Molecular studies of two strains of marine magnetotactic bacteria

by

Jose N. E. Neto

A thesis submitted to the graduate faculty

in partial fulfillment of the requirement for the degree of

MASTER OF SCIENCE

Major: Microbiology

Program of study committee:
Dennis A. Bazylinski, Major Professor
Gwyn A. Beattie
Louisa B. Tabatabai

Iowa State University

Ames, Iowa

2004

Graduate College
Iowa State University

This is to certify that the master's thesis of

Jose N. E. Neto

has met the thesis requirement of Iowa State University

Signatures have been redacted for privacy

This thesis is dedicated to those people who ask questions just for the satisfaction of learning the answers.

"You shall know the truth and the truth shall make you free"

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Many thanks to my major professor, Dr. Dennis A. Bazylinski, for all the time that he has spent teaching me laboratory techniques, supporting me psychologically, correcting my writings and making sure that I had the supplies and equipments required for my experiments. I consider myself very luck for being under his supervision during my master studies.

Thanks to the other professors in my post graduate committee. Thanks to Dr. Gwyn A. Beattie who has helped me since the beginning of my master studies with scientific advice. Thanks to Dr. Louisa Tabatabai who has introduced me to the 2-dimensional SDS-PAGE technique.

Special thanks to Dr. Aparecido Divino da Cruz for supporting my scientific career since I was an undergraduate student. Special thanks to Dr. Vasco Martins Cardoso, whose teachings I can't describe in words but I can say that he shaped much of what I am. Thanks to Dr. Gregory J. Phillips and Dr. Alan D. Dispirito for helpful discussions and scientific advice.

Thanks to my family for supporting my carrier. Special thanks to my parents, Dr. Claudio Netto Estrella and Dr. Maria Auxiliadora da Costa Netto Estrella for supporting me regardless of the situation and need. Thanks to my sister, Renata da Costa Netto Estrella, for psychological support.

Many thanks to my girlfriend, Dayde Lane Mendonça da Silva for psychological support and for help with most of the figures present in this work. Special thanks to Dayde for giving my life a new and wonderful perspective. Many thanks to her for teaching me

more about myself, for having patience with me and with our situation, for cheering me up when I was depressed, for sharing her moments, her thoughts and her life with me. Many thanks to her, most of all for her willingness to make very difficult decisions to fit me in her life.

Thanks to Bradley Dubbels, Karen Paul, Catherine Axtell, Amarjyoti Sandhu, Ginger Shipp and Timothy Williams, former graduate students, graduate students or post doctorates that helped me with laboratory techniques, scientific advice and psychological support. Thanks to Eric Matson for scientific advice regarding RNA techniques, Southern blots and SQUID equipment. Thanks to Mr. Robert Huber and Mrs. Kay L. Christiansen for support regarding laboratory teachings.

Many thanks to my friends who supported me in a way or another. Special thanks to Ioannis Rousochatzakis for helpful discussions about physics, for psychological support and for his presence as a friend during many different occasions. Special thanks to Georgi Batinov who introduced me to board games making my life in Ames more bearable and for his presence as a friend during many different occasions. Thanks to my room mate during part of my stay in Ames, Wasin Charerntantanakul for his patience with me and discussions about cultural differences among other topics.

Thanks to the technicians and doctorates at the DNA and the Protein facilities for invaluable work regarding DNA sequencing and N-terminal sequencing among other services supplied that made the study present here possible.

ABSTRACT

Different molecular techniques were applied to further characterize the magnetite magnetosome formation in two strains of magnetotactic bacteria. MamC, present in other magnetotactic bacteria as a major membrane protein, was identified as a magnetosome membrane protein in the marine, magnetotactic coccus strain MC-1. Cell protein profiles of strain MV-1, a marine magnetotactic vibrio, and a spontaneous non-magnetotactic mutant derived from it (that does not produce magnetosomes), strain MV-1nm1, showed that wild-type MV-1 produce a number of proteins not produced by MV-1nm1. Previous results have indicated the MV-1nm1 is a deletion mutant and a genomic subtraction library was used to confirm a deletion in MV-1nm1. Some putative genes absent in MV-1nm1 were identified. cDNA subtractive hybridization library of MV-1 and MV-1nm1 identified a "cytochrome a_1-like" hemoprotein cytochrome *c* oxidase previously implicated in magnetosome formation. Random amplification of polymorphic DNA experiments showed that although the formation of spontaneous non-magnetotactic MV-1 occurs relatively frequently, the mutants are not identical genetically. Preliminary electroporation experiments were unsuccessful but a buffer system was developed in which cells of MV-1 cells were able to survive the procedure. Finally, an experiment attempting to synthesize magnetite in vitro using cell-free extracts was attempted and resulted in formation not of magnetite but presumably of iron sulfide.

1. INTRODUCTION: THESIS ORGANIZATION

This thesis is organized in six sections. The first section is an up-to-date literature review. This section starts with a brief history of magnetism and continues with an in-depth description of the magnetotactic bacteria and magnetotaxis. The second section is a statement of the problem addressed in this thesis. The third and fourth sections are the materials and methods and the results, respectively. The results from the experiments described in this thesis are discussed in the fifth section. The papers used as references in this work are listed in the sixth section. An appendix is included at the end consisting of DNA sequences retrieved during the studies described here.

2. LITERATURE REVIEW

2.1 Magnetism and life: historical perspective

The first reported description of magnetism occurred about 3000 years ago in the Greek province Magnesia. It was observed that some mineral rocks had the capacity to attract other certain types of rocks without losing the property over time (1).

Plato (427-347 BC), Aristotle (384-322 BC), Lucrecio (98-55 BC) and Paracelso (1546-1642) all described magnetism in some way and its supposed "miraculous properties" towards the improvement of health. The first truly scientific account of magnetism occurred in 1600 when William Gilbert (1544-1603), an Englishman, described Earth as a "gigantic magnet".

Research and understanding of the scientific basis of magnetism increased greatly in the 18^{th} and 19^{th} centuries. Hans Christian Oersted (1777-1851) showed that an electrical current generated a magnetic field. Carl Friedrich Gauss (1777-1855) made the first measurement of the Earth's magnetic field. James Clerk Maxwell (1831-1879), with the help of previous work of Michael Faraday (1791-1867), was able to describe the relationship between electricity and magnetism.

Even today magnetism remains a mystifying force to some and is believed to have healing properties beyond scientific explanation. For instance, a United States based company (Oriental Medical Supplies, MA) has commercialized a number of magnetized personal accessories such as bracelets, insoles, belts and the like, suggesting that such items will improve the overall welfare of the wearer.

2.2 Introduction to biomagnetism

The age of the Earth is estimated at 4.5 billion years (2). Since its formation, the Earth's metal molten core has been primarily responsible for a ubiquitous magnetic field on the planet. This magnetic field is generally oriented along the Earth's longitudinal axis. Life here has been developing in this magnetic field since its beginnings estimated at about 3.5 billion years ago (3). Therefore, it is not surprising that some organisms take advantage of this magnetic field. In 1975 Blakemore (4) reported the existence of bacteria called magnetotactic bacteria capable of responding to magnetic fields. These motile prokaryotes align along the Earth's geomagnetic field lines while they swim (4). Magnetotactic bacteria are now known to be cosmopolitan in distribution and very common in most aquatic habitats (5). Navigation using magnetic fields has also been recognized in higher organisms such as ants, pigeons, bees and fish, among others (6, 7).

Magnetotactic bacteria respond to magnetic fields due to the presence of intracellular membrane-bounded magnetite (Fe_3O_4) or greigite (Fe_3S_4) crystals, structures known as magnetosomes. The magnetosome is a single-magnetic domain crystal of magnetite or greigite enveloped in a lipoproteic membrane. The crystals are synthesized de novo by the bacteria rather than being acquired from the environment (8). Interestingly, similar crystals have been found in the human brain (9) and in the rainbow trout peripheric nervous system (10). Morphologically similar crystals have also been found in the Martian meteorite ALH84001 dated to 4.5 billion years old (11). If the crystals in the meteorite are truly from magnetotactic bacteria, it may give insight not only into the origin of life in this planet but also about the similarities between the existence of past life on Earth and on Mars. Draft genome sequences of two strains of magnetotactic bacteria, a coccus called MC-1 and

Magnetospirillum magnetotacticum are currently available for inspection at the Joint Genome Institute (JGI) website (12).

A model for magnetic field direction in the rainbow trout was developed by Walker *et al.* (6). These researchers were able to establish a system that demonstrates behavioral responses by the rainbow trout in the presence and absence of a magnetic field in experimental tanks. Furthermore, specific cells that contained magnetic particles were identified with crystallographic and magnetic properties similar to those properties encountered in magnetotactic bacteria magnetite (13). Another similarity between some eukaryotes that respond to magnetic fields and the magnetotactic bacteria is the construction of chains of magnetite crystals. These have been found in the frontal tissue of Chinook salmon and closely resemble the chains of magnetosomes found in bacteria (14).

Despite the excitement of finding new sensory systems, biomagnetism sensing remains quite understudied in general terms around the globe, with fewer than 10 research groups examining biomagnetism in prokaryotes and eukaryotes (6).

2.3 Magnetotactic bacteria

As previously stated, Blakemore was the first to report the existence of magnetotactic bacteria (4). While examining samples of mud microscopically, Blakemore observed large numbers of bacterial cells swimming in a single direction and accumulating at the edge of the drop. Initially, he thought he was observing a phototactic response but it soon become apparent that that was not the case since changing the position of the light didn't affect the swimming direction of the bacteria. When he moved a small magnet near the slide it become

clear that the swimming direction of the cells was due to the local magnetic field from the magnet.

Following up on the first report of the discovery of magnetotactic bacteria, Blakemore *et al.* (15) isolated and cultivated the first magnetotactic bacterium in axenic culture in 1979. One previous crucial observation was that the magnetotactic bacteria in the mud sample he examined did not appear to tolerate high amounts of O_2. Cells of these bacteria appeared to thrive in the transition region between the anaerobic and aerobic zones (15). To isolate magnetotactic bacteria, jars were filled with mud and water to about two-thirds of their volume and left undisturbed for at least a month. Under these conditions, numbers of magnetotactic cells can increase dramatically. After a month or more, bacteria from the transition zone of the jar were collected with a syringe and injected into tubes containing semi-solid media with different compositions. After three successive transfers of a well-isolated surface colony in growth media containing a higher agar concentration than the original enrichment cultures, *Magnetospirillum magnetotacticum* strain MS-1 was isolated.

Blakemore noted that cells of *Magnetospirillum magnetotacticum* lost their magnetotactic behavior after continual transfer, and thus, to keep them magnetic, he attached a small magnet to the tube with an elastic band. The bacteria that accumulated near the magnet were used as the inoculum for the next tube. This way, cells of *M. magnetotacticum* in culture were kept magnetotactic (15). This strain, as are all the magnetotactic strains described to date, is a relatively slow growing bacterium. Magnetotactic cells of *M. magnetotacticum* have a doubling time of 6 to 15 h, a DNA G+C content of 65% and consist of about 2% of iron on a dry weight basis (15).

Magnetotactic bacteria are a very diverse group of bacteria, both morphologically and physiologically (5). Nevertheless, this group shares 5 common features. These features include: (I) all magnetotactic bacteria examined to date are gram-negative members of the Domain *Bacteria*; (II) all are motile by means of flagella; (III) all possess magnetosomes and their synthesis is dependent on very low to no free O_2, (IV) all have a respiratory form of metabolism; and (V) all show a negative tactic and/or growth response to atmospheric levels of O_2 (16).

2.3.1 Morphology, ecology and phylogenetic diversity of magnetotactic bacteria

Magnetotactic bacteria have been found in freshwater and marine aquatic environments, hemipelagic sediments and in some soils (17-22). Magnetotactic bacteria are ubiquitous in aquatic habitats around the globe (19). Many different cellular morphologies have been described including cocci, rods, spirilla and vibrios (17, 19) (Fig. 1). A multicellular magnetotactic prokaryote has also been described (23). All magnetotactic bacteria described thus far possess flagella, which are arranged as a single polar, bipolar, bundles or tufts (16, 24) (Fig. 1).

On a local ecological basis, most magnetotactic bacteria exist at the oxic-anoxic transition zone (OATZ) or just below it (15). The OATZ is usually located in freshwater habitats at the water-sediment interface but can occur in the water column of marine, chemically-stratified environments. Greigite-producing magnetotactic bacteria are generally found below the OATZ in chemically-stratified marine systems, where hydrogen sulfide (H_2S) is present, in the anaerobic zone (25). Magnetite producers are generally located within the OATZ itself (15).

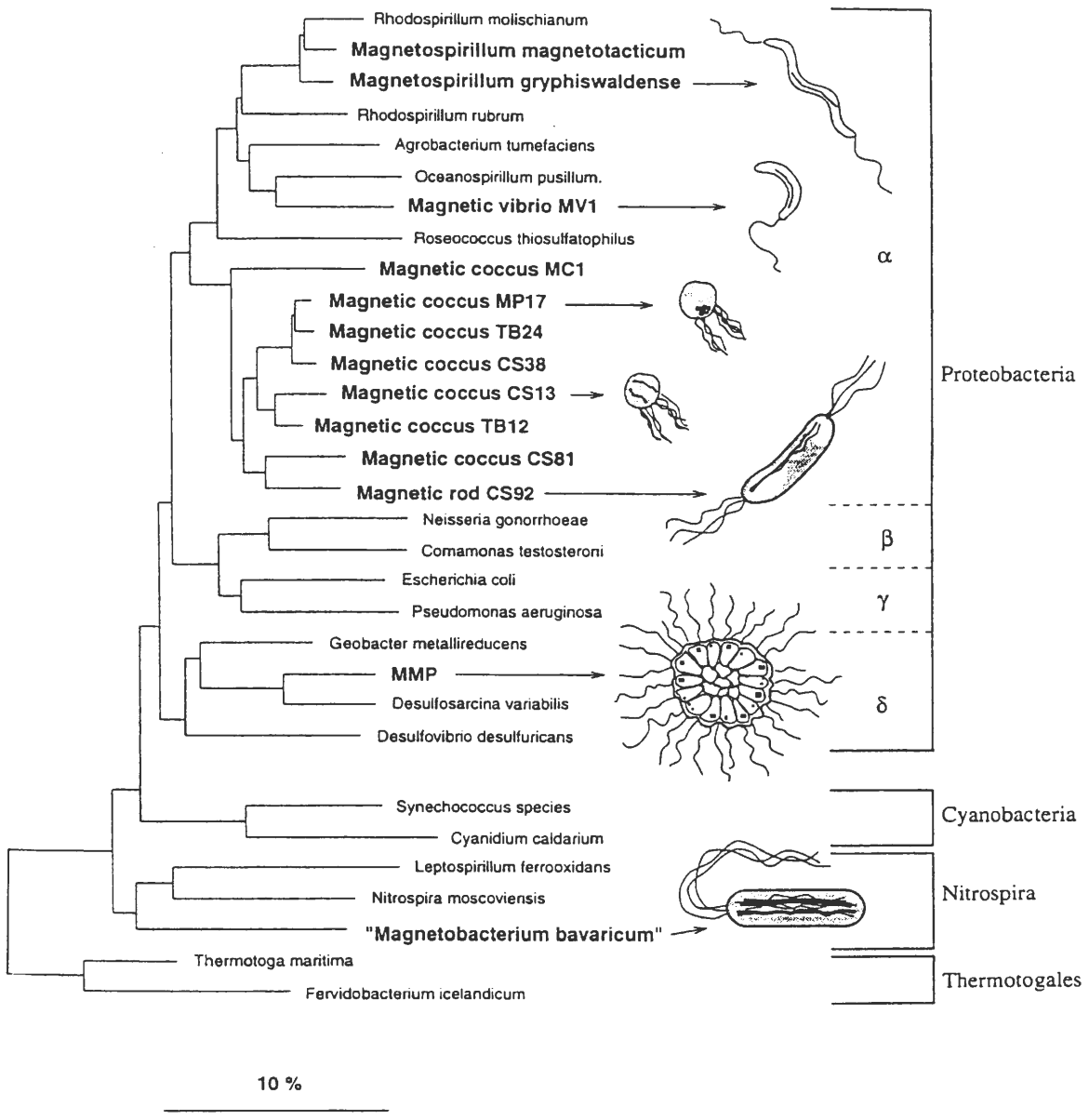


Fig. 1. Phylogenetic dendrogram based on 16S rDNA sequences showing the position of magnetotactic bacteria (bold-faced type) in three different phylogenetic groups of bacteria, including the α and the δ subgroups of the *Proteobacteria* and the *Nitrospira* group. Morphologies are represented by drawings, spots or black lines inside the bacteria represent magnetosomes or chains of magnetosomes respectively. Bar at the bottom represents sequence divergence. MMP is the only greigite producer shown. Figure taken from reference 24

The magnetotactic bacteria strains in pure culture are fastidious with regard to growth. Thus, it is not trivial to isolate these strains in axenic culture. Farina and co-workers (26) have described several types of magnetotactic bacteria from sediments by means of electron microscopy. This approach has the advantage that some bacterial features such as S layers and capsules which are often lost during cultivation can be studied (27). Such features could be related to iron absorption and thus magnetite formation. S-layers are secreted glycoprotein or protein assemblies, which completely cover the cell and interact with themselves and the cell wall surface via non-covalent forces, forming oblique, square or hexagonal crystal arrays (27-29). The function of S-layers remains unknown, but it has been suggested that they serve as ion trappers (28, 29). Farina's group described the presence of capsules and S-layers in uncultured cocci collected from the southern hemisphere (Fig. 2) (26). They also showed some evidence of the presence of membrane vesicles in the S-layer (Fig. 2) (26). In addition, an association and possible interaction between flagella and magnetosomes was described (Fig. 2) (26).

Thus far, magnetotactic bacteria have been found to be phylogenetically affiliated with two different phyla of the domain *Bacteria* (24). These include the *Proteobacteria*, which contains most of the described magnetotactic bacteria, and the *Nitrospira* phylum which contains one species of magnetotactic bacteria described so far, an uncultured rod-shaped bacterium, candidatus *Magnetobacterium bavaricum* (30). Subclasses α and δ of the phylum *Proteobacteria* contain magnetotactic representatives (Fig. 1) (31). Magnetite producers are found within these two subclasses and in the *Nitrospira* phylum (24). Only one species of greigite-synthesizing magnetotactic bacterium has been phylogenetically studied thus far; it is a multicellular prokaryote associated within the δ subclass of the *Proteobacteria*

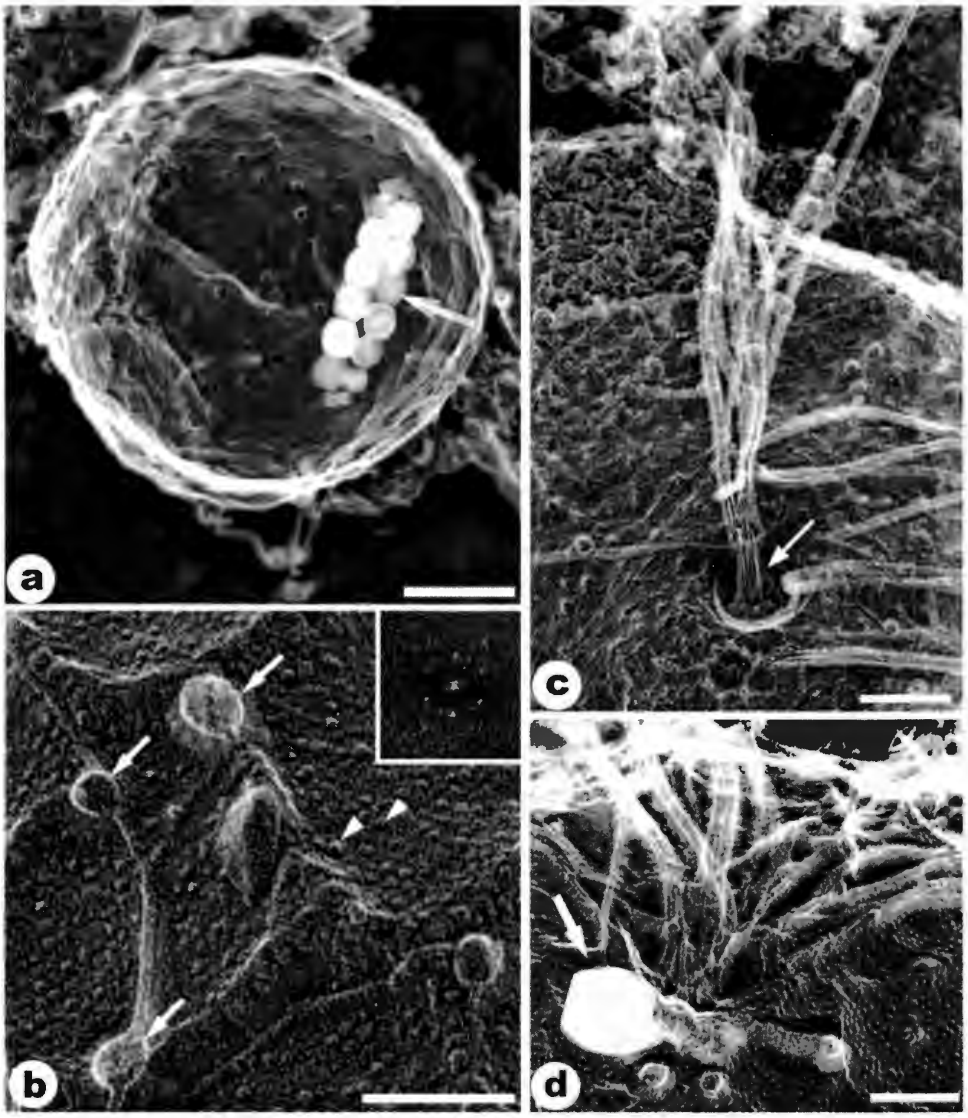


Fig. 2. Electron micrograph of freeze-etched uncultured magnetotactic cocci. a: S-layer, forming the outer layer of the surface of the cell and two chains of magnetosomes (arrow). b: high magnification of the cell surface showing numerous particles from the S-layer (arrowheads) and the presence of membrane vesicles (arrows). c and d: flagella structure, c: several flagella originating from the same pit (arrow), d: intimate association of magnetosome (indicated by arrow) and flagella bundle. Scale bars indicate respectively a, b, c and d: 0.5 μm, 0.1 μm, 300 nm, 120 nm. Figure taken from reference 26.

