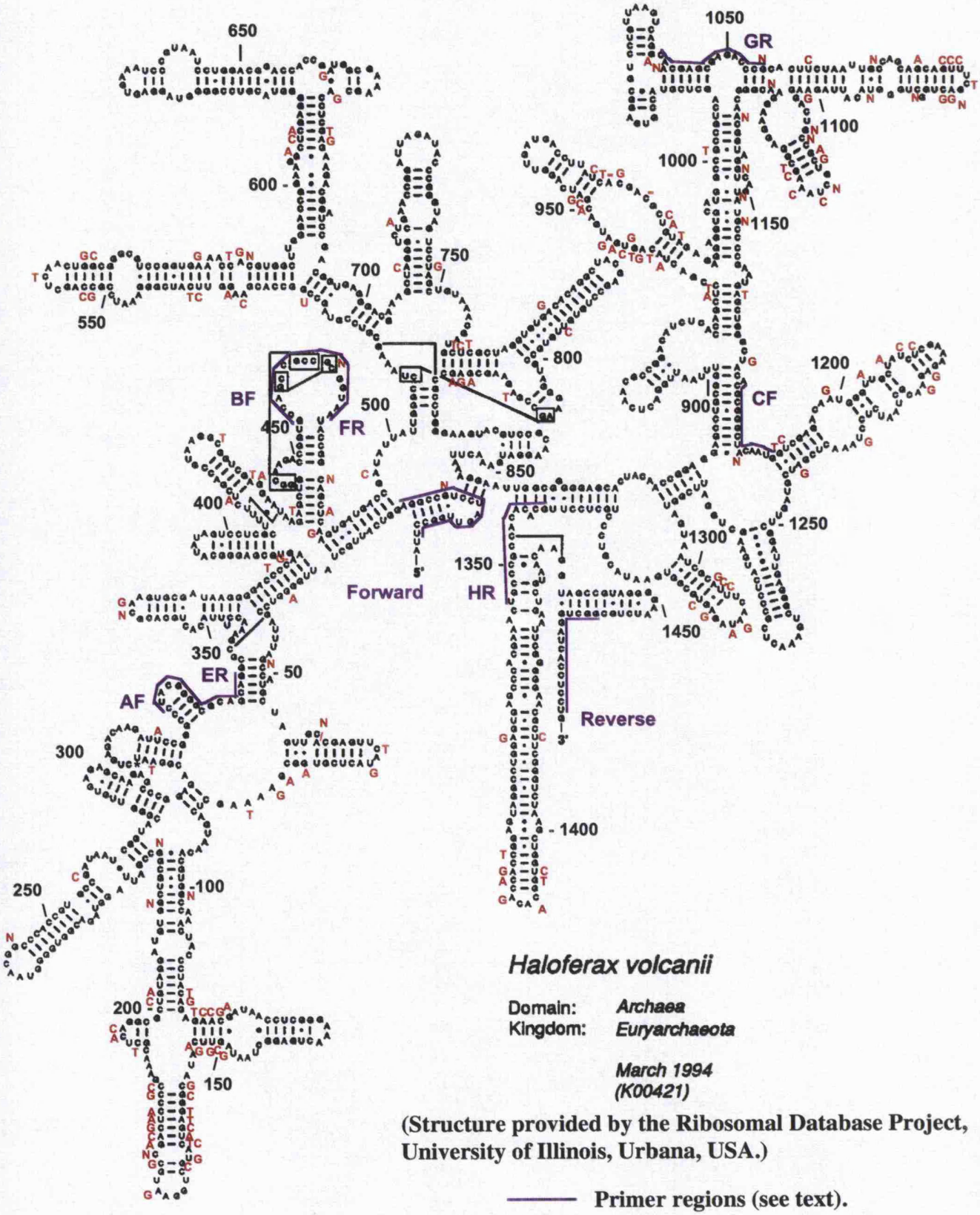
```
                              170
    T A      160         A    T G
       CGGCTCGC  CGC     G
       GCCGAGCG  GCG     A
    A G          180          T G
```

Shortened helix / widened loop (B1bra.5, PW5.4 and WM2C.13; *c.*160):

```
T A A                      170
    C 160      A       CTG
      GCTCC  CCC          G
      CGAGG  GGG          A
    T     180  C        CA
AGG
```

Figure 3.16. 16S rRNA Secondary Structure.

The 16S rDNA sequence of the NCIMB 777-like isolate, T5.7, has been superimposed on the 16S rRNA sequence of *Haloferax volcanii*, by marking (in red) bases which differ.



Haloferax volcanii

Domain: *Archaea*
Kingdom: *Euryarchaeota*

March 1994
(K00421)

(Structure provided by the Ribosomal Database Project, University of Illinois, Urbana, USA.)

——— **Primer regions (see text).**

The 57 - 80 helix (*Haloferax volcanii* numbering) is shortened in the case of Nh.2, by the internal deletion of two base-pairs.

```
G      60
  A  GC
  GTT  ACG..TT T
  CNA  TGC..AA .
  A              70
 C
```

This unusual feature is unique, not being found in any other of the sequences used in these alignments and analyses.

The distribution of constant and variable regions of the 16S rRNA molecule, which can be determined from the sequence alignment, has a more apparent pattern within the secondary structure.

The variable regions of the halobacterial sequences appear to be concentrated around 'hairpin bends', and the helices behind, at six main locations (Appendix 1, sequence alignment numbering; see also Figure 3.16):

1. 115-117 and 200-202; 119-122 and 150-153; 162-167 and 177-186;
2. 529-539 and 583-586; 551-555 and 562-565;
3. 941-958, 963-982;
4. 1075-1088;
5. 1191-1194 and 1238-1242; 1204-1224;
6. 1385-1396.

In the highly conserved regions of the halobacterial sequences, bases are predominantly unpaired, in bulges, loops, 'elbows', some 'hairpin bends', and regions thought to involve tertiary structure interactions (Figure 3.16):
41-47 (involves tertiary structure); 86-100; 129-143 ('hairpin bend'); 233-246 ('hairpin bend'); 258-272; 275-280; 300-308; 314-355 (a side bulge; involves tertiary structure; AF and ER primer region); 405-409 (involves tertiary structure); 431-437 (except the T at 434 is sometimes a C, although this region is involved in tertiary structure interactions); 440-454 (involves tertiary structure; this bulge loop character is conserved in Archaea and

Eukarya; Winker and Woese, 1991); 455-471 (loop involved in tertiary structure interactions; BF and FR primer region); 499-523 (involves tertiary structure); 622-638; 695-707; 718-745 (involves a bulge, short helix and 'hairpin bend'); 806-809 (involves tertiary structure); 822-839 (involves tertiary structure); 842-929 (large stretch including helical structures, bulges and 'hairpin bends'); 984-998 ('elbow'); 1000-1037; 1040-1055 (GR primer region); 1140-1153; 1156-1163; 1173-1190 (CF primer region); 1245-1254; 1270-1274; 1279-1297; 1316-1324; 1326-1339; 1340-1354 (involves tertiary structure; HR primer region); 1355-1358; 1360-1369; 1375-1383; 1401-1407; 1412-1421; 1424-1447 (involves tertiary structure); 1449-1456; 1458-1473 (if sequenced to this length). The 3'-terminal sequence CCUCC (1468-1472), which is present in Archaea, Bacteria and plant mitochondria, is is thought to be involved in mRNA selection (Gutell *et al.*, 1985).

Most of these highly conserved regions are within the secondary structure "core" which appears to be universal across kingdoms (Gutell *et al.*, 1985). For a comprehensive analysis of the comparative morphology of the 16S rRNA secondary structure, see Gutell *et al.* (1985).

3.5.3 Signature Analysis.

The identities of the halobacteria isolated from ancient salt deposits, and their groupings, as determined by phylogenetic analysis (discussed below) were confirmed by the signature sequences of McGenity (1994). These sub-divided the *Halorubrum* isolates appropriately into the *Hr. saccharovorum* / *Hr. lacusprofundi* and *Hr. sodomense* / *Hr. trapanicum* groups. The unknown status of Nh.2, T1.3 and T1.6 was maintained. The PIG.5 signature places it with '*Halobacterium cutirubrum*' NCIMB 763, confirming its identity by polar lipid analysis. T4.2 shares signatures with *Halorubrum*, *Haloarcula* (strongest affinity), *Haloferax*, the '777' group and *Natronobacterium magadii*, and remains *incertae sedis*. T3.1 and T3.6 share signatures with *Halobacterium*, the '777' group, *Halococcus* and *Haloarcula*, and again remain *incertae sedis*.

It is possible to identify additional signature sequences from the alignments, at for example, 550-556 (Table 3.16).

Table 3.16. Signature Sequences for Halobacteria. (Positions 550-556; Appendix 1).

Halorubrum (*saccharovorum* and *lacusprofundi*)*	TCCACCC
Halorubrum (*trapanicum* and *sodomense*)	TCCACCT
B2s.26 and BbpA.3	TCCACGT
Halobaculum	TCTACCA
Nh.2	TCCACTC
Haloferax volcanii	TCCGCCA
Halobacterium salinarium	TCTGTCC
004.1, PW5.4, WM2C.13, B1bra.5, BbpA.1	TCTGTTC
PIG.5	TCTGCTC
'Halobacterium' spp. 777, 784, 786; T5.7, SR1.5	TCCGCGC
'Halobacterium trapanicum' NCIMB 767	ATTCCGC
Natronobacterium magadii	TCTCTTC
Natronococcus occultus	TCTCCCC
Haloarcula marismortui rrnB; T3.2 and T3.5	TCGACCA
T1.3	TCCACGC
T1.6	TCCACAC
T3.1 and T3.6	TCTGCCC
T4.2	TCGGCGC

**Halococcus* ATCC 17082 has an identical signature to *Halorubrum* (*saccharovorum* and *lacusprofundi*), but can easily be differentiated by cell morphology.

Halococcus spp. ATCC 17082 and NRC 16008 have different signatures. The two *Haloarcula marismortui* signature sequences differ, and also differ from *Haloarcula sinaiiensis*. This region is not therefore suitable as a signature for *Halococcus* spp. and *Haloarcula* spp.

3.5.4 Phylogenetic Trees.

3.5.4.1 Tree Topology.

The 16S rRNA gene sequences of the halobacteria isolated from modern and ancient salt deposits, determined for the purposes of this study (Table 3.17), and 16S rRNA gene sequences of halobacteria identified and published in other studies (Table 3.18), were together used to construct halobacterial phylogenetic trees. The *Haloarcula* isolates, T3.2 and T3.5, were included in the analysis, despite the excessive number of Ns in these sequences, which inevitably reduced the information content, but did not result in any misinformation.

Phylogenetic trees were derived from the sequence data by two methods of phylogenetic analysis, as described in *Materials and Methods*, *i.e.* (i) the Distance Matrix method (DNADIST) of Jukes and Cantor (1969), with the tree inference method (FITCH) of Fitch and Margoliash (Fitch and Margoliash, 1967), and (ii) the Maximum Likelihood method (DNAML) of Felsenstein (1981). Four sets of sequences were used to produce a main tree and three sub-trees, all rooted by including the 'sister' outgroup *Methanospirillum hungatei* (Burggraf *et al*, 1991). These were:

(i) a main tree, including all sequences produced in this study and five published genera (Figures 3.17-3.19);
(ii) a sub-tree, with the *Halorubrum* group and outlying strains (Figures 3.20 and 3.21);
(iii) a sub-tree, with the '777' and '*Natrono*' groups, and outliers (Figures 3.22 and 3.23);
(iv) a sub-tree with the *Halobacterium* and *Haloarcula* groups, and outliers (Figures 3.24 and 3.25).

The *Halorubrum* group:

Halorubrum saccharovorum, *Halorubrum sodomense*, *Halorubrum trapanicum*, *Halorubrum lacuprofundi*, Ch2, 54R, GluBr1.1, Malk006.3, GSL5.48, SC1.2, PW1.1, WB4E.8, PIG.8 and Nh.5.

The *Halobacterium* group:

Halobacterium salinarium, 004.1, BbpA.1, B1bra.5, WM2C.13 and PW5.4.

The *Haloarcula* group:

Haloarcula marismortui rrnA, *Haloarcula marismortui rrnB*, *Haloarcula sinaiiensis* Major rRNA gene, *Haloarcula sinaiiensis* Minor rRNA gene, T3.2 and T3.5.

Table 3.17. Halobacteria Chosen for Phylogenetic Analysis.

Isolate	Source	Age of Salt Deposit (Ma)	Identity by Polar Lipid Analysis
B2S.26	'sediment' from brine sump 202, Boulby	250-260	*Halorubrum sp.*
B1bra.5	brine entering sump120, Boulby	250-260	*Halobacterium sp.*
PIG.5	solution mine, Gora Poland	250-260	*'H. cutirubrum'* sp. NCIMB 763
PIG.8	solution mine, Gora Poland	250-260	*Halorubrum sp.*
Nh.2	brine well, Netherlands	250-260	?
Nh.5	brine well, Netherlands	250-260	*Halorubrum sp.*
WB4E.8	efflorescence beside brine pool B4,Winsford	235-240	*Halorubrum sp.*
WM2C.13	salt crystals on brine pool M, Winsford	235-240	*Halobacterium sp.*
T1.3	brine sedimentation tank, Thailand	*c.* 65	?
T1.6	brine sedimentation tank, Thailand	*c.* 65	?
T3.1	solution cavern brine, Thailand	*c.* 65	?
T3.2	solution cavern brine, Thailand	*c.* 65	*Haloarcula sp.*
T3.5	solution cavern brine, Thailand	*c.* 65	*Haloarcula sp.*
T3.6	solution cavern brine, Thailand	*c.* 65	?
T4.2	solution cavern brine, Thailand	*c.* 65	?
T5.7	disused saltern, Thailand	*c.* 65	*'H. halobium'* 777 sp.
PW1.1	Crystal Lake, Wieliczka, Poland	*c.* 20	*Halorubrum sp.*
PW5.4	inflow brine, Wieliczka, Poland	*c.* 20	*Halobacterium sp.*
SC1.2	salt pond, Spain	<1	*Halorubrum sp.*
SR1.5	salt pond, Spain	<1	?
GSL5.48	Great Salt Lake USA	<1	*Halorubrum sp.*
'H. halobium' NCIMB 777	salted cod fish	<1	*'H. halobium'* 777 sp.
'H.trapanicum' NCIMB 784	'Red Heat' in salted hides	<1	-
'H.salinarium' NCIMB 786	'Red Heat' in salted hides	<1	-

Table 3.18. Published Strains Included in the Phylogenetic Analysis of Isolates from Ancient Salt Deposits.

Name	Source of Original Isolation	Strain	Accession no.	Reference
Methanospirillum hungatei		DSM 864 ^{T}	M60880	Yang *et al.*, 1985
Halorubrum saccharovorum	salterns, San Francisco Bay	NCIMB 2081 ^{T}	X82167	McGenity and Grant, 1995
Halorubrum lacusprofundi	Deep Lake, Antarctica	ACAM 34 ^{T}	X82170	McGenity and Grant, 1995
Halorubrum trapanicum	Trapani solar salt, Sicily	NRC 34021	X82168	McGenity and Grant, 1995
Halorubrum sodomense	Dead Sea	ATCC 33755 ^{T}	X82169	McGenity and Grant, 1995
Haloferax volcanii (DS-2)	mud, Dead Sea	ATCC 29605 ^{T}	K00421	Gupta *et al.*, 1983
Haloarcula marismortui rrnA	Dead Sea	Ginzburg strain	X61688	Mylvaganam and Dennis, 1992
Haloarcula marismortui rrnB	Dead Sea	Ginzburg strain	X61689	Mylvaganam and Dennis, 1992
Haloarcula sinaiiensis Major	salt pool, Gulf of Eilat, Sinai	ATCC 33800 ^{T}	D14129	Kamekura and Seno, 1993
Haloarcula sinaiiensis Minor	salt pool, Gulf of Eilat, Sinai	ATCC 33800 ^{T}	D14130	Kamekura and Seno, 1993
Halobacterium salinarium (R1)		DSM 671	M38280	Mankin *et al.*, 1985
Halococcus morrhuae		NRC 16008	D11106	Kamekura and Seno, 1992
Halococcus morrhuae	Dead Sea	ATCC 17082 ^{T}	X00662	Leffers and Garrett, 1984
Halobaculum gomorrense	Dead Sea	DSM9297 ^{T}	L37444	Oren *et al.*, 1995
'Halobacterium trapanicum' 767	Trapani solar salt, Sicily	NCIMB 767	D14125	Kamekura and Seno, 1993
Natronobacterium magadii	Lake Magadi, Kenya	NCIMB 2190 ^{T}	X72495	Lodwick *et al.*, 1991
Natronococcus occultus	Salt pan liquor, Lake Magadi, Kenya	NCIMB 2192 ^{T}	Z28378	McGenity and Grant, 1993
172P1	sea sands, Japan Sea	JCM 9576 ^{T}	D14123	Kamekura and Seno, 1993
Ch2	marine solar saltern, Australia	ACM 3911 ^{T}	L00922	Nuttall and Dyall-Smith, 1993 + 1995
54R	rock salt, Winsford salt mine			McGenity, 1994
GluBr1.1	solution mining brine, Lostock, Cheshire			McGenity, 1994
Malk006.3	brine pool, Winsford salt mine			McGenity, 1994
BbpA.1	brine pool, panel 202 Boulby salt mine			McGenity, 1994
BbpA.3	brine pool, panel 202 Boulby salt mine			McGenity, 1994
004.1	brine pool, Winsford salt mine			McGenity, 1994

Figures 3.17-3.25. Phylogenetic Trees of Extremely Halophilic Archaea from Ancient Salt Deposits. 16S rRNA gene sequences were aligned in CLUSTAL V (or W) and manually adjusted. The phylogenetic trees were derived using either the Distance Matrix method (Figures 3.17, -.18, -.20, -.22 -.24) or the Maximum Likelihood method (Figures 3.19, -.21, -.23, -.25). The outgroup *Methanospirillum hungatei* was included to root the trees. Details are given in the text, Section 2.6.4. Trees were calibrated, with a phylogenetic distance of 0.015 Knuc being equivalent to 50 million years (Moran *et al.*, 1993).

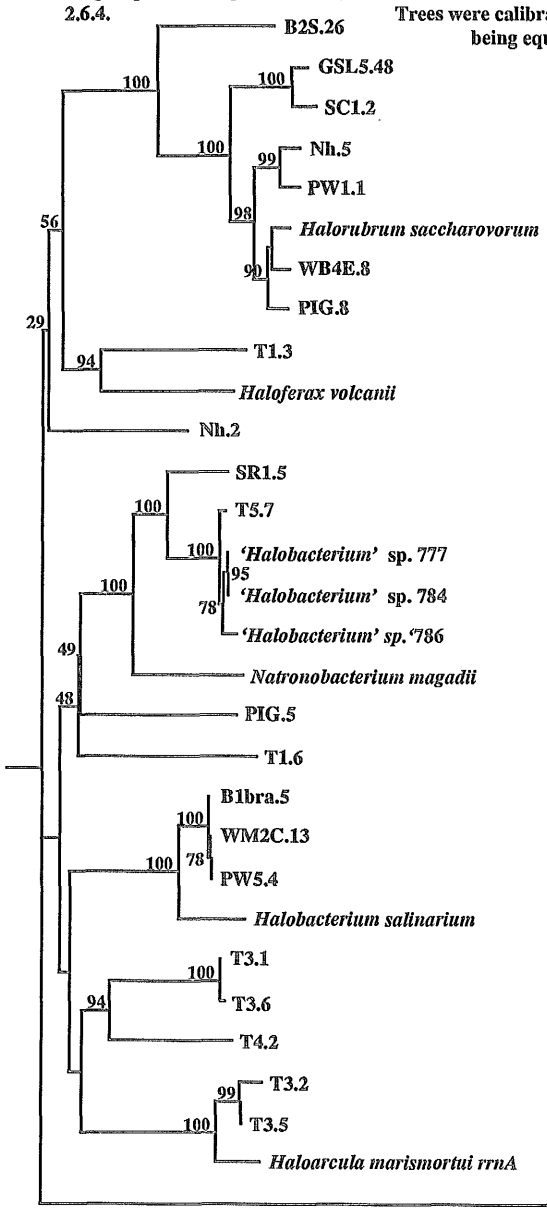


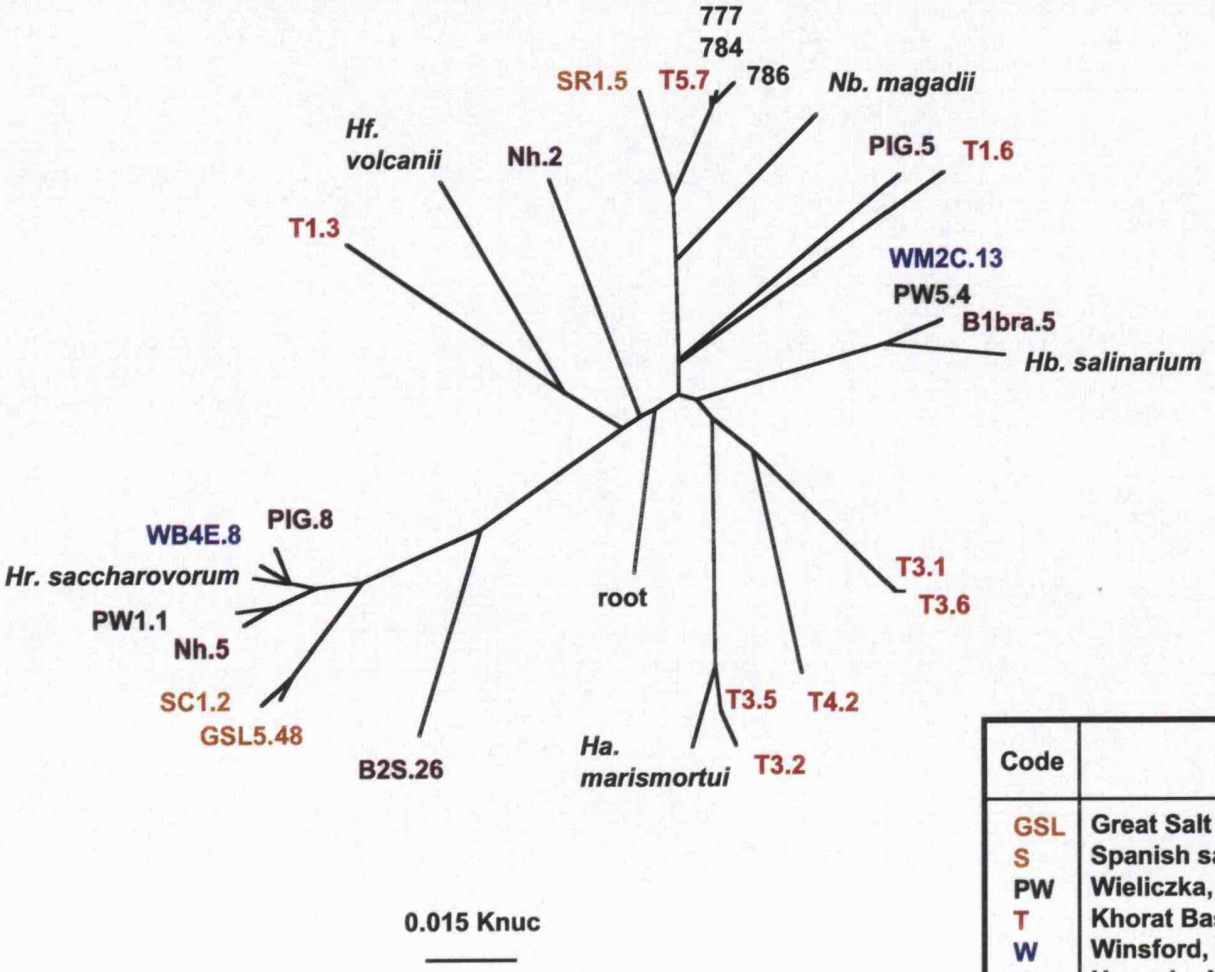
Figure 3.17.

Distance tree of all isolates in *Drawgram* format.
Bootstrap values are given at nodes.

Scale: 0.015 Knuc

Figure 3.18. Phylogenetic Tree of Extremely Halophilic Archaea from Ancient Salt Deposits.

This tree includes all sequenced isolates and the published genera *Halobacterium salinarium*, *Halorubrum saccharovorum*, *Haloarcula marismortui rrnA*, *Haloferax volcanii* and *Natronobacterium magadii* (see text for details).



Code	Site	Geological Age (Ma)
GSL	Great Salt Lake, USA	Recent
S	Spanish salterns	Recent
PW	Wieliczka, Poland	*c.* 20
T	Khorat Basin, Thailand	*c.* 65
W	Winsford, Cheshire, UK	235-240
Nh	Hengelo, Netherlands	250-260
PIG	Gora, Inowroclaw, Poland	250-260
B	Boulby, Cleveland, UK	250-260

Figure 3.19. Phylogenetic Tree of Extremely Halophilic Archaea from Ancient Salt Deposits.
A Maximum Likelihood tree, which includes all sequenced isolates and the published genera, *Halorubrum saccharovorum*, *Haloarcula marismortui rrnA*, *Haloferax volcanii* and *Natronobacterium magadii*.

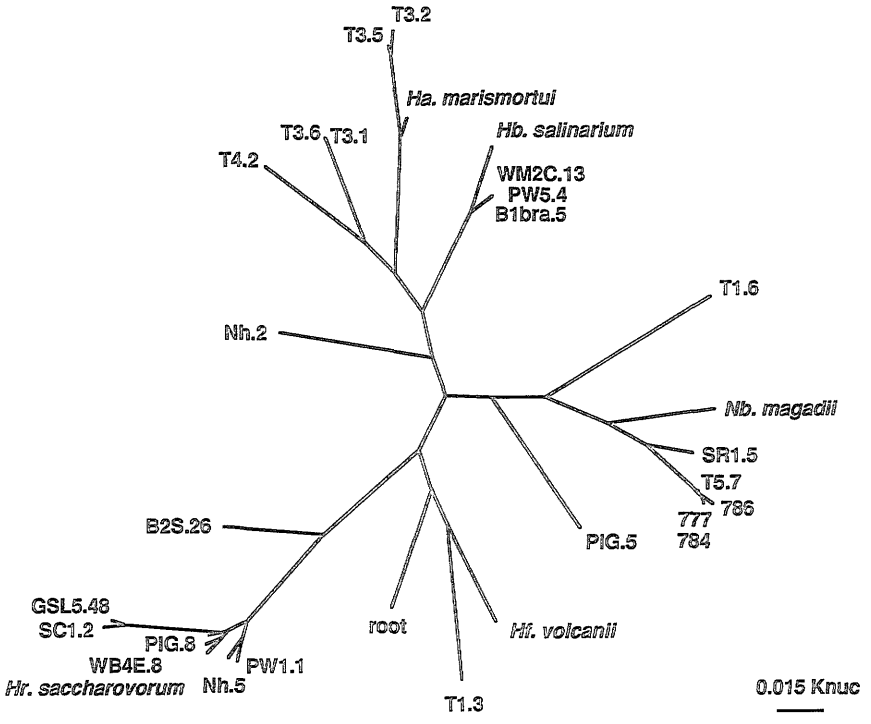


Figure 3.20. Phylogenetic Tree of *Halorubrum*-like Isolates from Ancient Salt Deposits.

A Distance tree, using the distance matrix method (DNADIST) of Jukes and Cantor (1969), with the tree inference method (FITCH) of Fitch and Margoliash (1967). (See Tables 3.17 and 3.18 for information on strains, and text for details.)

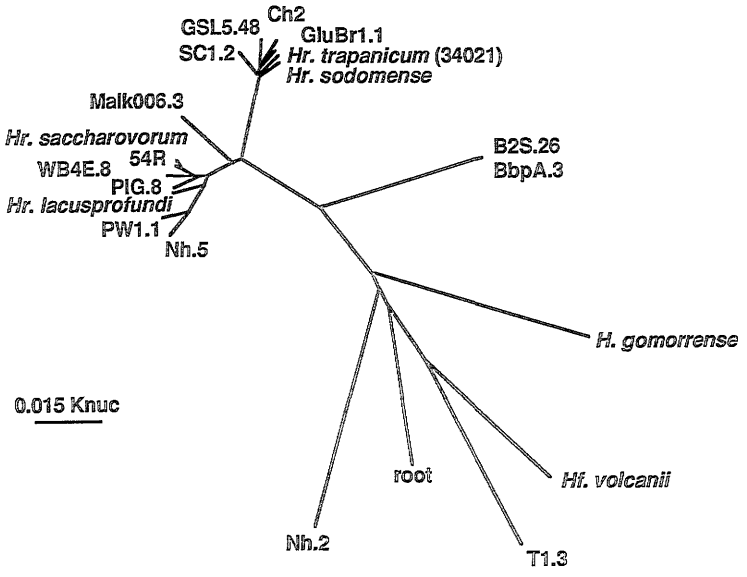


Figure 3.21. Phylogenetic Tree of *Halorubrum*-like Isolates from Ancient Salt Deposits.
A Maximum Likelihood tree of the strains included in Figure 3.20.

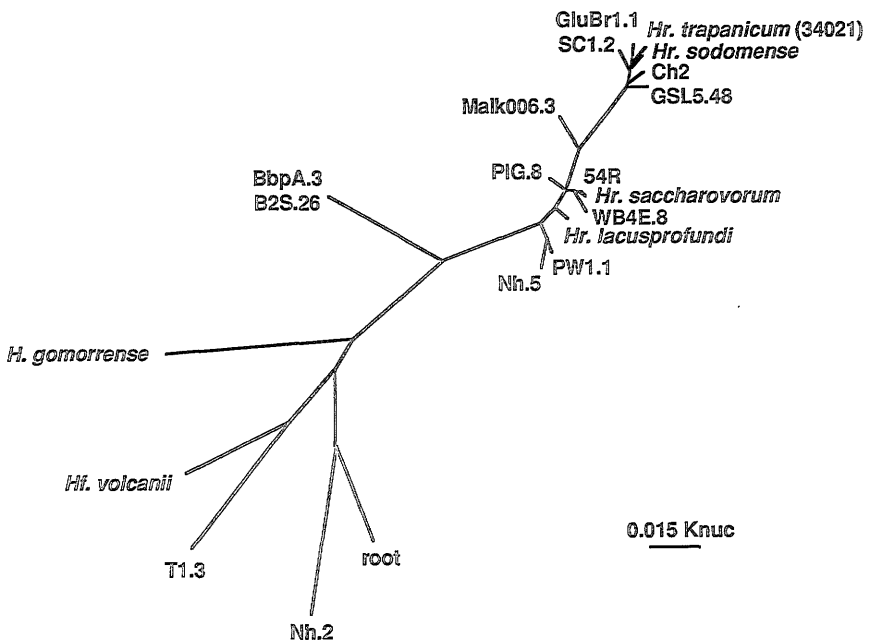


Figure 3.22. Phylogenetic Tree of NCIMB 777-like Isolates (and others) from Ancient Salt Deposits. A Distance tree, using the distance matrix method (DNADIST) of Jukes and Cantor (1969), with the tree inference method (FITCH) of Fitch and Margoliash (1967).
(See Tables 3.17 and 3.18 for information on strains, and text for details.)

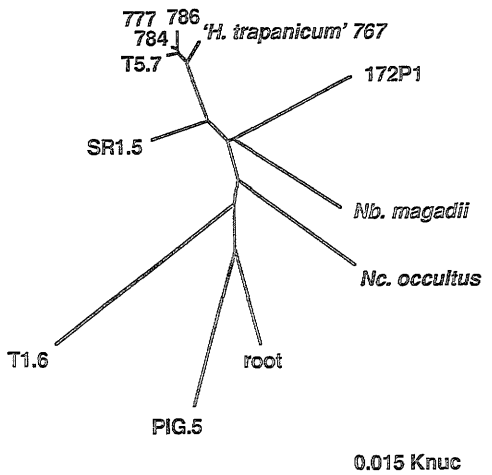


Figure 3.23. Phylogenetic tree of NCIMB 777-like Isolates (and others) from Ancient Salt Deposits. A Maximum Likelihood tree of the strains included in Figure 3.22.