(Fig. 1) (32). Because magnetotactic bacteria appear in different subclasses and even different phyla, it has been suggested that magnetotaxis as a trait is polyphyletic and evolved independently in different groups of bacteria (33). However, the presence of magnetotaxis in phylogenetically distant species may also be explained by horizontal gene transfer. The latter possibility is becoming a more plausible explanation since some of the genes thought to be required for magnetotaxis are flanked by insertion elements (transposons) in one species, *Magnetospirillum gryphiswaldense* (34), and homologous genes to those thought to be required for magnetotaxis in *Magnetospirillum gryphiswaldense* have also been found in strain MC-1, a bacterium not closely related to *M. gryphiswaldense* (35).

2.3.2 The function of magnetotaxis

Magnetotaxis has been used to describe the behavior of magnetotactic bacteria in magnetic fields but the term is actually a misnomer since cells are neither attracted nor repelled by magnetic fields (4, 36). To really comprehend and appreciate the meaning and function of magnetotaxis, one must understand the magnetism of the planet Earth. Therefore this section begins with an introduction to geomagnetism and then magnetotaxis is explained through the major experiments that led to our understanding of its presumed function.

The Earth's geomagnetic field has both a horizontal and vertical component and the overall direction of the field lines at any location on the Earth is the vectorial sum of these components. At the Equator, the vertical component is zero and the magnetic field direction is completely horizontal (parallel to Earth's longitudinal axis, Fig. 3). The deviation from the horizontal, called the angle of dip, increases from the Equator, where it is 0°, to the poles, where it is 90°. This is illustrated in Figure 3. The Earth's North Pole is, in reality, a south

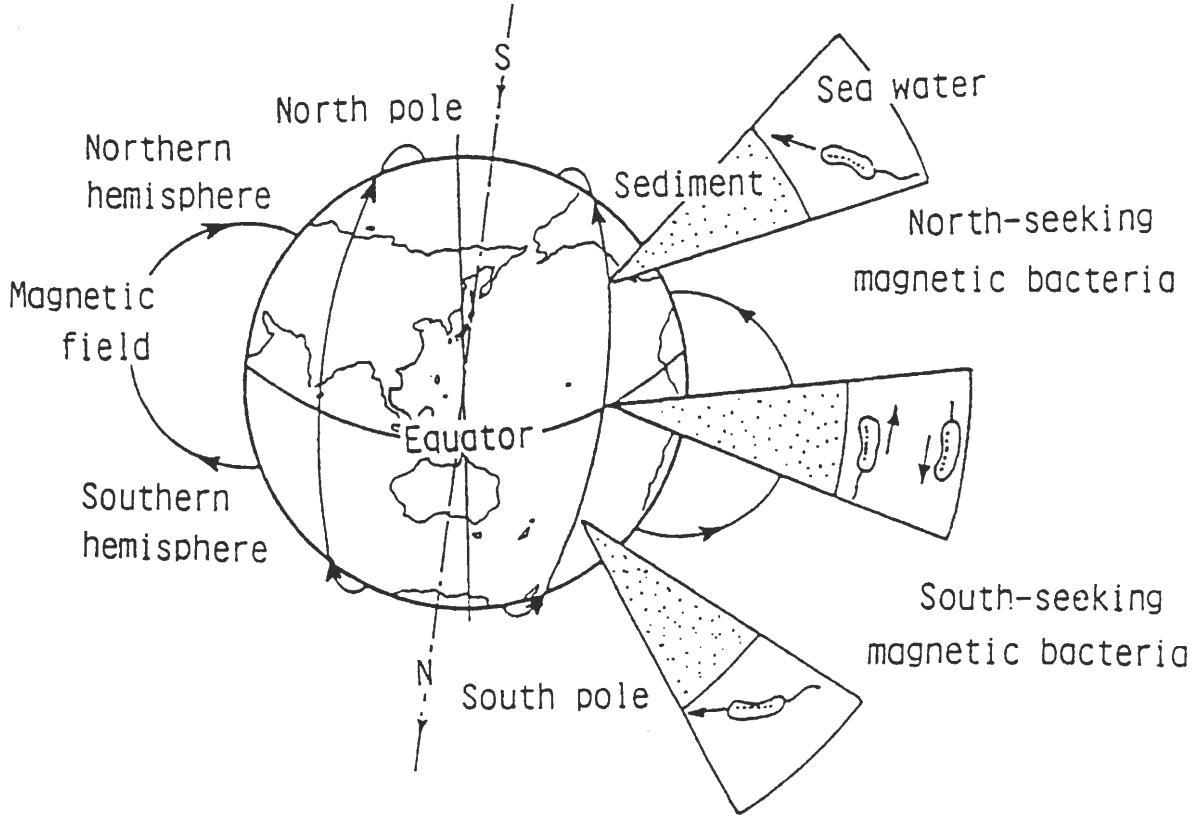


Fig. 3. Schematic representation of the Earth's geomagnetic field. The swimming direction of magnetotactic bacteria collected at that site is indicated by arrows. Generally, in samples collected from natural aquatic environments in the Northern Hemisphere, greater than 99.9% of the magnetotactic cells are north-seeking and the opposite is true for samples collected in the Southern Hemisphere. Sample taken at the Equator contain about equal numbers of both types of cells. Under both conditions magnetotactic bacteria migrate away from high O_2 concentration [O_2]. Figure taken from reference 38.

magnetic pole which is why the north end of a compass needle points in that direction. In sum, the vertical component of the Earth's magnetic field results in the direction north pointing downward in the northern hemisphere and the south pointing downward in the southern hemisphere. Generally, in samples collected from natural aquatic environments in the northern hemisphere, greater than 99.9% of the magnetotactic cells are north-seeking and south-seekers predominate in samples collected from the southern hemisphere (37). Because of the vertical component of the Earth's magnetic field, north-seeking magnetotactic bacteria are not only directed northwards but also downwards (37). The reverse is true in the southern hemisphere. North- and south-seeking magnetotactic bacteria appear to be present in equal numbers at the Equator (39). This is not surprising since there is no vertical component to the magnetic field direction which is horizontal and neither polarity is selected for or against (39).

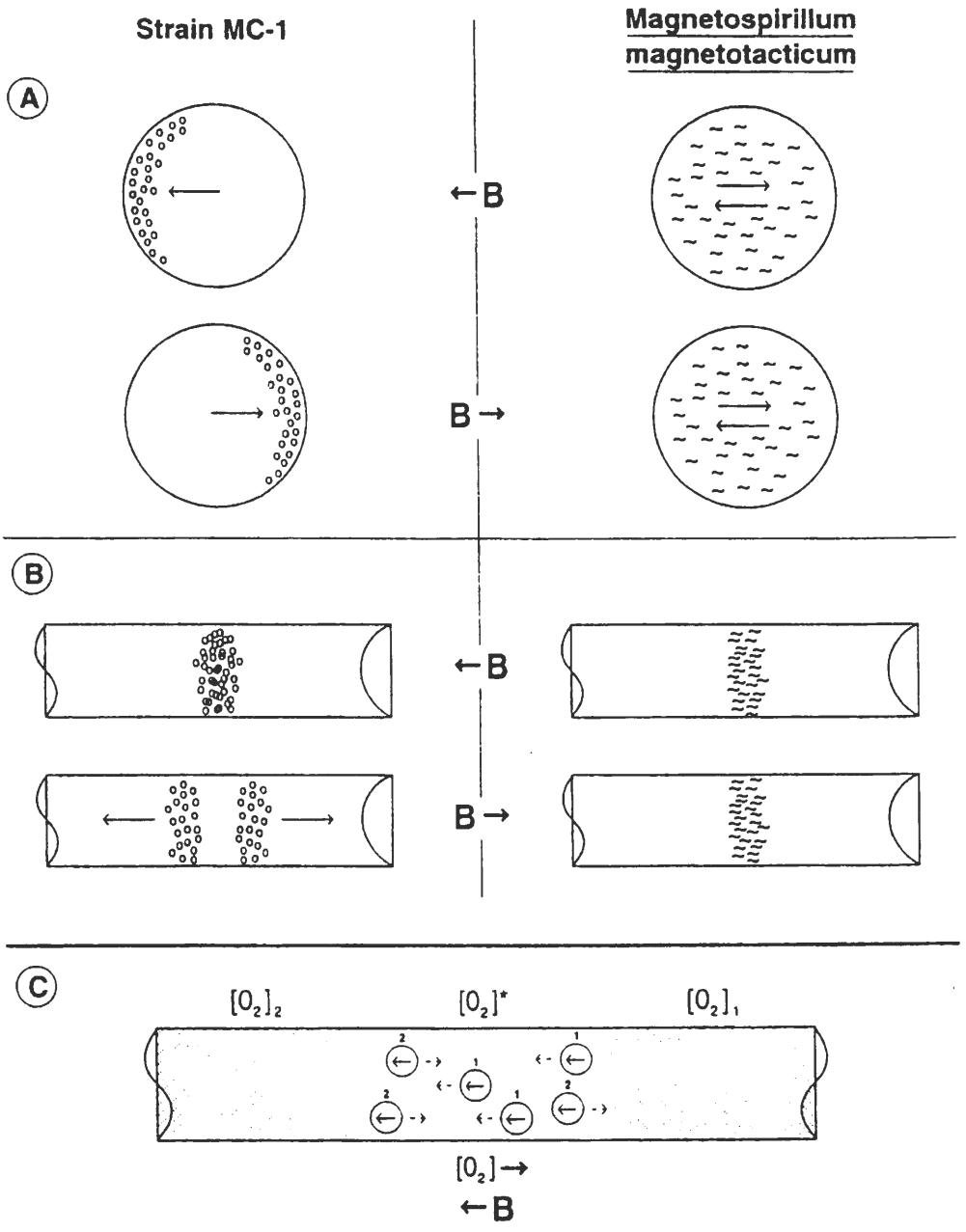
The original theory regarding the magnetotactic behavior of bacteria proposed by Blakemore (8) took into account two facts: 1) that magnetotactic bacteria are anaerobic or microaerobic; and 2) that magnetotactic bacteria in the northern hemisphere are generally north-seeking and those in the southern hemisphere are generally south-seeking. Therefore, regardless of the location, except at the Equator, magnetotactic bacteria would be directed downward, until they reached the sediment (where O_2 is absent or in low concentrations). This theory accounts for both observations (8).

Several interesting experiments were performed in the attempt to further understand the above observations (37). Mud and water samples collected from the northern hemisphere containing high numbers of magnetotactic cocci were placed in a mumetal enclosure (37). Mumetal is used to virtually eliminate local and the Earth's geomagnetic fields. Over a

period of about a month, numbers of south-seeking cells increased until the ratio of south-seeking and north-seeking cells became close to 1. This experiment appears to support the original idea of Blakemore but also lead to additional speculation regarding the appearance of south-seeking cells (37). It is currently thought that if two daughter cells receive magnetite crystals of the parent cell, they will retain the polarity of the parent cell. But if a daughter cell does not receive a magnetosome and synthesizes them de novo, it has a 50:50 chance of being north-seeking or south-seeking. Thus, if there is no selective pressure on magnetotactic populations as in the mumetal enclosure, one would eventually expect to observe equal numbers of north- and south-seeking magnetotactic bacteria (37).

However, the theory described above does not explain two relative recent findings: 1) the presence of large populations of magnetotactic bacteria in the water columns of chemically-stratified aquatic habitats (40); and 2) a newly isolated magnetotactic coccus (strain MC-1) grows as a microaerobic band of cells at some point below the meniscus in oxygen-gradient cultures (41). The latter situation is particularly notable because the culture contains >99.9% north-seeking cells and one would expect the cells to be growing or present at the bottom of the tube (in the northern hemisphere).

Frankel *et al.* (41) revised the original theory of magnetotaxis based on the behavior of cells of strain MC-1 and *Magnetospirillum magnetotacticum* under different conditions. When cells of strain MC-1 are viewed microscopically, >99.9% of them swim in one direction, northward, and accumulate on the north side of the drop (Fig. 4A). They do not back up under these conditions and have been called "one-way" swimmers (41). Cells of *M. magnetotacticum*, on the other hand, align along magnetic field lines and swim backwards and forwards and do not accumulate at the sides of the drop (Fig. 4 A). Organisms like *M.*


Strain MC-1
Magnetospirillum magnetotacticum
A
←B
B→
B
←B
B→
C
$[O_2]_2$
$[O_2]^*$
$[O_2]_1$
$[O_2]\rightarrow$
←B

magnetotacticum are called "two-way" swimmers.

Cells of *Magnetospirillum* do not have a polar preference, they swim bidirectionally and show abrupt changes in their swimming direction (41). When the magnetic field is reversed for cells of *M. magnetotacticum*, they rotate 180°, aligning along the direction of the field. If the field is rotated 90°, they rotate 90°. Cells of this bacterium, like those of MC-1, are strongly aerotactic and grow as a microaerophilic band of cells in oxygen gradients. Cells of *M. magnetotacticum* suspended in growth medium were placed in flattened capillary tubes and allowed to grow and form bands at the OATZ some point below the meniscus. The entire capillary was placed in a magnetic field with north towards the sealed end of the capillary. Cells in the band are generally oriented along the magnetic field lines parallel to the long axis of the capillary and spend much of the time within the band but are continually swimming and take excursions above and below the band. When the field is reversed, the cells rotate 180° but the band remains in the same position and does not disperse (Fig. 4B) (41). This behavior fits the temporal model of chemotaxis (42), where cells sense the $[O_2]$ over time and change the direction of flagellar rotation accordingly.

To understand how cells of MC-1 form microaerophilic bands, north-seeking cells were selected from a culture tube and introduced into the same type of flattened capillary tube set up described above. Under air, the cells initially swam northward to the sealed end of the capillary. As the cells appeared to use up the O_2, they moved as a band in the opposite direction (southward). Eventually they formed a stable band at the OATZ at some distance below the meniscus. This experiment shows clearly that cells of MC-1 can, in fact, swim in both directions; however, because cells cannot physically turn around due to the magnetic field, the reversal must be due to a reversal in the direction of the rotation of the flagella.

Thus, under oxic conditions (like those in the drops under the microscope) cells of MC-1 swim northward (downward) and when conditions become anoxic, cells reverse direction (via a reversal of the direction of the flagella motor) and swim southward (upward). These observations are significant because they are not consistent with the well described temporal model of chemotaxis (aerotaxis in this case) (43). Instead, the behavior is consistent with a two-state, switch-like mechanism of aerotaxis. Under oxic conditions, cells are in state 1 and swim northward (downward) seeking the OATZ. If they go past the OATZ into the anoxic, reduced zone, they "switch" to state 2 where the flagella motors reverse and cells swim upwards again seeking the OATZ. Cells of MC-1 also have other strategies to maintain their position in O_2 gradients; for example, cells of MC-1 possess pili which they use to attach to surfaces and to one other forming raft of cells at the OATZ (41).

Individual cells of strain MC-1, like those of *Magnetospirillum magnetotacticum*, present as a band at the OATZ in flattened capillaries set up as described for *M. magnetotacticum*, take linear excursions above and below the band. When a band of MC-1 at the OATZ was exposed to a magnetic reversal, two bands formed, one going upward from the OATZ (direction of high [O_2]) and the other going downward from the OATZ (direction of lower [O_2]) (Fig. 4 B). The cells did not change their swimming direction and the band did not reform in the short term. The results of this experiment support the two-state, "switch" mechanism of aerotaxis. When the field was reversed, the bacteria in state 2 now migrating downwards, presumably attempting to find the OATZ, would not change direction until they reach the oxic zone which is not possible. Those cells in state 1, now migrating upwards (still northward) towards the meniscus, will not change their swimming direction until they reach the anoxic zone which is also not possible (Figs. 4B and 4C). Thus cells appear to remain in

their theoretical physiological states. What causes the cells to be in these states is unknown, but it is possible that some cellular component can be oxidized and reduced leading to a change in the direction of the rotation of the flagella motors. In one state cells swim in one direction and the flagella rotate in a given direction, clockwise for instance. In the other state, the flagella rotate counter clockwise and the cells swim in an opposite direction. The direction of rotation of the flagella appears to be controlled by $[O_2]$ and/or redox conditions. The "two-way" switch model elegantly explains why some magnetotactic bacteria don't migrate persistently downward but remain at the OATZ (41).

Frankel *et al.* (41) describe the magnetotactic behavior of strain MC-1 and *Magnetospirillum magnetotacticum* as two forms of magneto-aerotaxis (41). The first, polar magneto-aerotaxis, is typified by cells of MC-1. These cells have a polar preference in their direction of motility and use the magnetic field for direction and as an axis for motility. Cells of *M. magnetotacticum* exhibit what is referred to as axial magneto-aerotaxis in which the magnetic field is used by these cells solely as an axis for motility. Cells of both species align along the inclined geomagnetic field lines and swim bidirectionally relying on two different forms of aerotaxis to locate the OATZ. In either case, the biological advantage is that the magnetic dipole of the cell aids the bacterial cell in finding and maintaining an optimal position (the OATZ) (41). Bacteria without a magnetosome chain have a three dimensional search task to locate the OATZ. Once magnetotactic bacteria are aligned along geomagnetic field lines, there is a limitation to the motility of the cell resulting in a single dimensional search task thus increasing the efficiency of aerotaxis in vertical concentration gradients, which helps the cells locate and maintain position at the OATZ more efficiently (41).

2.3.3 The magnetosome

The term magnetosome was first proposed in 1980 by Balkwill *et al.* (43) who examined cells of *Magnetospirillum magnetotacticum* in detail by electron microscopy. The term "magnetosome" refers to an intracellular magnetite crystal surrounded by a lipoproteic membrane (Fig. 5) (44). Magnetic minerals found in magnetosomes include magnetite and greigite and, in all but one case, each magnetotactic bacterial strain or species biomineralizes only one type of mineral. An uncultured bacterium has been found to contain crystals of both minerals (40). Cells synthesize magnetosomes de novo in an ordered fashion, with all or most magnetosomes aligned in a chain (44) that functions essentially as a single magnetic dipole rather than a collection of individual dipoles. The magnetosomes provide the bacterial cell with a permanent magnetic dipole moment. In cells of *M. magnetotacticum* grown microaerobically with 1-3% O_2 and nitrate as the nitrogen source and additional terminal electron acceptor (43), the magnetite particles are about 42 nm in diameter and a chain consists of an average of 18 magnetosomes. In other bacteria, magnetosomes ranging from 19 to 136 nm in diameter have been described (45) and more than 300 magnetosomes can be found in a single cell (45). However, magnetosome chains in most magnetotactic bacteria usually consist of about 10 to 20 magnetosomes (46). The chain generally transverses the cell along its long axis if the organism is rod- or helically-shaped and a single cell can have more than one magnetosome chain. Cells generally form mature magnetosome crystals of a very narrow size range as well as forming smaller immature magnetosome crystals that are usually found at the end of the chain (44). Electron micrographs depicting empty vesicles suggest that the magnetosome membrane vesicle forms first and then the mineral crystal forms inside the vesicle (44).

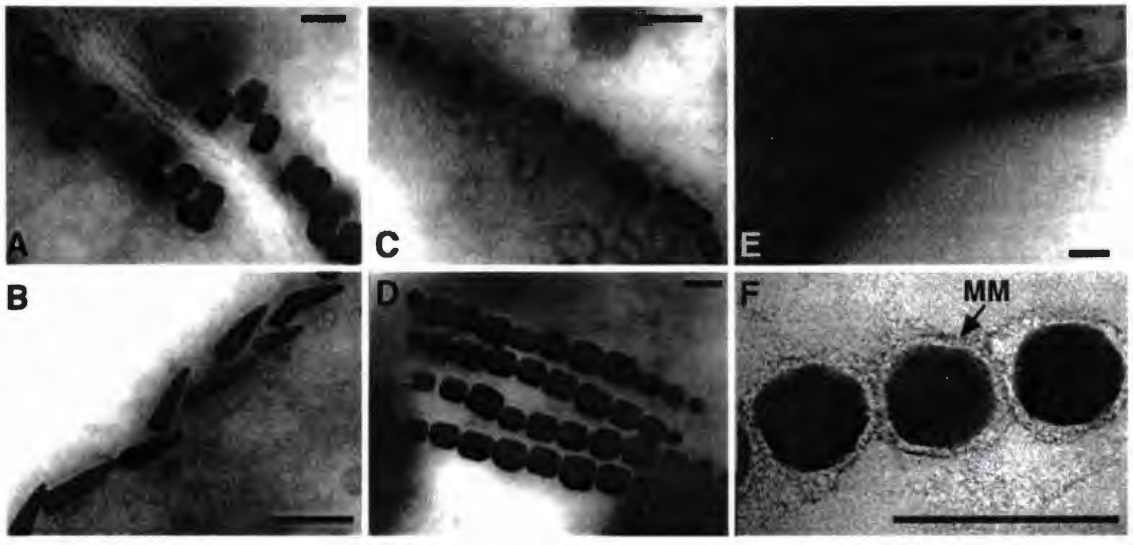


Fig. 5. Electron micrographs of magnetosomes showing the diversity of crystal morphologies in different species of magnetotactic bacteria. A-D, magnetite magnetosomes from several different uncultivated magnetotactic bacteria. E, cubo-octahedral magnetite crystals in *Magnetospirillum gryphiswaldense.* F, thin section of magnetite magnetosomes of *M. gryphiswaldense* showing magnetosome membrane (MM). Bar indicates 100 nm. Note that in all cases, a single cell produces one morphological type of crystals. Figure taken from reference 49.