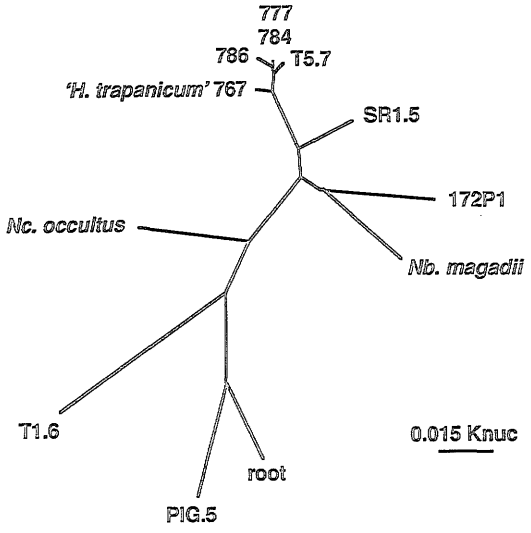


Figure 3.24. Phylogenetic Tree of *Halobacterium* and *Haloarcula*-like Isolates from Ancient Salt Deposits. A Distance tree, using the distance matrix method (DNADIST) of Jukes and Cantor (1969), with the tree inference method (FITCH) of Fitch and Margoliash (1967).
(See Tables 3.17 and 3.18 for information on strains, and text for details.)

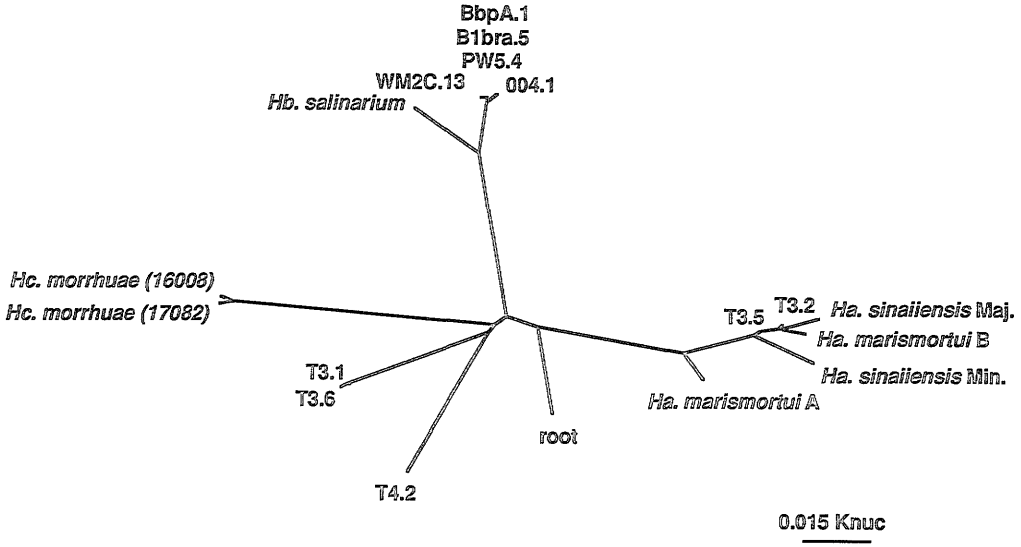
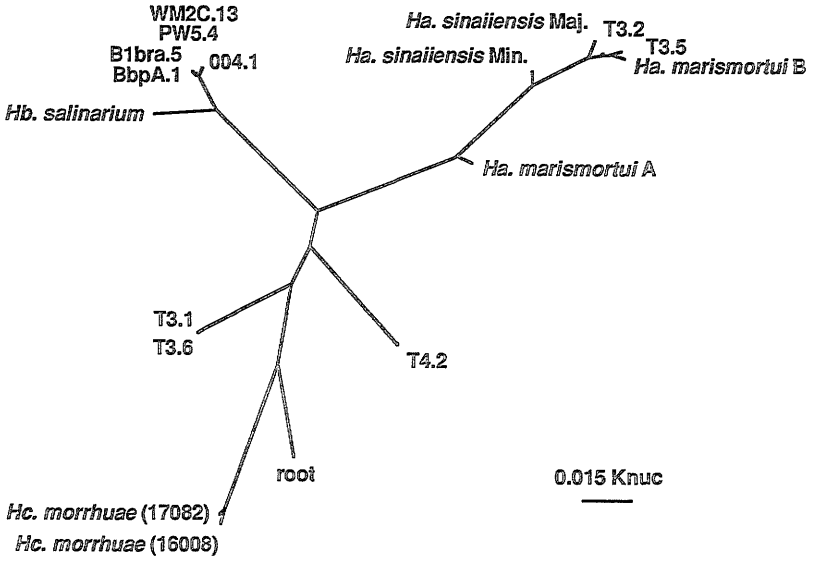


Figure 3.25. Phylogenetic Tree of *Halobacterium* and *Haloarcula*-like Isolates from Ancient Salt Deposits. A Maximum Likelihood tree of the strains included in Figure 3.24.



The '777' group:
Halobacterium sp. 777, *Halobacterium* sp. 784, *Halobacterium* sp. 786, '*Halobacterium trapanicum*' 767, T5.7 and (on the margin) SR1.5.

The '*Natrono*' group:
Natronobacterium magadii, *Natronococcus occultus* and 172P1.

Outliers to the sub-groups, but included, were:
BbpA.3, B2S.26, *Halobaculum gomorrense*, Nh.2, *Haloferax volcanii*, T1.3, T1.6, PIG.5, T3.1, T3.2, T4.2, *Halococcus morrhuae* (NRC 16008) and *Halococcus morrhuae* (ATCC 17082).

The main tree was intended to show the phylogenetic relatedness of these halobacteria from ancient salt deposits, establishing their group affiliation and revealing those isolates which may represent new lineages. The additional halobacterial sequences, from other sources (Table 3.18), were included in the sub-trees with the aim of showing the relationships within these groups in greater detail. Origins of individual strains are given in Tables 3.17 and 3.18.

As a result of the phylogenetic analyses, most of these organisms, for which the 16S rRNA gene has been sequenced, can be grouped with established halobacterial genera. A Distance tree of these halobacteria (Figure 3.17) drawn in a different format (DRAWGRAM; PHYLIP 3.54c), showed clearly that: SC1.2, GSL5.48, PW1.1, WB4E.8, PIG.8 and Nh.5 group with *Halorubrum saccharovorum*, with B2S.26 as an outlier; SR1.5 and T5.7 group with the '777' group; B1bra.5, WM2C.13 and PW5.4 group with *Halobacterium salinarium* (R1); T3.2 and T3.5 are *Haloarcula*-like; T4.2, T3.1 and T3.6 appear to form a new group; T1.3 is closest to *Haloferax* sp., but branches deeply; Nh.2, T1.6 and PIG.5 are not affiliated with any of the groups represented here. The reliability of the tree topology was tested by bootstrapping the sequence data (SEQBOOT) and deriving a consensus tree (CONSENSE), using the procedure detailed in *Materials and Methods*. Bootstrap values indicate that these group affiliations are strong (94-100%); it is the deeper branching orders which are less certain ($\leq$ 56 %).

The sub-trees (Figures 3.20-25) which included halobacteria from other sources, verified these relationships and indeed were able to define affiliation with more clarity. The phylogenetic tree of the *Halorubrum* group (Figures 3.20 and 3.21), for example, shows a bifurcation between the *Hr. saccharovorum* / *Hr. lacusprofundi* lineage and the *Hr. sodomense* / *Hr. trapanicum* lineage, with two separate clusters of organisms (see also

McGenity and Grant, 1995). The appearance of Malk006.3 between these sub-groups, however, suggests that this apparent dichotomy may not be sustained when additional *Halorubrum* species are included in the tree.

Tree topology was generally consistent, whether a Distance method (Figures 3.17, 3.18, 3.20, 3.22, 3.24) or a Maximum Likelihood approach (Figures 3.19, 3.21, 3.23, 3.25) was used. Regions of inconsistency lay at deep branch points and, in some cases, between highly-related species at branch termini. For example, either the deeply-branching Nh.2 (Figures 3.18 and 3.21) or the *Haloferax* / T1.3 lineage (Figures 3.19 and 3.20) lay closest to the outgroup root. At the other end of the scale, within the *Halorubrum* group (Figures 3.20 and 3.21), although there was no variation in sub-group membership (100 % bootstrapped Distance trees), the precise branching order within each sub-group was variable between trees. This was particularly marked within the *Hr. sodomense* / *Hr. trapanicum* group; only 25-31 % bootstrapped trees gave the branching order shown in Figure 3.20.

The 16S rRNA phylogenetic trees of the salt mine isolates reveal the diversity of the halobacteria. Whilst some strains are clearly affiliated to described strains, others are distinctly different and may represent 'new' genera. Lineages represented by isolates T1.3, Nh.2, T4.2 (with T3.1 and T3.6), T1.6 and PIG.5 may indeed be novel, or it could be that the appropriate sequences have not been included in the analysis. Not all currently-recognized members of the *Halobacteriaceae* have been sequenced (*i.e.* the 16S rRNA gene). PIG.5, for example, has a polar lipid pattern resembling '*Halobacterium cutirubrum*' NCIMB 763 (Section 3.2), but the 16S rRNA of this organism has not been sequenced.

3.5.4.2 Phylogenetic Distances and % Similarities.

The Distance Matrix obtained from sequence data as a result of running the Jukes-Cantor program, is a matrix of phylogenetic distances (given in the 'outfile'), based on the pairwise comparison and calculated dissimilarities (not given) of all sequences in the data set. From these phylogenetic distances one can calculate back to the dissimilarity values between sequences, and thus the percentage similarities. This information can be entered into the lower and upper halves, respectively, of a grid, or Matrix. Such Distance Matrices, for the main tree (Figures 3.17, 3.18) and each of the sub-trees (Figures 3.20, 3.22, 3.24) are given in Appendix 2.

Table 3.19. Phylogenetic Relatedness of Some Extremely Halophilic Archaea from Ancient Salt Deposits. % similarity values are taken from Appendix 2.

Isolate	Group affiliation	% similarity to closest relative/named strain
SC1.2	*Halorubrum*	98.8 (GSL5.48; *Halorubrum sodomense*)
GSL5.48	*Halorubrum*	98.8 (GluBr1.1; SC1.2; Ch.2) 98.3 (*Hr. sodomense; Hr. trapanicum*)
PW1.1	*Halorubrum*	98.7 (Nh.5) 98.1 (*Halorubrum lacusprofundi*)
WB4E.8	*Halorubrum*	98.9 (54R) 98.8 (*Halorubrum saccharovorum*)
PIG.8	*Halorubrum*	98.7 (54R; *Hr.saccharovorum*)
Nh.5	*Halorubrum*	98.7 (PW1.1) 97.6 (*Halorubrum lacusprofundi*)
B2S.26	-	100 (BbpA.3) 92.9 (*Halorubrum saccharovorum*)
PW5.4	*Halobacterium*	99.9 (WM2C.13; B1bra.5; BbpA.1) 96.9 (*Halobacterium salinarium*)
WM2C.13	*Halobacterium*	99.9 (PW5.4; B1bra.5; BbpA.1) 96.9 (*Halobacterium salinarium*)
B1bra.5	*Halobacterium*	100 (BbpA.1) 97.0 (*Halobacterium salinarium*)
T1.6	-	91.0 ('*Halobacterium*' sp. 786)
PIG.5	-	91.4 (*Nc.occultus*)
T3.2	*Haloarcula*	99.4 (T3.5) 99.2 (*Haloarcula marismortui* B)
T3.5	*Haloarcula*	99.4 (T3.2) 99.1 (*Haloarcula marismortui* B)
T4.2	'T4.2'	93.2 (T3.1) 91.6 (*Haloarcula marismortui* A)
T3.1	'T4.2'	99.8 (T3.6) 91.4 (*Haloarcula marismortui* A)
T3.6	'T4.2'	99.8 (T3.1) 91.1 (*Haloarcula marismortui* A)
T5.7	'777'	99.6 ('*Halobacterium*' sp. 777)
SR1.5	'777'	96.8 (T5.7) 96.5 ('*Halobacterium*' sp. 777)
T1.3	-	91.9 (*Hf. volcanii*)
Nh.2	-	91.5 (B1bra.5; PW5.4) 90.1 (*Halobacterium salinarium*)
'*Halobacterium*' sp. 777	'777'	100 ('*Halobacterium*' sp. 784)
'*Halobacterium*' sp. 784	'777'	100 ('*Halobacterium*' sp. 777)
'*Halobacterium*' sp. 786	'777'	99.4 ('*Halobacterium*' spp. 777 & 784)

A more objective expression of these relationships is given in Table 3.19, which summarizes the distance matrix data (Appendix 2), and gives the group affiliation for the isolate (if determined) and the % similarity to the closest relative. Those isolates with a 16S rRNA % similarity of less than 97.5 % are unlikely to be highly related species, whereas those of greater than 97.5 % similarity may or may not be related species (Stackebrandt and Goebel, 1994). Amongst the halobacteria, % similarities between the genera are approximately in the range 86-90 %, although the % similarity is a little higher between the haloalkaliphile genera and the '777' group, at about 92-94 % (Appendix 2). The data presented here also identify those isolates which may represent new lineages (Figures 3.18 and 3.19); *i.e.* those with % similarities of less than ~92 % to any other group, *e.g.* T1.6 and T4.2.

Isolate B2S.26 is unusual in that polar lipid analysis has identified it as a member of the *Halorubrum* group; but it is only about 92 % similar to this group (92.9 % similar to *Halorubrum saccharovorum*). It may well be that group delineation by polar lipid profiling circumscribes a larger area of relatedness than that determined by 16S rRNA gene sequence comparisons. B2S.26 is 100 % identical to BbpA.3, independently isolated from a Boulby brine pool, sampled some 18 months earlier, in 1990 (McGenity, 1994).

Within the *Halobacterium* group, isolates B1bra.5, WM2C.13 and PW5.4 are equally similar, although they are from sites which are both geographically and geologically distinct. B1bra.5 is 100 % identical to BbpA.1, also isolated from a Boulby brine pool, sampled independently in 1990 (McGenity, 1994).

3.5.4.3 Calibrating the Halobacterial Phylogenetic Tree.

The main purpose of these phylogenetic analyses was to see if there was any correlation between phylogenetic relatedness of two organisms of the same lineage, and the respective geological ages of the salt deposits from which they were isolated. *i.e.* in a phylogenetic tree would an isolate from a Recent salt deposit, such as a Spanish solar saltern, lie at a branch terminus, and one from a Permian (250-260 Ma) salt deposit lie deeper within the tree, with an isolate from a Cretaceous deposit (*c.* 65 Ma) somewhere between the two?

Phylogenetic distances (Knuc), which are given in the Distance Matrices (Appendix 2), can be converted to 'geological age' on the basis of the calibrations of Moran *et al.* (1993), who have estimated that 0.01-0.02 Knuc is equivalent to 50 million years. Using this measure, the chronological distances between organisms of the *Halorubrum* and *Halobacterium* groups have been estimated, for comparison with the actual age

differences between the salt deposits from which the organisms were isolated (Tables 3.20 and 3.21).

It is apparent from the phylogenetic trees that there is no clear correlation between phylogenetic relatedness and geological ages of the salt deposits. This can only realistically be considered with organisms of the same lineage, and within any group, isolates from the more ancient salt deposits do not emerge closer to the root of the tree than the isolates from more recent salt deposits. This is exemplified by the *Halobacterium* group (Figure 3.24), in which isolates PW5.4, WM2C.13 and B1bra.5 (from salt deposits *c.* 20 Ma, 235-240 Ma and 250-260 Ma, respectively), covering over a 200 million-year timespan, lie together at the terminus of the branch and are only 2-7 Ma apart in evolutionary time (Table 3.21). The phylogenetic distance between *Halorubrum saccharovorum* and PIG.8 is 0.0127 Knuc (Appendix 2); on the basis of the calibration of Moran *et al.* this is equivalent to 32-64 million years in geological/evolutionary time, whereas the actual geological timespan involved is 250-260 million years (Table 3.20).

The data summarized in Tables 3.20 and 3.21, for the *Halorubrum* group and the *Halobacterium* group, respectively, show clearly that there is no agreement between the estimated chronological distances between organisms and the actual geological timespans involved. One could assume that the calibration of Moran *et al.* is inaccurate or unsuitable for this analysis, and the estimated geological age is therefore incorrect, but this would merely change the scale of the tree and not affect any underlying correlation between the phylogenetic relatedness of these 16S rRNA gene sequences and the geological ages of the sources.

The great species diversity, not only between the halobacterial groups, but also within any one group, such as *Halorubrum* (Figure 3.20), also effectively masks any correlation there may be between the geological age of the salt sample and the phylogenetic 'age' of the halobacterium. It may be that the surface isolates included in this analysis are not of the same lineage as those isolated from ancient salt, explaining why these latter appear as branches, and not the expected 'blips' on the phylogenetic tree. Those isolates which have been suggested as representing new lineages (T1.3, T1.6, *etc.*) may not, in fact, be merely 'newly-discovered' halobacteria, but actually be unique 'living fossils' with no extant relatives. However, this form of analysis is unable to distinguish between the two interpretations.

Table 3.20. 'Relative Ages' of the *Halorubrum* Group. Geological ages, based on Knuc distance values, using the calibration of Moran *et al.* (1993). In brackets: the actual age difference between the salt deposits from which the organisms were isolated. Units are Ma (mega annons, or 10^6 years).

	H. sacc												
H. lac.	42-84 (<1)	*H. lac.*											
PW1.1	66-131 (20)	47-95 (20)	PW1.1										
WB4 E.8	30-60 (235)	51-102 (235)	62-124 (215)	WB4 E.8									
54R	7-14 (235)	38-77 (235)	60-120 (215)	26-53 (0)	54R								
Malk 0063	71-142 (235)	89-178 (235)	100-200 (215)	73-146 (0)	71-142 (0)	Malk 0063							
PIG8	32-64 (255)	46-92 (255)	64-128 (235)	37-74 (20)	32-64 (20)	77-154 (20)	PIG8						
Nh5	69-139 (255)	60-121 (255)	32-64 (235)	55-110 (20)	68-136 (20)	89-179 (20)	64-128 (<1)	Nh5					
H. trap	105-210 (<1)	121-243 (<1)	122-245 (20)	110-220 (235	109-218 (235)	114-228 (235)	106-212 (255)	127-254 (255)	*H. trap*				
H. sod	102-204 (<1)	114-229 (<1)	119-238 (20)	114-228 (235)	106-212 (235)	100-200 (235)	99-198 (255)	121-243 (255)	30-60 (<1)	*H. sod*			
Ch.2	120-240 (<1)	128-257 (<1)	129-259 (20)	124-249 (235)	123-247 (235)	105-210 (235)	113-226 (255)	134-268 (255)	37-74 (<1)	35-70 (<1)	Ch.2		
SC12	117-235 (<1)	130-260 (<1)	123-246 (20)	114-228 (235)	121-242 (235)	117-235 (235)	99-198 (255)	119-239 (255)	32-64 (<1)	30-60 (<1)	40-81 (<1)	SC12	
GSL5.48	108-216 (<1)	117-234 (<1)	123-246 (20)	108-216 (235)	112-224 (235)	106-212 (235)	97-194 (255)	108-216 (255)	44-88 (<1)	42-85 (<1)	31-63 (<1)	30-60 (<1)	GSL5.48
GluBr 1.1	109-218 (235)	118-236 (235)	130-260 (215)	114-228 (0)	113-226 (0)	107-214 (0)	101-202 (20)	134-269 (20)	30-60 (235)	28-56 (235)	30-60 (235)	37-75 (235)	30-60 (235)

Table 3.21. 'Relative Ages' of the *Halobacterium* Group. Geological ages, based on Knuc distance values, using the calibration of Moran *et al.* (1993). In brackets: the actual age difference between the salt deposits from which the organisms were isolated. Units are Ma (mega annons, or 10^6 years).

	Hb. salinar-ium				
PW5.4	78-156 (20)	PW5.4			
WM2C.13	78-156 (235)	2-4 (215)	WM2C.13		
004.1	81-163 (235)	9-18 (215)	10-21 (0)	004.1	
BbpA.1	78-156 (255)	2-4 (235)	3-7 (20)	7-14 (20)	BbpA.1
B1bra.5	76-152 (255)	2-4 (235)	3-7 (20)	7-14 (20)	0 (0)

IV

GENERAL DISCUSSION

The halobacterial habitat is not restricted to contemporary salty environments such as salted foods, salterns and salt lakes, but includes salt deposits which have been deeply-buried for millions of years. Previous work has shown that halobacteria and halophilic Bacteria can be isolated from the 235-240 and 250-260 million-year-old salt deposits at Winsford and Boulby, in the N.W. and N.E. of England, respectively (Norton *et al.*, 1993; McGenity, 1994). This study supports those previous results, and has shown that halophiles are found not only in the English salt mines, but in geographically distinct salt deposits, which represent a spectrum of geological ages, from the Recent (*e.g.* Great Salt Lake) to over 250 million years old (*e.g.* Góra, Poland). Although both halophilic Bacteria and the extremely halophilic Archaea (halobacteria) have been isolated from these salt deposits, this study has focused on the latter.

Salt deposits are the result of the evaporative concentration of salts in ancient seas or lakes (Section 1.3), and in the later stages of their genesis presumably supported rich communities of halophilic microorganisms, much like those found in present-day hypersaline environments (Section 1.2). In the artificial evaporation of seawater, in the final crystallizing ponds of salinas, where salt is harvested, the extremely halophilic Archaea usually predominate over halophilic Bacteria (Section 1.2), with *Halobacterium*, *Halorubrum* and *Haloarcula* as the predominant genera. This study has shown that these genera are also the predominant isolates from within ancient salt deposits.
On the basis of polar lipid analysis, approximately 23 % of identified isolates were *Halorubrum* spp., 19 % *Haloarcula* spp. and 15 % *Halobacterium* spp. Generally, 16S rRNA gene sequencing backed up the results of the polar lipid analysis, except that T1.3 had a *Hb. saccharovorum*-like profile and sequencing indicated it was different. *Halobacterium* spp. (from polar lipid patterns) B1bra.5, WM2C.13 and PW5.4, from Boulby, Winsford and Wieliczka, respectively, were 99.9-100 % similar based on 16S rRNA similarities, but they were only similar at the 73 % level based on protein patterns, and did not have identical ribopatterns, so they are probably different species of *Halobacterium*. Halobacteria described as the '777 group', which includes the Thai isolate T5.7, are phylogenetically distinct from other halobacteria, supporting the findings of Ross and Grant (1985), that this group constitutes a genus-level taxon (*"Halonema"*; Grant and Ross, 1986). From polar lipid analysis (Section 3.2) and signature sequence analysis (McGenity, 1994), PIG.5 can be grouped with *'Halobacterium cutirubrum'*

NCIMB 763, which was given genus status (*"Halothrix"*; Grant and Ross, 1986) by Ross and Grant (1985).
25 % of isolates could not be classified, although several polar lipid groups could be detected. 68 % of unclassified halobacteria were from residual brine (Boulby) or solutioned brine (Thailand, Góra, AKZO brines); they formed a significant proportion of isolates from Thailand (57 %). One may speculate that some of these unclassified halobacteria represent archaic lineages within the autochthonous community in ancient salt deposits and may have no representatives in modern hypersaline environments. Those unclassified isolates that were sequenced, namely Nh.2, T1.3, T1.6, T3.1, T3.6 and T4.2, from AKZO and Thai solution brines, formed distinct lineages on the phylogenetic tree (Figure 3.18), distant from known halobacterial genera, and may represent 'living fossils' of extinct genera, or new genera with as yet unrecognized contemporary relatives.

Viable counts of halophilic Bacteria and halobacteria present in these environments gave values in the region of 10^5 ml^{-1}, with halobacteria usually predominating after incubation at 40 °C. Mine brines originating as inflows from overlying younger rocks (PW3, PW4, PW5, B3, B7 and B8), with little salt contact, had much lower counts ($<10^3$ ml^{-1}). The halophiles isolated from these brines could have arrived from outside the deposit, via fissures, and multiplied in the favourable mine environment, the numbers being higher where residence had been longer. The problem with accepting an external source for these organisms, is that halobacteria (except *Halococcus* spp.) usually lyse in water, although they may survive very low salt concentrations. Halobacteria were isolated from two Spanish salt ponds (SR1 and SR2) which had only 1 % total salts, supporting earlier unusual findings that some halobacteria can survive exposure to such low salt concentrations (Rodriguez-Valera *et al.*, 1983; Torreblanca *et al.*, 1986).
Alternatively, the halophiles originated from within the salt deposit and not from extraneous sources. Low or negligable viable counts, in brines with a salt concentration sufficient to support halobacterial growth (PW4, PW5, B3, B7, B8, B350A), can be explained by the salt originating from a metamorphosed part of a deposit, or by insufficient recovery time for the resident organisms within the salt solutioned by the inflow. The B350A brine, for example, yielded no bacteria, unlike other brines from Boulby, but it was believed to be residual and has probably been subject to diagenetic events. Some isolates from PW4 inflow brine took three months to form visible colonies, presumably requiring time to adapt; these were identified by polar lipid analysis as a *Halorubrum* sp. The Kali und Salz brines did not yield any halophiles, although they were capable of supporting their growth, and their sterility has been explained as being possibly due to metamorphic events in the salt deposits, which would have killed any resident microorganisms. Halobacteria (and halophilic Bacteria) were isolated from

solution mines, where the buried salt deposits are excavated by pumping in water and pumping out the solutioned salt. In the Thai site, which had been in operation for only three years, brine was taken directly from the pipeline, at its exit from a salt dome which had been solutioned by sterile water. Also, it was observed that a large proportion of isolates from the T1-4 and AKZO solutioned brines (and less conspicuously the PIG brine) produced tiny, transparent, slow-growing colonies, which did not always survive sub-culturing (AKZO Gst), which suggested inadequate recovery (from dormancy?). Taken together, these results strongly support an autochthonous origin for the salt mine halophiles. The question then to ask, is: are these halophiles 'living fossils' which have been revived?

One can presume that organisms which have recovered from a state of suspended animation, after millions of years, would be genetically more primitive, or less evolved, than contemporary organisms of the same phylogenetic lineage. On this premise, a selection of halobacteria have been compared by phylogenetic analysis of the 16S rRNA gene. It was expected that for any one lineage, the more ancient the environment of the isolate, the closer to the root of the 16S rRNA phylogenetic tree it would appear, and isolates from contemporary hypersaline environments would lie at the branch termini. Also, if the isolate was a 'living fossil' that shared common origins with a present-day isolate, it should be a point on an extant lineage, representing one stage in the evolution of the contemporary isolate. The problem with this hypothesis is that it is necessary to have many representatives of any one lineage to see any true correlation between the phylogenetic distance, and estimated evolutionary age difference, of two strains, and the geological ages of the salt deposits from which they were isolated.

The relationship between phylogenetic distance and geological age was <u>not</u> seen and could be explained by one or more of the following hypotheses:

1. **The halobacteria are autochthonous 'living fossils' which have been revived, but 16S rRNA phylogenetic analysis is somehow not a suitable method to show this.**
2. **The great diversity of halobacterial species has effectively masked any correlation there might be between geological age and phylogenetic distance.**
3. **The halobacteria isolated from the ancient salt deposits are not 'ancient' microorganisms which have survived dormancy for millions of years, but contemporary isolates which have reached the deposits from outside.**

4. **The halobacteria are autochthonous and have continuously inhabited the salt deposit, but grown very slowly, perhaps in saltatory fashion, over the millenia; they are younger, therefore, than the salt.**
5. **The autochthonous halobacteria were entrapped over an extensive time period, during the depositional history of the salt deposit.**

Considering each of these hypotheses in turn:

Hypothesis 1.

Halobacteria have survived in a state of suspended animation, either within fluid inclusions, or encapsulated within the salt. In fluid inclusions depleted of nutrients halobacteria may enter and remain in a state of suspended animation, possibly undergoing some morphological change (Wais, 1985; Kostrikina *et al.*, 1991) as they enter the vegetative state. Alternatively, the halobacteria could have acted as living nucleation sites for crystal growth and become 'preserved' within a halite capsule. In this desiccated state of dormancy they would have a basal metabolic rate. There are difficulties to reconcile with suspended animation for millions of years, however. Apart from the possibility of radiation damage due to ^{40}K, which should not be a problem over a timespan of millions of years (Sneath, 1962), or cosmic rays, DNA is subject to endogenous hydrolytic depurination and deamination, oxidative damage, and nonenzymatic methylation, which are normally mended by excision-repair processes in metabolically active cells (Lindahl, 1993b). However, it is possible that cells in a dormant, survival state are biochemically protected against chemical degradation of this nature. The DNA in bacterial spores is protected against hydrolytic depurination and oxidative damage by its partially dehydrated state and the synthesis of DNA-binding proteins (Lindahl, 1993b). Encapsulted within halite crystals, halobacteria would be in a low-oxygen environment of very high ionic strength and low water activity; conditions which may act to protect and stabilize the DNA (Lindahl, 1993b). The combined properties of dehydration and salification may, as Dombrowski (1963) suggested, conserve bacteria for millions of years.

However, if these halobacteria are indeed 'living fossils', then this 16S rRNA phylogenetic analysis has not revealed it. Distance and Maximum Likelihood may not be suitable methods for inferring phylogenetic trees from the sequence data to show the relationship between phylogenetic distance and geological age. Distance methods are known to suffer some loss of information; multiple changes at one site are obscured and so branch lengths are too short (Section 1.5.5). With the 'older' strains this may underestimate phylogenetic distance, and therefore the calculated evolutionary age. The more powerful Maximum Likelihood method does not suffer from these restraints, however, and compared to other methods is least affected by different mutation rates

(Nei, 1991), which are a potential problem in the interpretation of phylogenetic relationships.

If the evolution rate within the halobacteria is slow, then the differences between 16S rRNA genes in any one lineage may not be sufficient for phylogenetic analysis to discriminate the highly-related strains. It is generally recognized that 16S rRNA sequence comparisons are unable to unequivocally delimit at the species and sub-species levels (Fox *et al.*, 1992; Stackebrandt and Goebel, 1994).

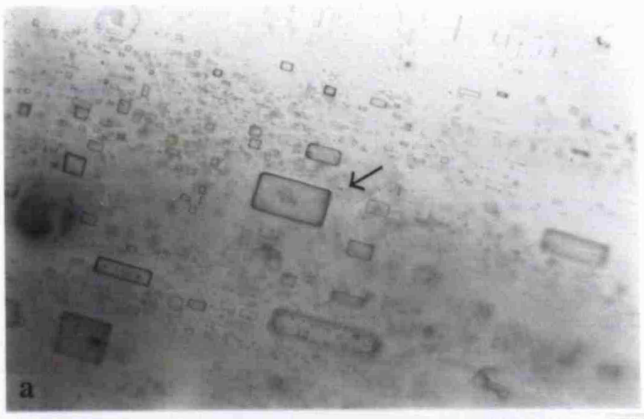
Hypothesis 2.

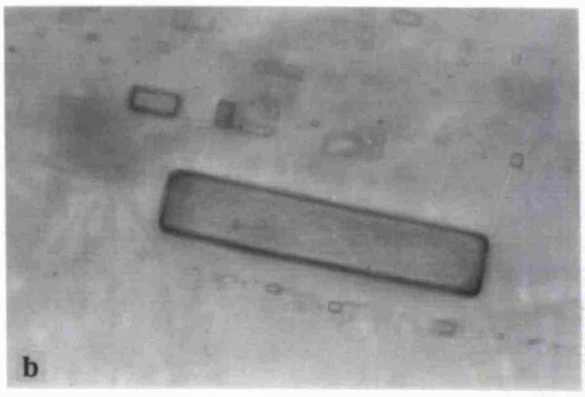
Certainly a diversity of halobacteria was found in the salt mine samples, as shown by polar lipid analysis and the two typing methods used: analyses of electrophoresed whole-cell protein patterns and 16S rDNA restriction patterns. The results of both typing methods indicated that although the same types of halobacteria were probably being isolated from different sites, these were not identical. For each site the halobacteria were unique, revealing a diversity which has yet to be acknowledged.

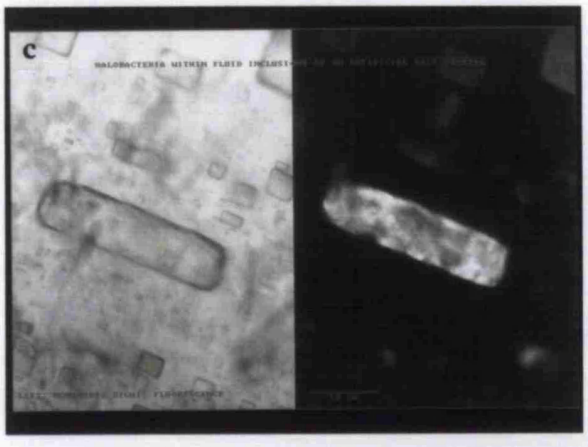
Because relatively few isolates were included in the phylogenetic analysis, and insufficient representatives of any one lineage, therefore the results did not produce a clear pattern. To clarify any relationship between phylogenetic distance and geological age, the 16S rRNA genes of many more of the salt mine isolates would need to be entered into the analysis.

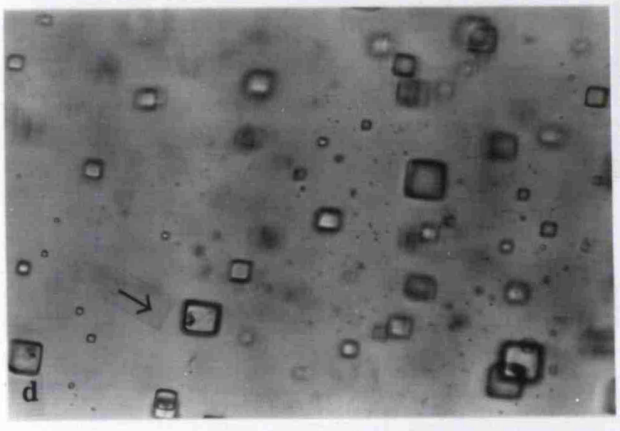
Hypothesis 3.

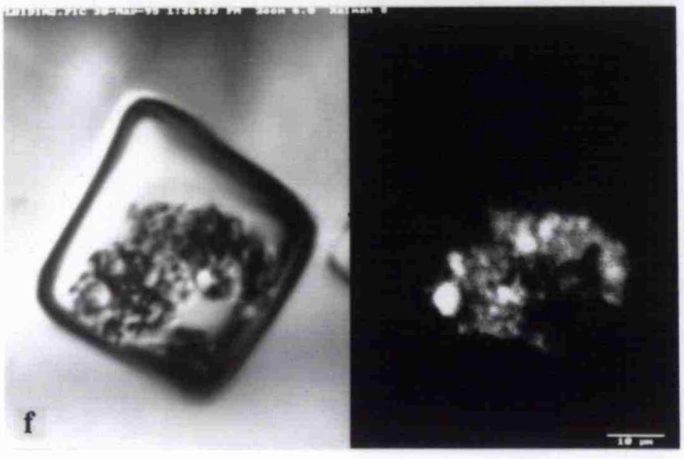
Halobacteria have a very specific, extreme and therefore localized environment - a hypersaline one, as described in Section 1.2. Halobacteria lyse in water and so entry into mines via groundwater is highly unlikely. Wind-blown salt crystals, harbouring halobacteria, could enter the traditionally-mined deposits via the air ventilation systems. The nature of the sampling in some instances precludes the possibility of contamination by surface isolates; the solutioned brines, particularly from Thailand, operate almost as closed systems. The Kali und Salz brines, although capable of supporting the growth of halobacteria, appeared to be sterile, indicating that halobacteria have not entered the salt at these sites since the metamorphic events which probably killed any resident microbes (see Section 3.1.10). Halobacteria were isolated from all of the other sites, and the diversity of these isolates suggests *in situ* halobacterial communities are being sampled. If these arose from more recent contamintion, such diversity would not be expected. Considering the intolerance of halobacteria to low salt concentrations, contamination by any means is highly unlikely.











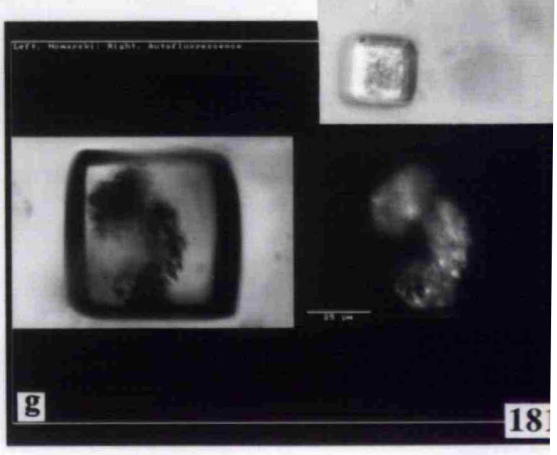


Figure 4.1 **Fluid Inclusions in Halite Crystals.**

Artificial salt crystals, prepared by Norton and Grant (1988):

a **array of fluid inclusions, viewed by phase contrast microscopy (low magnification) - central, large inclusion is 100 x 50 μm;**

b **fluid inclusion, 27 x 9 μm, which appeared to contain intact rod-shaped cells;**

c **when viewed by confocal fluorescence microscopy, the fluid inclusions exhibited autofluorescence;**

Primary Miocene salt crystals (Wieliczka):

d **phase contrast microscopy confirmed the presence of fluid inclusions (indicated inclusion is 50 x 50 μm);**

e **some fluid inclusions (125 x 75 μm) contained detritus;**

f,g **confocal microscopy of a fluid inclusion showed that the detrital material (probably organic) autofluoresces (left, Nomarski, and right, fluorescence microscopy).**

Hypothesis 4.

The evidence is against an extraneous origin for the salt mine isolates and in favour of them being autochthonous. Viable, believed autochthonous, Bacteria and Archaea have been isolated from various deep subsurface environments (Section 1.4). It is quite feasible that these halobacteria have lived in this subterranean habitat, in pockets of brine or fluid inclusions within salt crystals, since the deposit was laid down. It can be hypothesized that, trapped within fluid inclusions, they would initially utilize DOC in the pockets of brine, and possibly engage in 'cryptic growth' on dead cells, but as these nutrient sources became depleted, the halobacteria would enter a starvation-survival state. Endogenous substrates, such as redundant enzymes, would be utilized, followed by a slowing down of metabolic processes. The salt mine isolates could have existed in this survival-state, perhaps growing intermittently in saltatory spurts as nutrients became available. With the migration of fluid inclusions through the rock salt, due to temperature or pressure gradients, embedded organic matter could periodically be encountered.

These halobacteria would be younger therefore than the salt deposit they inhabit. Previous research has shown that halobacteria can be recovered following entrapment within fluid inclusions in artificial salt crystals (Figure 4.1a-c), and that this may be a mechanism for long-term survival within salt deposits (Norton and Grant, 1988). Primary salt crystals from Wieliczka (Figure 3.1b) were examined by light- and confocal microscopy, to search for entrapped microorganisms, but although there was evidence of organic matter within some fluid inclusions (Figures 4.1d-g), no microorganisms could be seen.

Hypothesis 5.

Halobacteria would be entrapped throughout the deposit. If the salt was deposited over several hundred thousands of years, or millions of years, the 'age' of an isolate would significantly differ according to its source within the deposit. To complicate matters further, there may even be a mingling of different ages due to migrating fluid inclusions. This would confuse any attempt at correlating geological age and phylogenetic distance.

Overall, the results indicate that inter-species 16S rRNA gene sequence comparison is not an appropriate method to show whether or not the halobacteria isolated from ancient salt deposits are 'living fossils' which have somehow survived within the salt since its deposition. Nonetheless, it has been useful in confirming group affiliations and identifying possible new phylogenetic lines.

This study has been very useful in identifying some of the organisms native to such hypersaline environments, but the relative proportions of types (Table 3.13), though providing some indication of the *in vivo* situation, can only be interpreted as proportions isolatable on the media and under the conditions used, and it should be remembered that not all bacteria are culturable. This may sometimes be due to a form of "culture shock", the bacteria being unable to tolerate the transfer from an equable environment to one which is nutrient rich, this perhaps causing some system overload. Oren has made some attempts to identify the true community structure of halophilic bacterial populations in the Dead Sea and saltern crystallizer ponds (see Oren, 1994a). Halobacteria can be enumerated by quantitative determination of the carotenoid pigments, but there are problems with this (Oren, 1994a). Polar lipid analysis may also provide some indication of the dominant species present or those absent in a hypersaline environment (Oren, 1994b). In a novel application of ribotyping, Martínez-Murcia *et al.* (1995) have studied the diversity of halophilic Bacteria and the extremely halophilic Archaea in salt ponds of a solar saltern, by analysing the electrophoresed restriction digests of 16S rRNA directly amplified by PCR from cells collected by filtration of brine samples.

The bacterial populations could be examined directly by the use of probes. The demonstration that halobacterial species are immunologically distinct (Conway de Macario *et al.*, 1986) provides one such method. The ecological tapestry of the hypersaline environment may be more richly described by *the rRNA approach* (Olsen *et al.*, 1986; reviewed in Amann *et al.*, 1995), which obviates the necessity to cultivate the organisms. In this approach, a thorough description of the composition and diversity of the bacterial community can be obtained by *full-cycle rRNA analysis*, whereby information from an initial analysis of rDNA sequences retrieved directly from the environmental sample is used to design oligonucleotide probes to rRNA signature sequences (see Sections 1.5.5.5 and 3.5.3). By whole-cell (*in situ*) hybridization of these fluorescently-labelled probes to cells in the original sample, the *in situ* abundance and spatial distribution of an organism can be determined (Amann *et al.*, 1995). As proposed by Liesack *et al.* (1990), with discriminate probe design, the DIG detection system could be used to identify specific organisms at all taxonomic levels; dyes of different fluorescence could be used in the anti-DIG immunological detection reactions, in an *in situ* hybridization procedure (Boehringer Mannheim). Some of the different culturable halobacterial populations found in the salt mines are now known from this present study, and may be enumerated using direct analysis of environmental rDNA sequences and probe technology. Hitherto uncultured halobacteria could be recognized and their taxonomic position determined by phylogenetic analysis. These new strains could be isolated for pure culture and characterization by employing the techniques used by Huber *et al.* (1995), who used 'optical tweezers' to isolate (from an enrichment) a single bacterial cell which they cloned and were able to identify with a sequence derived by *in situ* 16S rRNA analysis of environmental samples.

Studies of halobacteria from ancient salt deposits could be developed in several directions. Transport of halobacteria and seeding of surface sites requires methodical investigation. A useful control would be a 'salt pond' of sterile brine. Presumably, salinas and salt lakes are seeded by the sea, fish, crustaceans, birds and aeolian salt. The halobacterial community structure should be examined by applying probe technology, as discussed above. The problems relating to starvation, survival and longevity should be addressed. Physiological responses to starvation and on recovery, dehydration and salt-biomolecule interactions, should be investigated. The availability of organics in the halite, and assimilability by halobacteria, is pertinent to the questions of survival and longevity. Primary fluid inclusions, which contain the mother liquor and atmospheric gases of the original brine pool, should be individually investigated for autochthonous microorganisms by microscopy, and also by aseptically removing the primary brine and identifying any microorganisms by culture and the use of probes. This again would be possible by employing the 'optical tweezer' techniques used by Huber *et al.* (1995).

On the basis of the results gained in this study, the halobacteria in the salt mines are autochthonous populations of considerable diversity. It may be that they have grown slowly over the millenia, utilizing organic matter available in the deposit. That some of these halobacteria are 'living fossils', which remain locked within the salt in a state of dormancy until revived, is still an open question.

APPENDICES.

Appendix 1. 16S rRNA Gene Sequence Alignment.