There are three main crystal morphologies regardless of mineral composition. These morphologies include: (I) roughly cuboidal which have been determined to be cubo-octahedral, (II) elongated hexagonal or octagonal prisms which are parallelepipedal-shaped or roughly rectangular in projection or (III) tooth-, bullet- or arrowhead-shaped (Fig. 5) (47, 48). The crystal morphology is conserved within a bacterial strain or species, suggesting that the bacterium exerts a fine control over crystal biomineralization (50). To date, certain elongated crystals (types II) are known to be exclusively produced by organisms and their presence in sediments and meteorites as so-called "magnetofossils" have been used as an indicator of the past presence of magnetotactic bacteria (11). More recently, however, a group has been able to synthesize magnetite with similar crystal morphologies (51).

Magnetosome crystals, whether they are of greigite or magnetite, generally are of a narrow size range of 35 to 120 nm (36). Crystals of magnetite or greigite (52) in this size range are permanent single magnetic domains at ambient temperature (52). Single domain crystals represent the maximum magnetic moment that can be achieved. Theoretically, multiple domains would form in crystals larger than 120 nm lowering the magnetic moment of the crystal since the domains would likely not be perfectly aligned with each other. Crystals smaller than 35 nm are superparamagnetic; in these particles the magnetic direction shifts resulting in a crystal without a permanent magnetic moment at room temperature (52). The Earths' magnetic field strength averages about 0.5 gauss; taking this value into account and the value of the magnetic moment of the magnetosome chain, Frankel e*t al.* (53) demonstrated that the magnetic moment of an average-sized magnetosome chain (10 to 20 magnetosomes, each 50 nm in diameter) is sufficient to overcome the thermal energy which tends to randomize the position of the cell.

2.3.3.1 Biomineralization of the bacterial magnetosome

It is important to emphasize that the magnetosome is synthesized de novo by the cell. The crystallochemical characteristics of the particles produced by the magnetotactic bacteria including the species- and/or strain-specific crystal morphologies, the narrow size range and the chemical purity of the crystals indicate that they are formed by a biologically-controlled biomineralization process (BCM) (16). In BCM it is believed that mineral synthesis is under physiological and ultimately genetic control. In this case, BCM of magnetosomes is thought to employ specific proteins for iron acquisition and transport, membrane vesicle formation, and crystal nucleation and growth. It seems likely that coordinated events are required for the synthesis of the magnetosome chain. Crystals in magnetosomes generally consist of one mineral, either magnetite (Fe_3O_4) or greigite (Fe_3S_4), but not both with one notable exception (40). Moreover, magnetite crystals are formed even when the bacteria are grown under high concentrations of hydrogen sulfide (where greigite might be expected to form), thus demonstrating that the bacterial biomineralization is specific regardless of the composition of the growth medium (54). Proteins have not been found within the crystals and the presence of other metals has only been described at trace values with the exception of copper in greigite crystals (47, 55, 56).

Greigite synthesizers have not yet been isolated in pure culture. Selected area electron diffraction in the electron microscope showed that the non-magnetic minerals mackinawite (tetragonal FeS) and an unusual form of cubic FeS are present in greigite-producing magnetotactic bacteria and are likely precursors to ferrimagnetic greigite in these bacteria (25).

Magnetite and greigite can also be formed by bacteria by a very different kind of biomineralization called biologically induced biomineralization or BIM (16, 57). In this case, for example, dissimilatory iron-reducing bacteria reduce ferric iron to ferrous iron; the ferrous iron in the environment reacts with certain poorly crystalline forms of ferric oxides and hydroxides and magnetite forms extracellularly. Although the bacteria mediate the chemistry, they do not control the biomineralization process and magnetite crystals synthesized via BIM are similar to those produced abiotically (chemically). That is, they are not well-ordered (are amorphous), they have irregular shapes and, unlike the crystals of magnetotactic bacteria, their size distribution is log normal (16, 57).

2.3.3.2 Genes and proteins involved in magnetosome synthesis

Different genes and proteins believed to be directly involved in the magnetosome formation have been described in the literature (35, 55, 58, 59, 60, 61, 62). These genes and proteins are described in this section with their putative functions in the strains that they were identified in. Although many are magnetosome membrane proteins, some are not. Other genes and proteins possibly involved in biomineralization will be described later when possible mechanisms for magnetite biomineralization are described.

Magnetosomes from cultured bacteria can easily be magnetically separated from lysed cells, washed to remove contaminating proteins and studied biochemically (44). One of the usual approaches to address the question of how magnetosomes are synthesized is to characterize the magnetosome membrane proteins using sodium-dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Magnetosome membranes are easily solubilized with detergents such as SDS and the proteins separated by SDS-PAGE (44). Then

N-terminal sequences of specific proteins can be determined and used to design degenerate primers for polymerase chain reaction (PCR) experiments to identify the genes encoding for these proteins (35). Since this strategy goes from protein to gene, it has been referred to as "reverse genetics". This strategy or variations of it has lead to the identification of many magnetosome membrane proteins and genes from several different magnetotactic bacteria as described below.

Magnetospirillum magneticum strain AMB-1

magA. An iron regulated gene, *magA*, was identified in *Magnetospirillum magneticum* strain AMB-1 (58). Using transposon mutagenesis, a non-magnetotactic mutant strain incapable of magnetite synthesis was generated. A DNA fragment apparently required for magnetosome synthesis was found and sequenced. The fragment contained an open reading frame, a putative gene, whose expression was confirmed by northern blotting. This gene encodes a protein with a molecular mass of 46.8 kDa that shows 25.4% homology to the *Escherichia coli* gene *kefC* which encodes a cation efflux antiporter (a type of protein responsible for the exchange of cations, usually K^+ from inside the cell for protons from outside the cell (63)) (58). *magA* was expressed at higher levels by cells grown in iron-limited medium compared to those grown in iron-sufficient medium (58). *magA* was over expressed in *E. coli* and used to prepare membrane vesicles. These vesicles took up iron when ATP was supplied to them suggesting that *magA* might be important in iron transport to the magnetosome membrane vesicle (58). The Genbank accession number for *magA* is D32253.

mpsA. Using reverse genetics and anchored PCR, *mpsA*, a gene encoding for a magnetosome membrane protein, was identified and sequenced (59). This magnetosome membrane protein

has an apparent molecular weight of 35.6 kDa and shows 52% homology to an acetyl-CoA carboxylase of *E. coli*. Using gene fusions of *mpsA* to a luciferase gene, it was demonstrated that this protein is mainly located in the magnetosome membrane (59). Some acetyltransferases are known to be associated with membrane vesicle invagination suggesting a possible role in magnetosome vesicle formation (64, 65). The Genbank accession number for *mpsA* is D87827.

Mms16. Reverse genetics was also used to identify Mms16, another magnetosome membrane protein (60). The gene encoding for this protein was fused to a hemagglutinin tag and overexpressed in *E. coli*. A G-Sepharose column with immobilized antibody anti-hemagglutinin was used to separate the tagged protein. The purified protein showed GTPase activity. The GTPase inhibitor AlF_4^- was added to growth media and magnetosome formation was found to be was significantly impaired (60). Some GTPases in eukaryotes are known to be involved in priming and budding of trafficking vesicles (66, 67). It is possible that Mms16 might have a related function in the formation of the magnetosome membrane vesicles. The Genbank accession number for mms16 is AB051013.

Mms5, Mms6, Mms7 and Mms13. Reverse genetics was also used to identify the genes that encode for these four magnetosome membrane proteins (55). Mms7 and Mms13 appear to be virtually identical to the magnetosome membrane proteins MamC and MamD, respectively, from *Magnetospirillum gryphiswaldense* (35). Mms5 shows no homology to any known protein or nucleotide sequence. *mms*6, *mms*7 and *mms*13 are within the same genomic region in strain AMB-1 (within a 3.2 kb region). Mms6 shows significant homology to a putative protein present in the partially sequenced genome of strain *M. magnetotacticum*. The functions of these proteins and their homologs have not yet been determined, although Mms6

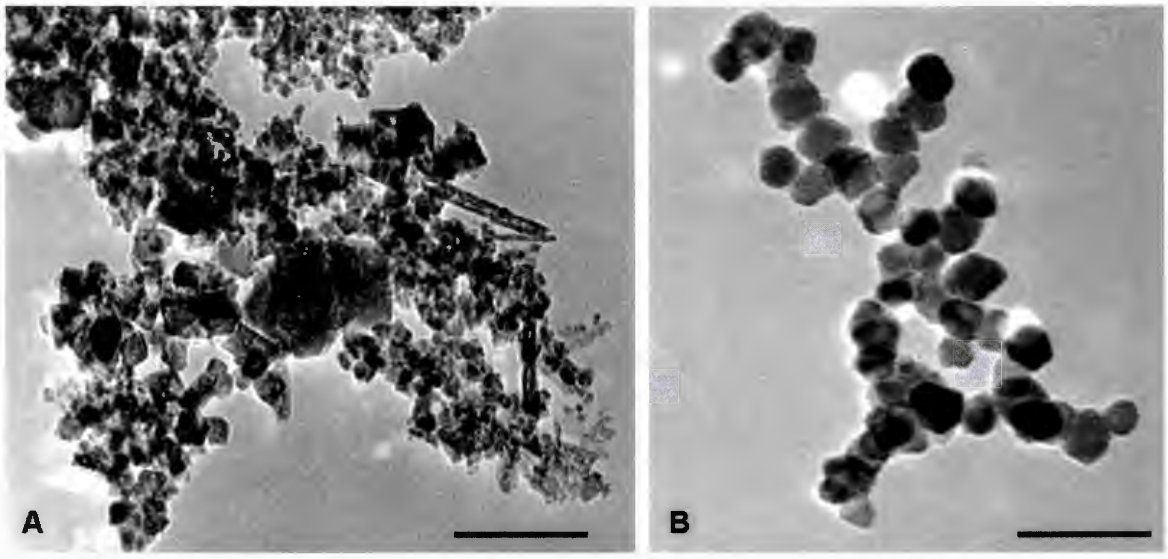


Fig. 6. Electron micrograph showing magnetite synthesized chemically in the absence of Mms6 (A) and in the presence of Mms6 (B). Scale corresponds to 100 nm. Figure taken from reference 55.

may be involved in magnetite crystal nucleation and growth. There are two pieces of evidence, besides the fact that this is a magnetosome membrane protein, supporting this conclusion: 1) the presence of this protein appears to cause the formation of cubo-octahedral magnetite crystals similar to those found in *M. magnetotacticum* as shown in Fig. 6; 2) this protein binds iron (53). The Genbank accession numbers for the sequences of the corresponding genes are AB096081 and AB096082.

NMA21 defective proteins. Applying transposon mutagenesis a non-magnetotactic mutant of strain *Magnetospirillum magneticum* strain AMB-1 strain NMA21 was generated (61). Using inverse PCR, the transposon inserted region was mapped. The transposon was found to have disrupted the gene for an aldehyde ferredoxin oxidoreductase (AOR). The AOR gene is present in an operon containing 5 ORFs. ORF-1 encodes for a ferredoxin, ORF-2 encodes for the AOR, ORF-3 encodes for a NADH oxidase, ORF-4 encodes for an assimilatory nitrate reductase electron transfer subunit and ORF-5 encodes for a molybdenum cofactor that is involved in the synthesis of molybdopterin. This last protein is required for the maturation of the AOR (61). AOR catalyses the oxidation of aldehydes to carboxylic acids while reducing ferredoxin. Wahyudi *et al.* suggest that electrons could flow from the ferredoxin to the NADH oxidase (via NAD(P) or NAD(P)H) and then to the assimilatory nitrate reductase electron transfer subunit and eventually be used to reduce ferric iron to ferrous iron which is required for magnetite synthesis. AOR was determined to be a cytoplasmic protein (61). The Genbank accession numbers for the ORF-1 to ORF-5 are AAL81603, Q51739, AAD35480, P42433 and E75252 respectively.

***Magnetospirillum gryphiswaldense* and its non-magnetic mutant**

MamA, MamB, MamC and mamD. Grünberg *et al* (35) using reverse genetics identified a number of specific genes (*mamA*, *mamB*, *mamC* and *mamD*) that made possible the discovery of clusters of genes (in an operon-like organization) encoding for magnetosome membrane proteins presumably responsible for magnetosome formation in *Magnetospirillum gryphiswaldense*. The relevance of this work is the fact that similar gene clusters were also found in strains *M. magnetotacticum* and strain MC-1 (Fig. 7). Most of the putative genes identified did not show homology to any known genes and others were homologs to cation diffusion facilitators (CDFs). CDFs are proteins that function as efflux pumps of toxic divalent cations such as some heavy metals (49). An N-terminal sequence result also indicates the presence of a *mamC* gene in the strain MV-1 (Bazylinski, D. A., unpublished results). The accession number for mamAB cluster is AF374354 and for mamCD cluster is AF374355.

Other genes and proteins involved in magnetosome synthesis. Spontaneous non-magnetic mutants form from *Magnetospirillum gryphiswaldense* with a relatively high frequency (62). One of these mutants designated strain *M. gryphiswaldense* MSR-1 is a deletion mutant that lacks 80 kb segment of genomic DNA (62). This mutant was isolated from serial subcultures of *Magnetospirillum gryphiswaldense* which were stored at 4°C between transfers. Many genes were found to be missing in this deletion including the *mamAB* and *mamCD* gene clusters (estimated to be 10 kb apart) and genes homologous to *mms5*, *mms6* and *mms7* from *Magnetospirillum magneticum* strain AMB-1 (Fig. 7). A striking feature of this region was the abundance of mobile genetic elements at one end of the deleted DNA fragment (62). Homologous of many of the missing genes are present in MC-1 and *Magnetospirillum*

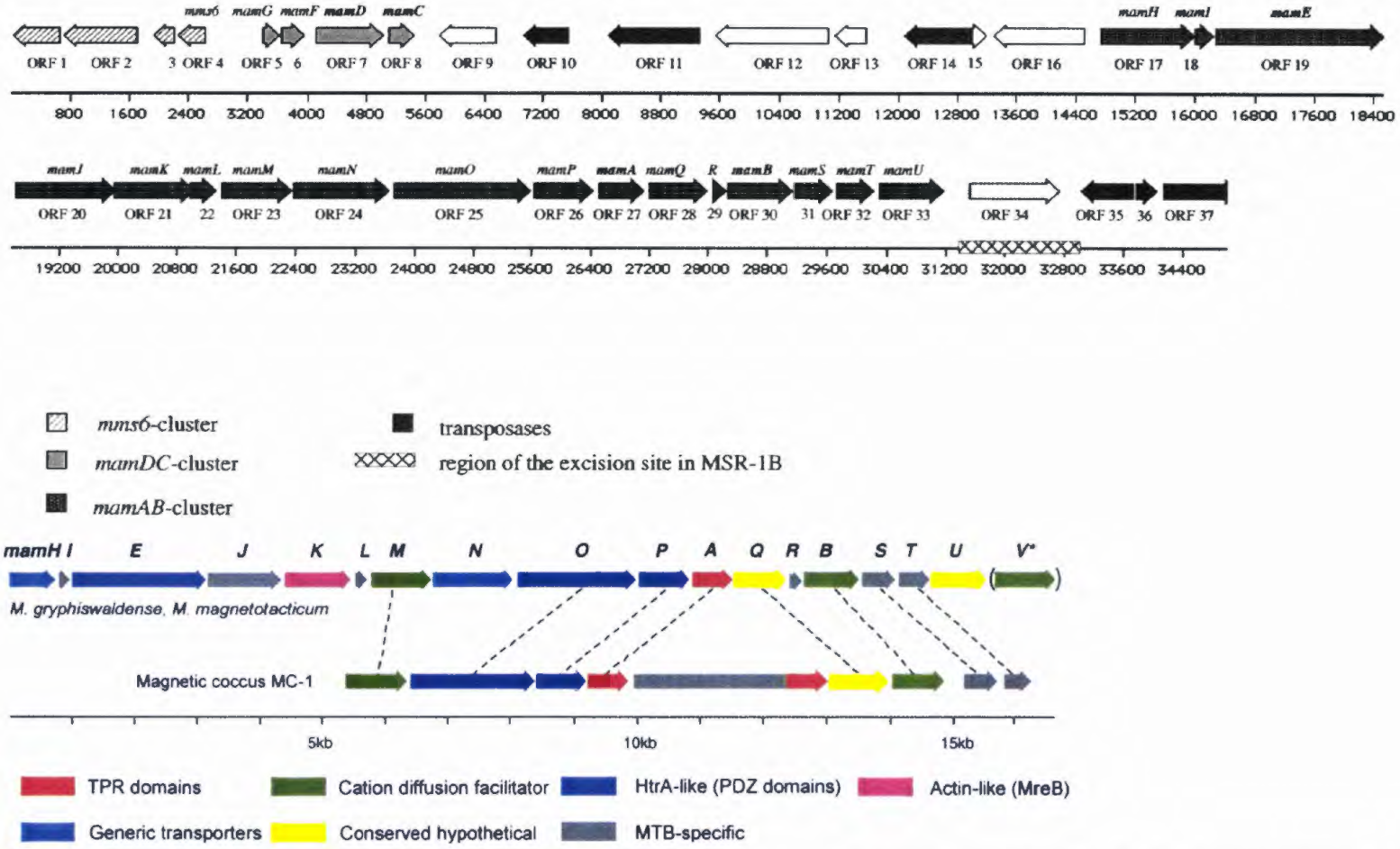


Fig. 7. Top and bottom: Representation of *mamAB* and *mamHV* gene clusters, respectively, of *M. gryphiswaldense* and similar homologs of *Magnetospirillum magnetotacticum* and unnamed strain MC-1. Arrows indicate direction of gene transcription. Filled arrows indicate homologous ORFs present in all the mentioned three species, dashed lines connecting them. Bars below the arrow indicate size of the gene in a 1000 base pairs (kb). CDF: cation diffusion facilitators, TRP: tetratricopeptide repeats. **mamV* is present in the genome data of *M. magnetotacticum* MS-1 but absent in the genome of *M. gryphiswaldense*. Figure extracted from references 35 and 49.

magnetotacticum (12, 62). Because of this and the fact that these genes are clustered together with genes known to encode magnetosome membrane proteins, many or most of these genes might have roles in magnetosome synthesis. The mutant appears to be sensitive to high levels of iron in the media which might indicate a link between magnetite formation and detoxification of iron (35). The Genbank accession numbers for the deleted region are BX571797, BX571782 and BX571783.

2.3.3.3 Possible physiological role for magnetosome synthesis and mechanism for magnetosome formation.

Magnetosome formation must involve a number of steps: these include iron uptake by the cell, transport of iron to the magnetosome vesicle, vesicle formation and mineral precipitation. Evidence for the sequential order of these steps remains obscure. However, because empty and partially-filled magnetosome vesicles have been observed in some organisms, it seems reasonable to assume that the magnetosome vesicle forms prior to the mineral precipitation (44). Alternatively, perhaps a small amount of mineral precipitation occurs in the periplasm and leads to an invagination of the plasma (cell) membrane which becomes the magnetosome membrane vesicle.

How the magnetosome mineral crystal is biomineralized biochemically and chemically is unknown. There is much speculation on the biomineralization process, however, based on putative functions of magnetosome membrane proteins and other proteins that may be missing from non-magnetotactic mutant strains (36). Therefore, a model is still missing that explains magnetosome formation and discerns different aspects of magnetosome formation, such as the steps described above, and links magnetosome synthesis to cellular

physiology and genetics. This section describes some of the speculation in the attempt at forming this model.

Magnetite magnetosome production by *Magnetospirillum magnetotacticum* has been shown to be optimal at an $[O_2]$ of 1% in the headspace of sealed liquid cultures (68). The low narrow range of $[O_2]$ under which magnetite is synthesized appears to apply to all microaerophilic magnetotactic bacteria described to date. Cells of *M. magnetotacticum* are capable of denitrification only under microaerobic conditions (as opposed to most denitrifiers that denitrify anaerobically) and cannot grow under strict anaerobic conditions (15, 68, 69, 70). Denitrifying cells of *M. magnetotacticum* appear to synthesize more magnetite than those grown with ammonium as the nitrogen source and O_2 as the terminal electron acceptor (68).