The 21 salt deposit isolates and 3 culture collection strains, sequenced as part of this study, were aligned with the following published sequences: *Haloarcula marismortui rrnA*, *Haloferax volcanii*, *Halobacterium salinarium*, *Halorubrum saccharovorum* and *Natronobacterium magadii*. *Methanospirillum hungatei* was included as an outgroup, with which to root the phylogenetic tree. The first 17 bases (5') and end 8 bases (3') were omitted from the alignment (in **bold**), to minimize the inclusion of N's. Amplification and sequencing primer regions are underlined. Sequences were first aligned in CLUSTAL V, with equal weight given to all sites, then adjusted manually with reference to the secondary structures, to optimize the alignments.

```
            5'

            1                                                  50
Methanosp.  ATTCTGNTTG ATCCTGCCAG AGGCCACTGC TATCGGGGTT TGACTAAGCC
       777  NNNNNGGTTG ATCCTCCCGG AGGTCATTNC TATTGGAGTC CGATTTAGCC
       784  NNNNNGGTTG ATCCTCCCGG AGGTCATTGC TATTGGAGTC CGATTTAGCC
       786  NNNNNNNNNN NNNNTCCCGG AGGTCATTNC TATTGGAGTC CGATTTANCC
   B1bra.5  NNNNNNNNNN NNNNNNNCGG AGGCCATTGC TATCGGAGTC CGATTTAGCC
   B2S.26   NNNNNGGTTG ATCCTGCCGG AGGCCATTGC TATCAGGGTC CGATTTANCC
   GSL5.48  NNNNNNNNNN ATCCTGCCGG AGGCCATTGC TATTGGGATC CGATTTAGCC
Haloarcula  ATTCCGGTTG ATCCTGCCGG AGGCCATTGC TATCGGAGTC CGATTTAGCC
      Nh.2  NNNNNGGTTG ATCCTGCCGG AGGCCATTGC TATCGGTGTC CGATTTANCC
      Nh.5  NNNNNNGTTG ATCCTGCCGG AGGCCATTGC TATTGGGATT CGATTTAGCC
      PIG.5 NNNNNNGTTG ATCCTNCCGG AGGTCATTGC TATTGGAGTT CGATTTANCC
      PIG.8 NNNNNGGTTG ATCCTNCCGG AGGCCATTGC TATTGGAATC CGATTTANCC
      PW1.1 NNNNNNNNNN NNCCTGCCGG AGGCCATTGC TATTGGGATT CGATTTAGCC
      PW5.4 NNNNNNGTTG ATCCTGCCGG AGNCCATTGC TATCGGAGTC CGATTTANCC
      SC1.2 NNNNNGGTTG ATCCTGCCGG AGGCCATTNC TATTGGGATC CGATTTANCC
      SR1.5 NNNNNNGTTG ATCCTGCNGG AGGTCATTNC TATTGGAGTC CGATTTANCC
       T1.3 NNNNNGGTTG ATCCTGCCGG AGGCCATTNC TATCGGAGTC CGATTTANCC
       T1.6 NNNNNNNNNN NTCCTNCCGG AGGCCATTGC TATGGGGATC CGATTTANCC
       T3.1 NNNNNGGTTG ATCCTGCCGG AGGCCATTGC TATCGGAGTC CGATTTANCC
       T3.2 NNNNNGGTTG ATCCTGCCGG AGGCCATTGC TATCGGAGTC CGATTTAGCC
       T3.5 NNNNNGGTTG ATCCTGCCGG AGGCCATTGC TATCGGAGTC CGATTTAGCC
       T3.6 NNNNNGGTTG ATCCTGCCGG AGGCCATTGC TATCGGAGTC CGATTTAGCC
       T4.2 NNNNNGGTTG ATCCTGCCGG AGGCCATTGC TATCGGAGTC CGATTTAGCC
       T5.7 NNNNNGGTTG ATCCTNCCGG AGGTCATTGC TATTGGAGTC CGATTTANCC
    WB4E.8  NNNNNNGTTG ATCCTNCCGG AGGCCATTNC TATTGGGATT CGATTTANCC
   WM2C.13  NNNNNGGTTG ATCCTGCCGG AGGCCATTGC TATCGGAGTC CGATTTAGCC
 Haloferax  ATTCCGGTTG ATCCTGCCGG AGGTCATTGC TATTGGGGTC CGATTTAGCC
 Halobact.  ATTCCGGTTG ATCCTGCCGG AGGCCATTGC TATCGGAGTC CGATTTAGCC
Halorubrum  NNNNNNNTTG ATCCTGCCGG AGGCCATTGC TATTGGGATT CGATTTAGCC
 Natronob.  ATTCCGGTTG ATCCTGCCGG AGGTCATTGC TATTGGAGTC CGATTTAGCC
```

```
            51                                                  100
Methanosp.  ATGCGAGTCG AGAG.GTGCA AGACCTCGGC GTACTGCTCA GTAACACGTG
       777  ATGCTAGTTG CACGAGTTNA GACTCGTAGC AGATAGCTCA GTAACACGTG
       784  ATGCTAGTTG NNCGAGTTCA GACTCGTAGC AGATAGCTCA GTNACACGTG
       786  ATGCTAGTTG CACGAGTTCA GACTCGTAGC AGATAGCTCA GTAACACGTG
   B1bra.5  ATGCTAGTTG TGCGGGTTTA GACCCGCAGC GGAAAGCTCA GTAACACGTG
    B2S.26  ATGCTAGTCG CACGAGTTCA TACTCGTGGC ATATAGCTCC GTAACACGTG
  GSL5.48   ATGCTAGTCG CACGAGTTCA GACTCGTGGC GAATAGCTCA GTAACACGTG
Haloarcula  ATGCTAGTCG CACGGGCTTA GACCCGTGGC ATATAGCTCA GTAACACGTG
      Nh.2  ATGCGAGTTG CACG..TTTA .A..CGTANC ACACTGCTCA GTAACACGTG
      Nh.5  ATGCTAGTCG CACGAGTTCA GACTCGTGGC GAATAGCTCA GTAACACGTG
     PIG.5  ATGCTAGTCG CACGAGTTCA GACTCGTGGC AGATAGCTCA GTAACACGTG
     PIG.8  ATGCTAGTCG CNCGAGTTCA GACTCGTGGC GAATAGCTCA GTAACACGTG
     PW1.1  ATGCTAGTCG CACGAGTTCA GACTCGTGGC GAATAGCTCA GTAACACGTG
     PW5.4  ATGCTAGTTG TGCGGGTTTA GACCCGCAGC GGAAAGCTCA GTAACACGTG
     SC1.2  ATGCTAGTCG CACGAGTTCA GACTCGTGGC GAATAGCTCA GTAACACGTG
     SR1.5  ATGCTAGTTG CNCGAGTTTA GACTCGTAGC AGATAGCTCA GTAACACGTG
      T1.3  ATGCTAGTCG TGCGAGTTCA CACTCGCGCC ANATAGCTCA GTAACACGTG
      T1.6  ATGCTAGTCG CACGAGTTTA GACTCGTGGC AGATAGCTCA GTAACACGTG
      T3.1  ATGCTAGTCG CACGGGTTTA GACCCGTGGC AGATAGCTCA GTAACACGTG
      T3.2  ATGCTAGTTG CACGAGTTTA GACTCGTAGC ATATAGCTCA GTAACACGTG
      T3.5  ATGCTAGTTG NACGAGTTTA GACTCGTNGC ATATAGCTCA GTAACACGTG
      T3.6  ATGCTAGTCG CACGGGTTTA GACCCGTGGC AGATAGCTCA GTAACACGTG
      T4.2  ATGCTAGTCG CACGGGTTTA GGCCCGTGGC GAAAAGCTCA GTAACACGTG
      T5.7  ATGCTAGTTG CNCGAGTTTA GACTCGTAGC AGATAGCTCA GTAACACGTG
    WB4E.8  ATGCTAGTCG CACGAGTTCA GACTCGTGGC GAATAGCTCA GTAACACGTG
   WM2C.13  ATGCTAGTTG TGCGGGTTTA GACCCGCAGC GGAAAGCTCA GTAACACGTG
 Haloferax  ATGCTAGTTG CACGAGTTCA TACTCGTGGC GAAAAGCTCA GTAACACGTG
 Halobact.  ATGCTAGTTG TGCGGGTTTA GACCCGCAGC GGAAAGCTCA GTAACACGTG
Halorubrum  ATGCTAGTCG CACGAGTTTA GACTCGTGGC GAATAGCTCA GTAACACGTG
 Natronob.  ATGCTAGTTG TACGAGTTTA GACTCGTAGC AGATAGCTCA GTAACACGTG


            101                                                 150
Methanosp.  GACAATCTGC CCTGAAGAGG AGGATAATCC CGGGAAACTG GGGGTAATAC
       777  NCCANACTAC CCTATGGATC CANATAACCT CGGGAAACTG AGGCTAATGT
       784  NCCANACTAC CCTATGGATC CACATAACCT CGGGAAACTG AGGCTAATGT
       786  NCCANACTAC CCTATGGATC CAAATAACCT CGGGAAACTG AGGCTAATGT
   B1bra.5  GCCAAACTAC CCTGTGGATG GGAACAATCT CGGGAAACTG AGGCTAATTC
    B2S.26  GCCANACTAC CCTTCGGAGA CGCATAACCT CGGGAAACTG AGGCTAATAC
  GSL5.48   NCCANACTAC CCTTCGGAAC ACAATACCCT CGGGAAACTG AGGCTAATAG
Haloarcula  GCCAAACTAC CCTACAGACC GCGATAACCT CGGGAAACTG AGGCCAATAG
      Nh.2  NCCACACTNC CCTATGGACT GGGATAACCT CGGGAAACTG AGGCTAATCC
      Nh.5  NCCACACTAC CCTTCGGACT ACAATACCCT CGGGAAACTG AGGCTAATAG
     PIG.5  NCCACACTAC CCTATGGAGA GCGATAACCT CGGGAAACTG AGNCTAATAG
     PIG.8  GCCANACTAC CCTTCGGAGC ACGATACCCT CGGGAAACTG AGGCTAATAG
     PW1.1  GCCAAACTAC CCTTCGGACC GTCATACCCT CGGGAAACTG AGGTTAATAA
     PW5.4  NCCAAACTAC CCTGTGGATG GGAACAATCT CGGGAAACTG AGGCTAATTC
     SC1.2  NCCACACTAC CCTTCGGAAC ACAATACCCT CGGGAAACTG AGGCTAATAG
     SR1.5  NCCAAACTAC CCTATGGATC CGGACAACCT CGGGAAACTG AGGCTAATCC
      T1.3  NCCACACTAC CCTACAGAGA CGAACAACCT CGGGAAACTG AGGCTAATTC
      T1.6  NCCACACTAC CCTCTGGACG GGAATACCCT CGGGAAACTG AGGCTAATCC
      T3.1  GCCAAACTGC CCTATGGATG GAGACAACCT CGGGAAACTG AGGCTAATCT
      T3.2  GCCAAACTAC CCTACAGACN GCGATAACCT CGGGAAACTG AGGCCAATAG
      T3.5  GCCAAACTAC CCTACAGACC GCAATAACCT CGGGAAACTG AGGCCAATAG
      T3.6  GCCAAACTGC CCTATGGATG GAGACAACCT CGGGAAACTG AGGCTAATCT
      T4.2  GCCAAACTGC CCTGTGGACG GAGATACCCT CGGGAAACTG AGGTCAATCT
      T5.7  NCCAAACTAC CCTATGGATC CGAATAACCT CGGGAAACTG AGGCTAATGC
    WB4E.8  NCCANACTAC CCTTCGGAGN ACGATACCCT CGGGAAACTG AGGCTAATAG
   WM2C.13  GCCAAACTAC CCTGTGGATG GGAACAATCT CGGGAAACTG AGGCTAATTC
 Haloferax  GCCAAACTAC CCTACAGAGA ACGATAACCT CGGGAAACTG AGGCTAATAG
 Halobact.  GCCAAGCTAC CCTGTGGACG GGAATACTCT CGGGAAACTG AGGCTAATCC
Halorubrum  GCCAAACTAC CNTTCGGAGC TCCATACCCT CGGGAAACTG AGGCTAATAG
 Natronob.  GCCAAACTAC CCTATGGATC CAGACAACCT CGG.AAACTG AGGCTAATCT
```

```
            151                                                     200
Methanosp.  TCCATAGTTC GTGCTGACTG GAA.TGTTAT GCGAACGAAA .GATCCGTCG
       777  GGAATAGCGT TCATCTCCTG GAG.TGGCAC GAACGCGAAA CGTTACGGCG
       784  GGAATAGCGT TCATCNCCTG GAG.TGNCAC GAACGCGAAA CGTTACGGCG
       786  GGAATAGCGT TCATCGCCTG GAG.TGGCAC GAACGCGAAA CGTTACGGCG
   B1bra.5  CCAATAACGC TCCACCCCTG GAA.CGGGCG GAGCTGGAAA CGCTACGGCG
   B2S.26   GTCACACGGC TTTCAGCCTG GAA.TGGCGA AAGCCCGAAA CGCTCCGGCG
   GSL5.48  TGTATACCAC ATCACCACTG GAA.TGAGTG ATGTGCCACA CGCTCCGGCG
Haloarcula  CGGATATAAC TCTCATGTTG GAG.TGCCGA GAGTTAGAAA CGTTCCGGCG
      Nh.2  CGGATATCAC ATCCANCCTG GAACTGGCGG ATGTGTGAAA CGCTCCGGCG
      Nh.5  TAGACACCAC AGTCCACCTG GAA.TGAGGA CTGTGCCACA CGCTCCGGCG
      PIG.5 TNCATACGAC TCTCTCCCTG GAAATGGGTT GAGNCNGAAA CGCTNCGGCG
      PIG.8 TGCATCCCAC ANCCCACCTG GAA.TGAGGN CTGTGTCANA CGCTCCGGCG
      PW1.1 CAGATACCAC AGTCCACCTG GAA.TGAGGA CTGTGCCAAA CGCTCCGGCG
      PW5.4 CCAATAACGC TCCACCCCTG GAA.CGGGCG GAGCTGGAAA CGCTACGGCN
      SC1.2 TGTATATCAT ACCACCACTG GAA.TGAGTG GTATGCCACA CGCTCCGGCG
      SR1.5 GGAATACGAC TCTCAGCCTG GAAGTGGCGA GAGTCAGAAA CGCTCCGGCG
      T1.3  GTCATACCCC TCTCACGCTG GAA.TGCCAA GAGGCACACA CGCTNNCGCG
      T1.6  CCGATACTAC TCNCTGCCTG GAA.TGGCGC GAGATCCANA CGCTCCGGCG
      T3.1  CCAATACAGC TCTCACGTTG GAGTTGCAGA GAGCTGGAAA TGCTCCGGCG
      T3.2  CGGATATNAC TCTCANNCTG GAG.TGNCGA GAGTGAGAAA CGTTCCGGCG
      T3.5  TGGATATAAC TCTCATGCTG GAG.TGCCGA GAGTTAGAAA CGTTCCGGCG
      T3.6  CCAATACAGC TCTCACGTTG GAGTTGCAGA GAGCTGGAAA TGCTCCGGCG
      T4.2  CCGATACGGC TCGCACGCTG GAG.TGCGGC GAGCCGGAAA CGCTCCGGCG
      T5.7  GGAATAGCGT TCATCGCCTG GAG.TGGNAC GAACGCGAAA CGTTACGGCG
    WB4E.8  TGCATACCAC AATCCAGCTG GAA.TGAGGA TTGTGCCANA CGCTCCGGCG
   WM2C.13  CCAATAACGC TCCACCCCTG GAA.CGGGCG GAGCTGGAAA CGCTACGGCG
Haloferax   TTCATACGGG AGTCATGCTG GAA.TGCCGA CTCCCCGAAA CGCTCAGGCG
Halobact.   CCGATAACGC TTTGCTCCTG GAA.GGGGCA AAGCCGGAAA CGCTCCGGCG
Halorubrum  AGCATACCAC AATCCAGCTG GAA.TGAGGA TTGTGCCAAA CGCTCCGGCG
Natronob.   GGAATACGAC TCTCATCCTG GAG.TGGAGA GAGTCCGAAA TGCTCCGGCG

            201                                                     250
Methanosp.  CTTCAGGATG AGTCTGCGGC CGATTAGGTA GTTGTTGGGG TAACGGCCCA
       777  CCATAGGATG TGNCTNCGCC CGATTAGGTA GACGGTGGGG TAACGNCCCA
       784  CCATAGGATG TGTCTNCGCC CGATTAGGTA GACGGTGGGG TAACGGCCCA
       786  CCATAGGATG TGNCTGCGCC CGATTAGGTA GACGGTGGGG TAACGGCCCA
   B1bra.5  CCACAGGATG TGGCTGCGGT CGATTAGGTA GACGGTGGGG TAACGGCCCA
   B2S.26   CCGAAGGATG TGNCTGCGNC CGATTAGGTA GACGGTGGGG TAACGGCCCA
   GSL5.48  CCGAAGGATG TGNCTGCGNC CGATTAGGTA GACGGTGGGG TAACGGCCCA
Haloarcula  CTGTAGGATG TGGCTGCGGC CGATTAGGTA GATGGTGGGG TAACGGCCCA
      Nh.2  CCATAGGATG TGTCTGCGCC CGATTAGGTA GACGGTGGGG TAACGGCCCA
      Nh.5  CCGAAGGATG TGTCTGCGNC CGATTAGGTA GACGGTGGGG TAACGGCCCA
      PIG.5 CCATAGGATG TGNCTCCGCC CGATTAGGTA GACGGTGGGG TAACGGCCCA
      PIG.8 CCGAAGGATG TGNCTNCGCC CGATTAGGTA GACGGTGGGG TAACGGCCCA
      PW1.1 CCGAAGGATG TGNCTGCGNC CGATTAGGTA GACGGTGGGG TAACGGCCCA
      PW5.4 CCACAGGATG TGNCTGCGGT CGATTAGGTA GACGGTGGGG TAACGGCCCA
      SC1.2 CCGAAGGATG TGTCTNCGCC CGATTAGGTA GACGGTGGGG TAACGGCCCA
      SR1.5 CCATAGGATG TGNCTGCGNC CGATTAGGTA GACGGTGGGG TAACGGCCCA
      T1.3  CTGTAGGATG TGNCTNCGCC CGATTAGGTA GACGGTGGGG TAACGGCCCA
      T1.6  CCAGAGNCTG TGNCTGCGNC CGATTAGGTA GACGGTGGGG TAACGGCCCA
      T3.1  CCATAGGATG TGNCTGCGGC CGATTAGGTA GACGGTGGGG TAACGGCCCA
      T3.2  CTGTAGGATG TGGCTGCGNC CGATTAGGTA GATGGTGGGG TAACGGCCCA
      T3.5  CTGTAGGATG TGGCTGCGNC CGATTAGGTA GATGGTGGGG TAACGGCCCA
      T3.6  CCATAGGATG TGNCTGCGGC CGATTAGGTA GACGGTGGGG TAACGGCCCA
      T4.2  CCACAGGATG TGGCTGCGGC CGATTAGGTA GACGGTGAGG TAACGGCTCA
      T5.7  CCATAGGATG TGNCTGCGNC CGATTAGGTA GACGGTGGGG TAACGNCCCA
    WB4E.8  CCGAAGGATG TGNCTNCGCC CGATTAGGTA GACGGTGGGG TAACGGCCCA
   WM2C.13  CCACAGGATG TGGCTGCGGT CGATTAGGTA GACGGTGGGG TAACGGCCCA
Haloferax   CTGTAGGATG TGGCTGCGGC CGATTAGGTA GACGGTGGGG TAACGGCCCA
Halobact.   CCACAGGATG CGGCTGCGGT CCATTAGGTA GACGGTGGGG TAACGGCCCA
Halorubrum  CCGAAGGATG TGGCTGCGGC CGATTAGGTA GACGGTGGGG TAACGGCCCA
Natronob.   CCATAGGATG TGGC..CGGC CGATTAGGTA GACGGTGGGG TAACGGCCCA
```

```
          251                                                       300
Methanosp. ACAAGCCTGT CATCGGTACG GGTTGTGGGA GCAAGAGCCC GGAGATGGAT
       777 CCGTGCCCAT AATCGGTACG GGTTGTGAGA GCAAGAGCCC GGAGACGGTA
       784 CCGTGCCCAT AATCGGTACG GGTTGTGAGA GCAAGAGCCC GGAGACGGTA
       786 CCGTGCCNAT AATCGGTACG GGTTGTGAGA GCAAGAGCCC GGAGACGGTA
   B1bra.5 CCGTGCCAAT AATCGGTACG GGTTGTGAGA GCAAGAGCCC GGAGACGGAA
   B2S.26 CCGTGCCGAT AATCGGTACG GGTTGTGAGA GCAAGAACCC GGAGACGGAA
   GSL5.48 CCGTGCCAAT AATCGGTACG GGTCATGAGA GTGAGAACCC GGAGACGGAA
Haloarcula CCATGCCGAT AATCGGTACG GGTTGTGAGA GCAAGAACCT GGAGACGGTA
     Nh.2 CCGTNCCGAT AATCGGTACG GGTCATGAGA GTGAGAACCC GGAGACGGAC
     Nh.5 CCGTGCCAAT AATCGGTACG GGTCATGAGA GTGAGAACCC GGAGACGGAA
     PIG.5 CCGTGCCAAT AATCGGTANG GGTTGTGAGA GCAAGAGCCN AGAGACGGAA
     PIG.8 CCGTGCCAAT AATCGGTACG GGTCATGAGA GTGAGAACCC GGAGACGGAA
     PW1.1 CCGTGCCAAT AATCGGTACG GGTCATGAGA GTGAGAACCC GGAGACGGAA
     PW5.4 CCGTGCCAAT AATCGGTACG GGTTGTGAGA GCAAGAGCCC GGAGACGGAA
     SC1.2 CCGTGCCAAT AATCGGTACG GGTCATGAGA GTGAGAACCC GGAGACGGAA
     SR1.5 CCGTGCCCAT AATCGGTACG GGTTGTGAGA GCAAGAGCCC GGAGACGGTA
      T1.3 CCGTGCCCAT AATCGGTACG GGTTGTGAGA GCAAGAGCCC GGAGACGGAA
      T1.6 CCGTGCCCAT AATCGGTACG GGTTGTGAGA GCAAGAGCCC GGAGACGGTA
      T3.1 CCGTGCCGAT AATCGGTACG GGTTGTGAGA GCAAGAGCCC GGAGACGGAA
      T3.2 CCATGCCGAT AATCGGTACN GGTTGTGAGA GCAAGAGCCT GGAGACGGTA
      T3.5 CCATGCCGAT AATCGGTACN GGTTGTGAGA GCAAGAACCT GGAGACGGTA
      T3.6 CCGTGCCGAT AATCGGTACG GGTTGTGAGA GCAAGAGCCC GGAGACGGAA
      T4.2 CCGTGCCGAT AATCGGTACG GGTCATGAGA GTGAGAACCC GGAGACGGTA
      T5.7 CCGTGCCCAT AATCGGTACG GGTTGTGAGA GCAAGAGCCC GGAGACGGTA
   WB4E.8 CCGTGCCAAT AATCGGTACG GGTCATGAGA GTGAGAACCC GGAGACGGAA
   WM2C.13 CCGTGCCAAT AATCGGTACG GGTTGTGAGA GCAAGAGCCC GGAGACGGAA
Haloferax CCGTGCCGAT AATCGGTACG GGTTGTGAGA GCAAGAGCCC GGAGACGGAA
Halobact. CCGTGCCCAT AATCGGTACG GGTTGTGAGA GCAAGAGCCC GGAGACGGAA
Halorubrum CCGTGCCAAT AATCGGTACG GGTCATGAGA GTGAGAACCC GGAGACGGAA
Natronob. CCGTGCCGAT AATCGGTACG GGTTGTGAGA GCAAGAGCCC GGAGACGGTA

          301                                                       350
Methanosp. TCTGAGACAC GAATCCAGGC CCTACGGGGC GCAGCAGGCG CGAAAACTTT
       777 TCTGAGACAA GATACCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
       784 TCTGAGACAA GATACCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
       786 TCTGAGACAA GATACCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
   B1bra.5 TCTGAGACAA GATTCCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
   B2S.26 TCTGAGACAA GATTCCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
   GSL5.48 TCTGAGACAA GATTCCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
Haloarcula TCTGAGACAA GATACCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
     Nh.2 TCTGAGACAC GAGTCCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
     Nh.5 TCTGAGACAA GATTCCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
     PIG.5 TCTGAGACAA GATTCCGGNC CCTACGGGGC GCAGCAGGCG CGAAAACTTT
     PIG.8 TCTGAGACAA GATTCCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
     PW1.1 TCTGAGACAA GATTCCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
     PW5.4 TCTGAGACAA GATTCCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
     SC1.2 TCTGAGACAA GATTCCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
     SR1.5 TCTGAGACAA GATACCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
      T1.3 TCTGAGACAA GATTCCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
      T1.6 TCTGAGACAA GATACCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
      T3.1 TCTGAGACAA GATTCCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
      T3.2 TCTGAGACAA GATACCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
      T3.5 TCTGAGACAA GATACCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
      T3.6 TCTGAGACAA GATTCCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
      T4.2 TCTGAGACAA GATGCCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
      T5.7 TCTGAGACAA GATACCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
   WB4E.8 TCTGAGACAA GATTCCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
   WM2C.13 TCTGAGACAA GATTCCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
Haloferax TCTGAGACAA GATTCCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
Halobact. TCTGAGACAA GATTCCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
Halorubrum TCTGAGACAA GATTCCGGGC CNTACGGGGC GCAGCAGGCG CGAAACCTTT
Natronob. TCTGAGACAA GATACCGGGC CCTACGGGGC GCAGCAGGCG CGAAACCTTT
```

```
            351                                                     400
Methanosp.  ACCATGCGGG CAACCGTGAT AAGGAAACCC CGAGTGCCAG ..CACA.GG.
        777 ACACTGCACG CGAGTGCGAT ANGGGGACTC CGAGTGCGAG GGCATATAGT
        784 ACACTGCACG CGAGTGCGAT ANGGGGACTC CGAGTGCGAG GGCATATAGT
        786 ACACTGCACG CGAGTGCGAT AAGGGGACTC CGAGTGCGAG GGCATATAGT
    B1bra.5 ACACTGTACG CAAGTGCGAT AAGGGGACTC CGAGTGTGAA GGCATAGAGC
    B2S.26 ACACTGCACG CCAGTGCGAT AAGGGGACCC TGAGTGCACA GGCATAGTGC
    GSL5.48 ACACTGCACG ACAGTGCGAT AGGGGGATCC CAAGTGCACA GGCATAGCGC
Haloarcula ACACTGCACG ACAGTGCGAT AGGGGGACTC CGAGTGTGAG GGCATATAGC
        Nh.2 ACACTGCACG ACAGTGCGAT AGGGGGACAC CGAGTGCGAG GGCATAGAGC
        Nh.5 ACACTGCACG ACAGTGCGAT AAGGGAATCC CAAGTGCGTA GGCATAGAGC
      PIG.5 ACACTGCACG ACAGTGCGAT AAGGGAACTC CAAGTGCGAG GGCATAGAGC
      PIG.8 ACACTGCACG CCAGTGCGAT AAGGGGATTC CAAGTGCGCA GGCATAGAGC
      PW1.1 ACACTGCACG CCAGTGCGAT AAGGGAATCC CAAGTGCGTA GGCATAGAGC
      PW5.4 ACACTGTACG CAAGTGCGAT AAGGGGACTC CGAGTGTGAA GGCATAGAGC
      SC1.2 ACACTGCACG ACAGTGCGAT AGGGGGATCC CAAGTGCACA GGCATAGCGC
      SR1.5 ACACTGCACG AAAGTGCGAT AAGGGGACTC CAAGTGCGAG GGCATATAGT
       T1.3 ACACTGCACG CACGTGCGAT AAGGGGACTC CGAGTGCGAG GGCATATCGT
       T1.6 ACACTGCACG CCAGTGCGAT AAGGGGATCC CCAGTGCGAG GGCATACAGT
       T3.1 ACACTGCACG ACAGTGCGAT AAGGGGACTC CGAGTGCGAG GGCATACAGT
       T3.2 ACACTGCACG ACAGTGCGAT AGGGGGACTC CGAGTGTGAG GGCATATAGC
       T3.5 ACACTGCACG ACAGTGCGAT AGGGGGACTC CGAGTGTGAG GGCATATAGC
       T3.6 ACACTGCACG ACAGTGCGAT AAGGGGACTC CGAGTGCGAG GGCATACAGT
       T4.2 ACACTGCACG ACAGTGCGAT AGGGGGACTC CGAGTGCGAG GGCATATAGT
       T5.7 ACACTGCACG NGAGTGCGAT AAGGGGACTC CGAGTGCGAG GGCATATAGT
      WB4E.8 ACACTGCACG CCAGTGCGAT AAGGGAATCC CAAGTGCGCA GGCATAGAGC
    WM2C.13 ACACTGTACG CAAGTGCGAT AAGGGGACTC CGAGTGTGAA GGCATAGAGC
Haloferax ACACTGCACG CAAGTGCGAT AAGGGGACCC CAAGTGCGAG GGCATATAGT
Halobact. ACACTGTACG AAAGTGCGAT AAGGGGACTC CGAGTGTGAA GGCATAGAGC
Halorubrum ACACTGCACG ACAGTGCGAT AAGGGAATCC CAAGTNCGCA GGCATAGAGC
Natronob. ACACTGCACG CCAGTGCGAT AAGGGGACTC CAAGTGCGAG GGCATATAGT

            401                                                     450
Methanosp.  .CTGGCTGTC CACCAGTGTA AATA.ACTGG TGAAGAAAGG GCCGGGCAAG
        777 CCTCGCTTTT CACGACCGTA AGGT.GGTCG TAGAATAAGT GCTGGGCAAG
        784 CCTCGCTTTT CACGACCGTA AGGT.GGTCG TAGAATAAGT GCTGGGCAAG
        786 CCTCGCTTTT CACGACCGTA AGGT.GGTCG TAGAATAAGT GCTGGNCAAG
    B1bra.5 CTTCACTTTT GTACACCGTA AGGT.GGTGT ACGAATAAGG GCTGGNCAAG
    B2S.26 CTGTGCTTTT CTCGACCGTA GGGA.GGTCG AGGAACAAGA GCTGGGCAAG
    GSL5.48 CTGTGCTTTT CGGTACCGTA AGGT.GGTAC CAGAATAAGG GCTGGNCAAG
Haloarcula CCTCGCTTTT CTGTACCGTA AGGT.GGTAC AGGAACAAGG ACTGGGCAAG
        Nh.2 CCTCGCTTTT GTAGACCGTA AGGA.GGTCT ACGAATAAGG GCTGGGCAAG
        Nh.5 CTACGCTTTT GTCCACCGTA GGGA.GGTGG ACGAATAAGG GCTGGGCAAG
      PIG.5 CCTCGCTTTT CTGAACCCTA GGGT.GGTTC AGGAACAAGA GCTGGGCAAG
      PIG.8 CTGCGCTTTT GTACACCGTA GGGA.GGTGT ACGAATAAGG GCTGGGCAAG
      PW1.1 CTACGCTTTT GTCCACCGTA GGGA.GGTGG ACGAATAAGG GCTGGGCAAG
      PW5.4 CTTCACTTTT GTACACCGTA AGGT.GGTGT ACGAATAAGG GCTGGGCAAG
      SC1.2 CTGTGCTTTT CGGTACCGTA AGGT.GGTAC CAGAATAAGG GCTGGGCAAG
      SR1.5 CCTCGCTTTT CACTACCGTA AGGT.GGTAG TAGAATAAGT GCTGGGCAAG
       T1.3 CCTCGCTTTT GTGTACCGTA GGGC.GGTAC ACGAACAAGA GCTGGGCAAG
       T1.6 CCTCGCTTTT CACCACCGTA AGGT.GGTGG TGGAACAAGA GCTGGGCAAG
       T3.1 CCTCGCTTTT CTGTACCGTA AGGT.GGTAC AGGAACAAGT GCTGGGCAAG
       T3.2 CCTCGCTTTT CTGTACCGTA AGGT.GGTAC ANGAANAAGG ACTGGGCAAG
       T3.5 CCTCGCTTTT CTGNACCGTA AGGT.GGTNC ANGAANAAGG ACTGGGCAAG
       T3.6 CCTCGCTTTT CTGTACCGTA AGGT.GGTAC AGGAACAAGT GCTGGGCAAG
       T4.2 CCTCGCTTTT CTGTACCGTA AGGT.GGTAC AAGAATAAGG GCTGGGCAAG
       T5.7 CCTCGCTTTT CACGACCGTA AGGT.GGTCG TAGAATAAGT GCTGGGCAAG
      WB4E.8 CTGCGCTTTT GTACACCGTA GGGA.GGTGT ACGAATAAGG GCTGGGCAAG
    WM2C.13 CTTCACTTTT GTACACCGTA AGGT.GGTGT ACGAATAAGG GCTGGGCAAG
Haloferax CCTCGCTTTT CTCGACCGTA AGGC.GGTCG AGGAATAAGA GCTGGGCAAG
Halobact. CTTCACTTTT GTACACCGTA AGGT.GGTGC ACGAATAAGG ACTGGGCAAG
Halorubrum CTGCGCTTTT GTACACCGTA GGGA.GGTGT ACGAATAAGG GCTGGGCAAG
Natronob. CCTCGCTTTT TGCGACCGTA AGGTTGGTCG CGGAATAAGT GCTGGGCAAG
```

```
          451                                                        500
Methanosp. ACCGGTGCCA GCCGCCGCGG TAATACCGGC GGCTCGAGTG GTGGCCGCTA
       777 ACCGGTGNCA GCCGCCGNGG TAATACCGGN AGCACGAGTG ATGACCGCTA
       784 ACCGGTGNCA GCCGNCGNGG TAATACCGGC AGCACGAGTG ATGACCGCTA
       786 ACCGGTGNCA GCCGCCGNGG TAATACCGGN AGCACGAGTG ATGACCGCTA
   B1bra.5 ACCGGTGCCA GCCGCCGCGG TAATACCGGN AGCCCGAGTG ATGGCCGATA
   B2S.26 ACCGGTGCCA GCCGCCGNGG TAATACCGGC AGCTCGAGTG ATGGCCGATC
   GSL5.48 ACCGGTGACA GCCGCCGNGG TAATACCGNN AGCCCAAGTG ATGGNCGATC
Haloarcula ACCGGTGCCA GCCGCCGCGG TAATACCGGC AGTCCGAGTG ATGGCCGATA
     Nh.2 ACCGGTGCCA GCCGCCGCGG TAACACCGGC AGCCCAAGTG ATGGCCGATA
     Nh.5 ACCGGTGACA GCCGCCGNGG TAATACCGNN AGCCCAAGTG ATGGNCGATC
     PIG.5 ACCGGTGCCA GCCGCCGCGG TAACACCGGC AGCTCGAGTG ATGACCGATA
     PIG.8 ACCGGTGCCA GCCGCCGCGG TAATACCGGC AGCCCGAGTG ATGGNCGATC
     PW1.1 ACCGGTGCCA GCCGCCGCGG TAATACCGGC AGCCCGAGTG ATGGNCGATC
     PW5.4 ACCGGTGCCA GCCGCCGCGG TAATACCGGC AGCCCGAGTG ATGGCCGATA
     SC1.2 ACCGGTGCCA GCCGNCGCGG TAATACCGGN AGCCCAAGTG ATGGCCGATC
     SR1.5 ACCGGTGCCA GCCGCCGNGG TAATACCGGN AGCACGAGTG ATGACCGCTC
     T1.3 ACCGGTGCCA GCCGCCGCGG TAATACCGGC AGCTCAAGTG ATGGCCGATC
     T1.6 ACCGGTGNCA GCCGCCGCGG TAATACCGGC AGCTCGAGTG ATGGCCGTTC
     T3.1 ACCGGTGCCA GCCGCCGCGG TAATACCGGC AGCACGAGTG ATGGCCGCTA
     T3.2 ACCGGTGNCA GCCGCCGNGG TAATACCGGC AGTCCAAGTG ATGGCCGATA
     T3.5 ACCGGTGNCA GCCGCCGCGG TAATACCGGC AGTCCAAGTG ATGGNCGATA
     T3.6 ACCGGTGACA GCCGCCGNGG TAATACCGGC AGCACGAGTG ATGGCCGCTA
     T4.2 ACCGGTGCCA GNCGCCGCGG TAATACCGGC AGCCCGAGTG ATGGCCGCTA
     T5.7 ACCGGTGCCA GCCGCCGNGG TAATACCGGN AGCACGAGTG ATGACCGCTA
   WB4E.8 ACCGGTGCCA GCCGCCGCGG TAATACCGGC AGCCCGAGTG ATGGNCGATC
  WM2C.13 ACCGGTGCCA GCCGCCGCGG TAATACCGGN AGNCCGAGTG ATGGNCGATA
Haloferax ACCGGTGCCA GCCGCCGCGG TAATACCGGC AGCTCAAGTG ATGACCGATA
Halobact. ACCGGTGCCA GCCGCCGCGG TAATACCGGC AGTCCGAGTG ATGGCCGATC
Halorubrum ACCGGTGCCA GCCGCCGCGG TAATACCGGC AGCCCGAGTG ATGGCCGATC
Natronob. ACCGGT.CCA GCCGCCGCGG TAA.ACCGGC AGCACGA..G ATGACCGGTG

          501                                                        550
Methanosp. TTACTGGGCT TAAAGGGTCC GTAGCTGGAT ATACAAGTCC CTTGAGAAAT
       777 TTATTGGGCC TAAAGCGTCC GTAGCTGGCC ACGCAAGTCT ATCGGGAAAT
       784 TTATTGGGCC TAAAGCGTCC GTAGCTGGCC ACGCAAGTCT ATCGGGAAAT
       786 TTATTGGGCC TAAAGCGTCC GTAGCTGGCC ATGCAAGTCC ATCGGGAAAT
   B1bra.5 TTATTGGGCC TAAAGCGTCC GTAGCTGGCC GGACAAGTCC GTTGGGAAAT
   B2S.26 TTATTGGGCC TAAAGCGTCC GTAGCTGGNC AGGNAAGTCC GTCGGGAAAT
   GSL5.48 TTATTGGGCC TAAAGCGTCC GTAGCTGGCC GCGCAAGTCC ATCGGGAAAT
Haloarcula TTATTGGGCC TAAAGCGTCC GTAGCCGGCC GAACAAGTCC GTTGGGAAAT
     Nh.2 TTATTGGGCC TAAAGCGTCC GTAGCTGGCC GCGCAAGTTC GTCGGGAAAT
     Nh.5 TTATTGGGCC TAAAGCGTCC GTAGCTGGCC GCACAAGTCT ATCGGAAAAT
     PIG.5 TTATTGGGCC TAAAGCGTCC GTAGCTGGCN GAGAAGGTCT GTCGGNAAAT
     PIG.8 TTATTGGGCC TAAAGCGTCC GTAGCTGGCC GCACAAGTCC ATCGGAAAAT
     PW1.1 TTATTGGGCC TAAAGCGTCC GTAGCTGGCC GCGCAAGTCT ATCGGAAAAT
     PW5.4 TTATTGGGCC TAAAGCGTCC GTAGCTGGCC GGACAAGTCC GTTGGGAAAT
     SC1.2 TTATTGGGCC TAAAGCGTCC GTAGCTGGCC GCGCAAGTCC ATCGGGAAAT
     SR1.5 TTATTGGGCC TAAAGCGTCC GTAGCTGGCC ACGCAAGTCT ATCGGGAAAT
     T1.3 TTATTGGGCC TAAAGCGTCC GTAGCTGGCC TCGAAAGTCC GTCGGGAAAT
     T1.6 TTATTGGGCC TAAAGCGTCC GTAGCCGGCC AGGTAAGTTC GTCGGGAAAT
     T3.1 TTATTGGGCC TAAAGCGTCC GTAGCCGGCC GAGCAAGTTC GTTGGGAAAT
     T3.2 TTATTGGGCC TAAAGCGTCC GTAGCTTGCT GTNTAAGTCC ATTGGGAAAT
     T3.5 TTATTGGGCC TAAAGCGTCC GTAGCNNGCN GNNNAAGTCC ATTGGGAAAT
     T3.6 TTATTGGGNC TAAAGCGTCC GTAGCCGGNC GAGCAAGTTC GTTGGGAAAT
     T4.2 TTATTGGGCC TAAAGCGTCC GTAGCCGGCC AGACAAGTCC GTTGGGAAAT
     T5.7 TTATTGGGCC TAAAGCGTCC GTAGCTGGCC ACGCAAGTCT ATCGGGAAAT
   WB4E.8 TTATTGGGCC TAAAGCGTCC GTAGCTGGCC GCACAAGTCC ATCGGAAAAT
  WM2C.13 TTATTGGGCC TAAAGCGTCC GTAGNTGGNC GGACAAGTCC GTTGGGAAAT
Haloferax TTATTGGGCC TAAAGCGTCC GTAGCCGGCC ACGAAGGTTC ATCGGGAAAT
Halobact. TTATTGGGCC TAAAGCGTCC GTAGCTGGCT GAACAAGTCC GTTGGGAAAT
Halorubrum TTATTGGGCC TAAAGCGTCC GTAGCTGGCC GCACAAGTCC ATCGGAAAAT
Natronob. TTATTGGGCC TAAAGCGTCC GTAGCTGGCC GCGCAAGTCT ATCGGGAAAT
```

```
            551                                                        600
Methanosp.  CCGCCGGCTT AACCGGTGGG CG.TTCAGGG GAAACTGTAT TTCTAGGGAC
       777  CCGCGCGCTT AACGCGCGGG CG.TCCGGTG GAAACTGCGT GGCTTGGGAC
       784  CCGCGCGCTT AACGCGCGGG CG.TCCGGTG GAAACTGCGT GGCTTGGGAC
       786  CCGCGCGCTT AACGCGCGGG CG.TCCGGTG GAAACTGCAT GGCTTGGGAC
   B1bra.5  CTGTTCGCTT AACGAGCAGG CG.TCCAGCG GAAACTGTTC GGCTTGGGAC
    B2S.26  CCACGTGCTC AACGNGTGGG CG.NCCGGNG GAAACTGNCT AGCTAGGGAC
   GSL5.48  CCACCTGCTC AACAGGTGGG CG.CCCGGTA GAAACTGCGT GGCTTGGGAC
Haloarcula CGACGCGCTC AACGCGTCGG CG.TCCAGCG GAAACTGTCC GGCTTGGGGC
       Nh.2  CCACTCGCTT AACGAGTGGG CG.TCCGGCG AAAACTGTGC GGCTTGGGAC
       Nh.5  CCACCCGCNC AACGGGTGGG CG.TCCGGTA GAAACTGTGT GGCTTGGGAC
      PIG.5  CTGCTCGCTC AACGAGCAGG CG.TCCGGCA GAAACNANTC NGCTTGGGAC
      PIG.8  CCACCCGCTC AACGGGTGGG CG.TCCGGTG GAAACTGTGT GGCTTGGGAC
      PW1.1  CCACCCGCNC AACGGGTGGG CG.TCCGGTA GAAACTGCGT GGCTTGGGAC
      PW5.4  CTGTTCGCTT AACGAGCAGG CG.TCCAGCG GAAACTGTTC GGCTTGGGAC
      SC1.2  CCACCTGCTC AACAGGTGGG CG.CCCGGTA GAAACTGCGT GGCTTGGGAC
      SR1.5  CCGCGCGCTC AACGCGCGGG CG.TCCGGTG GAAACTGTGT GGCTTGGGAC
       T1.3  CCACGCGCCC AACGCGTGGG CG.TCCGGCG GAAACTCCGA GGCTTGGGAC
       T1.6  CCACACGCTC AACGTGTGGG CG.TCCGGCG AAACCTGNCT GGCTTGGGAC
       T3.1  CTGCCCGCTT AACGGGCAGG CG.TCCGGCG AAAACTGTTC GGCTTGGGGC
       T3.2  CGACCAGCTC AACTGGTCGG CG.TCCAGTG GAAACTACAC AGCTTGGGGC
       T3.5  CGACNAGCTC AACNGGTCGN CG.TCCAGCG GAAACTANAC AGCTTGGGGC
       T3.6  CTGCCCGCTT AACGGGCAGG CG.TCCGGCG AAAACTGTTC GGCTTGGGGC
       T4.2  CGGCGCGCTC AACGCGTCGG CG.TCCGGCG GAAACTGTCT GGCTTGGGGC
       T5.7  CCGCGCGCTT AACGCGCGGG CG.TCCGGTG GAAACTGNGT GGCTTGGGAC
     WB4E.8  CCACCCGCCC AACGGGTGGG CG.TCCGGTG GAAACTGTGT GGCTTGGGAC
    WM2C.13  CTGTTCGCTT AACGAGCAGG CG.TCCAGCG GAAACTGTTC GGCTTGGGAC
 Haloferax  CCGCCAGCTC AACTGGCGGG CG.TCCGGTG AAAACCACGT GGCTTGGGAC
 Halobact.  CTGTCCGCTT AACGGGCAGG CG.TCCAGCG GAAACTGTTC AGCTTGGGAC
Halorubrum  CCACCCGCTC AACGGGTGGG CG.TCCGGTG GAAACTGTGT GGCTTGGGAC
 Natronob.  CTCTTCGCTT AACGGAGAGG CGATCCGGTG GAA.AAATGT GGCTTGGGAC

            601                                                        650
Methanosp.  CGGGAGAGGT GAGAGGTACT GCCGGGGTAG GAGTGAAATC CTGTAATCCC
       777  CGGAAGACCA GAGGGGTACG TCCGGGGTAG GAGTGAAATC CCGTAATCCT
       784  CGGAAGACCA GAGGGGTACG TCCGGGGTAG GAGTGAAATC CCGTAATCCT
       786  CGGAAGACCA GAGGGGTACG TCCGGGGTAG GAGTGAAATC CCGTAATCCT
   B1bra.5  CGGAAGACCT GAGGGGTACG TCCGGGGTAG GAGTGAAATC CTGTAATCCT
    B2S.26  CGGAAGGCGC GACGGGTACG TCTAGGGTAG GAGTGAAATC CCGTAATCCT
   GSL5.48  CGGAAGGCGC GACGGGTACG TCCGGGGTAG GAGTGAAATC CCGTAATCCT
Haloarcula CGGAAGACCT GAGGGGTACG TCCGGGGTAG GAGTGAAATC CTGTAATCCT
       Nh.2  CGGAAGATCC AAGGGGTACG TCCGGGGTAG GAGTGAAATC CTGTAATCCT
       Nh.5  CGGAAGGCGC GACGGGTACG TCCGGGGTAG GAGTGAAATC CCGTAATCCT
      PIG.5  CGGAAGATCT GAGGGGTACG TCTGGGGTAG GAGTGAAATC CCGTAATCCT
      PIG.8  CGGAAGGCGC GACGGGTACG TCCGGGGTAG GAGTGAAATC CCGTAATCCT
      PW1.1  CGGAAGGCGC GACGGGTACG TCCGGGGTAG GAGTGAAATC CCGTAATCCT
      PW5.4  CGGAAGACCT GAGGGGTACG TCCGGGGTAG GAGTGAAATC CTGTAATCCT
      SC1.2  CGGAAGGCGC GACGGGTACG TCCGGGGTAG GAGTGAAATC CCGTAATCCT
      SR1.5  CGGAAGATCC AGAGGGTACG TCCGGGGTAG GAGTGAAATC CCGTAATCCT
       T1.3  CGGAAGGCTC GAGGGGTACG TCCGGGGTAG GAGTGAAATC CCGTAATCCC
       T1.6  CGGACGGCCC GAGGGGTACG GCCGGGGTAG GAGTGAAATC CTGTAATCCC
       T3.1  CGGGAGATCT GAGGGGTACG TCCGGGGTAG GAGTGAAATC CCGTAATCCT
       T3.2  CGAGAGACTC AACGGGTACG TCCGGGGTAG GAGTGAAATC CTNTAATCCT
       T3.5  CGNNAGACNN AACGGGTACG TCCGGGGTAG GAGTGAAATC CTNTAATCCT
       T3.6  CGGGAGATCT GAGGGGTACG TCCGGGGTAG GAGTGAAATC CCGTAATCCT
       T4.2  CGGAAGACCT GAGGGGTACG TCCGGGGTAG GAGTGAAATC CCGTAATCTC
       T5.7  CGGAAGACCA GAGGGGTACG TCCGGGGTAG GAGTGAAATC CCGTAATCCT
     WB4E.8  CGGAAGGCGC GACGGGTACG TCCGGGGTAG GAGTGAAATC CTGTAATCCT
    WM2C.13  CGGAAGACCT GAGGGGTACG TCCGGGGTAG GAGTGAAATC CTGTAATCCT
 Haloferax  CGGAAGGCTC GAGGGGTACG TCCGGGGTAG GAGTGAAATC CCGTAATCCT
 Halobact.  CGGAAGACCT GAGGGGTACG TCTGGGGTAG GAGTGAAATC CTGTAATCCT
Halorubrum  CGGAAGGCGC GACGGGTACG TCCGGNGTAG GAGTGAAATC CCGTAATCCT
 Natronob.  CGGAAGACCA GAGGGGTACG TCTGGGGTAG GAGTGAAATC CCGTAATCCT
```

```
          651                                                       700
Methanosp. GGTGGGACCA CCTATGGCGA AGGCATCTCA CCAGAACGGG TCCGACAGTG
       777 GGACGGACCA CCGGTGGCGA AAGCGCCTCT GGAAGACGGA TCCGACGGTG
       784 GGACGGACCA CCGGTGGCGA AAGCGCCTCT GGAAGACGGA TCCGACGGTG
       786 GGACGGACCA CCGGTGGCGA AAGCGCCTCT GGAAGACGGA TCCGACGGTG
   B1bra.5 GGACGGACCA CCGGTGGCGA AAGCNCCTCA GGAGGACGGA TCCGACAGTG
   B2S.26 GGACGGACCG CCGATGGGGA AACCACGTCG CGAAGACGGA TCCGACAGTG
   GSL5.48 GGACGGACCG CCGATGGCGA AAGCACGTCG CGAGGACGGA TCCGACAGTG
Haloarcula GGACGGACCA CCAATGGGGA AACCACCTCA GGAAGACGGA CCCGACGGTG
     Nh.2 GGACGGACCA CCGATGGCGA AAGCACCTTG GAAGGACGGA TCCGACAGTG
     Nh.5 GGACGGACCG CCGATGGCGA AAGCACGTCG CGAAGACGGA TCCGACAGTG
   PIG.5 GGACGGACCA CCGGTGGCGA AAGCACCTCA GGAAGACGGA TCCGACAGTG
   PIG.8 GGACGGACCG CCGATGGCGA AAGCACGTCG CGAGGACGGA TCCGACAGTG
   PW1.1 GGACGGACCG CCGATGGCGA AAGCACGTCG CGAGAACGGA TCCGACAGTG
   PW5.4 GGACGGACCA CCGGTGGCGA AAGCGCCTCA GGAGGACGGA TCCGACAGTG
   SC1.2 GGACGGACCG CCGATGGCGA AAGCACGTCG CGAGAACGGA TCCGACAGTG
   SR1.5 GGACGGACCA CCGGTGGCGA AAGCGCTCTG GAAAGACGGA TCCGACGGTG
     T1.3 GGACGGACCA CCGATGGCGA AAGCACCTCG AGAAAACGGA TCCGACAGTG
     T1.6 GGCCGGACCA CCGGTGGCGA AAGCGCCTCG GGAGGACGGA TCCGACGGTG
     T3.1 GGACGGACCA CCGGTGGCGA AAGCGCCTCA GAAAAACGGA CCCGACGGTG
     T3.2 GGACGGACCA CCAATGGGGA AACCACNTTG AGANACCGGA CCCGACAGTG
     T3.5 GGACGGACCA CCAATGGGGA AACCACCTNN NGANNNCGGA CCCGACAGTG
     T3.6 GGACGGACCA CCGGTGGCGA AAGCGCCTCA GAAAAACGGA CCCGACGGTG
     T4.2 GGACGGACCG CCGGTGGCGA AAGCGCCTCA GGAAGACGGA CCCGACGGTG
     T5.7 GGACGGACCA CCGGTGGCGA AAGCGCCTCT GGAAGACGGA TCCGACGGTG
   WB4E.8 GGACGGACCG CCGATGGCGA AAGCACGTCG CGAGGACGGA TCCGACAGTG
  WM2C.13 GGACGGACCA CCGGTGGCGA AAGCGCCTCA GGAGGACGGA TCCGACAGTG
Haloferax GGACGGACCA CCGATGGCGA AAGCACCTCG AGAAGACGGA TCCGACGGTG
Halobact. GGACGGACCG CCGGTGGCGA AAGCGCCTCA GGAGAACGGA TCCGACAGTG
Halorubrum GGACGGACCG CCGATGGCGA AAGCACGTCG CGAGGACGGA TCCGACAGTN
Natronob. GGACGGACCA CCGGTGGCGA AAGCGCCTCT GGAAGACGGA TCCGACGGTG

          701                                                       750
Methanosp. AGGGACGAAA GCTGGGGGAG CAAACCGGAT TAGATACCCG GGTAGTCCCA
       777 AGGGACGAAA GCTCGGGTCA CGAACCGGAT TAGATACCCG GGTAGTCCGA
       784 AGGGACGAAA GCTCGGGTCA CGAACCGGAT TAGATACCCG GGTAGTCCGA
       786 AGGGACGAAA GCTCGGGTCA CGAACCGGAT TAGATACCCG GGTAGTCCGA
   B1bra.5 AGGGACGAAA GCTAGGGTCT CGAACCGGAT TAGATACCCG GGTAGTCCTA
   B2S.26 AGGGNCGAAA GCTAGGGTCT CGAACCGGAT TAGATACCCG GGTAGTCCTA
   GSL5.48 AGGGACGAAA GCCAGGGTCT CGAACCGGAT TAGATACCCG GGTAGTCCTG
Haloarcula AGGGACGAAA GCTAGGGTCT CGAACCGGAT TAGATACCCG GGTAGTCCTA
     Nh.2 AGGGACGAAA GCTAGGGTCT CGAACCGGAT TAGATACCCG GGTAGTCCTA
     Nh.5 AGGGACGAAA GCTAGGGTCT CGAACCGGAT TAGATACCCG GGTAGTCCTA
   PIG.5 AGGGACGAAA GCTGGGGTCA CGAACCGGAT TAGATACCCG GGTAGTCCCA
   PIG.8 AGGGACGAAA GCCAGGGTCT CGAACCGGAT TAGATACCCG GGTAGTCCTG
   PW1.1 AGGGACGAAA GCTAGGGTCT CGAACCGGAT TAGATACCCG GGTAGTCCTA
   PW5.4 AGGGACGAAA GCTAGGGTCT CGAACCGGAT TAGATACCCG GGTAGTCCTA