In order to discern a possible link between nitrate reduction, denitrification and magnetite magnetosome formation, Fukumori and co-workers (70-75) partially characterized the respiratory chain of *Magnetospirillum magnetotacticum*.

Tamegai *et al.* (70) reported a cytochrome a_1-like hemoprotein in *Magnetospirillum magnetotacticum*. This hemoprotein lacks heme *a* even though it has absorption peaks similar to those of heme *a*-containing cytochrome a_1 molecules (70). Cytochrome *a* proteins are part of complex IV of the respiratory chain, and are responsible for oxidizing cytochrome *c* (76, 77). The respiratory chain is comprised of five complexes that transport electrons across the inner membrane and drive ATP synthesis (76, 77). Briefly, complex I and II are entry points for electrons from NADH and succinate, respectively (76, 77). Complex III contains intermediate electron carriers that reduce cytochrome *c* and components of complex IV are responsible for the reduction of dioxygen (which results in water formation), for

oxidizing cytochrome *c* and for pumping protons from the inside to the outside of the cell (76, 77). The flow of protons creates a negative charge inside the cell which is used to drive ATP synthesis by complex V (76, 77). Bacterial cytochrome *c* oxidases usually have 3 or 4 different subunits (76, 77). Subunit I, which contains the heme group and subunit II are responsible for electron transport (76, 77). The cytochrome a_1-like hemoprotein from *M. magnetotacticum* showed very little cytochrome *c* oxidase activity and larger amounts of it were found in magnetic cells versus non-magnetic cells (70). The protein belongs to the heme-copper oxidase superfamily since it has copper and heme binding sites in its subunits (76, 77). Subunits I (41 kDa) and II (17 kDa) are homologous to subunits present in typical cytochrome *c* oxidases (76, 77). However the amino acid residues responsible for O_2 reduction are not conserved in subunit I, as they are in other cytochrome a_1 proteins, suggesting that this molecule is not a terminal oxidase in O_2 respiration in *M. magnetotacticum* (70). The genes encoding for the subunits of the hemoprotein were later identified and sequenced (78). The gene encoding for subunit I is 377 bases upstream of a gene highly homologous with other genes encoding for heme *o* synthases. Heme molecules of A-, O- or B-type are generally found in the subunit I of cytochrome *c* oxidases. An authentic cytochrome *c* oxidase was later purified from *M. magnetotacticum* (71) and it was shown to be a different protein from the cytochrome a_1-like protein described earlier.

A cytochrome cd_1 was also described in *Magnetospirillum magnetotacticum* (73). This protein was found in a concentration five times higher in the wild-type than in a spontaneous non-magnetotactic mutant obtained when *M. magnetotacticum* was cultured microaerobically in the presence of nitrate (same growth conditions used for the wt) (73). Like all cd_1-type cytochromes, this proteins shows nitrite and oxygen reductase activities (76,

77). However, the spectral properties of this cytochrome showed to be different of other cytochromes containing *cd* heme groups, suggesting that the environment surrounding the heme binding site for the *M. magnetotacticum* cytochrome cd_1 may be different from other *cd*-heme containing cytochromes (73). Furthermore, different electron donors, including a ferrocytochrome *c*-550 (from *M. magnetotacticum*) and succinate (both were expected to reduce described cytochrome cd_1 proteins), were tested and none reduced this cytochrome cd_1. The facts that this cytochrome is not expressed at the same levels in magnetotactic cells versus non-magnetotactic ones, that it does not present some spectral properties of cytochromes containing *cd* heme groups and that this protein does not oxidize usual electron donors such as ferrocytochrome *c*-550 or succinate suggest that this cytochrome cd_1 may not be involved in the respiratory chain used for denitrification. This cytochrome cd_1 exhibits a novel Fe^{2+} : nitrite oxidoreductase activity suggesting that its main role might be Fe^{2+} oxidation to Fe^{3+} (73). According to this theory, nitrite would be a sink for electrons not from the electron transport chain for the generation of energy for the cell (the normal function of this enzyme) but from Fe^{2+} (73). Nitrite reduction could be used for energy generation as well via a different cytochrome cd_1, as there is strong evidence that indeed cells of *M. magnetotacticum* use nitrate for energy production. Cytochrome cd_1 involved in denitrification generally oxidizes not Fe^{2+} but ferrocytochrome *c*-550 or succinate (76, 77)). Magnetite contains both oxidized and reduced iron and this enzyme might provide Fe^{3+}.

An Fe^{3+} reductase was purified from *Magnetospirillum magnetotacticum* (74). The amount of enzyme activity in the soluble fraction of the cells was reduced when the amount of iron quinate was decreased in the growth medium. Cells grown with an initial Fe^{3+} – quinate concentration of 20 μM in the growth medium produced an average of about 15

magnetosomes per cell and the soluble fraction from these cells had a specific Fe^{3+} – reductase activity of 10 nm Fe^{2+} –ferrozine formed per min per mg of protein. When the initial concentration of Fe^{3+} –quinate was dropped to 1 μM, the number of magnetosomes decreased to an average of about 9 magnetosomes per cell and the specific Fe^{3+} –reductase activity of the soluble fraction to 7 nm Fe^{2+} –ferrozine formed per min per mg of protein. Final cell yields were not affected. Ferric iron reductases are common in prokaryotes but their activity is generally not dependent on extracellular iron concentration (79, 80). Non-magnetotactic cells of *M. magnetotacticum* separated from magnetotactic cells grown in the same culture showed only 50% of the Fe^{3+} reductase activity of that of magnetotactic cells (74). Fe^{3+} reductase activity in the soluble fraction of cells was inhibited by zinc. The addition of Zn^{2+} to the growth media resulted in decreased Fe^{3+} reductase activity and magnetosome production. At 1 μM of Zn^{2+} (and 20 μM Fe^{3+} –quinate), cells produced an average of about 18 magnetosomes per cell and 16% of the cells had no magnetosomes. At 75 μM Zn^{2+} the number of magnetosomes per cell decreased to about 6 and 48% of the cells had no magnetosomes (74). Very high concentrations of Zn^{2+} (>200 μM) inhibited cell growth (74). These observations suggest that the Fe^{3+} reductase is required for magnetosome synthesis as well as for vital metabolic functions.

In *Magnetospirillum magnetotacticum,* the ferric reductase could be involved in the reduction of Fe^{3+} to Fe^{2+} after iron internalization by the cell which may involve siderophores (36). The cytochrome cd_1 could be the link between denitrification and Fe^{2+} to Fe^{3+} conversion since magnetite is composed of Fe^{2+} and Fe^{3+} and cytochrome cd_1 presents Fe^{2+} : nitrite oxidoreductase activity. Components of the magnetite synthesis pathway could serve as electron sinks. Supporting this evidence is a study conducted earlier where denitrification,

at optimal levels, in *M. magnetotacticum*, was shown to be dependent in the presence of low levels of dioxygen (68). Several other magnetosome membrane proteins in *M. magnetotacticum* with molecular weight of 12 kDa, 22 kDa, 28 kDa and possibly others might be responsible for controlling the chemical environment favoring magnetite synthesis inside the magnetosome membrane vesicle thereby directing crystal growth and nucleation (81). This model takes into account several factors and observations including denitrification, the microaerobic respiratory chain, microaerobic cell requirements and magnetosome synthesis and places them within a single model as summarized in Fig. 8 (81).

A simpler explanation for the lack of magnetosomes synthesis at high levels of $[O_2]$ is based on what is known about synthesizing magnetite inorganically. Inorganic magnetite can be formed at high pH (9-10) under anaerobic conditions by the addition of Fe^{2+} and Fe^{3+} (82). At high levels of $[O_2]$, magnetite can be synthesized inorganically but the high levels of O_2 result in the formation of amorphous, poorly-ordered and partly oxidized crystals (82). These crystals do not show the characteristics nor the magnetic properties seen in the crystals produced when O_2 is not present. Therefore, one can imagine that high O_2 levels inhibit the formation of crystals found in the magnetosome since these are well formed.

Even taking into account some known physiological features of *Magnetospirillum magnetotacticum*, magnetosome synthesis remains poorly understood in this species as well as in others. Functions of the proteins implicated in magnetosome synthesis must be determined to further refine a model for magnetosome synthesis. There is little agreement on which proteins are required for magnetite synthesis for different strains. Thus there is a need to characterize more proteins and to elucidate their functions if we are to better understand magnetosome synthesis and its importance for the magnetotactic bacteria.

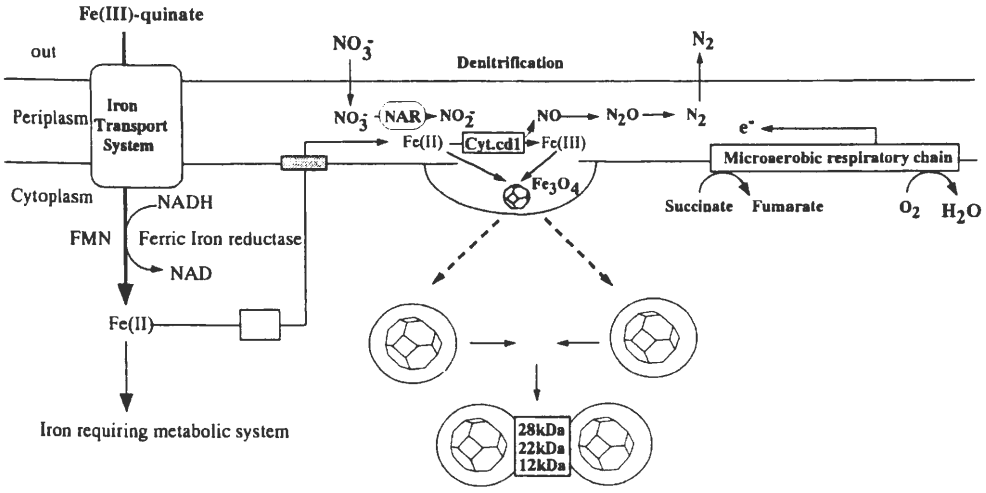


Fig. 8. Representation of magnetite pathway synthesis proposed by Fukumori (81). On the left, Fe^{3+} would be internalized from outside the cell to the cytoplasm by an iron transport system. Fe^{3+} would then be reduced to Fe^{2+} by a ferric iron reductase. In the upper center, the cytochrome cd_1 coupling denitrification with iron oxidation, cytochrome cd_1 would oxidize Fe^{2+} while reducing nitrite. In the bottom center, 28 kDa, 22 kDa and 12 kDa represent magnetosome membrane proteins possibly involved in magnetite synthesis, by regulating crystal nucleation and morphology. On the right, electrons for Fe^{3+} reduction could serve as an alternative electron sink for the respiratory chain instead reducing O_2. NAR: nitrate reductase. Figure extracted from 81.

2.4 Strains used in this study

2.4.1 Strain MC-1

Strain MC-1 was the first and is the only magnetotactic coccus isolated in pure culture (54) (Fig. 9). This strain is an obligately microaerophilic, Gram-negative, marine bacterium that was isolated water collected from an estuary in Rhode Island, USA. Cells of MC-1 are bilophotrichous; that is, they are motile by means of two flagella bundles inserted in one side of the cell (54). The bacterium grows chemolithoautotrophically with thiosulfate or sulfide as an electron source under microaerobic conditions. O_2 is the only known terminal electron acceptor for this strain and cells do not grow anaerobically. Magnetite crystals synthesized by MC-1 have an elongated hexahedral prismatic morphology (54). MC-1 cells contain an average of about 14 magnetosomes per cell and each magnetosome is approximately 80 nm in diameter on average (54). MC-1 is phylogenetically affiliated with the α-subdivision of the Proteobacteria (Fig. 1). As previously discussed, strain MC-1 exhibits polar magneto-aerotaxis (39).

2.4.2 MV-1 strain

Strain MV-1 is a vibrioid marine bacterium isolated from mud collected from a salt marsh pool near the Neponset River Estuary (Boston, MA, USA) (Fig. 9). Cells of MV-1 are motile by means of a single polar unsheathed flagellum (83). The bacterium grows microaerobically with O_2 or anaerobically with N_2O as the terminal electron acceptor. This strain grows chemoorganoheterotropically with organic or amino acids as carbon and energy sources and chemolithoautotrophically with thiosulfate and sulfide as electron sources. Cells of MV-1 contain an average of about 10 magnetosomes per cell. Magnetite crystals produced

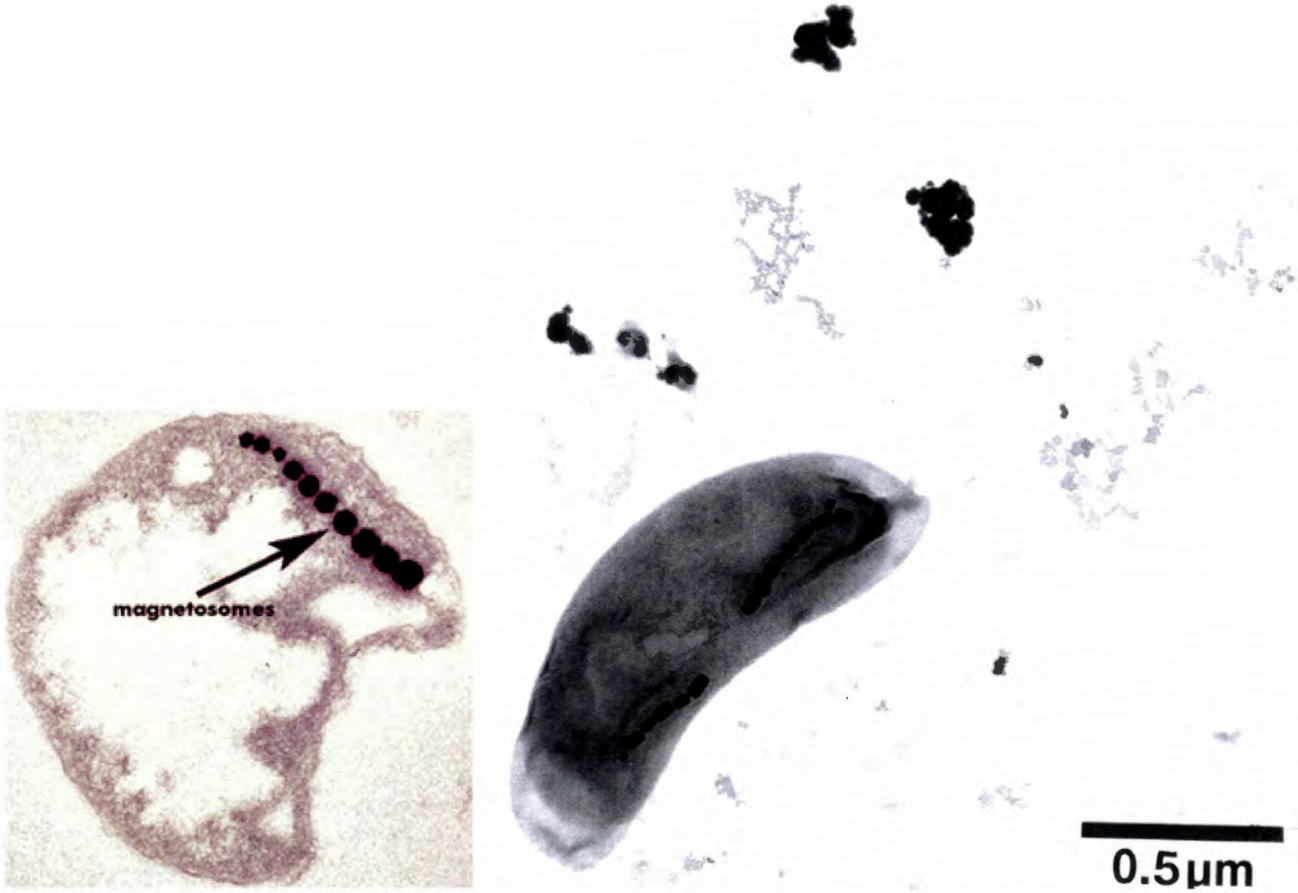


Fig. 9. Electron micrograph of bacterial strains used in the study here described. On the left, strain MC-1. On the right, strain MV-1.

by this strain have an elongated, hexa-octahedral morphology and are approximately 53 × 35 nm on average (57). MV-1 is phylogenetically classified in the α-subdivision of the Proteobacteria (Fig. 1).

Strain MV-1nm1 is spontaneous a non-magnetotactic mutant derived from strain MV-1 that does not produce magnetosomes. This mutant was isolated from a cream-colored colony grown in shake tubes with N_2O under anaerobic conditions. Strains MV-1nm2 to MV-1nm7 are also non-magnetotactic mutants isolated from agar plates (84)

A recent work on strain MV-1 unveiled a copper binding protein named p19 (84). This protein was present in the soluble fraction of wt MV-1 but not in the soluble fraction of MV-1nm1. This protein was hypothesized to be involved in iron transport contributing to magnetosome formation in this strain (84).

A study initiated by a former graduate student (Karen K. Paul) used strain MV-1 as the model organism. A brief description of these previous experiments with MV-1 and MV-1nm1 is here described. Karen constructed a genomic subtraction library using the procedures outlined in the CLONTECH PCR-Select Bacterial Genome Subtraction Kit (Clontech In., CA, Palo Alto). Genomic DNA from both strains was digested with the restriction endonuclease *Rsa*I (New England Biolabs, Beverly, MA). An adaptor was ligated to the wt MV-1 DNA fragments and the DNA from both strains allowed to hybridize under stringent conditions according to manufacturer's instructions. After mixing and hybridization of DNA sequences common to both strands, the DNA ends were filled in, thus adding distinct priming sites to strands that were exclusively composed of wt DNA. PCR was used to amplify the DNA fragments of the wt that did not hybridize to the genomic DNA of the mutant. The amplified DNA was purified and ligated to pGem-T Easy Vector (Promega,

Madison, WI) following the manufacture's instructions and the constructs transformed into competent *Escherichia coli* DH5α cells using a standard procedure (91). Transformed *E. coli* was plated on the appropriate selective media and about 200 transformants were randomly selected and their plasmids extracted according to Carter *et al.* (93).

Plasmids were digested with an appropriate endonuclease and the presence of DNA inserts was verified by gel electrophoresis. DNA fragment inserts were purified from the gel and were used as probes in Southern blot experiments using genomic DNA of wt MV-1 and MV-1nm1. The probes that annealed exclusively to the wt DNA were then used as probes to screen a cosmid wt MV-1 DNA library (86). Out of seven DNA fragments that hybridized only to wt MV-1 genomic DNA in Southern blot experiments, six hybridized to a single cosmid (designated clone 5C3) in the genomic library of wt MV-1. The cosmid was purified and the insert digested with *Hind*III (New England Biolabs), gel-separated and used as a target for Southern blot experiments using the six DNA fragments as probes. A *Hind*III DNA fragment of about 4 kb hybridized to all six probes and was sequenced. Cosmid 5C3 was sequenced by primer walking from the sequence of 4 kb *Hind*III fragment. DNA sequences were assembled using Vector NTI software (InforMax Inc.)

3. STATEMENT OF THE PROBLEM

The main focus of the work presented in this thesis is directed at how magnetite magnetosomes are synthesized in two marine, magnetotactic bacterial strains: the vibrio MV-1 and the coccus MC-1.

I used several different approaches to attempt to determine how magnetite is biomineralized in magnetotactic bacteria and what the important biochemical components are in this process. In general, using molecular and biochemical techniques, I made comparisons between wt MV-1 and the spontaneous, non-magnetotactic mutant derived from it, strain MV-1nm1 (84), which does not produce magnetosomes. These comparisons can be made at different levels. Protein profiles of the different cell fractions of the strains can be compared using either 1- or 2-D SDS-PAGE. Therefore, 2-D SDS-PAGE of wt MV-1 and MV-1nm1 cell fractions was performed. 2-D SDS-PAGE of magnetosome membrane proteins of wt MV-1 and 1-D SDS-PAGE of magnetosome membrane proteins of MC-1 were also performed. Specific proteins, such as those present only in wt MV-1 or MV-1nm1, were chosen for N-terminal sequencing and were used in designing DNA primers or probes to identify the corresponding genes.

Studies involving mRNA can be used to compare expression between wt MV-1 and MV-1nm1. This is a reasonable approach in that it has already been shown that specific proteins produced by wt MV-1 are not synthesized by MV-1nm1 (84). Disadvantages to this approach are that RNA is very unstable and the preparation can be contaminated with DNA. However, I attempted to construct a subtractive cDNA library to determine specific genes expressed in wt MV-1 and not MV-1nm1.

Finally, comparisons can also be made at the genomic level. This approach is especially desirable when a deletion is suspected to be responsible for the magnetotactic phenotype. However, this approach overlooks changes in the regulation of gene expression. A previous graduate student (Karen K. Paul) constructed a subtractive genomic library between wt MV-1 and MV-1nm1 and determined that several DNA fragments were missing from the genome of MV-1nm1. I continued this work.

Random amplification of polymorphic DNA (RAPD) analysis was also carried out to compare the genetic relatedness between wt MV-1, MV-1nm1 and several other non-magnetotactic mutant strains isolated by a former graduate student (Bradley L. Dubbels). This technique and the related technique, arbitrarily primed PCR (AP-PCR), are used to distinguish closely related strains of bacteria (85).

Two other unrelated experiments were performed. A novel experiment aiming to synthesize magnetite (or its preceding chemical or mineral intermediates) in vitro using cell-free extracts of MV-1 was performed. Finally, in order to establish a tractable genetic system in MV-1, preliminary experiments with electroporation were carried out.

4. MATERIALS AND METHODS

4.1 Bacteria, growth conditions. Cells of wt MV-1 and a non-magnetotactic mutant derived from it, strain MV-1nm1, were grown heterotrophically as previously described (86).

Strain MC-1, a marine magnetotactic coccus, was grown autotrophically in liquid cultures under microaerobic conditions as previously described (41) except that the ferric quinate was replaced with 25 μM $FeSO_4 \cdot 6H_2O$ added as an anaerobic stock solution (10 mM in 0.01 N HCl) after the bottles of growth medium became reduced (turned colorless). Sterile pure O_2 was then injected into the cultures to about 1% of the headspace. Cultures were not shaken so an [O_2] gradient formed in the liquid media. As cells increased in numbers and consumed O_2, they tolerated higher [O_2] and additional O_2 was injected by syringe to higher headspace values.

4.2 Preparation of cell-free extracts and soluble and membrane fractions. Cells of wt MV-1, MV-1nm1 and MC-1 were harvested in late log phase by centrifugation at 10,000 x g at 4°C for 20 min. Cells were then washed once with dilute artificial sea water (50) buffered with 20 mM Tris•HCl, pH 7.2. Cells were recentrifuged and resuspended to about 5×10^{11} cells/ml in 50 mM Tris•HCl (pH 7.2). Cells were lysed by three passages in a French pressure cell at 18,000 lb/in^2. Unbroken cells and cell debris were removed by centrifuging the cell lysate as described above. The cell-free extract (supernate) was either used directly or ultracentrifuged at 150,000 x g at 4°C for 2 h to separate soluble and membrane fractions. After ultracentrifugation, the supernate was taken as the soluble fraction and the membranes

(the resultant pellet) were Dounce homogenized and resuspended in 50 mM Tris•HCl, pH 7.2, to a volume corresponding to the volume of the soluble fraction.

4.3 Isolation of magnetosomes. Cells of MC-1 or MV-1 were harvested from approximately 70 l of culture by centrifugation at 10,000 x g at 4°C for 20 min. Cells were washed once with ice cold dilute artificial sea water buffered with 20 mM Tris•HCl, pH 7.1, recentrifuged and stored at -20°C. Cells were lysed as described above. Magnetosomes were separated from lysed cells using a gap magnet as previously described (44). Magnetosomes in tubes within the gap magnet accumulated at the magnetic poles and the cell lysate was carefully removed with a Pasteur pipette and used as a cell-free extract. Magnetosomes were then resuspended in ice cold 20 mM Tris•HCl, pH 7.1, and then placed again in the gap magnet until they accumulated at the magnetic poles at which time the buffer was removed and discarded. Magnetosomes were repeatedly resuspended and magnetically separated in this way at least 25 times. Magnetosomes were then resuspended in 1 M NaCl in 10 mM Hepes buffer, pH 7.1, to remove adventitious, electrostatically charged proteins (15). Magnetosomes were then washed in Hepes buffer without NaCl 15 more times. Magnetosomes were finally resuspended in 1% SDS in 10 mM Hepes buffer, pH 7.1, and allowed to sit with occasional shaking for approximately 2 h to allow for complete extraction of the magnetosome membranes. The mixture was centrifuged at 10,000 x g at 4°C for 20 min to remove the magnetosome magnetite crystals. The supernate, which contained the magnetosome membrane lipids and proteins, was removed and stored at -80°C.

4.4 Protein analysis and N-terminal sequencing. Proteins were separated using the NuPAGE (Invitrogen Inc., Carlsbad, CA) gel system. Gels were either blotted onto polyvinylidene fluoride (PVDF) membranes (Millipore, Billerica, MA) or stained with Coomassie Brilliant Blue R-250. Electrophoresis and blotting were carried out per the manufacturers' instructions. N-terminal sequences of proteins were determined by the Protein Facility at Iowa State University, USA, using a Perkin Elmer Applied Biosystems Model 494 Procise protein/peptide sequencer (Foster City, CA) with an on-line Perkin Elmer Applied Biosystems Model 140C PTH Amino Acid Analyzer according to the method developed by Edman (87). Protein was measured by the method of Bradford (88).

Reagents and equipment used for 2-D SDS-PAGE were from Amersham Biosciences unless specified otherwise. Cells of wt MV-1 and MV-1nm1 were grown as described earlier and washed once with artificial sea water (50), resuspended in lysis solution (7 M urea containing 4% CHAPS, 2% pharmalyte, 2 M thiourea and 60 mM dithiothreitol and 0.002% bromophenol blue), lysed using the French press and centrifuged at 10,000 x g at 4°C for 20 min to remove magnetosomes and unbroken cells. For 2-D SDS-PAGE gels of magnetosome membrane proteins, magnetosomes were purified as described earlier with the exception of that the SDS used for extraction of the magnetosome membranes was replaced by the lysis solution above. Protein concentrations of cell fractions were determined with the PlusOne 2-D Quant Kit. 2-D SDS-PAGE was carried out using the method of O'Farrell (89). The first dimension was performed with immobilized dry strip gels (pH 3-10; 7 cm, Bio-Rad, Hercules, CA) on an IPGPhor electrophoresis system. The strips were rehydrated in 125 µl of 8 M urea containing 0.5% CHAPS, 0.2% dithiothreitol containing 0.5% IPG buffer, 0.002% bromophenol blue and 60 µg of cell protein at 20°C for 10 h. The strips were run at 20°C

using the following program: step 1, 500 V for 1.5 h; step 2, 1,000 V for 1.5 h; step 3, 2,000 V for 1 h; step 4, 4,000 V for 1 h; step 5, 6000 V for 1 h; step 6, 8000 V (hold) for 7 h. The final volt/hours was about 50,000. IPG strips were stored at -80°C until used for running the gel in the second dimension. IPG strips were equilibrated in 10 ml of 1 X MOPS buffer containing 2% SDS, 6 M urea, 30% glycerol, 0.002% bromophenol and 100 mg dithiothreitol for 15 min at room temperature. A second equilibration step was carried out in the same solution for the same incubation time but 2.5% (w/v) of iodoacetamide replaced the dithiothreitol. Strips were loaded onto NuPAGE 4-12% Bis-Tris ZOOM Gels (Invitrogen Inc.). A sealing gel solution consisting of 0.5% agarose, 0.002% bromophenol blue and 1X MOPS buffer was used to seal IPG strip in place. BenchMark Pre-Stained Protein molecular weight markers (Invitrogen Inc.) were used to estimate the molecular weight of cell proteins. Electrophoresis in the second dimension was carried out as described above. Blotting, staining and N-terminal sequencing were carried out as described earlier. Blotted membranes and gels were scanned with an Amersham Image Scanner and analyzed using the PDQuest software program (Biorad, Hercules, CA).

4.5 DNA extraction. Genomic DNA was extracted from all bacterial strains according to the method of Kimble *et al.* (90)

4.6 PCR, cloning and sequencing of *mamC* from strain MC-1. PCR primers and annealing temperature were designed and predicted from the putative *mamC* gene of strain MC-1 (12) using Vector NTI software (InforMax Inc., Frederick, MD). PCR reactions (15 µl) contained: 16 mM $(NH_4)_2SO_4$, 67 mM Tris•HCl (pH 8.8), 0.01% Tween-20, 10 mM β-

mercaptoethanol, 100 μg bovine serum albumin per ml, 4 mM $MgCl_2$, 300 μM deoxynucleoside triphosphate mix, 100 ng of genomic DNA, 0.375 units of *Taq* polymerase (Bioline, Randolph, MA) and 30 pmol of primers. Primers used were: Sense 5'-GTCGACTCACAGATGGTTAGCGGGAA-3' and Antisense 5'-GTCGACGCTTGGCGCTATAAAGGAGA-3'. PCR conditions were as follows: first cycle: 3 min at 94°C, 1 min at 55°C and 1 min at 72°C. The next 23 cycles were 30 s at 94°C, 1 min at 55°C and 1 min at 72°C. The last cycle was 30 s at 94°C, 1 min 55°C and 7 min at 72°C. PCR was performed in a Stratagene (La Jolla, CA) RoboCycler Gradient 96 Temperature Cycler with Hot Top Assembly.

PCR products were electrophoresed in an agarose gel, separated, and then ligated into pGEM T Easy Vector (Promega, Madison, WI) following the manufacture's instructions. Plasmids were introduced into competent *Escherichia coli* DH5α cells by transformation using a heat shock protocol (91). Clones were selected for antibiotic resistance and mass cultured. Plasmids were isolated as described (91). Sequencing was performed by the DNA Facility at Iowa State University, USA, using an ABI Prism 377 sequencer (Foster City, CA) following the manufacturer's instructions.

4.7 Testing for in vitro magnetite synthesis by cell-free extracts. Soluble iron, as an anaerobic solution of 10 mM $FeSO_4 \cdot 7H_2O$ dissolved in 0.01 N HCl, was added to cell-free extracts to a final concentration of 450 μM Fe. Reactions were carried out in closed 1.5 ml microfuge tubes under aerobic and anaerobic conditions (under O_2-free N_2). Tubes were incubated at 25°C for varying amounts of time. When necessary, reaction mixtures were centrifuged for 10 min at 18,000 x g at 25°C to concentrate products. Approximately 20 μl of

the reaction mixtures were loaded into an appropriate chamber in a Super Quantum Interference Device (SQUID) magnetometer to determine the magnetic properties of the precipitates. The positive control was magnetosome magnetite collected during cell lysis. The material was tested at temperatures between 1.8° K and 300° K under a magnetic field strength of 0.1 T (tesla).

4.8 Random amplification of polymorphic DNA (RAPD) products. RAPD was carried out using standard procedures (92). Genomic DNA was quantified using a fluorometer (Turner Quantech Digital Filter Fluorometer, Dubuque, IA) and PicoGreen dsDNA Quantitation Kit (Molecular Probes, Eugene, OR). Three primers were used: GTAAGCCTAC (RAPD-1), GGTGAAGCGG (RAPD-2) and GCCATATTGC (RAPD-3) were used. PCR reactions were set up as described as in the *mamC* section with the following exceptions: 25 ng of genomic DNA, 2.5 mM $MgCl_2$, 1 unit of *Taq* polymerase and 12.5 pmol of primer were used and the reaction had a final volume of 25 µl. PCR was performed in the Stratagene RoboCycler thermocycler under the following conditions: the reactions were heated for 2 min at 94°C and then exposed to 45 cycles of 30 s at 93°C, 1 min at 36°C and 1 min at 72°C, followed by an elongation period of 5 min at 72°C. PCR products were loaded in a 2% agarose gel and electrophoresed under a field of 6 V/cm for 2 h 45 min. Gels were stained with ethidium bromide (1 µg/ml) for 5 min and washed for 15 min in distilled water. DNA fragments were visualized using a Gel Documentation System (Biorad).

4.9 Electroporation. Cells were grown and harvested by centrifugation as described earlier with the following differences. MV-1 cells were harvested at mid log phase at a cell

concentration of about 8 x 10^8 cells/ml. Cells were washed twice in cold electroporation buffer (10 mM TES (pH 7.3) containing 1 mM $MgCl_2$ and 272 mM sucrose). Cells were resuspended in electroporation buffer to a concentration of about 8 x 10^{11} cells/ml after centrifugation. About 200 ng of plasmid DNA was added to the cells which were then incubated for 15 min at 4°C in the electroporation cuvette. Electroporation was carried out in a 1 mm gap cuvette with 18-18.5 kV/cm (voltage per cm), 1475 Ω (resistance) and 25 μF (capacitance). Following electroporation, cells were incubated at 30°C for 1 h and inoculated into appropriate growth media (86). Appropriate antibiotics were added after 1, 2, 3, 6, 12, 18 or 24 h. The plasmids used in electroporation experiments are shown in Table 1. Plasmids were isolated using the protocol described by Carter *et al.* (93).

4.10 Construction of a subtractive cDNA library of MV-1. Total RNA was isolated and handled as described by Hepinstall (91, 99). Formaldehyde gel electrophoresis was used to confirm the integrity of the RNA (99). PCR reactions using RNA were carried out as described by Becker *et al.* (100) to check for DNA contamination. RNA was quantified spectrophometrically and used to synthesize cDNA (reagents and enzymes necessary were purchased from Invitrogen Inc.) according to the manufacture's instructions. cDNA was quantified using a fluorometer as described earlier (RAPD section) and used for the hybridization procedure as described in (100). cDNA from wt MV-1 and MV-1nm1 was digested with *Dpn*II (New England Biolabs) and ligated (Invitrogen Inc.) to the specific adapters. PCR reactions of the ligated DNA were carried out separately for each strain. After this step, amplified DNA was digested again with *Dpn*II and the wt cDNA fragments ligated to a different adapter and hybridized with MV-1nm1 cDNA. This was followed by a

Table 1.

Plasmids used in the electroporation experiments of strain MV-1.

Plasmids	Antibiotic Selection marker	Size	Reference
pDSK519	Kanamycin	8.1 kb	94
pCPP46	Tetracycline	10.8 kb	95
pME6010	Tetracycline	8.3 kb	96
pME6041	Kanamycin	5.6 kb	96
pFAJ1701	Tetracycline	12.4 kb	97
pFAJ1702	Tetracycline	9.7 kb	97
pKD46	Ampicillin	6.1 kb	98
pKD20	Ampicillin	6.1 kb	98

selective PCR step, where wt DNA fragments were amplified using primers complementary to the previously ligated adapters. The product of the PCR reaction was called differential product one, which was further digested, ligated, hybridized and amplified by PCR, yielding differential product two. Another round was carried out. The differential products were electrophoresed in 2% agarose gels. cDNA fragments that were present in the third differential product but not in the other two were sequenced. Sequences were BLAST searched and used to design primers for PCR amplification to test for the presence of the same sequences in the cDNA and the genomic DNA of both wt MV-1 and MV-1nm1.

5. RESULTS

5.1 Protein analyses of cell fractions of strain MC-1. 1-D SDS-PAGE of solubilized proteins from the magnetosome membrane and cell membrane fractions of strain MC-1 showed somewhat similar patterns and seemed to contain some of the same proteins based on apparent molecular weight. The soluble fraction clearly contained a number of proteins not present in either membrane fraction.

The protein profile of the magnetosome membrane protein fraction contained at least several proteins unique to that fraction, one with an apparent molecular weight of about 14 kDa (Fig. 10). The N-terminal sequence of this 14 kDa protein from the magnetosome membrane fraction was SIFNLALYLS which matched the amino acid sequence derived from the *mamC* gene of strain MC-1 (35).

PCR amplification using primers designed from this putative *mamC* gene (12) yielded an expected product of approximately 500 bp as shown in Fig. 10. The sequence of the product was identical to the sequence of the putative *mam*C gene (35).

5.2 2-Dimensional gel electrophoresis of strain MV-1. Comparisons between results from 2-D SDS-PAGE of wt MV-1 and MV-1nm1 cell-free extracts revealed the presence of several proteins unique to the wt (Fig. 11, A and C). One of these proteins had an apparent molecular weight of 20 kDa and an isoelectric point of about 4.5 (similar to a Cu-binding protein known as p19 described from the same strain (84)) and appeared to be produced in relatively high amounts based on its staining intensity. Another protein with an apparent molecular weight of 20 kDa but with an isoelectric point of about 6.75 was also unique to the

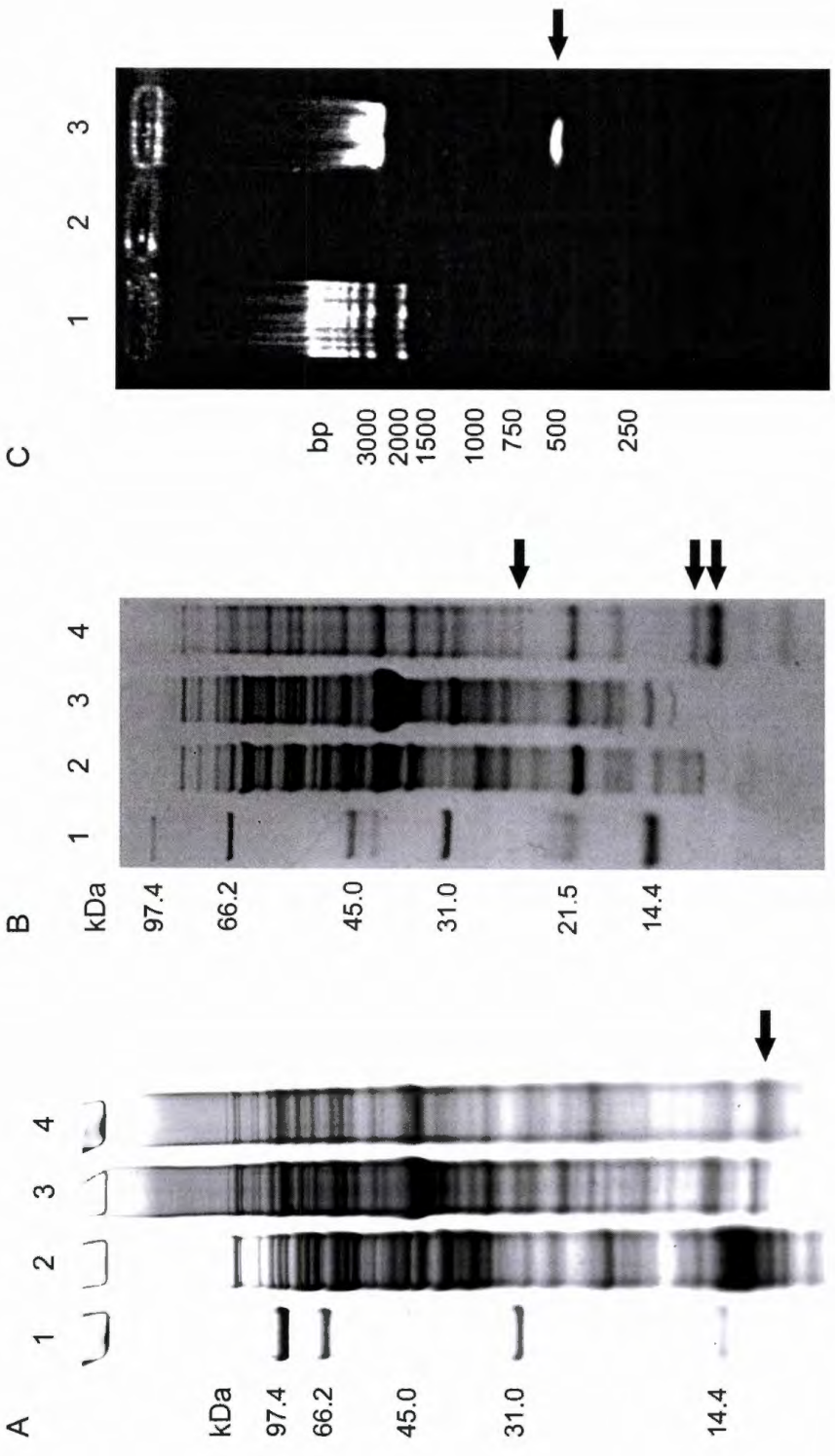


Fig. 10. **A** and **B**: Protein profiles of cell fractions of strain MC-1 (this work) and MV-1 (84), respectively, using 1-D SDS-PAGE. Lanes 1, molecular markers; lanes 2, soluble fractions; lanes 3, membrane fractions; and lanes 4, magnetosome membrane fractions. Arrow in **A** indicates the magnetosome membrane protein selected for N-terminal sequencing, ultimately shown to be MamC. Arrows in **B** show unique bands to the magnetosome membrane fraction of MV-1 **C**: PCR product resulting from use of primers designed for the *mamC* gene (lane 3). Lanes 1 and 2: molecular weight markers and nothing, respectively.