   SC1.2 AGGGACGAAA GCCAGGGTCT CGAACCGGAT TAGATACCCG GGTAGTCCTG
   SR1.5 AGGGACGAAA GCTCGGGTCA CGAACCGGAT TAGATACCCG GGTAGTCCGA
     T1.3 AGGGACGAAA GCTAGGGTCT CGAACCGGAT TAGATACCCG GGTAGTCCTA
     T1.6 AGGGACGAAA GCCAGGGTCA CGAACCGGAT TAGATACCCG GGTAGTCCTG
     T3.1 AGGGACGAAA GCTAGGGTCT CGAACCGGAT TAGATACCCG GGTAGTCCTA
     T3.2 AGGGACGAAA GCNAGGGTCT CGAACCGGAT TAGATACCCG GGTAGTCCTN
     T3.5 AGGGACGAAA GCNAGGGTCT CGAACCGGAT TAGATACCCG GGTAGTCCTN
     T3.6 AGGGACGAAA GCTAGGGTCT CGAACCGGAT TAGATACCCG GGTAGTCCTA
     T4.2 AGGGACGAAA GCCAGGGTCT CGAACCGGAT TAGATACCCG GGTAGTCCTG
     T5.7 AGGGACGAAA GCTCGGGTCA CGAACCGGAT TAGATACCCG GGTAGTCCGA
   WB4E.8 AGGGACGAAA GCCAGGGTCT CGAACCGGAT TAGATACCCG GGTAGTCCTG
  WM2C.13 AGGGACGAAA GCTAGGGTCT CGAACCGGAT TAGATACCCG GGTAGTCCTA
Haloferax AGGGACGAAA GCTAGGGTCT CGAACCGGAT TAGATACCCG GGTAGTCCTA
Halobact. AGGGACGAAA GCTAGGGTCT CGAACCGGAT TAGATACCCG GGTAGTCCTA
Halorubrum AGGGACGAAA GCCAGGGTCT CGAACCGGAT TAGATACCCG GGTAGTCCTG
Natronob. AGGGACGAAA GCTCGGGTCA CGAACCGGAT TAGATACCCG GGTAGTCCGA
```

```
            751                                                       800
Methanosp.  GCTGTAAACG ATGCGCGTTA GGTGTGTCAG TGACCACGTG TCACTGAGGT
       777  GCTGTAAACG ATGTCTGCTA GGTGTGGCAC AGGCTACGAG CCTGTGCTGT
       784  GCTGTAAACG ATGTCTGCTA GGTGTGGCAC AGGCTACGAG CCTGTGCTGT
       786  GCTGTAAACG ATGTCTGCTA GGTGTGGCAC AGGCTACGAG CCTGTGCTGT
   B1bra.5  GCTGTAAACG ATGCCCGCTA GGTGTGGCGC AGGCTACGAG CCTGCGCTGT
   B2S.26  GCCGTAAACG ATGCCTGTTA GGTGTGNCCC CCGCTACGAG CGGGTGCTGT
   GSL5.48  GCCGTAAACA ATGTCTGCTA GGTGTGGCTC CCACTACGAG TGGGTGCTGT
Haloarcula  GCTGTAAACG ATGCTCGCTA GGTGTGCCGT AGGCTACGAG CCTGCGCTGC
     Nh.2  GCCGTAAACG ATGCTCGCTA GGTGTGGCGC AGGCTACGAG CCTGCGCTGT
     Nh.5  GCCGTAAACA ATGCCTGCTA GGTGTGGCTC CCACTACGAG TGGGTGCTGT
   PIG.5  GCTGTAAACG ATGCTCACTA GGTGTGGCAC TGGCTACGAG CNAGTGCTGT
   PIG.8  GCCGTAAACA ATGTCTGTTA GGTGTGGCTC CCACTACGAG TGGGTGCTGT
   PW1.1  GCCGTAAACA ATGCCTGCTA GGTGTGGCTC CCACTACGAG TGGGTGCTGT
   PW5.4  GCTGTAAACG ATGCCCGCTA GGTGTGGCGC AGGCTACGAG CCTGCGCTGT
   SC1.2  GCCGTAAACA ATGTCTGCTA GGTGTGGCTC CCACTACGAG TGGGTGCTGT
   SR1.5  GCTGTAAACG ATGTCTGCTA GGTGTGCCAC AGGCTACGAG CCTGTGGTGT
   T1.3  GCCGTAAACG ATGTTCGCTA GGTGTGACAC AGGCTACGCG CCTGTGTTGT
   T1.6  GCCGTAAACG ATGNCCGTTA GGTGTGGCGC AGGCTACGAG CCTGCGCTGT
   T3.1  GCCGTAAACG ATGCTCGCTA GGTGTGGCGC AGGCTACGAG CCTGCGCTGT
   T3.2  GCTGTAAACA ATGCTCGCTA GGTATGTCAC NNGCNATGAG CNNGTGNTGT
   T3.5  GCTGTAAACN ATGCTCGCTA GGTNTGNCNN NNGCNANGAG CNNGNGNTGT
   T3.6  GCCGTAAACG ATGCTCGCTA GGTGTGGCGC AGGCTACGAG CCTGCGCTGT
   T4.2  GCCGTAAACG ATGCTCGCTA GGTGTGCCGC AGGCTACGAG CCTGCGCTGT
   T5.7  GCTGTAAACG ATGTCTGCTA GGTGTGGCAC AGGCTACGAG CCTGTGCTGT
   WB4E.8  GCCGTAAACA ATGCCTGCTA GGTGTGGCTC CCGCTACGAG CGGGTGCTGT
   WM2C.13  GCTGTAAACG ATGCCCGCTA GGTGTGGCGC AGGCTACGAG CCTGCGCTGT
Haloferax  GCTGTAAACG ATGCTCGCTA GGTGTGACAC AGGCTACGAG CCTGTGTTGT
Halobact.  GCTGTAAACG ATGTCCGCTA GGTGTGGCGC AGGCTACGAG CCTGCGCTGT
Halorubrum  GCCGTAAACA ATGTCTGTTA GGTGTGGCTC CCACTACGAG TGGGTGCTGT
Natronob.  GCTGTAAACG ATGTCTGCTA GGTGTGACAC AGGCTACGAG CCTGTGTTGT

            801                                                       850
Methanosp.  GCCGAAGGGA AACCGTGAAA CGCGCCGCCT GGGGAGTACG GTCGCAAGGC
       777  GCCGTAGGGA AGCCGTGAAG CAGACCGCCT GGGAAGTACG TCCGCAAGGA
       784  GCCGTAGGGA AGCCGTGAAG CAGACCGCCT GGGAAGTACG TCCGCAAGGA
       786  GCCGCAGGGA AGCCGTGAAG CAGACCGCCT GGGAAGTACG TCCGCAAGGA
   B1bra.5  GCCGTAGGGA AGCCGAGAAG CGGNCCGCCT GGGAAGTACG TCTGCAAGGA
   B2S.26  GCCGTAGGGA AGCCGCTAAA CAGNCCGCCT GGGAAGTACG TCTNCAAGGA
   GSL5.48  GCCGTAGGGA AGCCGCTAAG CAGACCGCCT GGGAAGTACG TCCGCAAGGA
Haloarcula  GCCCTAGGGA AGCCGAGAAG CGAGCCGCCT GGGAAGTACG TCTGCAAGGA
     Nh.2  GCCGTAGGGA AGCCGTGAAG CGANCCGCCT GGGAAGTACG TCTGCAAGGA
     Nh.5  GCCGTAGGGA AGCCGCTAAG CAGGCCGCCT GGGAAGTACG TCCGCAAGGA
   PIG.5  GCCGTAGTGA AGCCGTGAAG TGAGCCGCCT GGGAAGTACG TCCGCAAGGA
   PIG.8  GCCGTAGGGA AGCCGCTAAA CAGACCGCCT GGGAAGTACG TCCGCAAGGA
   PW1.1  GCCGTAGGGA AGCCGCTAAG CAGGCCGCCT GGGAAGTACG TCCGCAAGGA
   PW5.4  GCCGTAGGGA AGCCGAGAAG CGGNCCGCCT GGGAAGTACG TCTGCAAGGA
   SC1.2  GCCGTAGGGA AGCCGCTAAG CAGACCGCCT GGGAAGTACG TCCGCAAGGA
   SR1.5  GCCGTAGGGA AGCCGTGAAG CAGACCGCCT GGGAAGTACG TCCGCAAGGA
   T1.3  GCCGTAGGGA AGCCGAGAAG CGAACCGCCT GGGAAGTACG TCTGCAAGGA
   T1.6  GCCGCAGTGA AGACGCTAAA CGGGCCGCCT GGGAAGTACG TCCCCAAGGA
   T3.1  GCCGTAGGGA AGCCGTGAAG CGAGCCGCCT GGGAAGTACG TCTNCAAGGA
   T3.2  GCCGTAGTGA AGNCGANAAG CGAGCCGCCT GGGAAGTACG TCNNCAAGGA
   T3.5  GCCNTAGTGA AGNCGANAAG CGAGCCGCCT GGGAAGTACG TCNGCAAGGA
   T3.6  GCCGTAGGGA AGCCGTGAAG CGAGCCGCCT GGGAAGTACG TCTGCAAGGA
   T4.2  GCCGTAGGGA AGCCGCGAAG CGAGCCGCCT GGGAAGTACG TCTGCAAGGA
   T5.7  GCCGTAGGGA AGCCGTGAAG CAGACCGCCT GGGAAGTACG TCCGCAAGGA
   WB4E.8  GCCGTAGGGA AGCCGCTAAG CAGGCCGCCT GGGAAGTACG TCCGCAAGGA
   WM2C.13  GCCGTAGGGA AGCCGAGAAG CGGNCCGCCT GGGAAGTACG TCTGCAAGGA
Haloferax  GCCGTAGGGA AGCCGAGAAG CGAGCCGCCT GGGAAGTACG TCCGCAAGGA
Halobact.  GCCGTAGGGA AGCCGAGAAG CGGACCGCCT GGGAAGTACG TCTGCAAGGA
Halorubrum  GCCGTAGGGA AGCCGCTAAA CAGACNGCCT GGGAAGTACG TCCGCAAGGA
Natronob.  GCCGTAGGGA AGCCGTGAAG CAGACCGCCT GGGAAGTACG TCCGCAAGGA
```

```
            851                                                       900
Methanosp.  TGAAACTTAA AGGAATTGGC GGGGGAGCAC CACAACGGGT GGAGCCTGCG
       777  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
       784  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
       786  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
   B1bra.5  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
   B2S.26   TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
   GSL5.48  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
Haloarcula  TGAAACTTAA AGGAATTGGC GGGGGAGCAC CACAACCGGA GGAGCCTGCG
      Nh.2  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
      Nh.5  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
     PIG.5  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
     PIG.8  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
     PW1.1  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
     PW5.4  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
     SC1.2  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
     SR1.5  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
      T1.3  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
      T1.6  TGAAACTTAT AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
      T3.1  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
      T3.2  TGAAACTTAA AGGAATTGGC GGGGGAGCAC CACAACCGGA GGAGCCTGCG
      T3.5  TGAAACTTAA AGGAATTGGC GGGGGAGCAC CACAACCGGA GGAGCCTGCG
      T3.6  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
      T4.2  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
      T5.7  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
   WB4E.8  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
  WM2C.13  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
Haloferax  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
Halobact.  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG
Halorubrum  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCNTGCG
Natronob.  TGAAACTTAA AGGAATTGGC GGGGGAGCAC TACAACCGGA GGAGCCTGCG


            901                                                       950
Methanosp.  GTTTAATCGG ACTCAACGCC GGAAATCTCA CCGGATAAGA CAGCT.G.AA
       777  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCATCGA CAATGTGCAG
       784  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCATCGA CAATGTGCAG
       786  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCATCGA CAATGTGCAG
   B1bra.5  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCTCCGA CAGTA.GTAA
   B2S.26   GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCATCGA CTGTA.GTAA
   GSL5.48  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCATCGA CTGTA.ATAA
Haloarcula  GTTTAATTGG ACTCAACGCC GGACATCTCA CCGGTCCCGA CAGTA.GTAA
      Nh.2  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCTCCGA CAGCA.GTAA
      Nh.5  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCATCGA CTGTA.ATAA
     PIG.5  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCTCCGA CAGTATGCAG
     PIG.8  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCATCGA CTGTA.ATAA
     PW1.1  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCATCGA CTGTA.ATAA
     PW5.4  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCTCCGA CAGTA.GTAA
     SC1.2  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCATCGA CTGTG.ATAA
     SR1.5  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCATCGA CAATGTGCAG
      T1.3  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCTCCGA CTACA.GTAA
      T1.6  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCATCGA CAGTGTGCAG
      T3.1  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCACCGA CAGTGTGCAG
      T3.2  GTTTAATTGG ACTCAACGCC GGACATCTCA CCGGTCCCGA CAGTA.GTAA
      T3.5  GTTTAATTGG ACTCAACGCC GGAAATCTCA CCGGTCCCGA CAGTA.GTAA
      T3.6  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCACCGA CAGTGTGCAG
      T4.2  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCACCGA CAGTA.GCAG
      T5.7  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCATCGA CAATGTGCAG
   WB4E.8  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCATCGA CTGTA.GTAA
  WM2C.13  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCTCCGA CAGTA.GTAA
Haloferax  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCTCCGA CTACA.GTGA
Halobact.  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCCCCGA CAGTA.GTAA
Halorubrum  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCATCGA CTGTA.GTAA
Natronob.  GTTTAATTGG ACTCAACGCC GGACATCTCA CCAGCACCGA CAATA.GTAA
```

```
           951                                                  1000
Methanosp. TGATAGTCGG GATGAAGACT CTACT.TGA. CTAGCTGAGA GGAGGTGCAT
       777 TGAAAGTCAG GTTGATGACC TTACT.GGA. GCCATTGAGA GGAGGTGCAT
       784 TGAAAGTCAG GTTGATGACC TTACT.GGA. GCCATTGAGA GGAGGTGCAT
       786 TGAAAGTCAG GTTGATGACC TTACT.GGA. GCCATTGAGA GGAGGTGCAT
   B1bra.5 TGACGGTCAG GTTGATGACC TTACC.CGAT GCTACTGAGA GGAGGTGCAT
   B2S.26 TGACGGTCAG GTTGATGACC TCNCCACGA. GCTACAGAGA GGAGGTGCAT
   GSL5.48 TGACGACCAG GTTGATGACC TTGTC.CGA. GTTTCAGAGA GGAGGTGCAT
Haloarcula TGACAGTCAG GTTGACGACT TTACT.CGAC GCTACTGAGA GGAGGTGCAT
       Nh.2 TGACAGTCAG GTTGATGACC TTACT.CGAC GCTGCTGAGA GGAGGTGCAT
       Nh.5 TGACGACCAG GTTGATGACC TTATC.CGA. GTTTCAGAGA GGAGGTGCAT
     PIG.5 TGACAGTCAG TGTGATGAGC TTACT.TGA. GCTACTGAGA GGAGGTGCAT
     PIG.8 TGACGATCAG GTTGATGACC TTATC.CGA. GTTTCAGAGA GGAGGTGCAT
     PW1.1 TGACGATCAG GTTGATGACC TTATC.CGA. GTTTCAGAGA GGAGGTGCAT
     PW5.4 TGACGGTCAG GTTGATGACC TTACC.CGAT GCTACTGAGA GGAGGTGCAT
     SC1.2 TGACGATCAG GTTGATGACC TTATC.CGA. GCCTCAGAGA GGAGGTGCAT
     SR1.5 TGAANGTCAG GTTGATGACC TTACT.GGA. GCCATTGAGA GGAGGTGCAT
      T1.3 TGACGGTCAG GTTGATGACC TCGCCACGAC GCTGTAGAGA GGAGGTGCAT
      T1.6 TGACAGTCAG TGTGATGAGC TTACT.AGA. GCCACTGAGA GGAGGTGCAT
      T3.1 TGAAGGTCAG GTTGATGACC TTACT.GGA. GCCACTGAGA GGAGGTGCAT
      T3.2 TGACAGTCAG GTTGACGACT TTACT.CGAC GCTACTGAGA GGAGGTGCAT
      T3.5 TGACAGTCAG GTTGACGACT TTACT.CGAC GCTACTGAGA GGAGGTGCAT
      T3.6 TGAAGGTCAG GTTGATGACC TTACT.GGA. GCCACTGAGA GGAGGTGCAT
      T4.2 TGACGGTCAG TCTGATGAGC TTACC.TGA. GCTACTGAGA GGAGGTGCAT
      T5.7 TGAACGTCAG GTTGATGACC TTACT.GGA. GCCATTGAGA GGAGGTGCAT
   WB4E.8 TGACGACCAG GTTGATGACC TTGTC.CGA. GCTTCAGAGA GGAGGTGCAT
  WM2C.13 TGACGGTCAG GTTGATGACC TTACC.CGAT GCTACTGAGA GGAGGTGCAT
Haloferax TGACGATCAG GTTGATGACC TTATCACGAC GCTGTAGAGA GGAGGTGCAT
Halobact. TGACGGTCAG GTTGATGACC TTACC.CGAG GCTACTGAGA GGAGGTGCAT
Halorubrum TGACGATCAG GTTGATGACC TTATC.CGA. GCTTCAGAGA GGAGGTGCAT
Natronob. TGACGGTCAG TGTGATGAGC TTACCTGGAG GCTATTGAGA GGAGGTGCAT

           1001                                                 1050
Methanosp. GGCCGTCGTC AGTTCGTACT GTGAAGCATC CTGTTTAGTC AGGCAACGAG
       777 GNCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGACANCGAG
       784 GNCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGACAACGAG
       786 GTCCNCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGNCAACGAG
   B1bra.5 GNCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGGCAACGAG
   B2S.26 GNCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGGCAACGAG
   GSL5.48 GNCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGNCAACGAG
Haloarcula GGCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGGCAACGAG
       Nh.2 GTCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGGCAACGAG
       Nh.5 GTCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGGCAACGAG
     PIG.5 GNCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGNCAACGAG
     PIG.8 GGCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGGCAACGAG
     PW1.1 GNCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGGCAACGAG
     PW5.4 GGCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGGCAACGAG
     SC1.2 GNCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGGCANCGAG
     SR1.5 GTCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGGCAACGAG
      T1.3 GTCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGGCAACGAG
      T1.6 GNCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGGCAACGAG
      T3.1 GTCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGNCANCGAG
      T3.2 GNCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGGCANCGAG
      T3.5 GTCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGGCANCGAG
      T3.6 GTCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGGCANCGAG
      T4.2 GNCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGGCANCGAG
      T5.7 GTCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGACANCGAG
   WB4E.8 GNCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGNCAACGAG
  WM2C.13 GNCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGGCAACGAG
Haloferax GGCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGGCAACGAG
Halobact. GGCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGGCAACGAG
Halorubrum GGCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGGCAACGAG
Natronob. GGCCGCCGTC AGCTCGTACC GTGAGGCGTC CTGTTAAGTC AGGCAACGAG
```

```
          1051                                                      1100
Methanosp. CGAGACCCAC GCGAGCAGTT GCCAGCTTGA CCTTCG.GGT TGATGGGGAC
       777 CGAGACCCGC ACTCCTAATT GNCAGCAACA CCCTCGTGGT GGNTGGGTAC
       784 CGAGACCCGC ACTCCTAATT GNCAGCAACA CCCTCGTGGT GGTTGGGTAC
       786 CGAGACCCGC ACTCCTAATT GNCAGCAACA CC.TAGCGGT GGNTGGGTAC
    B1bra.5 CGAGACCCGC ACTCCTAATT GCCAGCAGTA CCCTTTGGGT AGCTGGGTAC
    B2S.26 CGAGACCCGC ATCCGTACTT GCCAGCAGCA CTC.AT.CGT GNCTGGGGAC
    GSL5.48 CGAGACCCGC ACCCTTACTT GCCAGCAGTA .TCGCGAGAT AGCTGGGGAC
Haloarcula CGAGACCCGC ACTTCTAGTT GCCAGCAATA CCCTTGAGGT AGTTGGGTAC
      Nh.2 CGAGACCCGC ACCTCTAGNT GCCAGCAGCA .TCTCNCGAT GNCTGGGNAC
      Nh.5 CGAGACCCGC ATCCTTACTT GCCAGCAGCA CTN.CGAAGT GNCTGGGGAC
    PIG.5 CGAGACCCGC ACTCTTAATT GCCAGCATAG NCCTTTGGCC TGATGGGTAC
    PIG.8 CGAGACCCGC ATCCTTACTT GCCAGCAGCA CTC.CGAAGT GNCTGGGGAC
    PW1.1 CGAGACCCGC ATCCTTACTT GCCAGCAGCA CTC.CGAAGT GNCTGGGGAC
    PW5.4 CGAGACCCGC ACTCCTAATT GCCAGCAGTA CCCTTTGGGT AGCTGGGTAC
    SC1.2 CGAGACCCGC ACCCTTACTT NCCANCAGTA .CCCCGAGGT ANCTGGGGAC
    SR1.5 CGAGACCCGC ACTCCTAATT GCCAGCAANA CCCTTGTGGT NGTTGGGTAC
     T1.3 CGAGACCCGC ACTCCTAATT GCCAGCAGCA GTTCCG.ACT GGCTGGGTAC
     T1.6 CGAGACCCAC ATCCCTAATT GCCAGCACTG G.CGCGAGCC AGGTGGGTAC
     T3.1 CGAGACCCGC ATCCCTAGNT GNCAGCAGCA ACCTTGTGTT GNCTGGGNAC
     T3.2 CGAGACCCAC ACTTCTAGTT GNCAGCAANA CCCCTGNGGT NNTTGGGTAC
     T3.5 CGAGACCCAC ACTTCTAGTT GNCAGCAANA CCCCTGNGGT NNTTGGGTAC
     T3.6 CGAGACCCGC ATCCCTAGNT GNCAGCAGCA ACCTTGTGTT GNCTGGGNAC
     T4.2 CGAGACCCGC ACTCCTAGTT GNCAGCAGCA TCTCN.CGAT GNCTGGGTAC
     T5.7 CGAGACCCNC ACTCCTAATT GNCAGCAACA CCCTTGNGGT GGNTGGGNAC
   WB4E.8 CGAGACCCGC ATCCTTACTT GCCAGCAGCA CTC.CGAAGT GNCTGGGGAC
   WM2C.13 CGAGACCCGC ACTCCTAATT GCCAGCAGTA CCCTTTGGGT AGCTGGGTAC
Haloferax CGAGACCCGC ACTTCTAATT GCCAGCAGCA GTTTCG.ACT GGCTGGGTAC
Halobact. CGAGACCCGC ACTCCTAATT GCCAGCGGTA CCCTTTGGGT AGCTGGGTAC
      sacch CGAGACCCGC ATCCTTACTT GCCAGCAGCA CT.GCGAAGT GGCTGGGGAC
Natronob. CGAGACCCGC TCTCCTAATT GCCAGCAACG ACCTTT.GGT GGTTGGGTAC

          1101                                                      1150
Methanosp. ACTGCTCGGA CCGCCTCTGC TAAAGGGGAG GAAGGAATGG GCAACGGTAG
       777 ATTAGGAGGA CTCCCAGTGN CACACTGGAG GAAGGAACGG GCACCGGTAG
       784 ATTAGGAGGA CTNCCAGTGC CANACTGGAG GAAGGAACGG GCANCGGTAG
       786 ATTAGGAGGA CTCCCAGTGN CACACTGGAG GAAGGANCGG GCACCGGTAG
    B1bra.5 ATTAGGTGGA CTGCCGCTNC CAAAGCGGAG GAAGGAACGG GCAACGGTAG
    B2S.26 AGTACGGAGA CCNCCGTGNC TAACACGGAG GAAGGAACGG GCAACGGTAG
    GSL5.48 AGTAGGGGGA CCNCCGTGNC TAACACGGAG GAAGGAACGG GCAACGGTAG
Haloarcula ACTAGGAGGA CTGCCGCTGC TAAAGCGGAG GAAGGAACGG GCAACGGTAG
      Nh.2 ACTAGGGGGA CTCCCGTCGC CAAGACGGAG GAAGGANCGG GCACCGGTAG
      Nh.5 AGTAGGGAGA CCNCCGTGNC CAACACGGAG GAAGGANCGG NCANCGGTAG
    PIG.5 ATTAGGAGGA CTGCCAGTGC TAAACTGGAG GAAGGAACGG GCAACGGTAG
    PIG.8 AGTAGGGAGA CCNCCGTGNC CAACACGGAG GAAGGAACGG NCANCGGTAG
    PW1.1 AGTAGGGAGA CCNCCGTGTC CAACACGGAG GAAGGAACGG NCANCGGTAG
    PW5.4 ATTAGGTGGA CTNCCGCTGC TAAAGCGGAG GAAGGAACGG NCANCGGTAG
    SC1.2 AGTAGGGGGA CCNCCGTGTC TAACACGGAG GAAGGAACGG NCANCGGTAG
    SR1.5 ATTAGGAGGA CTNCCAGTGC CANACTGGAG GAAGGAACGG GCANCGGTAG
     T1.3 ATTAGGAGGA CTGCCAGTGC CANACTGGAG GAAGGAACGG GCANCGGTAG
     T1.6 ATTAGGGAGA CTNCCGTGNC CAACACGGAG GAAGGAATGG GCANCGGTAG
     T3.1 ACTAGGGAGA CTNCCNCTCN CANANNGGAG GAAGGANCGG GCACCGGTAG
     T3.2 ACTAGGAGGA CTNCCANTGC TAAAATGGAG GAAGGAATGG GCANCGGTAG
     T3.5 ACTAGGAGGA CTNCCANTNN TAAANTGGAG GAAGGAANGG GCANCGGTAG
     T3.6 ACTAGGGAGA CTCCCNCTCN CANANNGGAG GAAGGANCGG GCANCGGTAG
     T4.2 ACTAGGAGGA CTNCCNCTNN CANANNGGAG GAAGGANCGG GCACCGGTAG
     T5.7 ATTAGGAGGA CTNNCAGTGN CACACTGGAG GAAGGANCGG GCANCGGTAG
   WB4E.8 AGTAGGGAGA CCNCCGTGNC CAACACGGAG GAAGGAACGG NCANCGGTAG
   WM2C.13 ATTAGGTGGA CTNCCGCTGC TAAAGCGGAG GAAGGAACGG NCANCGGTAG
Haloferax ATTAGAAGGA CTGCCGCTGC TAAAGCGGAG GAAGGAACGG GCAACGGTAG
Halobact. ATTAGGTGGA CTGCCGCTGC CAAAGCGGAG GAAGGAACGG GCAACGGTAG
Halorubrum AGTAGGGAGA CCGCCGTGGC CAACACGGAG GAAGGAACGG GCAACGGTAG
Natronob. ATTAGGAGGA CTGCCGCTGC CAAAGCGGAG GAAGGAACGG GCAACGGTAG
```

```
            1151                                                    1200
Methanosp.  GTCAGCATGC CCCGAATTAT CCGGGCTACA CGCGGGCTAC AATGGACAGG
       777  GTCAGTATGC CCCGAATGTG CTGGGCGACA CGCGGGCTAC AATGNCCGAG
       784  GTCAGTATNC CCCGAATGTG CTGGGCGACA CGCGGGCTAC AATGNCCGAG
       786  NTCAGTATNC CCCGAATGTG CTGGGCGACA CGCGGGCTAC AATGNCCGAG
   B1bra.5  GTCAGTATNC CCCGAATGAG CTGGGCAACA CGCGGGCTAC AATGGTCGAG
   B2S.26   GTCAGTATNC CCCGAATGTG CTGGGCAACA CGCGGGCNAC AATGGTCAGG
   GSL5.48  GTCAGTATGC CCCGAATGTG CTGGGCAACA CGCGGGCTAC AATGGTCGAG
Haloarcula GTCAGTATGC CCCGAATGGA CCGGGCAACA CGCGGGCTAC AATGGCTATG
       Nh.2 GTCAGTATNC CCCGAATGAG CTGGGCGACA CGCGGGCTAC AATGNTCGAG
       Nh.5 GTCAGTATNC CCCGAATGTG CTGGGCAACA CGCGGGCTAC AATGGTCGAG
     PIG.5  GTCAGTATGC CCCGAATGAG CTGGGCTACA CGCGGGCTAC AATGGCCGAG
     PIG.8  GTCAGTATGC CCCGAATGTG CTGGGCAACA CGCGGGCTAC AATGGTCGAG
     PW1.1  GTCAGTATNC CCCGAATGTG CTGGGCAACA CGCGGGCTAC AATGGTCGAG
     PW5.4  GTCAGTATGC CCCGAATGAG CTGGGCAACA CGCGGGCTAC AATGGTCGAG
     SC1.2  GTCAGTATNC CCCGAATGTG CTGGGCAACA CGCGGGCTAC AATGGTCGAG
     SR1.5  GTCAGTATGC CCCGAATGTG CTGGGCGACA CGCGGGCTAC AATGGCCGAG
      T1.3  GTCAGTATNC CCCGAATGAG CTGGGCTACA CGCGGGCTAC AATGGCCAGG
      T1.6  GTCAGTATNC CCCGAATGTG CTGGGCGACA CGCGGGCTAC AATGGCCGAG
      T3.1  GTCAGTATAC CCCGAATGTG CTGGGCAACA CGCGGGCTAC AATGTCCGGG
      T3.2  GTCAGTATNC CCCGAATGGA CCGGGCAACA CGCGGGCTAC AATGGCTNTG
      T3.5  GTCAGTATNC CCCGAATGGA CCGGGCAACA CNCGGGCTAC AATGGCTNTG
      T3.6  GTCAGTATAC CCCGAATGTG CTGGGCAACA CNCGGGCTAC AATGTCCGGG
      T4.2  GTCAGTATNC CCCGAATGTG CTGGGCGACA CGCGGGCTAC AATGNCCGGG
      T5.7  NTCAGNATNC CCCGAATGTG CTGGGCGACA CGCGGGCTNC AATGNCCGAG
   WB4E.8   GTCAGTATNC CCCGAATGTG CTGGGCAACA CGCGGGCTAC AATGGTCGAG
   WM2C.13  GTCAGTATNC CCCGAATGAG CTGGGCAACA CGCGGGCTAC AATGGTCGAG
Haloferax  GTCAGTATGC CCCGAATGAG CTGGGCTACA CGCGGGCTAC AATGGTCGAG
Halobact.  GTCAGTATGC CCCGAATGGG CTGGGCAACA CGCGGGCTAC AATGGTCGAG
Halorubrum GTCAGTATGC CCCGAATGTG CTGGGCAACA CGCGGGCTAC AATGGTCGAG
Natronob.  GTCAGTATGC CCCGAATGTG CTGGGCAACA CGCGGGCTAC AATGGCCACG

            1201                                                    1250
Methanosp.  ACAATGGGTT TCGACACCGA GAGGTGAGGA TAATCTCCTA AACCTGTCCG
       777  ACAGTGGGAT GCAACCCCGA AAGGGGACGC TAATCTCCGA AACTCGGTCG
       784  ACAGTGGGAT GCAACCCCGA AAGGGGACGC TAATCTCCGA AACTCGGTCG
       786  ACAGTGGGAT GCAACCCCGA AAGGGGGCGC TAATCTCCGA AACTCGGTCG
   B1bra.5  ACAATGGGAT GCAACTCCGA GAGGAGGCGC TAATCTCCTA AACTCGATCG
   B2S.26   ACAATGGGTT CCAACTCCGA AAGGAGACGG TAATCTCACA AACCTGATCG
   GSL5.48  ACAAAGGGTT CCTACTCCGA AAGGAGACGG TAATCTCAGA AACTCGATCG
Haloarcula ACAGTGGGAT GCAACGCCGA GAGGCGACGC TAATCTCC.A AACGTAGTCG
       Nh.2 ACAGTGGGAT GCGACACCGA GAGGTGAAGC TAATCTCCTA AACTCGATCG
       Nh.5 ACANAGGGTT CCTACTCCGA AAGGAGACGG TAATCTCAGA AACTCGATCG
     PIG.5  ACAGTGGGAT GCAACCCCGA GAGGGGGCGC TAATCTCCGA AACTCGNTCG
     PIG.8  ACAAAGGGTT CCAACTCCGA AAGGAGACGG TAATCTCAGA AACTCGATCG
     PW1.1  ACAAAGGGTT CCTACTCCGA AAGGAGACGG TAATCTCAGA AACTCGATCG
     PW5.4  ACAATGGGAT GCAACTCCGA GAGGAGGCGC TAATCTCCTA AACTCGATCG
     SC1.2  ACAAAGGGTT CCTACTCCGA AAGGAGACGG TAATCTCAGA AACTCGATCG
     SR1.5  ACAGTGGGAT GCAACCCCGA AAGGGGACGC TAATCTCCGA AACTCGGTCG
      T1.3  ACAATGGGTT GCAACCTCGA AAGAGGGCGC TAATCTCCGA AACCTGGTCG
      T1.6  ACAGTGGGAT GCAACCCCGA AAGGGGGCGC TAATCTCCGA AACTCGGTCG
      T3.1  ACAGTGGGAC GCGACCCCGA GAGGGAGAGC TAATCTCCTA AACCCGGTCG
      T3.2  ACAGTGGGAT GCAACGCCGA GAGGCGAAGC TAATCTCC.A AACGNAGTCG
      T3.5  ACAGTGGGAT GCAACGCCGA GAGGCGAAGC TAATCTCC.A AACGNAGTCG
      T3.6  ACAGTGGGAC GCGACCCCGA GAGGGAGAGC TAATCTCCTA AACCCGGTCG
      T4.2  ACAGTGGGAC GCCAGTCCGA GAGGACGCGC TAATCCCCGA AACCCGGTCG
      T5.7  ACAGTGGGAT GCAACCCCGA AAGGGGACGC TAATCTCCGA AACTCGGTCG
   WB4E.8   ACAAAGGGTT CCAACTCCGA AAGGAGACGG TAATCTCAGA AACTCGATCG
   WM2C.13  ACAATGGGAT GCAACTCCGA GAGGAGGCGC TAATCTCCTA AACTCGATCG
Haloferax  ACAATGGGTT GCTATCTCGA AAGAGAACGC TAATCTCCTA AACTCGATCG
Halobact.  ACAATGGGAA GCCACTCCGA GAGGAGGCGC TAATCTCCTA AACTCGATCG
Halorubrum ACAAAGGGTT CCAACTCCGA AAGGAGACGG TAATCTCAGA AACTCGATCG
Natronob.  ACAGTGGGAT GCAACGCCGA AAGGCGACGC TAATCTCCTA AACGTGGTCG
```

```
            1251                                                1300
Methanosp.  AAGTTCGGAT TGCGGGTTGT AACTCACCCG CATGAAGCTG GAATCCGTAG
       777  TAGTTCGGAT TGAGGGCTGA AACTCGCCCT CATGAAGCTG GATTCGGTAG
       784  TAGTTCGGAT TGAGGGCTGA AACTCGCCCT CATGAAGCTG GATTCGGTAG
       786  TAGTTCGGAT TGAGGGCTGA AACTCGCCCT CATGAAGCTG GATTCGGTAG
   B1bra.5  TAGTTCGGAT TGAGGGCTGA AACTCGCCCT CATGAAGCTG GATTCGGTAG
   B2S.26   TAGTTCGGAT TGTGGGCTGC AACTCGCCCA CATGAAGCTG GATTCGGTAG
   GSL5.48  TAGTTCGGAT TGTGGGCTGC AACTCGCCCA CATGAAGCTG GATTCGGTAG
Haloarcula  TAGTTCGGAT TGCGGGCTGA AACCCGCCCG CATGAAGCTG GATTCGGTAG
      Nh.2  TAGTTCGGAT TGCGGGCTGA AACCCGCCCG CATGAAGCTG GATTCGGTAG
      Nh.5  TAGTTCGGAT TGTGGGCTGC AACTCGCCCA CATGAAGCTG GATTCGGTAG
      PIG.5 TAGTTCGGAT TGTGGGCTGA AACTCGCCCA CATGAAGCTG GATTCGGTAG
      PIG.8 TAGTTCGGAT TGTGGGCTGC AACTCGCCCA CATGAAGCTG GATTCGGTAG
      PW1.1 TAGTTCGGAT TGTGGGCTGC AACTCGCCCA CATGAAGCTG GATTCGGTAG
      PW5.4 TAGTTCGGAT TGAGGGCTGA AACTCGCCCT CATGAAGCTG GATTCGGTAG
      SC1.2 TAGTTCGGAT TGTGGGCTGC AACTCGCCCA CATGAAGCTG GATTCGGTAG
      SR1.5 TAGTTCGGAT TGAGGACTGA AACTCGNCCT CATGAAGCTG GATTCGGTAG
       T1.3 TAGTTCGGAC TGAGGGCTGA AACTCGCCCT CACGAAGCTG GATTCGGTAG
       T1.6 TAGTTCGGAT TGAGGGCTGA AACTCGCCCT CATGAAGCTG GATTCGGTAG
       T3.1 TAGTTCGGAT TGCGGGCTGA AACCCGCCCG CATGAAGCTG GATTCGGTAG
       T3.2 TAGTTCGGAT TGCGGGCTGA AACCCGCCCG CATGAAGCTG GATTCGGTAG
       T3.5 TAGTTCGGAT TGCGGGCTGA AACCCGCCCG CATGAAGCTG GATTCGGTAG
       T3.6 TAGTTCGGAT TGCGGGCTGA AACCCGCCCG CATGAAGCTG GATTCGGTAG
       T4.2 TAGTTCGGAT TGAGGGCTGA AACTCGCCCT CATGAAGCTG GATTCGGTAG
       T5.7 TAGTTCGGAT TGAGGGCTGA AACTCGCCCT CATGAAGCTG GATTCGGTAG
     WB4E.8 TAGTTCGGAT TGTGGGCTGC AACTCGCCCA CATGAAGCTG GATTCGGTAG
   WM2C.13  TAGTTCGGAT TGAGGGCTGA AACTCGCCCT CATGAAGCTG GATTCGGTAG
 Haloferax  TAGTTCGGAT TGAGGGCTGA AACTCGCCCT CATGAAGCTG GATTCGGTAG
 Halobact.  TAGTTCGGAT TGAGGGCTGA AACTCGCCCT CATGAAGCTG GATTCGGTAG
Halorubrum  TAGTTCGGAT TGTGGGCTGC AACTCGCCCA CATGAAGCTG GATTCGGTAG
 Natronob.  TACTTCGGAT TGAGGGCTGA AACTCGCCCT CATGAAGCTG GATTCGGTAG


            1301                                                1350
Methanosp.  TAATCGCGTT TCAACATAGC GCGGTGAATA TGTCCCTGCT CCTTGCACAC
       777  TAATCGCGCC TCAGAAGGGC GCGGTGAATA CGTCCCTGCT CCTTGCACAC
       784  TAATCGCGCC TCAGAAGGGC GCGGTGAATA CGTCCCTGCT CCTTGCACAC
       786  TAATCGCGCC TCAGAAGGGC GCGGTGAATA CGTCCCTGCT CCTTGCACAC
   B1bra.5  TAATCGCGTG TCAGAAGCGC GCGGTGAATA CGTCCCTGCT CCTTGCACAC
   B2S.26   TAATCGCGTG TCAGCAGCGC GCGGTGAATA CGTCCCTGCT CCTTGNACAC
   GSL5.48  TAATCGCGTG TCACAAGCGC GCGGTGAATA CGTCCCTGCT CCTTGCACAC
Haloarcula  TAATCGCGTG TCAGAAGCGC GCGGTGAATA CGTCCCTGCT CCTTGCACAC
      Nh.2  TAATCGCGTG TCAGAAGCGC GCGGTGAATA CGTCCCTGCT CCTTGNACAC
      Nh.5  TAATCGCGTG TCATAAGCGC GCGGTGAATA CGTCCCTGCT CCTTGCACAC
      PIG.5 TAATCGCCTC TCAGAAGGAG GCGGTGAATA CGTCCCTGCT CCTTGNACAC
      PIG.8 TAATCGCGTG TCACAAGCGC GCGGTGAATA GGTCCCTGCT CCTTGNACAC
      PW1.1 TAATCGCGTG TCATAAGCGC GCGGTGAATA CGTCCCTGCT CCTTGCACAC
      PW5.4 TAATCGCGTG TCAGAAGCGC GCGGTGAATA CGTCCCTGCT CCTTGCACAC
      SC1.2 TAATCGCGTG TCACAAGCGC GCGGTGAATA CGTCCCTGCT CCTTGCACAC
      SR1.5 TAATCGCGCC TCAGAAGGGC GCGGTGAATA CGTCCCTGCT CCTTGNACAC
       T1.3 TAGTCGCCTT TCACAAGAGG GCGGCGAATA CGTCCCTGCT CCTTGCACAC
       T1.6 TAATCGCCAC TCAGAAGGTG GCGGTGAATA CGTCCCTGCT CCTTGCACAC
       T3.1 TAGTCGCGTG TCATTAGCGC GCGGTGAATA CGTCCCTGCT CCTTGCACAC
       T3.2 TAATCGCGTG TCAGAAGCGC GCGGTGAATA CGTCCCTGCT CCTTGCACAC
       T3.5 TAATCGCGTG TCAGAAGCGC GCGGTGAATA CGTCCCTGCT CCTTGCACAC
       T3.6 TAGTCGCGTG TCATTAGCGC GCGGCGAATA CGTCCCTGCT CCTTGCACAC
       T4.2 TAATCGCGTG TCAGAAGCGC GCGGTGAATC CGTCCCTGCT CCTTGCACAC
       T5.7 TAATCGCGCC TCAGAAGGGC GCGGTGAATA CGTCCCTGCT CCTTGCACAC
     WB4E.8 TAATCGCGTG TCACAAGCGC GCGGTGAATA CGTCCCTGCT CCTTGCACAC
   WM2C.13  TAATCGCGTG TCAGAAGCGC GCGGTGAATA CGTCCCTGCT CCTTGCACAC
 Haloferax  TAATCGCATT TCAATAGAGT GCGGTGAATA CGTCCCTGCT CCTTGCACAC
 Halobact.  TAATCGCGTG TCAGCAGCGC GCGGTGAATA CGTCCCTGCT CCTTGCACAC
Halorubrum  TAATCGCGTG TCACAAGCGC GCGGTGAATA CGTCCCTGCT CCTTGCACAC
 Natronob.  TAATCGCGCC TCAGAAGGGC GCGGTGAATA CGTCCCTGCT CCTTGCACAC
```

```
            1351                                                    1400
Methanosp.  ACCGCCCGTC AAACCACCCG AGTGAGGTCT TGATGAGGAT GTATCATTGA
       777  ACCGCCCGTC AAAGCACCCG AGTGGGGTCC GGATGAGGCC GACGCA..AC
       784  ACCGCCCGTC AAAGCACCCG AGTGGGGTCC GGATGAGGCC GACGCA..AC
       786  ACCGCCCGTC AAAGCACCCG AGTGGGGTCC GGATGAGGCC GACGCA..AC
   B1bra.5  ACCGCCCGTC AAAGCACCCG AGTGGGGTTC GGATGAGGCC G..GCA.TGC
   B2S.26   ACCGCCCGTC AAAGCACCCG AGTGAGGTCC GGATGAGGCT CCCG.A..GA
   GSL5.48  ACCGCCCGTC AAAGCACCCG AGTGAGGTCC GGATGAGGCG TTC.CA...C
Haloarcula  ACCGCCCGTC AAAGCACCCG AGTGGGGTCC GGATGAGGCC GT..CA.TGC
     Nh.2   ACCGCCCGTC AAAGCACCCG AGTGGGGTCC GGATGAGGCC TGGTCA....
     Nh.5   ACCGCCCGTC AAAGCACCCG AGTGAGGTCC GGATGAGGCT CGT.TA...C
     PIG.5  ACCGCCCGTC AAAGCACCCG AGTGAGGTCC GGATGAGGCC CAGA....G.
     PIG.8  ACCGCCCGTC AAAGCACCCG AGTGAGGTCC GGATGAGGCG CGTTC....C
     PW1.1  ACCGCCCGTC AAAGCACCCG AGTGAGGTCC GGATGAGGCT CGT.T..T.C
     PW5.4  ACCGCCCGTC AAAGCACCCG AGTGGGGTTC GGATGAGGCC G..GCA.TGC
     SC1.2  ACCGCCCGTC AAAGCACCCG AGTGAGGTCC GGATGAGGCG TAC.C..T.C
     SR1.5  ACCGCCCGTC AAAGCACCCG AGTGGGGTCC GGATGAGGCC TCGAT..T.C
      T1.3  ACCGCCCGTC AAAGCACCCG AGTGAGGTCC GGATGAGGCT GTGTCA...C
      T1.6  ACCGCCCGTC AAAGCACCCG AGTGGGGTCC GGATGAGGCT CGG.T..T.C
      T3.1  ACCGCCCGNC AAAGCACCCG AGTGAGGTCC GGATGAGGNC AC..CA.TGC
      T3.2  ACCGCCCGTC AAAGCACCCG AGTGGGGTCC GGATGAGGCC GT..CA.TGC
      T3.5  ACCGCCCGNC AAAGCACCCG AGTGGGGTCC GGATGAGGCC GN..CA.TGC
      T3.6  ACCGCCCGNC AAAGCACCCG AGTGAGGTCC GGATGAGGCC AC..CA.TGC
      T4.2  ACCGCCCGNC AAATCACCCG AGTGGGGTCC GGATGAGGCC GT..CA.TGC
      T5.7  ACCGCCCGTC AAAGCACCCG AGTGGGGTCC GGATGAGGTC GACGCA..AC
     WB4E.8 ACCGCCCGTC AAAGCACCCG AGTGAGGTCC GGATGAGGCG CGTTC....C
   WM2C.13  ACCGCCCGTC AAAGCACCCG AGTGGGGTTC GGATGAGGCC G..GCA.TGC
 Haloferax  ACCGCCCGTC AAAGCACCCG AGTGAGGTCC GGATGAGGCC ACCACA...C
 Halobact.  ACCGCCCGTC AAATCACCCG AGTGGGGTTC GGATGAGGCC G..GCA.TGC
Halorubrum  ACCGCCCGTC AAAGCACCCG AGTGAGGTCC GGATGAGGCG CGTTC....C
 Natronob.  ACCGCCCGTC AAAGCACCCG AGTGGGGTCC GGATGAGGCC CGGTT..T.A

            1401                                                    1450
Methanosp.  TATGTTCGAA TCTGGGTTTT GCAAGGGGGG TTAAGTCGTA ACAAGGTAGC
       777  GTCGGTCGAA TCTGGGCTCC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
       784  GTCGGTCGAA TCTGGGCTCC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
       786  GTCGGTCGAA TCTGGNCTCC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
   B1bra.5  GCCGGTCGAA TCTGGGCTCC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
   B2S.26   GGGAGTCGAA TCTGGGCTTC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
   GSL5.48  GAACGTCGAA TCTGGGCTTC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
Haloarcula  GACGGTCGAA TCTGGGCTCC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
     Nh.2   CCAGGTCGAA TCTGGGCTCC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
     Nh.5   ACGAGTCGAA TCTGGGCTTC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
     PIG.5  ATGGGTCGAA TCTGGGCTTC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
     PIG.8  ACGCGTCGAA TCTGGGCTTC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
     PW1.1  ACGAGTCGAA TCTGGGCTTC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
     PW5.4  GCCGGTCGAA TCTGGGCTCC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
     SC1.2  GTACGTCGAA TCTGGGCTTC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
     SR1.5  CGAGGTCGAA TCTGGGCTCC GCAAGGGGGN TTAAGTCGTA ACAAGGTAGC
      T1.3  CACAGTCGAA TCTGGGCTTC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
      T1.6  CCGAGTCGAA TCTGGGCTCC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
      T3.1  GGNGGTCGAA TCTGGGCTTC GCAAGGGGGC TTANGTCGTA ACAAGGTAGC
      T3.2  GACGGTCAAA TCTGGGCTCC GCAAGGGGGC TTANGTCGTA ACAAGGTAGC
      T3.5  GACGGTCGAA TCTGGNCTCC GCAAGGGGGC TTANGTCGTA ACAAGGTAGC
      T3.6  GGTGGTCGAA TCTGGGCTTC GCAAGGGGGC TTANGTCGTA ACAAGGTAGC
      T4.2  GACGGTCGAA TCTGGNCTCC GCAAGGGGGA TTAAGTCGTA ACAAGGTAGC
      T5.7  GTCGGTCGAA TCTGGGCTCC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
     WB4E.8 ACGCGTCGAA TCTGGGCTTC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
   WM2C.13  GCCGGTCGAA TCTGGGCTCC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
 Haloferax  GGTGGTCGAA TCTGG.CTTC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
 Halobact.  GCTGGTCAAA TCTGGGCTCC GCAAGGGGGA TTAAGTCGTA ACAAGGTAGC
Halorubrum  ACGCGTCGAA TCTNGGCTTC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
 Natronob.  CCGGGTCGAA TCTGG.CTCC GCAAGGGGGC TTAAGTCGTA ACAAGGTAGC
```