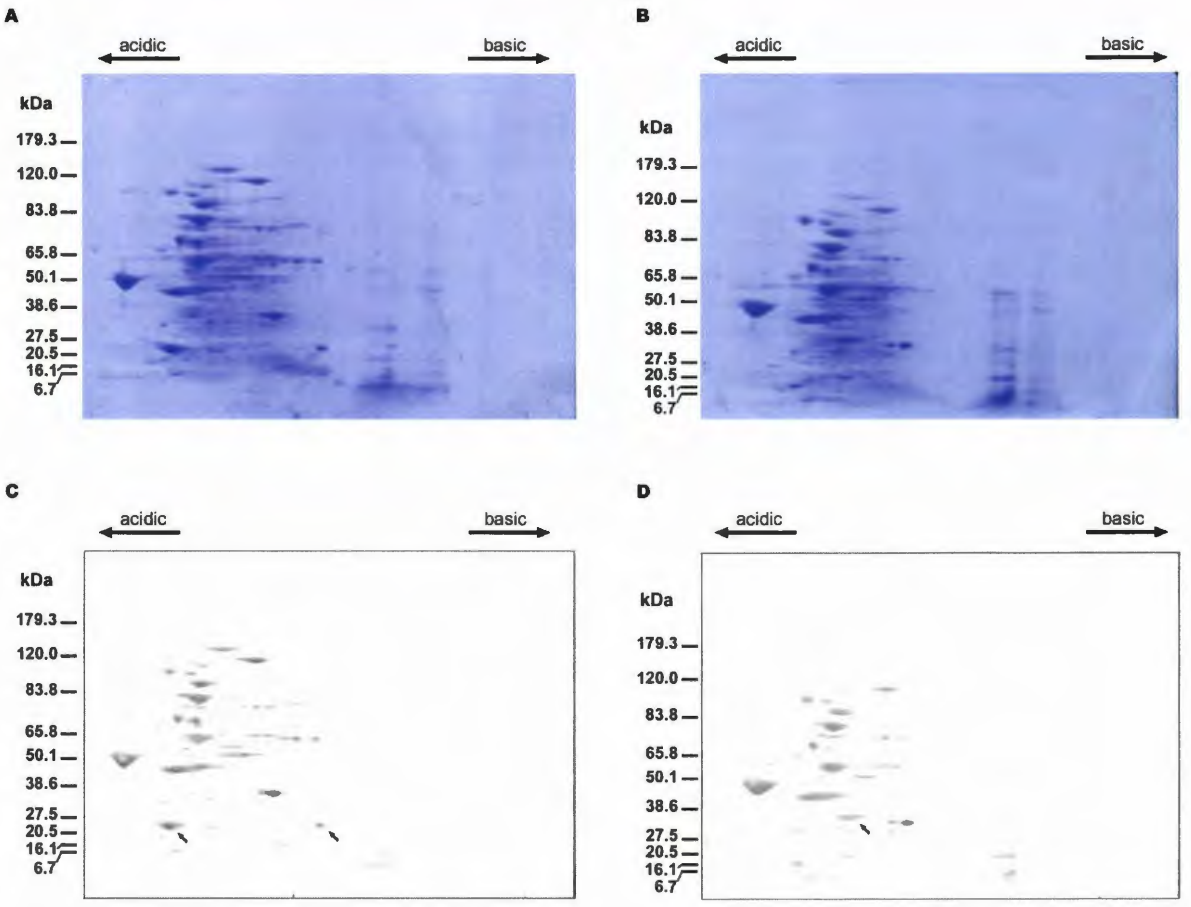


Fig. 11. Protein profiles of cell-free extracts resolved by 2-D SDS-PAGE of wt MV-1 and the spontaneous non-magnetotactic mutant MV-1nm1. **A** and **B**, protein profiles of wt MV-1 and MV-1nm1, respectively. **C** and **D** same as **A** and **B** with background filtered, respectively. Arrows in C and D indicate proteins unique to wt MV-1 and MV-1nm1, respectively, selected for N-terminal sequencing.

wt. The N-terminal sequence for this latter protein was determined to be: ?DMTISQPGARASA(?)GMAKKG?AFMTL. The protein profile of strain MV-1nm1 showed that it also produced some proteins not produced by wt MV-1. One of these appeared to be synthesized in relatively high amounts and had an apparent molecular weight of about 30 kDa and isoelectric point of about 5.2 (Fig. 11, B and D). The N-terminal sequence for this protein was determined to be MENAHSTPTPQ?Q(?)FF(?). Magnetosome membrane proteins of wt MV-1 were well resolved by 2-D SDS-PAGE (Fig. 12). Unfortunately, we did not have enough material for N-terminal sequencing of these proteins. BLAST searches did not yield significant homologies (with expected value below 10).

5.3 Confirmation of a deletion in MV-1nm1. Seven DNA fragments were identified from the selective amplification reaction of DNA present in wt MV-1 but not MV-1nm1. These were used as probes in Southern blot analyses and six of them hybridized to a single clone (designed 5C3) from a genomic cosmid library of wt MV-1. Cosmid 5C3 was digested with *Hind*III and the resulting DNA fragment used as a target for Southern blot experiments using the six DNA fragments that hybridized to clone 5C3 as probes. All six probes hybridized to the same ~4 kb *Hind*III DNA fragment. The 4 kb *Hind*III fragment was subcloned and sequenced. Once this was finished, further sequencing of cosmid 5C3 was determined using primer walking. Sequencing revealed transposases, hemerythrins and other genes as represented in Figure 13. The presumed deletion in the MV-1nm1 was demonstrated by PCR of the hybrid HK gene in both wt and mutant strains. Only wt DNA was positive for this gene. The end of the deletion was also determined by PCR of 300 bp near the ISS66 Family Tpase ORF3 which was present in both wt and mutant (Fig. 14).

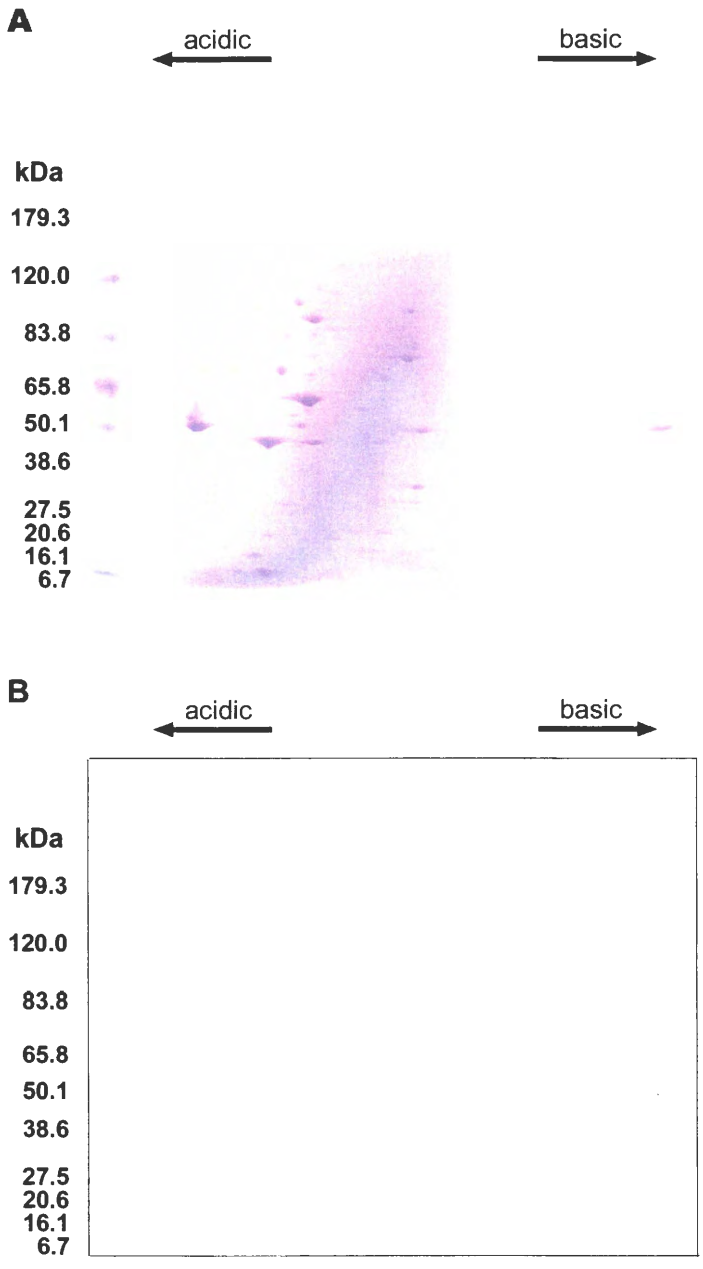


Fig. 12. 2-D protein profile of magnetosome membranes of strain MV-1. **A** and **B**, without and with background filtered.

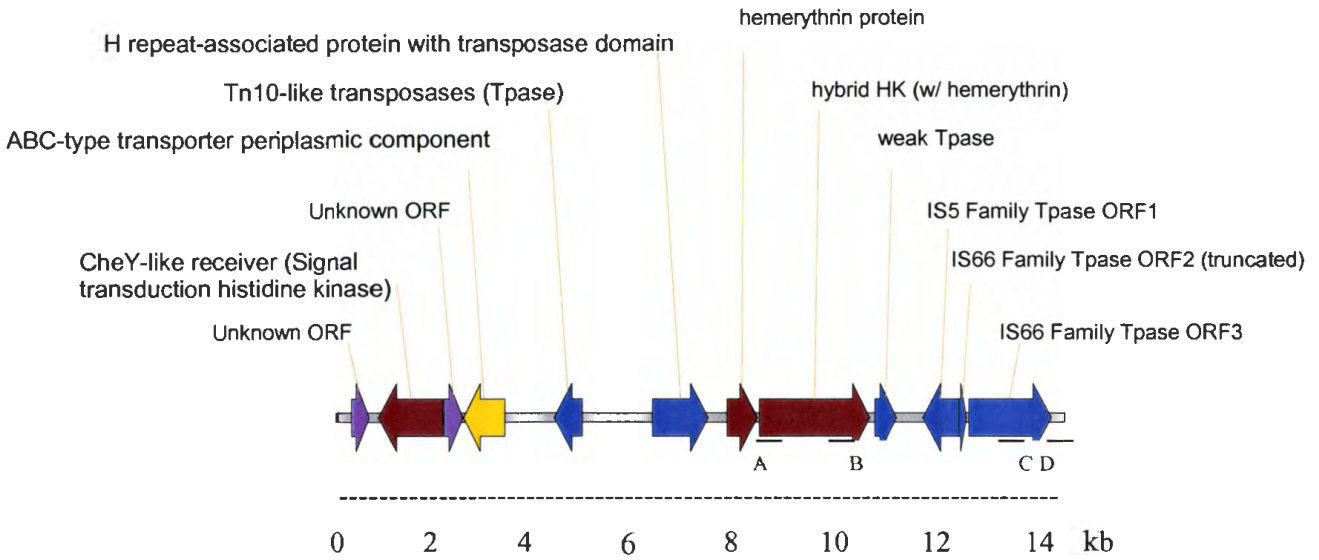


Fig. 13. Partial map of the putative deletion in wt MV-1 determined by primer walking of cosmid 5C3 and BLAST searches. Arrows depict open reading frames (ORFs). Scale bar indicates size of the sequenced area in kb. Blue arrows depict ORFs related to insertion sequences and transposases. Dark red arrows depict genes related to the histidine kinase signaling system. Yellow arrow depicts the ORF homologous to an ABC transport system, periplasmic component. Purple arrows depict ORFs with very low homology to known genes in databases. Small dashes above capital letters depict the locations where primers used to confirm the deletion (A and B) and its end (C and D) were targeted, PCR products from those reactions are shown in figure 14.

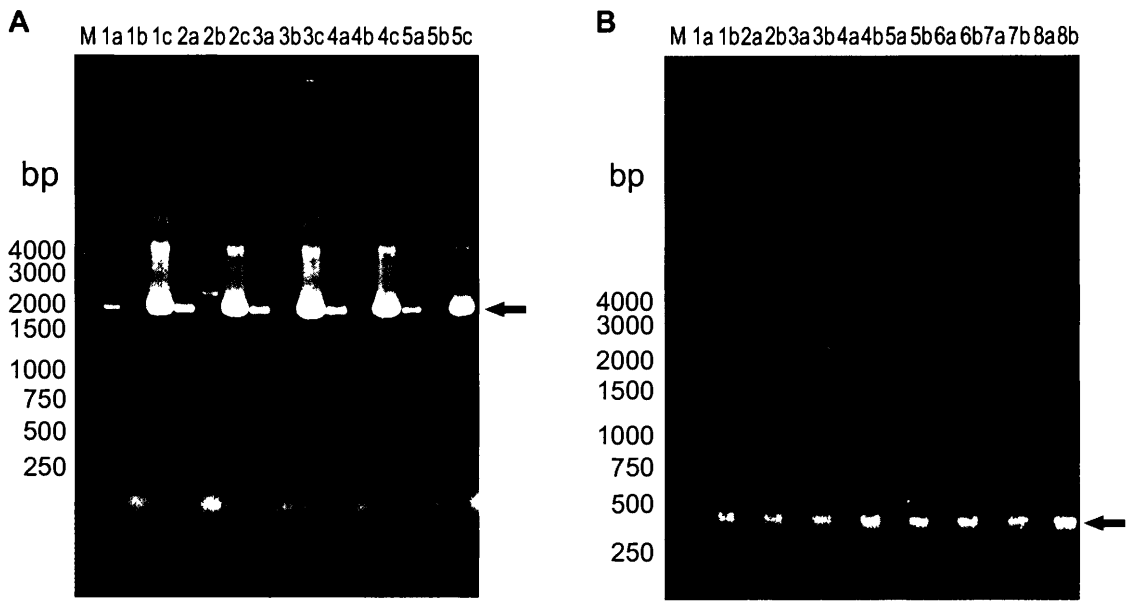


Fig. 14. **A**: PCR products of hemerythrin-like gene; *a*, *b* and *c*: MV-1 wt, MV-1nm1 and 5C3 cosmid DNAs respectively, temperature gradient ranging from 51°C to 65°C by 3.5°C is depicted using numbers from 1 (51°C) to 5 (65°C). Arrow depicts the expected band of 2kb. **B**: PCR products of one end of the putative deletion; *a* and *b*: wt MV-1 and MV-1nm1 DNA. Numbers 1 to 8 indicate the temperature gradient from 72°C to 61°C by 1.6°C. M: molecular markers.

A deletion in the MV-1nm1 strain was confirmed by PCR of a putative hemerythrin-like gene that appeared to be only in the wt MV-1. A DNA fragment near to the hemerythrin-like gene was demonstrated by PCR to be present in both strains. PCR reactions were carried out as described for the amplification of *mamC* with the following exceptions: primers 5'-TGAATATTAACCAACAACGC -3' and 5'- TCACAATGATGTCGTAAGGC – 3' were used to demonstrate the presence of the deletion and primers 5' – CCCCATTGTGCGGATTTGGAA - '3 and 5' – ATGGTTCAGGAGGCTGCTGGAC – '3 were used to demonstrate one end of the deletion. A positive control consisting of the cosmid DNA that was used as a template for the sequencing reactions was assessed in the PCR reactions with the non-magnetic mutant MV-1nm1 and the wt DNA and the hemerythrin primers.

5.4 cDNA subtractive library. The construction of a cDNA subtractive library between wt MV-1 and MV-1nm1 resulted in the selection of a number (four) of DNA fragments presumed to represent cDNA from mRNAs only formed by wt MV-1. The fragments were cloned and sequenced and most contained ribosomal DNA fragments. The only fragment that had sequence that was not homologous to ribosomal RNA genes showed homology to a gene encoding subunit I of "cytochrome a_1-like" hemoprotein (cytochrome *c* oxidase) from *Magnetospirillum magnetotacticum*. This hemoprotein has been implicated in magnetosome synthesis in this organism (71). Sequencing of the flanking regions of this DNA fragment resulted in the sequence of an open reading frame (ORF) that appeared to represent the complete gene for subunit I as well as a second ORF with a high homology to the subunit II gene of this protein. These putative genes were organized as in *Magnetospirillum*

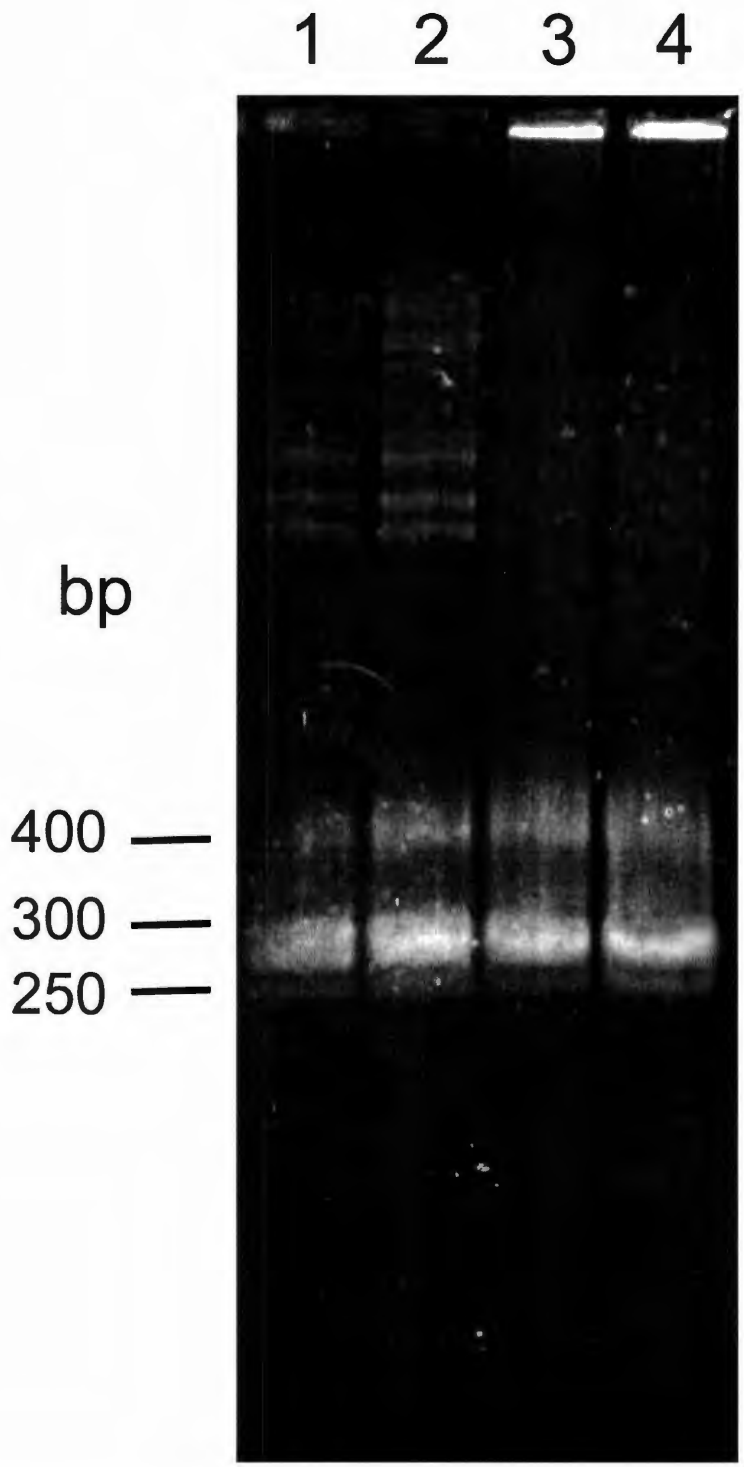


Fig. 15. Products from the amplification of a 300 bp DNA fragment from a subunit I "cytochrome a_1-like" hemoprotein using genomic DNA and cDNA from MV-1 and MV-1nm1 as templates. Lanes 1 and 2: wt MV-1 and MV-1nm1 genomic DNA, respectively; lanes 3 and 4: wt MV-1 and MV-1nm1 cDNAs, respectively.

magnetotacticum (71). Amplification by PCR using primers 5'-TCAAGTGGAACGCGGTTACC-'3 and 5'-ATCAAAGCGCCGACAT-'3 specific for a 300 bp fragment of subunit I of "cytochrome a_1-like" hemoprotein showed it to be present in both wt MV-1 and MV-1nm1 cDNA and genomic DNA (Fig. 15).

5.5 RAPD analyses of wt MV-1 and a series of MV-1 non-magnetotactic mutants.

RAPD analyses of wt MV-1 and the non-magnetotactic mutant strain, MV-1nm1 (Fig. 16A) resulted in different patterns of amplified DNA fragments by PCR using random primers. For example, a PCR product from MV-1 with an apparent size of a little greater than 750 bp was not amplified in MV-1nm1 when primer RAPD-1 was used in both directions (Fig. 16A, 1a and 2a). PCR amplification with primer RAPD-1 in both directions yielded few bands for wt MV-1 and MV-1nm1 although differences can still be observed (Fig. 16A, 1a and 2b).

RAPD patterns of three different sets of primers and genomic DNA from six spontaneous non-magnetic mutants (designed MV-1nm2 to MV-1nm7) as template are shown in Figure 16B. These spontaneous mutants were isolated from agar plates as previously described (84). Amplification using primer RAPD-1 distinguished four of the six strains used. MV-1nm2 appears to be different from all the other mutants since it lacks all bands present in the other strains but two bands of about 2 kb (Fig. 16B, lane 1a). MV-1nm3 differs from the others since it shows a band that has a little less than 750 bp, which is not present in any of the other strains (Fig. 16B, lane 2a). MV-1nm4 and MV-1nm5 appear to be similar to each other but different from the other four mutants since the two lack all bands but one of about 500 bp (Fig. 16B, lanes 3a and 4a). MV-1nm6 and MV-1nm7 have a similar band pattern but are different from the other four since they lack the band about 750 bp

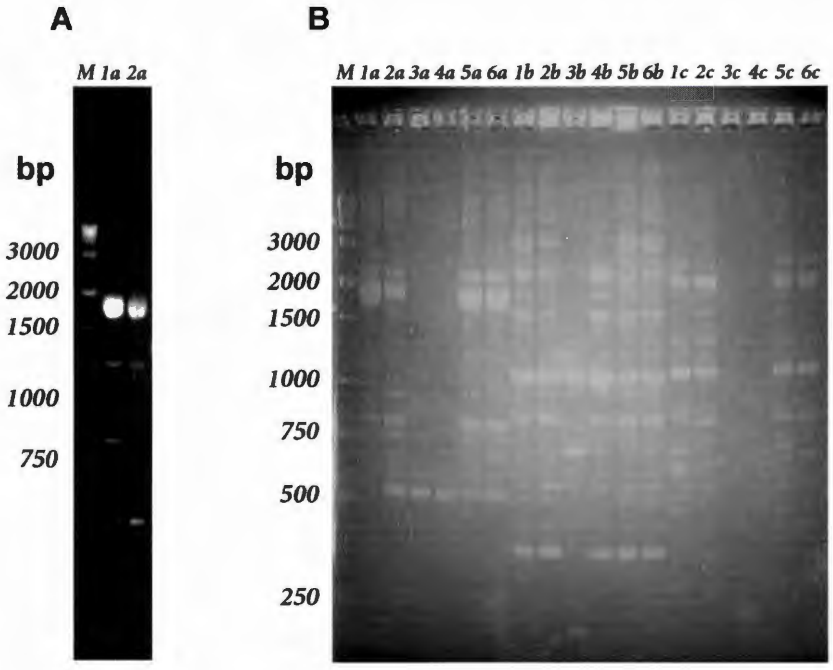


Fig. 16. Randomly amplified polymorphic DNA analysis (RAPD) of: **A**: Profiles of wt strain MV-1 (1*a*) and non-magnetotactic mutant strain MV-1nm1 (2*a*) using the primer GTAAGCCTAC. M represents DNA fragment size markers in bp. **B**: 1-6, 6 additional non-magnetotactic mutants isolated from agar plates named MV-1nm2 to MV-1nm7, respectively. The random primer GTAAGCCTAC (RAPD-1) was used both in the forward and reverse directions in reaction depicted in the *a* lanes, primer GGTGAAGCGG (RAPD-2) was used in the reactions depicted in the *b* lanes and primer GCCATATTGC (RAPD-3) was used in the reaction depicted in the *c*. Note that the mutant strains show several different patterns of PCR products indicating that some of the mutant strains are not genetically identical.