3'

	1451			1481
Methanosp.	CGTAGGGGAA	TCTGCGGCTG	GAT**CACCTCC**	**T**
777	CGTAGGGGAA	TCTGCGGCTG	GAT**CANNNNN**	**N**
784	CGTAGGGGAA	TCTGCGGCTG	GAT**CACCTCC**	**N**
786	CGTAGGGGAA	TCTGCGGCTG	GAT**CACCTCC**	**N**
B1bra.5	CGTAGGGGAA	TCTGCGGCTG	GAT**CACNNNN**	**N**
B2S.26	CGTAGGGGAA	TCTGCGGCTG	GAT**CANNNNN**	**N**
GSL5.48	CGTAGGGGAA	TCTGCGGCTG	GAT**CACCNNN**	**N**
Haloarcula	CGTAGAGGAA	TCTGTGGCTG	GAT**CACCTCC**	**T**
Nh.2	CGTAGGGGAA	TCTGCGGCTG	GAT**CACCTCC**	**N**
Nh.5	CGTAGGGGAA	TCTGCGGCTG	GAT**CACCTCC**	**N**
PIG.5	CGTAGGGGAA	TCTGCGGCTG	GAT**NNNNNNN**	**N**
PIG.8	CGTAGGGGAA	TCTGCGGCTG	GAT**CACCNNN**	**N**
PW1.1	CGTAGGGGAA	TCTGCGGCTG	GAT**CACCNNN**	**N**
PW5.4	CGTAGGGGAA	TCTGCGGCTG	GAT**CACCNNN**	**N**
SC1.2	CGTAGGGGAA	TCTGCGGCTG	GAT**CACCNNN**	**N**
SR1.5	CGTAGGGGAA	TCTGCGGNTG	GAT**CACCNNN**	**N**
T1.3	CGTAGGGGAA	TCTGCGGCTG	GAT**CACCNNN**	**N**
T1.6	CGTAGGGGAA	TCTGCGGCTG	GAT**CACCNNN**	**N**
T3.1	CGTAGGGGAA	TCTGCGGCTG	GAT**CACCTCN**	**N**
T3.2	CGTAGAGGAA	TCTGCGGCTG	GAT**CACCNNN**	**N**
T3.5	CGTAGAGGAA	TCTGTGGCTG	GAT**CACCTCN**	**N**
T3.6	CGTAGGGGAA	TCTGCGGCTG	GAT**CACCNNN**	**N**
T4.2	CGTAGGGGAA	TCTGCGGCTG	GAT**CACCNNN**	**N**
T5.7	CGTAGGGGAA	TCTGCGGCTG	GAT**CACCTCC**	**N**
WB4E.8	CGTAGGGGAA	TCTGCGGCTG	GAT**CACCNNN**	**N**
WM2C.13	CGTAGGGGAA	TCTGCGGCTG	GAT**CNNNNNN**	**N**
Haloferax	CGTAGGGGAA	TCTGCGGCTG	GAT**CACCTCC**	**T**
Halobact.	CGTAGGGGAA	TCTGCGGCTG	GAT**CACCTCC**	**T**
Halorubrum	CGTAGGGGAA	TCTGCGGCTG	GAT**CACCTNN**	**N**
Natronob.	CGTAGGGGAA	TCTGCGGCTG	GAT**CACCTCC**	**A**

	Hrp. hungatei	*Halobacterium' sp.777*	*Halobacterium' sp.784*	*Halobacterium' sp.786*	B1bra.5	B2S.26	GSL5.48	*Haloarcula marismortui rrnA*	Nh.2	Nh.5	PIG.5	PIG.8	PW1.1	PW5.4	SC1.2	SR1.5	T1.3	T1.4	T3.1	T3.2	T3.5	T3.6	T4.2	T5.7	WB4E.8	WM2C.13	*Haloferax volcanii*	*Halobacterium salinarium*	*Halorubrum saccharovorum*	*Natronobacterium magadii*
Hrp.hungatei		77.4	77.7	77.6	79.0	78.4	76.5	78.9	78.7	76.7	78.5	77.1	77.1	79.1	76.3	77.3	77.9	77.2	78.5	79.2	79.7	78.4	77.7	77.8	77.0	79.0	78.7	78.7	77.0	76.7
Halobacterium' sp.777	0.2688		100	99.4	90.7	87.6	87.1	88.8	89.5	87.1	90.9	87.5	87.3	90.7	87.9	96.3	89.4	90.5	91.0	89.0	90.0	90.8	90.2	99.6	87.2	90.7	89.8	89.1	87.0	93.9
Halobacterium' sp.784	0.2640	0		99.4	90.8	87.8	87.2	88.9	89.4	87.1	91.0	87.6	87.5	90.8	87.9	96.3	89.6	90.4	91.0	86.9	89.9	90.8	90.0	99.5	87.3	90.7	89.8	89.2	87.1	93.9
Halobacterium' sp.786	0.2664	0.0057	0.0057		90.8	87.8	86.8	88.5	89.2	86.8	90.8	87.2	86.9	90.8	87.6	95.9	89.4	90.7	91.0	89.0	89.9	90.9	90.3	99.2	86.9	90.8	89.5	89.1	86.6	93.4
B1bra.5	0.2459	0.0998	0.0983	0.0984		98.4	87.7	90.6	91.5	87.8	90.5	88.8	88.3	99.9	87.6	90.1	89.1	89.0	91.7	89.1	90.5	91.5	91.4	91.2	88.7	99.9	88.7	97.0	88.4	90.6
B2S.26	0.2543	0.1350	0.1333	0.1335	0.1265		91.7	87.4	88.6	92.5	87.9	92.6	92.7	88.3	91.3	88.0	89.3	88.5	88.0	87.4	88.2	87.8	87.3	87.7	92.8	88.3	89.2	88.0	92.8	87.9
GSL5.48	0.2825	0.1414	0.1399	0.1456	0.1347	0.0877		85.9	88.8	95.7	86.5	96.2	95.2	87.6	98.8	87.0	87.3	87.2	85.8	87.0	87.7	85.8	87.1	87.0	95.8	87.6	88.3	87.1	95.8	86.0
Haloarcula marismortui rrnA	0.2476	0.1214	0.1207	0.1250	0.1002	0.1381	0.1565		90.0	85.6	88.9	86.1	85.8	90.6	85.7	89.1	88.1	87.4	91.3	96.1	98.1	91.1	91.5	88.8	96.3	90.7	88.1	90.0	86.2	89.2
Nh.2	0.2507	0.1132	0.1147	0.1165	0.0905	0.1236	0.1212	0.1071		89.4	89.1	89.8	89.4	91.5	88.5	89.9	88.8	89.1	91.2	90.2	91.2	91.0	90.2	89.4	89.9	91.4	89.2	90.1	89.2	88.9
Nh.5	0.2792	0.1421	0.1411	0.1455	0.1326	0.0792	0.0441	0.1595	0.1138		87.5	97.5	98.7	87.7	95.1	87.3	87.5	88.1	86.5	86.5	87.4	86.5	86.8	87.0	97.8	87.7	88.8	87.3	97.3	86.7
PIG.5	0.2533	0.0968	0.0954	0.0986	0.1020	0.1322	0.1483	0.1196	0.1176	0.1362		87.4	87.4	90.5	86.7	91.2	89.5	89.8	90.5	89.0	90.1	90.2	89.2	90.9	87.7	90.4	89.6	89.1	87.1	91.1
PIG.8	0.2738	0.1368	0.1359	0.1406	0.1211	0.0775	0.0390	0.1543	0.1097	0.0258	0.1375		97.5	88.7	95.9	87.5	87.9	88.2	86.5	86.8	87.5	86.3	87.3	87.4	98.5	88.7	88.6	88.0	98.8	87.3
PW1.1	0.2737	0.1391	0.1373	0.1437	0.1271	0.0766	0.0493	0.1579	0.1144	0.0127	0.1380	0.0257		88.1	95.0	87.4	87.5	88.4	86.9	86.7	87.4	86.7	87.1	87.2	97.5	88.1	88.8	87.8	97.5	87.2
PW5.4	0.2453	0.0992	0.0984	0.0985	0.0907	0.1267	0.1351	0.1004	0.0906	0.1344	0.1020	0.1226	0.1296		87.6	90.0	88.9	88.9	91.5	89.1	90.4	91.3	91.3	91.1	88.5	100	88.7	96.9	88.3	90.5
SC1.2	0.2851	0.1324	0.1316	0.1354	0.1351	0.0921	0.0121	0.1585	0.1245	0.0503	0.1461	0.0420	0.0515	0.1351		87.3	87.1	87.3	86.2	87.1	87.8	86.0	86.9	87.7	95.4	87.6	87.9	86.8	95.3	85.9
SR1.5	0.2703	0.0377	0.0384	0.0422	0.1057	0.1309	0.1425	0.1183	0.1082	0.1394	0.0936	0.1366	0.1380	0.1072	0.1392		89.7	90.8	91.1	89.4	90.0	90.9	89.9	96.6	87.3	90.1	89.7	89.2	87.4	94.4
T1.3	0.2614	0.1146	0.1121	0.1148	0.1181	0.1152	0.1386	0.1291	0.1215	0.1363	0.1136	0.1319	0.1366	0.1204	0.1413	0.1105		88.6	88.8	88.5	89.3	88.7	88.5	89.3	88.1	88.9	91.8	87.7	87.5	88.4
T1.4	0.2718	0.1020	0.1027	0.0997	0.1193	0.1246	0.1404	0.1380	0.1172	0.1298	0.1097	0.1280	0.1266	0.1201	0.1389	0.0978	0.1302		89.6	87.0	88.0	89.4	89.4	90.5	88.3	88.9	87.3	88.0	89.2	89.5
T3.1	0.2535	0.0957	0.0959	0.0958	0.0881	0.1311	0.1576	0.0924	0.0936	0.1484	0.1019	0.1483	0.1443	0.0905	0.1525	0.0942	0.1208	0.1120		88.9	91.1	99.9	93.1	91.2	86.9	91.6	89.0	90.5	86.5	90.3
T3.2	0.2440	0.1189	0.1204	0.1193	0.1173	0.1383	0.1427	0.0402	0.1052	0.1489	0.1190	0.1449	0.1465	0.1175	0.1418	0.1140	0.1253	0.1427	0.1089		99.4	89.7	89.7	88.9	87.0	89.2	88.5	89.1	86.8	88.7
T3.5	0.2362	0.1073	0.1081	0.1083	0.1014	0.1283	0.1347	0.0191	0.0935	0.1382	0.1063	0.1357	0.1385	0.1024	0.1330	0.1074	0.1154	0.1310	0.0945	0.0959		91.0	91.0	89.9	87.9	90.6	89.2	90.0	87.6	89.8
T3.6	0.2552	0.0980	0.0982	0.0974	0.0905	0.1328	0.1576	0.0947	0.0960	0.1490	0.1045	0.1509	0.1468	0.0928	0.1551	0.0965	0.1224	0.1145	0.0014	0.1105	0.0960		92.9	90.8	86.7	91.4	88.9	90.4	86.3	90.1
T4.2	0.2642	0.1054	0.1070	0.1039	0.0918	0.1395	0.1420	0.0909	0.1050	0.1446	0.1170	0.1389	0.1415	0.0921	0.1445	0.1082	0.1250	0.1144	0.0725	0.1113	0.0959	0.0748		90.1	87.4	91.3	88.1	90.9	87.3	89.3
T5.7	0.2635	0.0949	0.0950	0.0978	0.0933	0.1339	0.1424	0.1215	0.1144	0.1425	0.0974	0.1384	0.1406	0.0948	0.1341	0.0948	0.1152	0.1015	0.0937	0.1204	0.1085	0.0976	0.1059		87.1	91.1	89.5	89.6	87.2	93.8
WB4E.8	0.2743	0.1401	0.1393	0.1438	0.1227	0.0760	0.0427	0.1511	0.1090	0.0222	0.1341	0.0149	0.0250	0.1243	0.0472	0.1394	0.1293	0.1271	0.1441	0.1422	0.1315	0.1468	0.1380	0.1411		88.5	89.1	87.8	98.9	87.0
WM2C.13	0.2457	0.0993	0.0992	0.0979	0.0907	0.1268	0.1352	0.0997	0.0908	0.1344	0.1024	0.1228	0.1296	0.0909	0.1350	0.1067	0.1200	0.1203	0.0891	0.1166	0.1007	0.0914	0.0921	0.0942	0.1245		88.7	97.0	88.3	90.5
Haloferax volcanii	0.2505	0.1100	0.1093	0.1134	0.1222	0.1162	0.1269	0.1298	0.1168	0.1214	0.1114	0.1238	0.1210	0.1224	0.1314	0.1110	0.0863	0.1390	0.1185	0.1254	0.1163	0.1201	0.1291	0.1131	0.1182	0.1225		87.8	88.6	89.5
Halobacterium salinarium	0.2512	0.1172	0.1165	0.1183	0.0305	0.1313	0.1420	0.1076	0.1062	0.1390	0.1179	0.1313	0.1333	0.0312	0.1448	0.1166	0.1340	0.1313	0.1012	0.1177	0.1068	0.1027	0.0971	0.1115	0.1332	0.0305	0.1330		87.9	89.6
Halorubrum saccharovorum	0.2742	0.1432	0.1416	0.1477	0.1255	0.0760	0.0434	0.1526	0.1168	0.0280	0.1413	0.0121	0.0257	0.1273	0.0487	0.1382	0.1366	0.1279	0.1488	0.1453	0.1355	0.1513	0.1393	0.1408	0.0114	0.1275	0.1232	0.1314		87.2
Natronobacterium magadii	0.2791	0.0640	0.0632	0.0695	0.1010	0.1321	0.1550	0.1166	0.1201	0.1459	0.0959	0.1396	0.1409	0.1019	0.1567	0.0587	0.1263	0.1130	0.1037	0.1228	0.1102	0.1060	0.1149	0.0644	0.1424	0.1021	0.1131	0.1117	0.1405	

Appendix 2a. Distance Matrix computed from the sequence alignment in Appendix 1, using the Jukes-Cantor model for nucleotide substitution, to derive the phylogenetic trees in Figures 3.17 and 3.18. Upper triangle, percentage similarities; lower triangle, phylogenetic distance values (Knuc) used in FITCH to derive the Distance tree.