(which is present in lane B 2a for MV-1nm3) but possess all other bands (Fig. 16B, lanes 5a and 6a). For primer RAPD-2 all mutant strains showed a similar pattern with the exception of MV-1nm4 which is similar to MV-1nm5 for primer RAPD-1, but for primer RAPD-2 there is a band shift from a 750 bp zone to a smaller band (Fig. 16B, lane 3b). This was the only strain to show this shift. Primer RAPD-3 failed to show any band for MV-1nm4 and MV-1nm5 and all other strains showed a similar pattern of amplified DNA fragments.

5.6 Electroporation of MV-1 cells with different plasmids. Cells of wt MV-1 survived in the TES buffer supplemented with sucrose and through the voltage applied in the electroporation procedure as evidenced by light microscopy and cell growth in medium without selective antibiotic. Moreover, the cells were used as an inoculum in growth medium and grew after 3-5 days of incubation. The cells which were inoculated into different antibiotic containing media grew after about a month, presumably after the antibiotics, which were bacteriostactic, broke down chemically. Plasmid isolation from electroporated cells of MV-1 failed to show any evidence that a plasmid was internalized and had replicated.

5.7 In vitro magnetite synthesis by cell-free extracts of strain MV-1. When an iron sulfate solution was added to crude cell-extracts of both the wt MV-1 and its non-magnetic MV-1nm1 a dark, almost black material precipitated. The reaction did not occur with cell-free extracts of *Escherichia coli* and *Pseudomonas aeruginosa*. This black material generally formed at the lower portion of the tube when the reaction was performed aerobically and the precipitate formed more quickly when the experiment was performed under anaerobic conditions. The black precipitate displayed no magnetic properties typical of magnetite or

any other ferrimagnetic mineral as determined using a SQUID magnetometer. The reaction mixtures had a relatively strong odor characteristic of hydrogen sulfide and thus the precipitate was likely some form of iron sulfide.

6. DISCUSSION

In this section, each set of experiments and results will be discussed and placed in a broader perspective. First, I will address the protein analyses of strain MC-1 and the identification of MamC as a major magnetosome membrane protein in this bacterium. This will be followed by a discussion of the protein analyses of strain MV-1 and the non-magnetotactic mutant MV-1nm1. Genomic and gene expression differences between MV-1 and MV-1nm1 will be then described, followed by a discussion of the results of a number of unrelated experiments including RAPD analyses of MV-1 strains, electroporation of MV-1 and magnetite formation in vitro by cell-free extracts of strain MV-1.

6.1 MC-1 protein profile.

The goal of analyzing protein profiles of MC-1 strain cell fractions was to identify and characterize proteins, particularly magnetosome membrane proteins, involved in the synthesis of the magnetosome. 1-D SDS-PAGE protein profiles of the soluble, membrane and magnetosome membrane were compared. The general membrane fraction (contains both the outer and plasma membranes) and the magnetosome membrane fraction were somewhat similar and appeared to share a number of the same proteins based on apparent molecular weights (Fig. 10). It is possible that some contamination of the magnetosome fraction might have occurred with the general membrane fraction although this seems doubtful considering how well the magnetosomes were washed. Alternatively, large pieces of outer and/or plasma membrane may have enveloped the magnetosomes after lysis. A final possibility is that both membrane fractions actually do share a relatively large number of proteins. This may not be

surprising since the magnetosome membrane most likely originates from the plasma membrane. A similar result was found for strain MV-1 (Fig. 10, adapted from 84) although protein profiles of the magnetosome membranes of *Magnetospirillum gryphiswaldense* do not seem to show this pattern (35). It is possible that the different methods of magnetosome purification might be responsible for these findings. Nonetheless, an abundant and apparently unique protein to the magnetosome membrane fraction with a molecular weight of about 14 kDa was selected for N-terminal sequencing. The N-terminal sequence of this protein showed homology to MamC proteins of *M. gryphiswaldense* and *M. magnetotacticum* (12) and matched the amino acid sequence encoded by a putative *mamC* gene from MC-1 (35, 55). The *mam*C gene has thus far only been detected in the genomes of strain MC-1, *M. gryphiswaldense* and *M. magneticum* strain AMB-1 (35, 55).

The finding that MamC is a magnetosome membrane protein in MC-1 supports the theory that there are conserved genes and proteins responsible for magnetite formation in unrelated magnetotactic bacteria. This is noteworthy in this case because strain MC-1 is not closely related to *M. magnetotacticum* and *M. gryphiswaldense* in which MamC was found to be present as a magnetosome membrane protein (35). They nevertheless share the common trait of magnetotaxis and the fact that they all contain MamC as a major magnetosome membrane protein might suggest that they all utilize the same process for magnetosome synthesis. Supporting this line of thought is that: 1) *mamC* appears to be present only in magnetotactic bacteria (35) and 2) *mamC* was found to be missing in a non-magnetotactic mutant of *M. gryphiswaldense* and MV-1nm1 along with many other genes implicated in magnetosome formation (62). The role of MamC in magnetite synthesis and possibly in other cellular functions is unknown, and it will likely take site-directed mutagenesis studies using a

genetic system like that developed by Schultheiss *et al.* (101) for *M. gryphiswaldense* to determine the function of MamC.

6.2 MV-1 2-D gel profiles.

2-D SDS-PAGE of cell-free extracts of strain wt MV-1 and MV-1nm1 was performed with the intention of identifying specific proteins produced by wt MV-1 and not by the non-magnetotactic mutant MV-1nm1 or vice-versa. Three proteins were detected on 2-D gels that were unique to wt MV-1 or MV-1nm1 (Fig. 11). Another protein unique to wt MV-1 has an apparent molecular weight of about 27 kDa. A protein with an apparent molecular weight of about 33 kDa was only found in MV-1nm1 (Fig. 11, C and D). The N-terminal sequences of these latter two proteins were determined. Neither sequence showed significant homology to any sequence in various protein and gene databases searched to date. Magnetosome membrane proteins of wt MV-1 were also resolved by 2-D gel electrophoresis and none of the proteins, according to their molecular weight and isoelectric focusing point were similar to a proteins of the magnetosome membrane of strain *Magnetospirillum magneticum* Strain AMB-1 (55). (Fig. 12).

6.3 MV-1 genomic subtractive library.

Dean (102), using pulsed-field gel electrophoresis, showed that at least one mutation in MV-1nm1 is likely the result of a deletion of part of the genome. The purpose of constructing the subtraction genomic library was to identify and sequence the deletion and to determine which genes are missing from wt MV-1. Since we detected a relatively large DNA

fragment absent from MV-1nm1, we confirmed that at least one of the mutations in MV-1nm1 is a result of deleted DNA from the genome.

A number of insertion sequences (including transposons and transposases), mobile genetic elements, were found at the presumed end of the deleted fragment (Fig. 13). This might explain why the deletion occurred in the first place and also why non-magnetotactic mutants occur in this strain (84). It also explains why we have never observed a reversion to the magnetotactic phenotype in MV-1nm1.

Since deletion of large amounts of DNA occur in other magnetotactic bacteria (35), it might explain why bacteria that are not magnetotactic cluster together phylogenetically with bacteria that are magnetotactic (24). It is not difficult to conceive a situation in the environment where the mobile genetic elements that flank the genes responsible for the magnetotactic trait jump from species to species via horizontal transfer (103). This might result in the observed wide morphological and physiological diversity of magnetotactic bacteria (5).

Besides the insertion sequences and transposases found at the end of the deletion, ORFs with strong homology to specific genes were also found. Two putative genes that show high homology to genes that contain hemerythrin domain were found next to one other. The first of the hemerythrin-like genes has sequences with homology to different domains conserved in different proteins (Fig. 17). One of these domains has high homology with the hemerythrin domain involved in interactions with O_2 interaction (104). An account of the known function of these domains is given below. Some hemerythrin genes encode a non-heme protein that binds O_2 reversibly (104). Hemerythrins have been mainly described in certain marine invertebrates (105) where they appear to function in O_2 transport (106). Only one

hemerythrin protein from a prokaryote, *Desulfovibrio vulgaris* (Hildenborough), has been studied to some extent (107). This protein has a motif with five histidines and two carboxylate ligands to an oxo-/hydroxo-bridged diiron active site and an O_2 binding pocket that is also conserved in hemerythrins from other organisms (104). This motif is also conserved in the hemerythrin-like gene found in MV-1. The protein from *Desulfovibrio vulgaris* (Hildenborough) has some homology to the hemerythrin gene described in invertebrates and was characterized as a chemoreceptor for O_2 and this receptor, upon stimulus, would ultimately cause a change in frequency of flagellar motor reversal which would lead the bacteria to swim up or down in an O_2 gradient (108, 109).

Downstream of the hemerythrin domain sequence there is a sequence that has high homology to a histidine protein kinase receiver (Fig. 17). Histidine protein kinase domains have been shown to be part of a signaling system involved in specific responses to specific changes in the environment (110). This suggests that this domain in the hemerythrin-like protein could function as a switch that might that might initiate a change in the protein in response to specific stimuli. Motifs homologous to a histidines kinase ATPase and histidine kinase motifs were found upstream of the hemerythrin motif (Fig. 17). Moreover, a gene with a high homology to a histidine kinase response regulator (cheY-like) was found within the region of the deletion in close proximity to the hemerythrin-like gene. Another motif in the hemerythrin-like gene encodes for a PAC subdomain (Fig. 17). PAC subdomains are always found in PAS domains (111). PAS domains are involved in monitoring changes of O_2, light, redox potentials and the overall energy of a cell (111). PAS domains, unlike other sensor modules, are usually located in the cytosol (111). Small motifs at the end of the hemorithrin-like gene codes for two transmembrane regions (Fig. 17).

Fig. 17. Representation of hemerythrin-like gene depicting its different coding regions. Blue, motifs homologous to motifs responsible for histidine kinase signaling. Red, hemerythrin-like motif. Black, PAC motif. Pink, transmembrane motifs. Total size of the open reading frame is 2244 bp.

The hemerythrin-like domain is likely involved in O_2 mediation/binding. The other domains of the hemerythrin-like gene likely encode a series of protein modules possibly involved in signaling and sensing specific changes. Finally, there is a transmembrane region at the end of the protein. Although one can only speculate the functions of the hemerythrin-like gene in wt MV-1, it seems probable that the molecule functions as an O_2 chemoreceptor in magnetotactic bacteria, where it helps the cell to find and maintain an optimal position in an O_2 gradient. A similar gene was found in the genome of strain MC-1, another magnetotactic bacterium (12). The actual function of the hemerythrin-like genes in the MV-1 strain remains to be determined.

Another putative gene within the deletion encodes for the periplasmic component of the ABC-type transport system. The periplasmic component of these types of transporters usually transport the molecule within the periplasm to an ATP dependent channel localized in the inner membrane (112). It is possible that this periplasmic protein transports the molecule that is sensed by the hemerythrin-like gene product if this product doesn't actually sense O_2 if those two genes are related at all.

Within the deletion there is also an ORF encoding a protein with an h-repeat. H-repeats are encoded by the rearrangement hotspot elements (*Rhs*). These elements were discovered in many *Escherichia coli* but not all strains. Evidence suggests that the proteins encoded by these elements are present in the cell wall membrane but their function remains unknown (113).

6.4 MV-1 cDNA subtractive library.

The goal of constructing the subtractive cDNA library was to identify genes expressed by wt MV-1 and not MV-1nm1, particularly those involved in magnetosome synthesis. Using the cDNA library protocol of Becker *et al.* (100), we were unable to isolate and identify cDNA fragments resulting from mRNAs produced solely by wt MV-1. The majority of the pieces sequenced were fragments of ribosomal RNA genes. Because we did not find any citations of this paper describing the use of this technique by others, it is difficult to evaluate the reliability of the technique. Only one group has applied the technique successfully and even in this case, the majority of the sequenced cDNA fragments were also of ribosomal RNA genes (100). It seems the most reliable method to study expression profiles of bacteria is the use of microarrays containing transcripts of a species (114). The technique we used might have provided a cost effective method to compare different expression profiles since the construction of microarrays is still relatively expensive.

Despite the problems involved with this technique, this experiment yielded a DNA sequence with significant homology to cytochrome a_1-like subunits I and II hemoprotein genes present in *Magnetospirillum magnetotacticum* (Fig. 18). These genes are of particular importance because they have been previously reported to be implicated in magnetite biomineralization (78). Since we were only supposed to retrieve mRNAs produced by the wt, it was surprising that the genes were present in the wt MV-1 and MV-1nm1 cDNAs (Fig. 15). Recovery of this gene fragment (belonging to subunit I) might have been possible if the gene for this hemoprotein was transcribed at higher levels in wt MV-1. This explanation is in accordance with previous work carried out with *M. magnetotacticum*, where a cytochrome

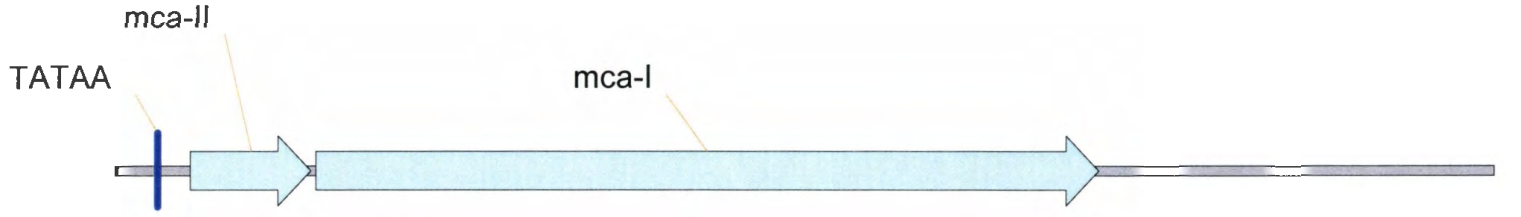


Fig. 18. Representation of DNA sequences from MV-1. Cytochrome a_1-like subunits I and II, mca-I and mca-II, respectively and a TATAA sequence. Grey bar represents total DNA sequenced (2980 bp).

a_1-like was found to be present in higher levels in the wt when compared to non-magnetotactic mutant cells (78).

Cytochrome a_1 subunits I and II hemoproteins are part of the cytochrome *c* oxidase (COX) protein that is involved in proton translocation across the inner membrane of respiring organisms. Cytochrome *c* is oxidized outside the inner membrane of bacteria and mitochondria, the electrons passed on to O_2 by cytochrome *c* oxidase resulting in the reduction of O_2 to water (76, 77). In exchange for the electrons pumped inside the cell to reduce O_2, protons are pumped outside the cell. This exchange creates a negative charge inside the cell, which is used to drive ATP formation.

As mentioned in the first section, Fukumori and co-workers isolated a cytochrome a_1-like protein from *Magnetospirillum magnetotacticum* (78) This cytochrome a_1 like protein lacked heme *a* (responsible for electron transfer) and showed almost no cytochrome *c* oxidase activity (78). Amino acids expected to be associated with heme *a* were conserved in this cytochrome a_1-like but lacked all amino acid residues proposed to participate in the oxygen-reduction reaction (78) as shown in Figure 19.

It is possible that this cytochrome a_1-like protein is pumping electrons in strain MV-1 not to O_2 but to Fe^{3+} reducing it to Fe^{2+}. Another possible scenario is that this molecule could be actually oxidizing Fe^{2+} to Fe^{3+}. Since magnetite is an oxide that requires both forms of iron, both scenarios are possible. Supporting this theory is the fact that none of the amino acid residues proposed to participate in the oxygen-reducing and the coupled proton pumping reaction in *Paracoccus denitrificans* cytochrome *c* oxidase were conserved in the MV-1 cytochrome a_1-like protein (Fig. 19). Nevertheless, those amino acid residues proposed to associate with metal in *P. denitrificans* in the same molecule are still conserved in this

Fig. 19. Comparison of amino acid sequences of MV-1 cytochrome a_1-like subunit I hemoprotein with those of heme-copper oxidase subunit I from different organisms. Amino acids that participate in heme binding are colored in red. Green residues have unknown function and are conserved in strain MV-1 and *Magnetospirillum magnetotacticum* replacing amino acids involved in oxygen-reducing. Residues responsible for coupling proton pumping reaction are depicted in blue. Conserved residues in all strains are colored in yellow. Residues shown in pink are conserved in MV-1 and *M. magnetotacticum* and possible other organisms. M. mag, *M. magnetotacticum* (accession number: ab022644); P. den, *Paracoccus denitrificans*; H. hal, *Halobacterium halobium*; E. col, *Escherichia coli*; B. sub, *Bacillus subtillis*; Bovine, *Bos taurus*

protein. Furthermore, the amino acids that are reported to be responsible for O_2 interactions in *P. denitrificans* are conserved in the species *Magnetospirillum magnetotacticum* and in the strain MV-1 as shown in Figure 19. Studies regarding functionality of this protein are required to prove or disprove this theory.

6.5 MV-1 RAPD analyses.

RAPD analyses of 6 different non-magnetotactic mutants clearly show that most are genetically different (Fig. 16). This suggests that different mutations can occur in strain MV-1 that lead to a non-magnetotactic phenotype. This is important since different experiments attempting to compare wt MV-1 with the mutant MV-1nm1 at different levels (DNA, cDNA and proteins) and the results of such comparisons should be analyzed with care. For instance, the deletion described previously is clearly one of several possible different deletions.

6.6 MV-1 genetic system.

Attempts at introducing foreign DNA as small, broad-host range plasmids into cells of strain MV-1 by electroporation were not successful. The reason for this is unknown although increasing the amount of plasmid DNA might be effective as published in a recent work (101). However, I was able to compose a buffer for electroporation in which the cells survived the procedure. This lays a foundation for future experiments with electroporation. There may be more effective methods, such as conjugation, for the introduction of foreign DNA into strain MV-1.

6.7 In vitro magnetite synthesis using MV-1 cell free extracts.

The goal behind the experiments designed to produce magnetite in vitro using cell-free extracts of strain MV-1 as a catalyst was to find specific components such as proteins that could be tested for their involvement in magnetite biomineralization. Different combinations of proteins could be tested and those that catalyzed the formation of magnetite could be screened and used in further studies. This approach could isolate the cellular components necessary to biomineralize magnetite.

Unfortunately, attempts to develop an assay for in vitro magnetite biosynthesis failed. The black precipitate that formed was likely FeS, which is black, resulting from the reaction of Fe in the iron sulfate solution and H_2S hypothetically produced by the cell-free extract. Because the reaction occurred with both wt MV-1 and MV-1nm1 extracts, it appears that this reaction is not relevant in magnetite biomineralization.

Nevertheless, it is interesting that H_2S was formed in the reaction since strain MV-1 does not grow with sulfate as a terminal electron acceptor. We did not determine whether the S from sulfide was derived from the S in SO_4^{2-} added as the ferrous sulfate solution. Thus, a significant finding from this experiment would be that cell-free extracts from strain MV-1 reduce sulfate to sulfide, which remains to be proven experimentally.

Bacteria reduce sulfate to sulfide in two separate pathways: 1) in dissimilatory sulfate reduction where sulfate is used as a terminal electron acceptor for energy conservation and 2) in assimilatory sulfate reduction, where sulfate is used as sulfur source for biomass (108). In assimilatory sulfate reduction, the sulfate is reduced to sulfide at the proper pH, which reacts with free protons forming hydrogen sulfide. In vivo, hydrogen sulfide does not form as a free intermediate and sulfide is assimilated as cysteine (108).

Alternatively, what we observed may be a new, undescribed reaction involving another biochemical pathway. The fact that cell-free extracts do not regularly form a blackish precipitate nor produce hydrogen sulfide even when iron is present (Dispirito, A. A, unpublished data) support this notion. The fact that the cell-free extracts of the mutant strain of MV-1nm1 reacts forming a blackish precipitate does not rule out that this reaction may be related to magnetite formation. This conclusion can be drawn since proteomic differences between wt MV-1 and MV-1nm1 are not yet established, thus maybe magnetite is not formed in the mutant because it lacks one or more components of the magnetite synthesis machinery not necessarily all of them. And these components that are still present maybe playing some role in the described reaction.