Appendix 2b	*M.hung.*	Ch2	GSL5.48	Nh.2	Nh.5	PIG8	PW1.1	SC1.2	WB4E.8	*Hr. sod.*	*Hr. sac.*	*Hr. tra.*	*Hr. lac.*	*H. gom.*	BbpA.3	54R	Malk006	GluBr1.1	B2S.26	*Hf.volc.*	T1.3
Msp.hungatei		76.9	77.0	78.8	77.1	77.5	77.4	76.8	77.5	77.4	77.5	77.1	77.7	78.6	78.6	77.6	77.4	77.3	78.9	78.8	77.9
Ch2	0.2760		98.8	87.9	94.8	95.6	95.0	98.4	95.2	98.6	95.4	98.5	95.0	89.8	91.4	95.2	95.9	98.8	91.4	88.0	87.0
GSL5.48	0.2748	0.0126		88.9	95.8	96.2	95.2	98.6	95.6	98.3	95.6	98.3	95.5	89.6	91.7	95.7	95.9	98.8	91.8	88.2	87.3
Nh.2	0.2493	0.1319	0.1197		89.4	89.6	89.2	88.7	89.7	88.5	89.1	88.1	88.8	88.8	88.0	89.0	88.1	88.6	88.4	89.0	88.9
Nh.5	0.2737	0.0535	0.0432	0.1148		97.5	98.7	98.4	97.6	95.3	97.3	95.1	97.6	89.0	92.1	97.3	96.5	94.8	92.5	88.8	87.9
PIG8	0.2677	0.0453	0.0389	0.1117	0.0257		97.5	98.1	98.5	96.1	98.7	95.9	98.2	89.5	92.4	98.7	97.0	96.1	92.7	88.5	88.2
PW1.1	0.2683	0.0518	0.0491	0.1162	0.0127	0.0256		95.2	97.6	95.4	97.4	95.3	98.1	89.4	92.5	97.6	96.1	95.0	92.7	88.8	87.8
SC1.2	0.2773	0.0162	0.0120	0.1230	0.0478	0.0396	0.0491		95.6	98.8	95.4	98.7	95.0	89.3	91.1	95.3	95.4	98.5	91.5	88.0	87.2
WB4E.8	0.2681	0.0498	0.0433	0.1109	0.0221	0.0148	0.0249	0.0455		95.6	98.8	95.7	98.0	89.0	92.7	96.9	97.1	95.6	92.9	89.0	88.3
Hr. sodomense	0.2691	0.0140	0.0170	0.1244	0.0486	0.0397	0.0476	0.012	0.0456		96.0	98.8	95.6	90.1	91.8	95.9	96.1	98.9	91.8	88.5	87.1
Hr. saccharovorum	0.2679	0.0480	0.0432	0.1178	0.0278	0.0127	0.0263	0.047	0.012	0.0408		95.9	98.3	89.4	92.7	99.7	97.2	95.8	92.9	88.5	87.7
Hr. trapanicum	0.2732	0.0147	0.0177	0.1291	0.0507	0.0425	0.0490	0.0127	0.0440	0.0119	0.0421		95.3	89.8	92.1	95.8	95.6	98.8	92.1	88.4	86.8
Hr. locusprofundi	0.2643	0.0514	0.0467	0.1213	0.0241	0.0184	0.0190	0.0520	0.0205	0.0458	0.0168	0.0486		89.3	92.3	98.5	96.5	95.4	92.4	88.7	87.7
H. gomorrense	0.2514	0.1098	0.1125	0.1212	0.1185	0.1126	0.1146	0.1150	0.1185	0.1056	0.1144	0.1102	0.1155		88.9	89.3	89.2	89.9	88.9	89.6	89.1
BbpA.3	0.2515	0.0916	0.0877	0.1306	0.0830	0.0800	0.0788	0.0949	0.0770	0.0873	0.0770	0.0831	0.0814	0.1203		92.6	92.4	91.6	100	88.9	89.2
54R	0.2656	0.0494	0.0447	0.1193	0.0271	0.0127	0.0241	0.0485	0.0106	0.0423	0.0028	0.0435	0.0154	0.1160	0.0754		97.2	95.6	93.0	88.7	88.0
Malk006.3	0.2686	0.0420	0.0424	0.1293	0.0358	0.0307	0.0401	0.0470	0.0293	0.0400	0.0283	0.0457	0.0355	0.1169	0.0800	0.0283		95.8	92.5	87.9	87.2
GluBr1.1	0.2701	0.0119	0.0120	0.1237	0.0538	0.0404	0.0521	0.0149	0.0457	0.0112	0.0437	0.0119	0.0473	0.1090	0.0888	0.0451	0.0429		91.6	88.0	87.0
B2S.26	0.2483	0.0917	0.0873	0.1255	0.0757	0.0765	0.0769	0.0907	0.0750	0.0867	0.0748	0.0833	0.0798	0.1199	0	0.0732	0.0794	0.0891		89.3	89.6
Hf. volcanii	0.2494	0.1319	0.1281	0.1187	0.1218	0.1243	0.1213	0.1309	0.1195	0.1244	0.1243	0.1258	0.1229	0.1119	0.1196	0.1226	0.1325	0.1311	0.1159		91.9
T1.3	0.2613	0.1429	0.1390	0.1202	0.1324	0.1282	0.1328	0.1405	0.1272	0.1421	0.1344	0.1452	0.1346	0.1177	0.1164	0.1310	0.1402	0.1431	0.1124	0.0858	

Appendix 2c	*Msp.hungatei*	*Ne.occultus*	*Nb.magadii*	T1.6	PIG.5	SR1.5	T5.7	784	786	777	*H.trap.*767	172P1
Msp.hungatei		77.8	76.9	77.4	78.6	77.4	77.8	77.6	77.5	77.4	77.3	77.1
Ne.occultus	0.2638		92.4	90.6	91.4	94.1	92.8	92.6	92.7	92.5	92.8	92.9
Nb.magadii	0.2758	0.0797		89.5	91.0	94.7	94.1	94.2	93.5	94.1	94.4	93.9
T1.6	0.2698	0.0999	0.1129		89.8	90.8	90.7	90.6	91.0	90.7	90.4	89.6
PIG.5	0.2519	0.0917	0.0964	0.1096		91.3	91.1	91.3	91.0	91.1	90.6	90.2
SR1.5	0.2688	0.0616	0.0548	0.0978	0.0920		96.8	96.4	96.0	96.5	96.2	94.2
T5.7	0.2638	0.0760	0.0613	0.0990	0.0950	0.0326		99.5	99.2	99.6	98.9	94.2
784	0.2666	0.0778	0.0608	0.1002	0.0929	0.0369	0.0050		99.4	100	99.0	93.9
786	0.2672	0.0764	0.0678	0.0964	0.0962	0.0406	0.0078	0.0057		99.4	98.5	93.7
777	0.2696	0.0794	0.0615	0.0995	0.0944	0.0361	0.0042	0	0.0056		98.9	93.8
*H.trap.*767	0.2702	0.0754	0.0578	0.1027	0.1000	0.0395	0.0113	0.0099	0.0156	0.0106		94.2
172P1	0.2732	0.0747	0.0639	0.1116	0.1048	0.0606	0.0609	0.0636	0.0660	0.0643	0.0606	

Appendix 2d	*Msp.hung.*	BbpA.1	OO4.1	*Hb.salin.*	PW5.4	B1bra.5	WM2C.13	*Ha.mariA*	*Ha.mariB*	*Ha.tin.Maj.*	*Ha.tin.Min.*	*Hc.17082*	*Hc.16008*	T3.2	T3.5	T4.2	T3.1	T3.6
Msp.hungatei		79.1	79.0	78.7	79.1	79.0	79.0	78.8	77.9	77.8	78.3	78.1	77.8	79.1	79.7	77.5	78.4	78.3
BbpA.1	0.2453		99.7	97.0	99.9	100	99.9	90.6	87.7	87.0	88.4	88.4	88.5	89.1	90.5	91.4	91.5	91.2
OO4.1	0.2459	0.0028		96.8	99.7	99.7	99.6	90.5	87.6	87.0	88.3	88.4	88.5	89.1	90.4	91.3	91.3	91.1
Hb.salinarium	0.2510	0.0311	0.0326		96.9	97.0	96.9	90.0	87.5	87.0	88.3	88.4	88.5	89.1	90.1	91.0	90.5	90.3
PW5.4	0.2451	0.0007	0.0035	0.0312		99.9	99.9	90.6	87.6	87.0	88.3	88.5	88.6	89.1	90.4	91.4	91.5	91.2
B1bra.5	0.2457	0.0000	0.0028	0.0304	0.0007		99.9	90.6	87.6	87.0	88.3	88.4	88.5	89.1	90.5	91.4	91.7	91.4
WM2C.13	0.2463	0.0014	0.0042	0.0312	0.0007	0.0014		90.6	87.6	86.9	88.3	88.4	88.5	89.1	90.5	91.3	91.5	91.3
Ha.marismortui rrnA	0.2493	0.1000	0.1002	0.1076	0.1003	0.1001	0.1004		94.9	94.0	96.0	89.1	89.0	96.1	98.1	91.6	91.4	91.1
Ha.marismortui rrnB	0.2620	0.1349	0.1360	0.1368	0.1353	0.1351	0.1354	0.0529		98.6	96.9	86.9	86.7	99.2	99.1	88.4	88.6	88.4
Ha.tinatlentis Maj.	0.2631	0.1429	0.1433	0.1424	0.1433	0.1431	0.1435	0.0621	0.0140		97.2	86.7	86.6	99.0	98.6	87.7	88.2	87.9
Ha.tinatlentis Min.	0.2556	0.1265	0.1268	0.1276	0.1268	0.1267	0.1269	0.0414	0.0312	0.0283		87.9	87.9	98.0	98.2	88.9	89.6	89.3
Hc. morrhuae17082	0.2586	0.1258	0.1261	0.1261	0.1245	0.1260	0.1263	0.1173	0.1436	0.1459	0.1319		99.3	88.0	89.4	89.8	90.8	90.6
Hc. morrhuae16008	0.2629	0.1247	0.1250	0.1251	0.1234	0.1249	0.1252	0.1186	0.1460	0.1471	0.1323	0.0070		87.8	89.3	89.7	90.6	90.4
T3.2	0.2450	0.1172	0.1175	0.1176	0.1174	0.1172	0.1173	0.0401	0.0079	0.0101	0.0203	0.1311	0.1326		99.4	89.6	89.9	89.8
T3.5	0.2372	0.1021	0.1032	0.1067	0.1023	0.1013	0.1014	0.0191	0.0088	0.0140	0.0185	0.1145	0.1158	0.0059		90.9	91.1	90.9
T4.2	0.2670	0.0910	0.0928	0.0963	0.0912	0.0910	0.0920	0.0891	0.1263	0.1342	0.1203	0.1099	0.1112	0.1121	0.0967		93.2	93.0
T3.1	0.2543	0.0903	0.0921	0.1011	0.0904	0.0881	0.0898	0.0916	0.1232	0.1287	0.1123	0.0979	0.0999	0.1081	0.0944	0.0709		99.8
T3.6	0.2568	0.0934	0.0952	0.1034	0.0935	0.0911	0.0928	0.0947	0.1263	0.1318	0.1154	0.1099	0.1030	0.1097	0.0967	0.0740	0.0021	

Appendix 2b-d. Distance Matrices computed from sequence alignments, using the Jukes-Cantor model for nucleotide substitution, to derive the phylogenetic trees in Figures 3.20, 3.22, and 3.24, respectively. Upper triangles, percentage similarities; lower triangles, phylogenetic distance values (Knuc) used in FITCH to derive the Distance trees.

REFERENCES.

Amann, R. I., W. Ludwig and K.-H. Schleifer. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Revs.* 59(1):143-169.

Amy, P. S., C. Durham, D. Hall and D. L. Haldeman. (1993). Starvation-survival of deep subsurface isolates. *Curr. Microbiol.* 26:345-352.

Anderson, H. (1954). The reddening of salted hides and fish. *Applied Microbiology.* 2(2):64-69.

Arthurton, R. S. (1973). Experimentally produced halite compared with Triassic layered halite-rock from Cheshire, England. *Sedimentology.* 20: 145-160.

Avise, J. C. (1994). *Molecular Markers, Natural History and Evolution.* Chapman & Hall, New York and London.

Baas-Becking, L. G. M. (1931). Historical notes on salt and salt-manufacture. *Scientific Monthly.* 32:434-446.

Benachenhou-Lahfa, N., P. Forterre, and B. Labedan. (1993). Evolution of glutamate dehydrogenase genes: evidence for two paralogous protein families and unusual branching patterns of the archaebacteria in the universal tree of life. *J. Mol. Evol.* 36:335-346.

Bertrand, J. C., M. Almallah, M. Acquaviva and G. Mille. (1990). Biodegradation of hydrocarbons by an extremely halophilic archaebacterium. *Letts. Appl. Microbiol.* 11:260-263.

Bibo, Von F.-J., R. Söngen and R. E. Fresenius. (1983). Vermehrungsfähige Mikroorganismen in Steinsalz aus primären Lagerstätten. *Kali und Steinsalz.* August, 1983; pp. 367-373.

Bien, E. and W. Schwartz. (1965). Über das Vorkommen konservierter toter und lebender Bakterienzellen in Salzgesteinen. *Zeitschrift für Allg. Mikrobiologie.* 5(3):185-205.

Bligh, E. G. and Dyer, W. J. (1959). A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol.* 37(8):911-917.

Bock, M., P. Kämpfer, K. Bosecker and W. Dott. (1994). Isolation and characterization of heterotrophic, aerobic bacteria from oil storage caverns in northern Germany. *Appl. Microbiol. Biotechnol.* 42:463-468.

Boehringer Mannheim. (1993). The DIG System User's Guide for Filter Hybridization. *Boehringer Mannheim GmbH, Biochemica.*

Bradley, W. H. and H. P. Eugster. (1969). Geochemistry and paleolimnology of the trona deposits and associated authigenic minerals of the Green River formation of Wyoming. *Geological Survey Professional Paper 496-B.* U. S. Government Printing Office, Washington.

Braitsch, O. *Salt Deposits: Their Origin and Composition.* (1971). Springer-Verlag.

Brown, H. J. and N. E. Gibbons. (1955). The effect of magnesium, potassium, and iron on the growth and morphology of red halophilic bacteria. *Can. J. Microbiol.* 1:486-494.

Brown, T. J. and C. A. Ison. (1993). Non-radioactive ribotyping of *Haemophilus ducreyi* using a digoxygenin labelled cDNA probe. *Epidemiol. Infect.* 110:289-295.

Burggraf, S., A. Ching, K. O. Stetter and C. R. Woese. (1991). The sequence of *Methanospirillum hungatei* 23S rRNA confirms the specific relationship between the extreme halophiles and the Methanomicrobiales. *System. Appl. Microbiol.* 14:358-363.

Cano, R. J., H. N. Poinar, N. J. Pieniazek, A. Acra and G. O. Poinar. (1993). Amplification and sequencing of DNA from a 120-135-million-year-old weevil. *Nature.* 363:536-538.

Cano, R. J. and M. K. Borucki. (1995). Revival and identification of bacterial spores in 25- to 40-million-year-old Dominican amber. *Nature.* 268:1060-1064.

Casas, E., T. K. Lowenstein, R. J. Spencer and Z. Pengxi. (1992). Carnallite mineralization in the non-marine, Qaidam Basin, China: evidence for the early diagenetic origin of potash evaporites. *J. Sed. Petrol.*62(5):881-898.

Chatton, E. (1938). Titres et travoux scientifiques (1906-1937) de Edouard Chatton. (E. Sottano, Sète, France).

Cimmino, C., G. L. Scoarughi and P. Donini. (1993). Stringency and relaxation among the halobacteria. *J. Bact.* 175(20):6659-6662.

Collins, M. D., M. Goodfellow and D. E. Minnikin. (1980). Fatty acid, isoprenoid quinone and polar lipid composition in the classification of *Curtobacterium* and related taxa. *J. Gen. Microbiol.* 118:29-37.

Conway de Macario, E., H. Konig and A. J. L. Macario. (1986). Immunologic distinctiveness of archaebacteria that grow in high salt. *J. Bact.* 168(1):425-427.

Costas, M. (1992). Classification, identification and typing of bacteria by the analysis of their one-dimensional polyacrylamide gel electrophoresis protein patterns. pp.351-408. In *Advances in Electrophoresis*. Volume 5. Chrambach, A., M. J. Dunn and B. J. Radola. (eds.). VCH Verlagsgesellschaft mbH.

Costas, M., S. L. W. On, R. J. Owen, B. Lopez-Urquijo and A. J. Lastovica. (1993). Differentiation of *Helicobacter* species by numerical analysis of their one-dimensional electrophoretic protein patterns. *System. Appl. Microbiol.* 16:396-404.

Cowie, J. W. and M. G. Bassett. (1989). International Union of Geological Sciences 1989 global stratigraphic chart. In *Palaeobiology: a Synthesis*. Briggs, D. E. G. and P. R. Crowther (eds.). Blackwell Scientific Publications (1990).

Darwin, C. R. (1845). *The Voyage of the Beagle*. Everyman's Library No. 104, J. M. Dent & Sons Ltd., London and E. P. Dutton & Co. Inc., New York.

De Ley, J., K. Kersters and I. W. Park. (1966). Molecular-biological and taxonomic studies on *Pseudomonas halocrenaea*, a bacterium from Permian salt deposits. *Anton. v. Leeuw.* 32:315-331.

DeLong, E. F. (1992). Archaea in coastal marine environments. *Proc. Natl. Acad. Sci. USA* 89:5685-5689.

DeLong, E. F., K. Y. Wu, B. B. Prézelin and R. V. M. Jovine. (1994). High abundance of Archaea in Antarctic marine picoplankton. *Nature*. 371:695-697.

Del Moral, A., E. Quesada and A. Ramos-Cormenzana. (1987a). Distribution and types of bacteria isolated from an inland saltern. *Ann. Inst. Pasteur / Microbiol.* 138:59-66.

Del Moral, A., E. Quesada, V. Bejar and A. Ramos-Cormenzana. (1987b). Evolution of bacterial flora from a subterranean saline well by graduated salinity changes in enrichment media. *J. Appl. Bact.* 62:465-471.

Del Moral, A., E. Roldan, J. Navarro, M. Monteoliva-Sanchez and A. Ramos-Cormenzana. (1987c). Formation of calcium carbonate crystals by moderately halophilic bacteria. *Geomicrobiol. J.* 5(1):79-87.

Del Moral, A., B. Prado, E. Quesada, T. García, R. Ferrer and A. Ramos-Cormenzana. (1988). Numerical taxonomy of moderately halophilic Gram-negative rods from an inland saltern. *J. Gen. Microbiol.* 134:733-741.

Denner, E. B. M., T. J. McGenity, H.-J. Busse, W. D. Grant, G. Wanner and H. Stan-Lotter. (1994). *Halococcus salifodinae* sp. nov., an archaeal isolate from an Austrian salt mine. *Int. J. Syst. Bact.* 44(4):774-780.

Dennis, P. P. (1993). The molecular biology of halophilic archaebacteria. Ch. 9:255-287. In *The Biology of Halophilic Bacteria.* Vreeland, R. H. and L. I. Hochstein (eds.). CRC Press, Inc.

De Rosa, M., A. Gambacorta, B. Nicolaus and W. D. Grant. (1983). A C_{25}, C_{25} diether core lipid from archaebacterial haloalkaliphiles. *J. Gen. Microbiol.* 129:2333-2337.

Dixon, M. T. And D. M. Hillis. (1993). Ribosomal RNA secondary structure: compensatory mutations and implications for phylogenetic analysis. *Mol. Biol. Evol.* 10(1):256-267.

Dobson, S. J., R. R. Colwell, T. A. McMeekin and P. D. Franzmann. (1993). Direct sequencing of the polymerase chain reaction-amplified 16S rRNA gene of *Flavobacterium gondwanense* sp. nov. and *Flavobacterium salegens* sp. nov., two new species from a hypersaline Antarctic lake. *Int. J. Syst. Bact.* 43(1):77-83.

Dolzani, L., E. Tonin, C. Lagatolla and C. Monti-Bragadin. (1994). Typing of *Staphylococcus aureus* by amplification of the 16S-23S rRNA intergenic spacer sequences. *FEMS Microbiol. Letts.* 119:167-174.

Dombrowski, H. J. (1960). Balneobiologische Untersuchungen der Nauheimer Quellen. *Zentralbl. Bakteriol. Parasitenkd. Infektionskr. Hyg. Abt. I.* 178:83-90.

Dombrowski, H. (1963). Bacteria from paleozoic salt deposits. *Ann. NY Acad. Sci.* 108:453-460.

Dussault, H. P. (1955). An improved technique for staining red halophilic bacteria. *J. Bact.* 70:484-485.

Earp, J. R. and B. J. Taylor. (1986). Geology of the country around Chester and Winsford. *Mem. Br. Geol. Surv.*, Sheet 109. HMSO, London.

Embley, T. M. (1991). The linear PCR reaction: a simple and robust method for sequencing amplified rRNA genes. *Letts. Appl. Microbiol.* 13:171-174.

Evans, W. B. (1970). The Triassic salt deposits of north-western England. *Q. J. Geol. Soc. Lond.* 126:103-123.

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* 17:368-376.

Felsenstein, J. (1982). Numerical methods for inferring phylogenetic trees. *Q. Rev. Biol.* 57(4):379-404.

Felsenstein, J. (1983). Parsimony in systematics: biological and statistical issues. *Ann. Rev. Ecol. Syst.* 14:313-333.

Felsenstein, J. (1988). Phylogenies from molecular sequences: inference and reliability. *Ann. Rev. Genet.* 22:521-565.

Fitch, W. M. and E. Margoliash. (1967). Construction of phylogenetic trees. *Science.* 155:279-284.

Fox, G. E., E. Stackebrandt, R. B. Hespell, J. Gibson, J. Maniloff, T. A. Dyer, R. S. Wolfe, W. E. Balch, R. S. Tanner, L. J. Magrum, L. B. Zablen, R. Blakemore, R. Gupta, L. Bonen, B. J. Lewis, D. A. Stahl, K. R. Luehrsen, K. N. Chen, C. R. Woese. (1980). The phylogeny of prokaryotes. *Science.* 209:457-463.

Fox, G. E., J. D. Wisotskey and P. Jurtshuk, Jr. (1992). How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. *Int. J. System. Bacteriol.* 42(1):166-170.

Franzmann, P. D., E. Stackebrandt, K. Sanderson, J. K. Volkman, D. E. Cameron, P. L. Stevenson, T. A. McMeekin and H. R. Burton. (1988). *Halobacterium lacusprofundi* sp. nov., a halophilic bacterium isolated from Deep Lake, Antarctica. *System. Appl. Microbiol.* 11:20-27.

Franzmann, P. D. and B. J. Tindall. (1990). A chemotaxonomic study of members of the family *Halomonadaceae*. *System. Appl. Microbiol.* 13:142-147.

Fuchtbaüer, H. and T. Peryt (eds.). (1980). *The Zechstein Basin*. E. Schweizerbartische Verlagsbuchhandlung (Nägele u. Obermiller). Stuttgart.

Fuhrman, J. A., K. McCallum and A. A. Davis. (1992). Novel major archaebacterial group from marine plankton. *Nature*.356:148-149.

Fuhrman, J. A., K. McCallum and A. A. Davis. (1993). Phylogenetic diversity of subsurface marine microbial communities from the Atlantic and Pacific Oceans. *Appl. Environ. Microbiol.* 59(5):1294-1302.

Garrett, R. A., J. Dalgaard, N. Larsen, J. Kjems and A. S. Mankin. (1991). Archaeal rRNA operons. *TIBS*. 16 (1):22-26.

Gauthier, M. J., B. Lafay, R. Christen, L. Fernandez, M. Acquaviva, P. Bonin and J.-C. Bertrand. (1992). *Marinobacter hydrocarbonoclasticus* gen. nov., sp. nov., a new, extremely halotolerant, hydrocarbon-degrading marine bacterium. *Int. J. Syst. Bact.* 42(4):568-576.

GCG. (1991). Genetics Computer Group, Program Manual for the GCG Package. Version 7, April 1991and version 7.3-UNIX, June 1993; 575 Science Drive, Madison, Wisconsin, USA 53711.

Gest, H. and Mandelstam, J. (1987). Longevity of microorganisms in natural environments. *Microbiol. Sci.* 4(3):69-71.

Gibbons, N. E. (1974). *Halobacteriaceae*. pp. 269-273. In *Bergey's Manual of Determinative Bacteriology*. Eighth Edition. Buchanan, R. E. and N. E. Gibbons (eds.). The Williams & Wilkins Company, Baltimore.

Gonzalez, C., C. Gutierrez and C. Ramirez. (1978). *Halobacterium vallismortis* sp. nov. An amylolytic and carbohydrate-metabolizing, extremely halophilic bacterium. *Can. J. Microbiol.* 24:710-715.

Grant, W. D. and Ross, H. N. M. (1986). The ecology and taxonomy of halobacteria. *FEMS Microbiol. Revs.* 39:9-15.

Grant, W. D. and H. Larsen. (1989). Extremely halophilic archaeobacteria. pp. 2216-2233. In *Bergey's Manual of Systematic Bacteriology.* Vol. 3. Staley, J. T., M. P. Bryant, N. Pfennig and J. G. Holt. (eds.). Williams & Wilkins, Baltimore.

Grant, W. D., R. T. Gemmell, T. J. McGenity. (1996). Halophiles. In *Extremophiles.* Horikoshi, K. and W. D. Grant (eds.). John Wiley and Sons, New York.

Grimont, F. and P. A. D. Grimont. (1986). Ribosomal ribonucleic acid gene restriction patterns as potential taxonomic tools. *Ann. Inst. Pasteur/Microbiol.* 137B:165-175.

Grzesiowski, J., A. Jodlowski, E. Kalwajtys, R. Kurowski, J. Piotrowicz and J. Wiewiórka. (1992). *Wieliczka Ancient Salt Mine - Cracow Salt-Works Museum.* Muzeum Zup Krakowskich Wieliczka, Poland.

Guillot, E. and H. Leclerc. (1993). Bacterial flora in natural mineral waters: characterization by ribosomal ribonucleic acid gene restriction patterns. *System. Appl. Microbiol.* 16:483-493.

Gupta, R., J. M. Lanter and C. R. Woese. (1983). Sequence of the 16S ribosomal RNA from *Halobacterium volcanii,* an archaebacterium. *Science.* 221:656-659.

Gutell, R. R., B. Weiser, C. R. Woese, and H. F. Noller. (1985). Comparative anatomy of 16-S-like ribosomal RNA. *Prog. Nucl. Acid Res. Mol. Biol.* 32:155-216.

Gutell, R. R. (1994). Collection of small subunit (16S- and 16S-like) ribosomal RNA structures: 1994. *Nucl. Acids Res.* 22(17):3502-3507.

Halbouty, M. T. (1967). *Salt Domes: Gulf Region, United States and Mexico.* Gulf Publishing Co., Houston, Texas.

Handford, C. R. (1991). Marginal marine halite: sabkhas and salinas. Ch.1:1-66. In *Evaporites, Petroleum and Mineral Resources*. Melvin, J. L. (ed.). Volume 50 in *Developments in Sedimentology*. Elsevier.

Hardie, L. A. and H. P. Eugster. (1970). The evolution of closed-basin brines. *Mineral. Soc. Amer. Spec. Pap.* 3:273-290.

Hartmann, R., H.-D. Sickinger and D. Oesterhelt. (1980). Anaerobic growth of halobacteria. *Proc. Natl. Acad. Sci. USA.* 77(6):3821-3825.

Harwood, G. M. and D. B. Smith (eds.). (1986). *The English Zechstein and Related Topics*. Geological Society Special Publication No.22. Blackwell Scientific Publications.

Hatch, F. H., R. H. Rastall and J. T. Greensmith. (1971). *Petrology of the Sedimentary Rocks*. 5th edition. George Allen and Unwin Ltd.

Hesselberg, M. and R. H. Vreeland. (1995). Utilization of protein profiles for the characterization of halophilic bacteria. *Curr. Microbiol.* 31:158-162.

Hillis, D. M., J. P. Huelsenbeck and D. L. Swofford. (1994). Hobgoblin of phylogenetics? *Nature*. 369:363-364.

Hite, R. J. and D. E. Anders. (1991). Petroleum and evaporites. Ch.4:349-411. In *Evaporites, Petroleum and Mineral Resources*. Melvin, J. L. (ed.). Volume 50 in *Developments in Sedimentology*. Elsevier.

Hofman, J. D., Lau, R. H. and Doolittle, W. F. (1979). The number, physical organization and transcription of ribosomal RNA cistrons in an archaebacterium: *Halobacterium halobium*. *Nucl. Acids Res.* 7:1321-1333.

Holland, T. H. (1911). The origin of desert salt deposits. *Proc. L'pool Geol. Soc.* Vol.XI:227-250.

Hori, H. and S. Osawa. (1979). Evolutionary change in 5S RNA secondary structure and a phylogenic tree of 54 5S RNA species. *Proc. Natl. Acad. Sci. USA.* 76(1):381-385.

Huber, R., S. Burggraf, T. Mayer, S. M. Barns, P. Rossnagel and K. O. Stetter. (1995). Isolation of a hyperthermophilic archaeum predicted by *in situ* RNA analysis. *Nature*. 376:57-58.

Hui, I. and P. P. Dennis. (1985). Characterization of the ribosomal RNA gene clusters in *Halobacterium cutirubrum*. *J. Biol. Chem.* 260(2):899-906.

Innis, M. A. and D. H. Gelfand. (1990). Optimization of PCRs. Ch.1:3-12. In *PCR Protocols. A Guide to Methods and Applications*. Innis, M. A., D. H. Gelfand, J. J. Sninsky and T. J. White. (eds.). Academic Press, Inc.

Innis, M. A. (1990). PCR with 7-deaza-2'-deoxyguanosine triphosphate. Ch.7:54-59. In *PCR Protocols. A Guide to Methods and Applications*. Innis, M. A., D. H. Gelfand, J. J. Sninsky and T. J. White. (eds.). Academic Press, Inc.

Iwabe, N., K. Kuma, M. Hasegawa, S. Osawa, and T. Miyata. (1989). Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. *Proc. Natl. Acad. Sci. USA.* 86:9355-9359.

Iwabe, N., K. Kuma, H. Kishino, M. Hasegawa, and T. Miyata. (1991). Evolution of RNA polymerases and branching patterns of the three major groups of archaebacteria. *J. Mol. Evol.* 32:70-78.

James, S. R., S. J. Dobson, P. D. Franzmann, and T. A. McMeekin. (1990). *Halomonas meridiana*, a new species of extremely halotolerant bacteria isolated from Antarctic saline lakes. *System. Appl. Microbiol.* 13:270-278.

Javor, B., C. Requadt and W. Stoeckenius. (1982). Box-shaped halophilic bacteria. *J. Bact.* 151(3):1532-1542.

Javor, B. (1989). *Hypersaline Environments: Microbiology and Biogeochemistry*. Springer-Verlag.

Johnson, J. (1991). The oldest DNA in the world. *New Scientist.* 11 May; 44-48.

Juez, G., F. Rodriguez-Valera, A. Ventosa and D. J. Kushner. (1986). *Haloarcula hispanica* spec. nov. and *Haloferax gibbonsii* spec. nov., two new species of extremely halophilic archaebacteria. *System. Appl. Microbiol.* 8:75-79.

Jukes, T. H. and C. R. Cantor. (1969). Evolution of protein molecules. Ch.24:21-132. In *Mammalian Protein Metabolism*. Vol.III. Munro, H. N. (ed.). Academic Press; New York and London.

Kamekura, M. and Y. Seno. (1992). Nucleotide sequences of 16S rRNA encoding genes from halophilic archaea *Halococcus morrhuae* NRC 16008 and *Haloferax mediterranei* ATCC 33500. *Nucl. Acid Res.* 20:3517.

Kamekura, M. and Y. Seno. (1993). Partial sequence of the gene for a serine protease from a halophilic archaeum *Haloferax mediterranei* R4, and nucleotide sequences of 16S rRNA encoding genes from several halophilic archaea. *Experientia.* 49:503-513.

Kamekura, M. and M. L. Dyall-Smith. (1995). Taxonomy of the family Halobacteriaceae and the description of two new genera *Halorubrobacterium* and *Natrialba. J. Gen. Appl. Microbiol.* 41:333-350.

Kanai, H., T. Kobayashi, R. Aono and T. Kudo. (1995). *Natronococcus amylolyticus* sp. nov., a haloalkaliphilic archaeon. *Int. J. Syst. Bact.* 45(4):762-766.

Kandler, O. and H. Konig. (1993). Cell envelopes of archaea: structure and chemistry. Ch. 8 223-259. In *The Biochemistry of the Archaea (Archaebacteria).* New Comprehensive Biochemistry. Volume 26. Kates, M., D. J. Kushner and A. T. Matheson (eds.). Gen. Eds. A. Neuberger and L. L. M. van Deenen. Elsevier.

Kandler, O. (1993). Archaea (archaebacteria). Review. *Progress in Botany.* 54:1-24.

Kaprelyants, A. S., J. C. Gottschal and D. B. Kell. (1993). Dormancy in non-sporulating bacteria. *FEMS Microbiol. Revs.* 104:271-286.

Kates, M. (1986). *Techniques of lipidology: isolation, analysis and identification of lipids.* Elsevier.

Kates, M. (1993). Membrane lipids of extreme halophiles: biosynthesis, function and evolutionary significance. In: Biology of halophilic bacteria, Part II (Multi-author Review). *Experientia.* 49:1027-1036.

Kendall, C. G. St. C. and J. K. Warren. (1988). Peritidal evaporites and their sedimentary assemblages. Ch.2:66-138. In *Evaporites and Hydrocarbons.* Schreiber B. C. (ed.). Columbia University Press.

Kennedy, M. J., S. L. Reader and L. M. Swierczynski. (1994). Preservation records of microorganisms: evidence of the tenacity of life. *Microbiology.* 140:2513-2529.

Kersters, K. and J. De Ley. (1975). Identification and grouping of bacteria by numerical analysis of their electrophoretic protein patterns. *J. Gen. Microbiol.* 87:333-342.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16:111-120.

Kimura, M. and T. Ohta. (1973). Eukaryotes-Prokaryotes divergence estimated by 5S ribosomal RNA sequences. *Nature New Biol.* 243:199-200.

Klenk, H-P. and W. Zillig. (1994). DNA-dependent RNA polymerase subunit B as a tool for phylogenetic reconstructions: branching topology of the archaeal domain. *J. Mol. Evol.* 38:420-432.

Kostrikina, N. A., I. S. Zvyagintseva and V. I. Duda. (1991). Cytological peculiarities of some extremely halophilic soil archaeobacteria. *Arch. Microbiol.* 156:344-349.

Kushner, D. J. (1978). Life in high salt and solute concentrations: halophilic bacteria. Ch.8:317-368. In *Microbial Life in Extreme Environments*. Kushner, D. J. (ed.). Academic Press.

Kushner, D. J. (1986). Molecular adaptation of enzymes, metabolic systems and transport systems in halophilic bacteria. *FEMS Microbiol. Revs.* 39:121-127.

Kushner, D. J. (1993). Growth and nutrition of halophilic bacteria. Ch.4:87-103. In *The Biology of Halophilic Bacteria.* Vreeland, R. H. and L. I. Hochstein (eds.). CRC Press.

Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature.* 227:680-685.

Lake, J. A. (1988). Origin of the eukaryotic nucleus determined by rate-invariant analysis of rRNA sequences. *Nature.* 331:184-186.

Lake, J. A. (1991). Tracing origins with molecular sequences: metazoan and eukaryotic beginnings. *TIBS.* 16:46-50.

Lanzotti, V., B. Nicolaus, A. Trincone, M. De Rosa, W. D. Grant and A. Gambacorta. (1989). A complex lipid with a cyclic phosphate from the archaebacterium *Natronococcus occultus. Biochim. et Biophys. Acta.* 1001:31-34.

Lappin-Scott, H. M. and J. W. Costerton. (1990). Starvation and penetration of bacteria in soils and rocks. *Experientia*. 46:807-812.

Larsen, N., G. J. Olsen, B. L. Maidak, M. J. McCaughey, R. Overbeek, T. J. Macke, T. L. Marsh and C. R. Woese. (1993). The ribosomal database project. *Nucl. Acids Res*. 21(13):3021-3023.

Lazar, B. and H. D. Holland. (1988). The analysis of fluid inclusions in halite. *Geochim. et Cosmochim. Acta*. 52:485-490.

Leffers, H. and R. A. Garrett. (1984). The nucleotide sequence of the 16S ribosomal RNA gene of the archaebacterium *Halococcus morrhua*. *The EMBO Journal*. 3(7):1613-1619.

Lefond, S. J. (1969). *Handbook of World Salt Resources*. Plenum Press, New York.

L'Haridon, S., A.-L. Reysenbach, P. Glénat, D. Prieur and C. Jeanthon. (1995). Hot subterranean biosphere in a continental oil reservoir. *Nature*. 377:223-224.

Liesack, W., M. A. O. H. Menke and E. Stackebrandt. (1990). Rapid generation of vector-free digoxygenin-dUTP labeled probes for nonradioactive hybridization using the polymerase chain reaction (PCR) method. *System. Appl. Microbiol*. 13:255-256.

Liesack, W., R. Söller, T. Stewart, H. Haas, S. Giovannoni, and E. Stackebrandt. (1992). The influence of tachytelically (rapidly) evolving sequences on the topology of phylogenetic trees - intrafamily relationships and the phylogenetic position of *Planctomycetaceae* as revealed by comparative analysis of 16S ribosomal RNA sequences. *System. Appl. Microbiol*. 15:357-362.

Lindahl, T. (1993a). Recovery of antediluvian DNA. *Nature*. 365:700.

Lindahl, T. (1993b). Instability and decay of the primary structure of DNA. *Nature*. 362:709-715.

Linkkila, T. P., and J. P. Gogarten. (1991). Tracing origins with molecular sequences: rooting the universal tree of life. *TIBS* 16:287-288.

Lipman, C. B. (1931). Living microörganisms in ancient rocks. *J. Bact*. 22(3):183-198.

Lodwick, D., H. N. M. Ross, J. E. Harris, J. W. Almond and W. D. Grant. (1986). *dam* methylation in the archaebacteria. *J. Gen. Microbiol.* 132:3055-3059.

Lodwick, D., H. N. M. Ross, J. A. Walker, J. W. Almond and W. D. Grant. (1991). Nucleotide sequence of the 16S ribosomal RNA gene from the haloalkaliphilic archaeon (Archaebacterium) *Natronobacterium magadii*, and the phylogeny of halobacteria. *System. Appl. Microbiol.* 14:352-357.

Lodwick, D., T. J. McGenity and W. D. Grant. (1994). The phylogenetic position of the haloalkaliphilic archaeon *Natronobacterium magadii*, determined from its 23S ribosomal RNA sequence. *System. Appl. Microbiol.* 17: 402-404.

Longley-Cook, W. A. (1989). Recent developments at Winsford salt mine, Cheshire, England. Technical note. *Trans. Instn. Min. Metall.(SectionA:Min. industry).* 98:A63-A67.

Lopez-Cortes, A., J.-L. Ochoa and R. Vazquez-Duhalt. (1994). Participation of halobacteria in crystal formation and the crystallization rate of NaCl. *Geomicrobiol. J.* 12:69-80.

Ludwig, W. and K. H. Schleifer. (1994). Bacterial phylogeny based on 16S and 23S rRNA sequence analysis. *FEMS Microbiology Reviews.* 15: 155-173.

Maniatis, T., E. F. Fritsch and J. Sambrook. (1982). *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbour Laboratory.

Mankin, A. S., V. K. Kagraminova, N. L. Teterina, P. M. Rubtsov, E. N. Belova, A. M. Kopylov, L. A. Baratova and A. A. Bogdanov. (1985). The nucleotide sequence of the gene coding for the 16S rRNA from the archaebacterium *Halobacterium halobium. Gene.* 37:181-189.

Margoliash, E. and Smith, E. L. (1965). Structural and functional aspects of cytochrome *c* in relation to evolution. pp. 221-242. In *Evolving Genes and Proteins.* Bryson, V. and H. J. Vogel (eds.). Academic Press, New York.

Marquez, M. C., A. Ventosa and F. Ruiz-Berraquero. (1987). A taxonomic study of heterotrophic halophilic and non-halophilic bacteria from a solar saltern. *J. Gen. Microbiol.* 133:45-56.

Martinez, J. D. (1991). Salt Domes. *American Scientist.* 79:420-431.

Martínez-Murcia, A. J., N. M. Harland and M. D. Collins. (1993). Phylogenetic analysis of some leuconostocs and related organisms as determined from large-subunit rRNA gene sequences: assessment of congruence of small- and large-subunit rRNA derived trees. *J. Appl. Bact.* 74:532-541.

Martínez-Murcia, A. J., S. G. Acinas and F. Rodriguez-Valera. (1995). Evaluation of prokaryotic diversity by restrictase digestion of 16S rDNA directly amplified from hypersaline environments. *FEMS Microbiol. Ecol.* 17:247-256.

Matheson, A. T., M. Yaguchi, R. N. Nazar, L. P. Visentin and G. E. Willick. (1978). The structure of ribosomes from moderate and extreme halophilic bacteria. pp. 481-501. In *Energetics and Structure of Halophilic Microorganisms.* Proceedings of an EMBO workshop, May 1978. The Weizmann Institute of Science, Rehovot. Caplan, S. R. and M. Ginzburg (eds.). Elsevier/North-Holland, Amsterdam, New York; Biomedical Press.

Matheson, A. T. (1992). Structure, function and evolution of the archaeal ribosome. pp.89-98. In *The Archaebacteria: Biochemistry and Biotechnology.* Biochemical Society Symposium 58. Danson, M. J., D. W. Hough and G. G. Lunt (eds.). Portland Press, London and Chapel Hill.

McCaffrey, M. A., B. Lazar and H. D. Holland. (1987). The evaporation path of seawater and the coprecipitation of Br^- and K^+ with halite. *J. Sed. Petrol.*57(5):928-937.

McGenity, T. J. (1994). *Halobacterial Phylogeny and Salt Mine Microbial Ecology.* PhD Thesis. University of Leicester.

McGenity, T. J. and W. D. Grant. (1995). Transfer of *Halobacterium saccharovorum, Halobacterium sodomense, Halobacterium trapanicum* NRC 34021 and *Halobacterium lacusprofundi* to the Genus *Halorubrum* gen. nov., as *Halorubrum saccharovorum* comb. nov., *Halorubrum trapanicum* comb. nov., and *Halorubrum lacusprofundi* comb. nov. *Syst. Appl. Microbiol.* 18(2):237-243.

Mindell, D. P. (1991). Aligning DNA sequences: homology and phylogenetic weighting. Ch. 5:73-89. In *Phylogenetic Analysis of DNA Sequences.* Miyamoto, M. M. and J. Craycraft (eds.). Oxford University Press, Inc.

Montero, C. G., A. Ventosa, F. Rodriguez-Valera, M. Kates, N. Moldoveanu and F. Ruiz-Berraquero. (1989). *Halococcus saccharolyticus* sp. nov., a new species of extremely halophilic non-alkaliphilic cocci. *Syst. Appl. Microbiol.* 12:167-171.

Moran, N. A., M. A. Munson, P. Baumann, and H. Ishikawa. (1993). A molecular clock in endosymbiotic bacteria is calibrated using the insect hosts. *Proc. R. Soc. Lond. B.* 253:167-171.

Morita, R. Y. (1988). Bioavailability of energy and its relationship to growth and starvation survival in nature. *Can. J. Microbiol.* 34:436-441.

Mullakhanbhai, M. F. and H. Larsen. (1975). *Halobacterium volcanii* spec. nov., a Dead Sea halobacterium with a moderate salt requirement. *Arch. Microbiol.* 104:207-214.

Mullis, K. B. (1990). The unusual origin of the polymerase chain reaction. *Scientific American.* April, 1990:36-43.

Murray, V. (1989). Improved double-stranded DNA sequencing using the linear polymerase chain reaction. *Nuc. Acids Res.* 17(21):8889.

Mwatha, W. E. and W. D. Grant. (1993). *Natronobacterium vacuolata* sp. nov., a haloalkaliphilic archaeon isolated from Lake Magadi, Kenya. *Int. J. Syst. Bact.* 43(3):401-404.

Mylvaganam, S. and P. P. Dennis. (1992). Sequence heterogeneity between the two genes encoding 16S rRNA from the halophilic archaebacterium *Haloarcula marismortui. Genetics.* 130:399-410.

Nakamura, S., R. Aono, S. Mizutani, T. Takashina, W. D. Grant and K. Horikoshi. (1992). The cell surface glycoprotein of *Haloarcula japonica* TR-1. *Biosci. Biotech. Biochem.* 56(6):996-998.

Namyslowski, M. B. (1913). Über unbekannte halophile mikroorganismen aus dem innern des salzbergwerkes Wieliczka. *Bulletin of the Academy of Sciences, Krakow.* Vol.B. No.3/4:88-104.

Nehrkorn, A. and W. Schwartz. (1961). Untersuchungen über lebensgemeinschaften halophiler mikroorganismen. I Mikroorganismen aus Salzseen der Californischen

Wüstengebiete und aus einer Natriumchlorid-sole. *Zeitschrift für Allg. Mikrobiologie.* 1(2) 121-141.

Nei, M. (1991). Relative efficiencies of different tree-making methods for molecular data. Ch.6:90-128. In *Phylogenetic Analysis of DNA Sequences.* Miyamoto, M. M. and J. Craycraft (eds.). Oxford University Press, Inc.

Newton, C. R. and A. Graham. (1994). *PCR.* Introduction to Biotechniques series. BIOS Scientific Publishers Ltd.

Nicolaus, B., V. Lanzotti, A. Trincone, M. De Rosa, W. D. Grant and A. Gambacorta. (1989). Glycine betaine and polar lipid composition in halophilic archaebacteria in response to growth in different salt concentrations. *FEMS Microbiol. Letts.* 59:157-160.

Norton, C. F. and W. D. Grant. (1988). Survival of halobacteria within fluid inclusions in salt crystals. *J. Gen. Microbiol.* 134:1365-1373.

Norton, C. F., T. J. McGenity and W. D. Grant. (1993). Archaeal halophiles (halobacteria) from two British salt mines. *J. Gen. Microbiol.* 139:1077-1081.

Nuttall, S. D. and M. L. Dyall-Smith. (1993). Ch2, a novel halophilic archaeon from an Australian solar saltern. *Int. J. Syst. Bact.* 43(4):729-734.

Ochman, H. and A. C. Wilson. (1987). Evolution in bacteria: evidence for a universal substitution rate in cellular genomes. *J. Mol. Evol.* 26:74-86.

Oliver, J. D. (1993). Formation of viable but nonculturable cells. Ch. 11:239-272. In *Starvation in Bacteria.* Kjelleberg, S. (Ed.). Plenum Press.

Ollivier, B., P. Caumette, J.-L. Garcia and R. A. Mah. (1994). Anaerobic bacteria from hypersaline environments. *Microbiol. Revs.* 58(1):27-38.

Olsen, G. J., D. J. Lane, S. J. Giovannoni, and N. R. Pace. (1986). Microbial ecology and evolution: a ribosomal RNA approach. *Ann. Rev. Microbiol.* 40:337-365.

Oren, A. (1983). *Halobacterium sodomense* sp. nov., a Dead Sea halobacterium with an extremely high magnesium requirement. *Int. J. Syst. Bact.* 33(2):381-386.

Oren, A., B. J. Paster and C. R. Woese. (1984). *Haloanaerobiaceae*: a new family of moderately halophilic obligatory anaerobic bacteria. *Syst. Appl. Microbiol.* 5:71-80.

Oren, A. (1990). Starch counteracts the inhibitory action of Bacto-peptone and bile salts in media for the growth of halobacteria. *Can. J. Microbiol.* 36: 299-301.

Oren, A., M. Ginzburg, B. Z. Ginzburg, L. I. Hochstein and B. E. Volcani. (1990). *Haloarcula marismortui* (Volcanii) sp. nov., nom. rev., an extremely halophilic bacterium from the Dead Sea. *Int. J. Syst. Bact.* 40(2):209-210.

Oren, A. (1993). Availability, uptake and turnover of glycerol in hypersaline environments. *FEMS Microbiol. Ecol.* 12:15-23.

Oren, A. (1994a). The ecology of the extremely halophilic archaea. *FEMS Microbiol. Revs.* 13:415-440.

Oren, A. (1994b). Characterization of the halophilic archaeal community in saltern crystallizer ponds by means of polar lipid analysis. *Int. J. Salt Lake Res.* 3:15-29.

Oren, A., P. Gurevich, R. T. Gemmell and A. Teske. (1995). *Halobaculum gomorrense* gen. nov., sp. nov., a novel extremely halophilic archaeon from the Dead Sea. *Int. J. Syst. Bact.* 45(4):747-754.

Owen, R. J., M. Desai and S. Garcia. (1993). Molecular typing of thermotolerant species of *Campylobacter* with ribosomal RNA gene patterns. *Res. Microbiol.* 144:709-720.

Parkes, K. and A. E. Walsby. (1981). Ultrastructure of a gas-vacuolate square bacterium. *J. Gen. Microbiol.* 126:503-506.

Payne, J. I., S. N. Sehgal and N. E. Gibbons. (1960). Immersion refractometry of some halophilic bacteria. *Can. J. Microbiol.* 6:9-15.

Pedersen, K. and S. Ekendahl. (1990). Distribution and activity of bacteria in deep granitic groundwaters of southeastern Sweden. *Microbial Ecology.* 20:37-52.

Penny, D., M. D. Hendy and M. A. Steel. (1992). Progress with methods for constructing evolutionary trees. *TREE* 7(3)73-79.

Pitcher, D. G., N. A. Saunders and R. J. Owen. (1989). Rapid extraction of bacterial genomic DNA with guanidium thiocyanate. *Letts. Appl. Microbiol.* 8:151-156.

Pliny. (*c.* 77 A.D.). *The Natural History of Pliny*. Vol. V; Book 31; Chapter 41. Translated by J. Bostock and H. T. Riley. Pub. Henry G. Bohn, London. (1856).

Poinar, G. and R. Poinar. (1994). *The Quest for Life in Amber*. Addison-Wesley.
Potts, M. (1994). Desiccation tolerance of prokaryotes. *Microbiol. Revs.* 58(4):755-805.

Prado. B., A. Del Moral, E. Quesada, R. Ríos, M. Monteoliva-Sanchez, V. Campos and A. Ramos-Cormenzana. (1991). Numerical taxonomy of moderately halophilic Gram-negative rods isolated from the Salar de Atacama, Chile. *Syst. Appl. Microbiol.* 14:275-281.

Priest, F. G., D. A. Kaji, Y. B. Rosato and V. P. Canhos. (1994). Characterization of *Bacillus thuringiensis* and related bacteria by ribosomal RNA gene restriction fragment length polymorphisms. *Microbiology*. 140:1015-1022.

Quesada, E., A. Ventosa, F. Rodriguez-Valera and A. Ramos-Cormenzana. (1982). Types and properties of some bacteria isolated from hypersaline soils. *J. Appl. Bact.* 53:155-161.

Ramos-Cormenzana, A. (1993). Ecology of moderately halophilic bacteria. Ch.3:55-86. In *The Biology of Halophilic Bacteria*. Vreeland, R. H. and L. I. Hochstein (eds.). CRC Press, Inc.

Ree, H. K., N.Larsen., R. R. Gutell, and R. A. Zimmermann. (1993). The primary and secondary structures of the 23S ribosomal RNA from *Thermoplasma acidophilum* define an ancient archaeal divergence. *System. Appl. Microbiol.* 16:333-341.

Reiser, R. and P. Tasch. (1960). Investigation of the viability of osmophile bacteria of great geological age. *Trans. Kans. Acad. Sci.* 63(1):31-34.

Rodriguez-Valera, F., F. Ruiz-Berraquero and A. Ramos-Cormenzana. (1979). Isolation of extreme halophiles from seawater. *Appl. Env. Microbiol.* 38:164-165.

Rodriguez-Valera, F., F. Ruiz-Berraquero and A. Ramos-Cormenzana. (1980). Isolation of extremely halophilic bacteria able to grow in defined inorganic media with single carbon sources. *J. Gen. Microbiol.* 119:535-538.

Rodriguez-Valera, F., F. Ruiz-Berraquero and A. Ramos-Cormenzana. (1981). Characteristics of the heterotrophic bacterial populations in hypersaline environments of different salt concentrations. *Microb. Ecol.* 7:235-243.

Rodriguez-Valera, F., A. Ventosa, E. Quesada and F. Ruiz-Berraquero. (1982). Some physiological features of a *Halococcus* sp. at low salt concentrations. *FEMS Microbiol. Letts.* 15:249-252.

Rodriguez-Valera, F., G. Juez and D. J. Kushner. (1983). *Halobacterium mediterranei* spec. nov., a new carbohydrate-utilizing extreme halophile. *System. Appl. Microbiol.* 4:369-381.

Rodriguez-Valera, F., A. Ventosa, G. Juez and J. F. Imhoff. (1985). Variation of environmental features and microbial populations with salt concentrations in a multi-pond saltern. *Microb. Ecol.* 11:107-115.

Roedder, E. (1984). The fluids in salt. *American Mineralogist.* 69:413-439.

Ross, H. N. M., M. D. Collins, B. J. Tindall and W. D. Grant. (1981). A rapid procedure for the detection of archaebacterial lipids in halophilic bacteria. *J. Gen. Microbiol.* 123:75-80.

Ross, H. N. M. and W. D. Grant. (1985). Nucleic acid studies on halophilic archaebacteria. *J. Gen. Microbiol.* 131:165-173.

Ross, H. N. M., W. D. Grant and J. E. Harris. (1985). Lipids in archaebacterial taxonomy. pp.289-300. In *Chemical Methods in Bacterial Systematics*. Goodfellow, M. and D. E. Minnikin (Eds.). Academic Press.

Roszak, D. B. and R. R. Colwell. (1987). Survival strategies of bacteria in the natural environment. *Microbiol. Revs.* 51(3):365-379.

Rouvière, P., L. Mandelco, S. Winker, and C. R. Woese. (1992). A detailed phylogeny for the *Methanomicrobiales*. *System. Appl. Microbiol.* 15:363-371.

Rowlands, T., P. Baumann and S. P. Jackson. (1994). The TATA -binding protein: a general transcription factor in eukaryotes and archaebacteria. *Science.* 264:1326-1329.

Saitou, N. and M. Nei. (1987). The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4(4):406-425.

Sanger, F., S. Nicklen, and A. R. Coulson. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74(12):5463-5467.

Sanz, J. L., I. Marín, L. Ramirez, J. P. Abad, C. L. Smith and R. Amils. (1988). Variable rRNA copies in extreme halobacteria. *Nucl. Acids Res.* 16:7827-7832.

Schreiber, B. C. (1988). Subaqueous evaporite deposition. Ch.4:182-255. In *Evaporites and Hydrocarbons*. Schreiber B. C. (ed.). Columbia University Press.

Severin, J., A. Wohlfarth and E. A. Galinski. (1992). The predominant role of recently discovered tetrahydropyrimidines for the osmoadaptation of halophilic eubacteria. *J. Gen. Microbiol.* 138:1629-1638.

Shlichta, P. J. (1968). Growth, deformation, and defect structure of salt crystals. In *Saline Deposits* A symposium based on papers from the International Conference on Saline Deposits, Houston, Texas, 1962. Special GSA (The Geological Society of America Inc.) Papers. No.88. *The Geol. Soc. of Amer., Inc.* Mattox, R. B. (ed.).

Sidow, A. and A. C. Wilson. (1991). Compositional statistics evaluated by computer simulations. Ch. 7. In *Phylogenetic Analysis of DNA Sequences*. Miyamoto, M. M. and J. Craycraft (eds.). Oxford University Press, Inc.

Smith, A. B. (1989). RNA sequence data in phylogenetic reconstruction: testing the limits of its resolution. *Cladistics*. 5:321-344.

Smith, D. B. (1973). The origin of the Permian Middle and Upper Potash deposits of Yorkshire: an alternative hypothesis. *Proc. Yorks. Geol. Soc.* 39(3;16):327-346.

Smith, M. W., Feng, D-F., and R. Doolittle. (1992). Evolution by acquisition: the case for horizontal gene transfers. *TIBS* 17:489-493.

Smoot, J. P. and T. K. Lowenstein. (1991). Depositional environments of non-marine evaporites. Ch.3:189-347. In *Evaporites, Petroleum and Mineral Resources*. Melvin, J. L. (ed.). Volume 50 in *Developments in Sedimentology*. Elsevier.

Smyk, B. (1992). Dlaczego Wieliczce grozi zaglada? *Problemy*. Nr. 2:11-15.

Sneath, P. H. A. (1962). Longevity of microorganisms. *Nature*. 195(4842):643-646.

Sogin, M. L. (1990). Amplification of ribosomal RNA genes for molecular evolution studies. Ch.37:307-314. In *PCR Protocols. A Guide to Methods and Applications*. Innis, M. A., D. H. Gelfand, J. J. Sninsky and T. J. White. (eds.). Academic Press, Inc.

Sonnenfeld, P. (1984). *Brines and Evaporites*. Academic Press, Inc.

Stackebrandt, E. and B. M. Goebel. (1994). Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int. J. Syst. Bact.* 44(4):846-849.

Stan-Lotter, H., M. Sulzner, E. Egelseer, C. F. Norton and L. I. Hochstein. (1993). Comparison of membrane ATPases from extreme halophiles isolated from ancient salt deposits. *Origins of Life and Evolution of the Biosphere*. 23:53-64.

Stein, C. L. and J. L. Krumhansl. (1988). A model for the evolution of brines in salt from the lower Salado Formation, southeastern New Mexico. *Geochim. et Cosmochim. Acta*. 52:1037-1046.

Stewart, C.-B. (1993). The powers and pitfalls of parsimony. *Nature*. 361:603-607.

Strong, M. W. (1956). Marine iron bacteria as rock forming organisms. *Advanced Science*. 12(49):583-585.

Suggs, S. V., T. Hirose, T. Miyake, E. H. Kawashima, M. J. Johnson, K. Itakura and R. B. Wallace. (1981). Use of synthetic oligodeoxyribonucleotides for the isolation of specific cloned DNA sequences. ICN-UCLA Symp. Mol. Cell Biol. 23: 683-693.

Suwanich, P. (*c*.1992). Potash-rock salt, NE of Thailand. Historical geology and structural evolution of Maha Sarakham evaporites and their reserves. Economic Geology Division, Department of Mineral Resources, Ministry of Industry, Thailand.

Swofford, D. L. and G. J. Olsen. (1990). Phylogeny reconstruction. Ch.11:411-501. In *Molecular Systematics*. Hillis, D. M. and C. Moritz (eds.). Sinauer Associates Inc., Sunderland, MA.

Takashina, T., T. Hamamoto, K. Otozai, W. D. Grant and K. Horikoshi. (1990). *Haloarcula japonica* sp. nov., a new triangular halophilic archaebacterium. *System. Appl. Microbiol.* 13:177-181.

Takashina, T., K. Otozati, T. Hamamoto and K. Horikoshi. (1994). Isolation of halophilic and halotolerant bacteria from a Japanese salt field and comparison of the partial 16S rRNA gene sequence of an extremely halophilic isolate with those of other extreme halophiles. *Biodiversity and Conservation.* 3:632-642.

Tew, R. W. (1980). Halotolerant *Ectothiorhodospira* survival in mirabilite: experiments with a model of chemical stratification by hydrate deposition in saline lakes. *Geomicrobiol. J.* 2(1):13-20.

Thongthai, C., T. J. McGenity, P. Suntinanalert and W. D. Grant. (1992). Isolation and characterization of an extremely halophilic archaeobacterium from traditionally fermented Thai fish sauce (nam pla). *Letts. Appl. Microbiol.* 14:111-114.

Thorpe, J. P. (1982). The molecular clock hypothesis: biochemical evolution, genetic differentiation and systematics. *Ann. Rev. Ecol. Syst.* 13:139-168.

Tindall, B. J., A. A. Mills and W. D. Grant. (1980). An alkaliphilic red halophilic bacterium with a low magnesium requirement from a Kenyan soda lake. *J. Gen. Microbiol.* 116:257-260.

Tindall, B. J., H. N. M. Ross and W. D. Grant. (1984). *Natronobacterium* gen. nov. and *Natronococcus* gen. nov., two new genera of haloalkaliphilic archaebacteria. *System. Appl. Microbiol.* 5:41-57.

Tindall, B. J. (1985). Qualitative and quantitative distribution of diether lipids in haloalkaliphilic archaebacteria. *System. Appl. Microbiol.* 6:243-246.

Tindall, B. J. (1990). A comparative study of the lipid composition of *Halobacterium saccharovorum* from various sources. *System Appl. Microbiol.* 13:128-130.

Tindall, B. J. (1992). The family Halobacteriaceae. Ch. 34:768-808. In *The Prokaryotes.* 2nd Edn. Vol.1. Springer-Verlag.

Tomlinson, G. A. and L. I. Hochstein. (1976). *Halobacterium saccharovorum* sp. nov., a carbohydrate-metabolizing, extremely halophilic bacterium. *Can. J. Microbiol.* 22:587-591.

Tomlinson, G. A., L. L. Jahnke and L. I. Hochstein. (1986). *Halobacterium denitrificans* sp. nov., an extremely halophilic denitrifying bacterium. *Int. J. Syst. Bact.* 36(1):66-70.

Torreblanca, M., F. Rodriguez-Valera, G. Juez, A. Ventosa, M. Kamekura and M. Kates. (1986). Classification of non-alkaliphilic halobacteria based on numerical taxonomy and polar lipid composition, and description of *Haloarcula* gen. nov. and *Haloferax* gen. nov. *System. Appl. Microbiol.* 8:89-99.

Torrella, F. and R. Y. Morita. (1981). Microcultural study of bacterial size changes and microcolony and ultramicrocolony formation by heterotrophic bacteria in seawater. *Appl. Environ. Microbiol.* 41(2):518-527.

Tucker, R. M. and M. E. Tucker. (1981). Evidence of synsedimentary tectonic movements in the Triassic halite of Cheshire. *Nature.* 290:495-496.

Tucker, M. E. (1991). Evaporites. Ch.5:167-181. In *Sedimentary Petrology.* 2nd edition. Blackwell Scientific Publications.

Upasani, V. N., S. G. Desai, N. Moldoveanu and M. Kates. (1994). Lipids of extremely halophilic archaeobacteria from saline environments in India: a novel glycolipid in *Natronobacterium* strains. *Microbiol.* 140:1959-1966.

Vauterin, L., J. Swings and K. Kersters. (1993). Protein electrophoresis and classification. Ch.6:251-280. In *Handbook of New Bacterial Systematics.* Academic Press Ltd.

Venkataraman, R. and A. Sreenivasan. (1954). Studies on the red halophilic bacteria from salted fish and salt. *Proc. Indian Acad. Sci.* Section B; 39(1):17-23.

Ventosa, A., E. Quesada, F. Rodriguez-Valera, F. Ruiz-Berraquero and A. Ramos-Cormenzana. (1982). Numerical taxonomy of moderately halophilic Gram-negative rods. *J. Gen. Microbiol.* 128:1959-1968.

Ventosa, A., F. Rodriguez-Valera, J. S. Poindexter and W. S. Reznikoff. (1984). Selection for moderately halophilic bacteria by gradual salinity increases. *Can. J. Microbiol.* 30:1279-1282.

Verissimo, A., P. V. Morais, A. Diogo, C. Gomes and M. S. DaCosta. (1996). Characterization of *Legionella* species by numerical analysis of whole-cell protein electrophoresis. *Int. J. Syst. Bact.* 46(1):41-49.

Wais, A. C. (1985). Cellular morphogenesis in a halophilic archaebacterium. *Curr. Microbiol.* 12:191-196.

Walsby, A. E. (1980). A square bacterium. *Nature* (London). 283:69-71.

Weast, R. C. (Ed. in Chief). *Handbook of Chemistry and Physics*. 45th Edition. 1964-1965. The Chemical Rubber Co., Cleveland, Ohio.

Wheeler, D. A. (1985). An analysis of the aeolian dustfall on eastern Britain, November 1984. *Proc. Yorks. Geol. Soc.*45(4):307-309.

Wheeler, W. C. and R. L. Honeycutt. (1988). Paired sequence difference in ribosomal RNAs: evolutionary and phylogenetic implications. *Mol. Biol. Evol.* 5(1):90-96.

Winker, S. and Woese, C. R. (1991). A definition of the domains *Archaea, Bacteria* and *Eucarya* in terms of small subunit ribosomal RNA characteristics. *System. Appl. Microbiol.* 14:305-310.

Woese, C. R., and G. E. Fox.(1977).Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. USA.*74(11):5088-5090.

Woese, C. R., L. J. Magrum, and G. E. Fox. (1978). Archaebacteria. *J. Mol. Evol.* 11:245-252.

Woese, C. R. (1981). Archaebacteria. *Scientific American.* June. 94-106.

Woese, C. R., R. Gutell, R. Gupta, and H. F. Noller. (1983). Detailed analysis of the higher-order structure of 16S-like ribosomal ribonucleic acids. *Microbiology Reviews.* 47(4):621-669.

Woese, C. R. (1987). Bacterial evolution. *Microbiological Reviews*. 51(2):221-271.

Woese, C. R., O. Kandler, and M. L. Wheelis. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. USA*. 87:4576-4579.

Woese, C. R., L. Achenbach, P. Rouvière, and L. Mandelco. (1991). Archaeal phylogeny: reexamination of the phylogenetic position of *Archaeoglobus fulgidus* in light of certain composition-induced artifacts. *System. Appl. Microbiol.* 14:364-371.

Woods, P. J. E. and P. J. Powell. (1979). The geology of Boulby mine. *Economic Geology*. 74:409-418.

Yang, D., B. P. Kaine, and C. R. Woese. (1985). The phylogeny of archaebacteria. *System. Appl. Microbiol.* 6:251-256.

Young, S. The dry life. (1985). *New Scientist*. 31 October; 40-44.

Zuckerkandl, E. and L. Pauling. (1965a). Evolutionary divergence and convergence in proteins. pp. 97-166. In *Evolving Genes and Proteins*. Bryson, V. and H. J. Vogel (eds.). Academic Press, New York.

Zuckerkandl, E. and L. Pauling. (1965b). Molecules as documents of evolutionary history. *J. Theoret. Biol.* 8:357-366.

Zvyagintseva, I. S. and A. L. Tarasov. (1987). Extreme halophilic bacteria from saline soils. *Mikrobiologiya*. 56(5):839-844. (English translation, *Microbiology*; 1988).