6.8 Conclusion.

The experiments here described served as a step towards the general goal of deciphering how magnetosomes are synthesized. Different research groups around the globe are working in this same question and since magnetotactic bacteria were described in 1975, this question has posed as one of the main enigmas to these research groups. Different mechanisms have been proposed taking into consideration different research results but the theories proposed are far from explaining the whole picture that one is beginning to appreciate. The RAPD and the subtractive experiments proved that the nature of the MV-1nm1 non-magnetic phenotype is not conserved among different mutants. The cDNA library with cytochrome a_1-like and the protein profile of strain MC-1 revealed genes and proteins homologous to genes and proteins that have been implicated in magnetosome synthesis. The 2 dimensional gels have pointed out different expressed proteins in a strain wt MV-1 and

MV-1nm1. And finally the electroporation experiment and the attempt to synthesize magnetite in vitro have not produced palpable results but may have opened some doors for future work.

7. REFERENCES

1. **Fernandes, M. R., S. A. Sanchez,** 2001. O magnetismo animal e seus aspectos ao longo do desenvolvimento da biologia e do magnetismo. Publicacao Interna (Universidade Estadual de Campinas, Campinas, SP, Brazil).

2. **York, D., R. M. Farquhar,** 1972. The Earth's Age and Geochronology, Oxford: Pergamon Press.

3. **Myers, J. S., J. L. Crowley,** 2000. Vestiges of life in the oldest rocks? A review of early Archean geology in the Godthabsfjord region, and reappraisal of field evidence for >3850 Ma life on Akilia. Precambrian Res; 103(3-4): 101-24.

4. **Blakemore, R. P.,** 1975. Magnetotactic bacteria. Science 190: 377-379

5. **Mann, S., N. H. C. Sparks, R. G. Board,** 1990. Magnetotactic bacteria: Microbiology, Biomineralization, Paleomagnetism and Biotechnology. Advances in Microbiol Physiology 31: 125-179.

6. **Walker, M. M., C. E. Diebel, C. V. Haugh, P. M. Pankhurst, J. C. Montgomery, C. R. Green,** 1997. Structure and function of the vertebrate magnetic sense. Nature 390: 371-376.

7. **Kirschvink, J. L., J. L. Gould,** 1981. Biogenic magnetite as a basis for magnetic field detection in animals. Biosystems; 13: 181-201.

8. **Frankel, R. B., R. P. Blakemore,** 1989. Magnetite and magnetotaxis in microorganisms. Bioelectromagnetics; 10: 223-37.

9. **Kirschvink, J. L., A. Kobayashi-Kirschvink, B. J. Woodford,** 1992. Magnetite biomineralization in the human brain. Proc Natl Acad Sci USA; 89:7683-7.

10. **Kirschvink,J. L., M. M. Walker, C. E. Diebel,** 2001. Magnetite-based reception. Curr Opin Neurobiol 11:462-467.

11. **Thomas-Keprta, K. L., D. A. Bazylinski, J. L. Kirschvink, S. J. Clemett, D. S. McKay, S. J. Wentworth, H. Vali, E. K. Jr. Gibson, C. S. Romanek,** 2000. Elongated prismatic magnetic crystals in ALH84001 carbonate globules: potential Martian magnetofossils. Geochim Cosmochim Acta; 64: 4049-81.

12. **Joint Genome Institute, Department of Energy, United States of America Government,** 2004. Date retrieved: May 18th http://www.jgi.doe.gov/JGI_microbial/html/index.html

13. **Diebel, C. E., R. Proksch, R. C. Green, P. Neilson, M. M. Walker,** 2000. Magnetite defines a vertebrate magnetoreceptor. Nature 406, 20: 299-302.

14. **Mann, S., N. H. C. Sparks, M. M. Walker, J. L. Kirschvink,** 1988. Ultrastructure, morphology and organization of biogenic magnetite from sockeye salmon, *Oncohynchus nerk*: implications for magnetoreception. J Exp Biol; 140:35-49

15. **Blakemore, R. P., D. Maratea, R. S. Wolfe,** 1979. Isolation and pure culture of a freshwater magnetic spirillum in chemically defined medium. J Bacteriol; 140:720-729

16. **Bazylinski, D. A., B. M. Moskowitz,** 1997. Microbial biomineralization of magnetic iron minerals: microbiology, magnetism and environmental significance. Rev Mineral; 35:181-223.

17. **Bazylinski, D. A.,** 1995. Structure and function of the bacterial magnetosome. ASM News; 61:337-343.

18. **Fassbinder, J. W. E., H. Stanjek, H. Vali,** 1990. Ocurrance of magnetotactic bacteria in soil. Nature; 343:161-163.

19. **Mann, S., N. H. C. Sparks, R. G. Board,** 1990. Magnetotactic bacteria: microbiology, biomineralization, paleomagnetism and biotechnology. Adv Microb Physiol; 31:125-181.

20. **Petermann, H., U. Bleil,** 1993. Detection of live magnetotactic bacteria in south Atlantic deep-sea sediments. Earth Planet Sci Lett; 117:223-228.

21. **Sparks, N. H. C., L. Courtaux, S. Mann, R. G. Board,** 1986. Magnetotactic bacteria are widely distributed in sediments in U.K. FEMS Microbiol Lett; 37:308-308.

22. **Vali, H., J. L. Kirschvink,** 1990. Observations of magnetosome organization, surface structure and iron biomineralization of undescribed magnetotactic bacteria: evolutionary speculations. Iron Biominerals. New York: Plenum Press.

23. **Rodgers, F. G., R. P. Blakemore, N. A. Blakemore, R. B. Frankel, D. A. Bazylinski, D. Maratea, C. Rodgers,** 1990. Intracellular structure in a many-celled magnetotactic prokaryote. Arch. Microbiol; 154:18-22.

24. **Spring, S., K. H. Schleifer,** 1995. Diversity of magnetotactic bacteria. System Appl Microbiol 18:147-153.

25. **Pósfai, M., P. R. Buseck, D. A. Bazylinski, R. B. Frankel,** 1988. Iron sulfides from magnetotactic bacteria: structure, compositions, and phase transitions. *Am. Mineral.* **83**, 1469-1481.

26. **Fretas, F., C. N. Keim, B. Kachar, M. Farina, U. Lins,** 2003. Envelope ultrastructure of uncultured naturally occurring magnetotactic cocci. FEMS Microbiol Lett; 219:33-38.

27. **Sleytr, U. B., T. J. Beveridge,** 1999. Bacterial S-layers. Trends Microbiol; 7:253-260.

28. **Sleytr, U. B.,** 1997. Basic and applied S-layer research: an overview. FEMS Microbiol Rev; 7:253-260

29. **Dieluweit, S., D. Pum, U. B. Sleytr,** 1998. Formation of a gold superlattice on an S-layer with square lattice symmetry. Supramol Sci; 5:15-19.

30. **Spring, S., R. Amann, W. Ludwig, K. H. Schleifer, H. van Gemerden, N. Petersen,** 1993. Dominating role of an unusual magnetotactic bacterium in the microaerobic zone of a freshwater sediment. Appl Environ Microbiol; 59:2397-2403.

31. **Kawaguchi, R., J. G. Burgess, T. Sakaguchi, H. Takeyama, R. H. Thornhill, T. Matsunaga,** 1995. Phylogenetic analysis of a novel sulfate-reducing magnetic bacterium, RS-1, demonstrates its membership of the δ-Proteobacteria. FEMS Microbiol Lett; 126: 277-282.

32. **DeLong, E. F., R. B. Frankel, D. A. Bazylinski,** 1993. Multiple evolutionary origins of magnetotaxis in bacteria. Science; 259:803-806.

33. **Zavarzin, G. A., E. Stackebrandt, R. G. E. Murray,** 1991. A correlation of phylogenetic diversity in the Proteobacteria with the influences of ecological forces. Can J Microbiol; 37:1-6.

34. **Schübbe, S., M. Kube, A. Scheffel, C. Wawer, U. Heyen, A. Meyerdierks, M. H. Madkour, F. Mayer, R. Reinhardt, D. Schüler,** 2003. Characterization of a spontaneous nonmagnetic mutant of *Magnetospirillum gryphiswaldense* reveals a large deletion comprising a putative magnetosome island. J Bacteriol; 185:5779-5790.

35. **Grunberg, K., C. Wawer, B. M. Tebo, D. Schuler,** 2001. A large gene cluster encoding several magnetosome proteins is conserved in different species of magnetotactic bacteria. Appl Environ Microbiol;67:4573-82

36. **Bazylinski, D. A., R. B. Frankel,** 2004. Magnetosome formation in prokaryotes. Nat Rev Microbiol; 2:1-14.

37. **Blakemore, R. P.,** 1982. Magnetotactic bacteria. Ann Rev Microbiol; 36:217-238.

38. **Matsunaga, T.,** 1991. Applications of bacterial magnets. Trends Biotechnol; 9:91-95.

39. **Frankel, R. P., R. P. Blakemore, F. F. Torres de Araujo, D. M. S. Esquivel,** 1981. Magnetotactic bacteria at the geomagnetic equator. Science; 212:1269-1270.

40. **Bazylinski, D. A., R. B. Frankel, B. R. Heywood, S. Mann, J. W. King, P. L. Donaghay, A. K. Hanson,** 1995. Controlled biomineralization of magnetite (Fe3O4) and greigite (Fe3S4) in a magnetotactic bacterium. Appl. Environ. Microbiol; 61:3232-3239.

41. **Frankel, R. B., D. A. Bazylinski, M. Johnson, B. L. Taylor,** 1997. Magneto-aerotaxis in marine, coccoid bacteria. Biophys J; 73:994-1000.

42. **Taylor, B. L.,** 1983. How do bacteria find the optional concentration of oxygen? Trends Biochem Sci; 8:438-441.

43. **Balkwill, D. L., D. Maratea, R. P. Blakemore,** 1980. Ultrastructure of a magnetotactic spirillum. J Bacteriol; 141:1399-1408.

44. **Gorby, Y. A., T. J. Beveridge, R. P. Blakemore,** 1988. Characterization of the bacterial magnetosome membrane. J Bacteriol; 170:834-841.

45. **Lins, U., M. Farina,** 1998. Magnetosome size distribution in uncultured rod-shaped bacteria as determined by electron microscopy and electron spectroscopic imaging. Microsc Res Tech; 42:459-464.

46. **Frankel, R. B., R. P. Blakemore,** 1980. Navigational compass in magnetic bacteria. J Magn Magn Mater; 15-18:1562-1564.

47. **Heywood, B. R., D. A. Bazylinski, A. J. Garratt-Reed, S. Mann, R. B. Frankel,** 1990 Controlled biosynthesis of greigite (Fe3S4) in magnetotactic bacteria. Naturwiss; 77: 536-538.

48. **Towe, K. M., T. T. Moench,** 1981. Electron-optical characterization of bacterial magnetite. Earth Planet Sci Lett; 52:213-220.

49. **Schüler, D.,** 2004. Molecular analysis of a subcellular compartment: the magnetosome membrane in *Magnetospirillum gryphiswaldense*. Arch Microbiol; 181:1-7.

50. **Bazylinski, D. A., A. J. Garratt-Reed, R. B. Frankel,** 1994. Electron microscopic studies of magnetosomes in magnetotactic bacteria. Microsc Res Tech; 27:389-401.

51. **Golden, D. C., D. W. Ming, C. S. Schwandt, H. V. Laue, R. A. Socki,** 2001. A simple inorganic process for formation of carbonates, magnetite and sulfides in the Martian meteorite. ALH84001. Amer Mineral; 8:370-375.

52. **Butler, R. F., S. K. Banerjee,** 1975. Theoretical single-domain grain size range in magnetite and titanomagnetite. J Geophys Res; 80:4049-4058.

53. **Frankel, R. B.,** 1984. Magnetic guidance of organisms. Ann Rev Biophys Bioeng; 13:85-103.

54. **Meldrum, F. C., S. Mann, B. R. Heywood, R. B. Frankel, and D. A. Bazylinski,** 1993. Electron microscopy study of magnetosomes in a cultured coccoid magnetotactic bacterium, Proc R Soc Lond B, 251:231-236.

55. **Arakaki, A., J. Webb, T. Matsunaga,** 2003. A novel protein tightly bound to bacterial magnetite particles in *Magnetospirillum magneticum* strain AMB-1. J Biol Chem; 278: 8745-8750.

56. **Keim, C. N., U. Lins, M. Farina,** 2003. Elemental analysis of uncultured magnetotactic bacteria exposed to heavy metals. Can J Microbiol; 47:1132-1136.

57. **Sparks, N. H. C., L. Courtaux, S. Mann, D. A. Bazylinski, D. R. Lovley, H. W. Jannasch, R. B. Frankel,** 1990. Structure and morphology of magnetite anaerobically-produced by a marine magnetotactic bacterium and a dissimilatory iron-reducing bacterium. Earth Planet Sci Lett; 37:305-308.

58. **Nakamura, C., J. G. Burgess, K. Sode, T. Matsunaga,** 1995An iron-regulated gene, *magA,* encoding an iron transport protein of *Magnetospirillum* AMB-1. J Biol Chem 270; 28392-28396.

59. **Matsunaga, T., N. Tsujimura, Y. Okamura, H. Takeyama,** 2000. Cloning and characterization of a gene, *mpsA*, encoding a protein associated with intracellular magnetic particles from *Magnetospirillum* sp. strain AMB-1. Biochem Biophys Res Commun; 268:932-937.

60. **Okamura, Y., H. Takeyama, T. Matsunaga,** 2001. A magnetosome specific GTPase from the magnetic bacterium *Magnetospirillum magneticum* AMB-1. J Biol Chem; 276:48183-48188.

61. **Wahyudi, A. T., H. Takeyama, Y. Okamura, Y Fukuda, T. Matsunaga,** 2003. Characterization of aldehyde ferredoxin oxidoreductase gene defective mutant in *Magnetospirillum magneticum* AMB-1. Biochem Biophys Res Commun; 303:223-229.

62. **Schübbe, S., M. Kube, A. Scheffel, C. Wawer, U. Heyen, A. Meyerdierks, M. H. Madkour, F. Mayer, R. Reinhardt, D. Schüler,** 2003. Characterization of a spontaneous nonmagnetic mutant of *Magnetospirillum gryphiswaldense* reveals a large deletion comprising a putative magnetosome island. J Bacteriol; 185:5779-5790.

63. **Munro, A. W., G. Y. Ritchie, A. J. Lamb, R. M. Douglas, I. R. Booth,** 1991. The cloning and DNA sequence of the gene for the glutathione-regulated potassium-efflux system KefC of Escherichia coli. Mol Microbiol; 5:607-16.

64. **Pfanner, N., B. S. Glick, S. R. Arden, J. E. Rothman,** 1990. Fatty acylation promotes fusion of transport vesicles with Golgi cisternae. J Cell Biol; 110:955-61.

65. **Randazzo, P. A., Y. C. Yang, C. Rulka, R. A. Kahn,** 1993. Activation of ADP-ribosylation factor by Golgi membranes. Evidence for a brefeldin A- and protease-sensitive activating factor on Golgi membranes. J Biol Chem; 268:9555-63.

66. **Ostermann, J., L. Orci, K. Tani, M. Amherdt, M. Ravazzola, Z. Elazar, J. E. Rothman,** 1993. Stepwise assembly of functionally active transport vesicles. Cell; 75:1015-25.

67. **Rothman, J. E.,** 1994. Mechanisms of intracellular protein transport. Nature; 372:55-63.

68. **Blakemore, R. P., K. A. Short, D. A. Bazylinski, C. Rosenblatt and R. B. Frankel,** 1985. Microaerobic conditions are required for magnetite formation within *Aquaspirillum magnetotacticum*. Geomicrobiol J; 4:53-71.

69. **Bazylinski, D. A., R. P. Blakemore,** 1983. Denitrification and assimilatory nitrate reduction in *Aquaspirillum magnetotacticum*. Appl Environ Microbiol; 46:1118-1124.

70. **Tamegai, H., T. Yamanaka, Y. Fukumori,** 1993. Purification and properties of a 'cytochrome a1'-like hemoprotein from a magnetotactic bacterium, Aquaspirillum magnetotacticum. Biochim Biophys Acta; 1158:237-43.

71. **Tamegai, H., Y. Fukumori,** 1994. Purification, and some molecular and enzymatic features of a novel ccb-type cytochrome c oxidase from a microaerobic denitrifier, Magnetospirillum magnetotacticum. FEBS Lett; 347:22-6.

72. **Yoshimatsu, K., T. Fujiwara, Y. Fukumori,** 1995. Purification, primary structure, and evolution of cytochrome c-550 from the magnetic bacterium, Magnetospirillum magnetotacticum. Arch Microbiol; 163:400-6.

73. **Yamazaki, T., H. Oyanagi, T. Fujiwara, Y. Fukumori,** 1995. Nitrite reductase from the magnetotactic bacterium Magnetospirillum magnetotacticum. A novel cytochrome cd1 with Fe(II):nitrite oxidoreductase activity. Eur J Biochem; 233:665-71.

74. **Noguchi, Y., T. Fujiwara, K. Yoshimatsu, Y. Fukumori,** 1999. Iron reductase for magnetite synthesis in the magnetotactic bacterium Magnetospirillum magnetotacticum. J Bacteriol; 181:2142-7.

75. **Taoka, A., K. Yoshimatsu, M. Kanemori, Y. Fukumori,** 2003. Nitrate reductase from the magnetotactic bacterium Magnetospirillum magnetotacticum MS-1: purification and sequence analyses. Can J Microbiol; 49:197-206

76. **Kannt, A., H. Michel,** 2004. Handbook of Metalloproteins, Bacterial cytochrome *c* oxidase, U.S.A.: John Wiley & Sons.

77. **Garcia-Horsman, J. A., B. Barquera, J. Rumbley , J. Ma, R. B. Gennis,** 1994. The superfamily of heme-copper respiratory oxidases. J Bacteriol; 17618):5587-600.

78. **Tanimura, Y., Y. Fukumori,** 2000. Heme-copper oxidase family structure of Magnetospirillum magnetotacticum 'cytochrome a1'-like hemoprotein without cytochrome c oxidase activity. J Inorg Biochem; 82:73-8.

79. **Huyer, M., W. J. Page,** 1989. Ferric reductase activity in Azotobacter vinelandii and its inhibition by Zn^{2+}. J Bacteriol; 171:4031-7

80. **Leong, S. A., J. B. Neilands,** 1981. Relationship of siderophore-mediated iron assimilation to virulence in crown gall disease. J Bacteriol; 147:482-91.

81. **Baeuerlein, E.,** 2000. Biomineralization, Weinheim (Federal Republic of Germany): Wiley-VCH press.

82. **Schwertmann, U., R. M. Cornell,** 2000. Iron oxides in the laboratory, preparation and characterization, Weinheim (Federal Republic of Germany): Wiley-VCH press

83. **Bazylinski, D. A., R. B. Frankel, H. W. Jannasch,** 1988. Anaerobic magnetite production by a marine magnetotactic bacterium. Nature; 334:518-519.

84. **Dubbels, B. L., A. A. DiSpirito, J. D. Morton, D. J. Semrau, J. N. E. Neto, D. A. Bazylinski,** Evidence for a Copper-Dependent Iron Transport System in the Marine, Magnetotactic Bacterium Strain MV-1. Submitted.

85. **Baldy-Chudzik, K.,** 2001. Rep-PCR - a variant to RAPD or an independent technique of bacteria genotyping? A comparison of the typing properties of rep-PCR

with other recognised methods of genotyping of microorganisms. Acta Microbiol Pol; 50:189-204.

86. **Dean, A. J., D. A. Bazylinski,** 1999. Genome analysis of several marine, magnetotactic bacterial strains by pulsed-field gel electrophoresis. Curr Microbiol; 39:219-25.

87. **Edman P.,** 1950. Edman degradation. Acta Chem Scand; 4:283.

88. **Bradford, M. M.,** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem; 72:248-54.

89. **O'Farrel, P. H.,** 1975. High resolution two-dimensional electrophoresis of proteins. J Biol Chem; 250:4007-4021.

90. **Kimble, L. K. L., L. Mandelco, C. R. Woese, M. T. Madigan,** 1995. *Heliobacterium modestocaldum*, sp. nov., a thermophilic heliobacterium of hot springs and volcanic soils. Arch Microbiol; 163:259-267.

91. **Sambrook, J., E. F. Fritsch, T. Maniatis,** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press. New York, N.Y.

92. **Dassanayake, R. S., L. P. Samaranayake,** 2003. PCR Protocols, Random amplified polymorphic DNA fingerprinting: the basics. Humana Press, NJ.

93. **Carter, M. J., I. D. Milton,** 1993. An inexpensive and simple method for DNA purifications on silica particles. Nucleic Acids Res; 21:1044

94. **Keen, N. T., S. Tamaki, D. Kobayashi, D. Trollinger,** 1988. Improved broad-host-range plamids for DNA cloning in Gram-negative bacteria. Gene; 70:191-197

95. **Bauer, D. W., A. Collmer,** 1997. Molecular cloning, characterization, and mutagenesis of a pel gene from Pseudomonas syringae pv. lachyrmans encoding a member of the Erwinia chrysanthemi pelADE family of pectate lyases. Mol Plant Microbe Interact;10:369-79.

96. **Heeb, S., Y. Itoh, T. Nishijyo, U. Schnider, C. Keel, J. Wade, U. Walsh, F. O'Gara, D. Haas,** 2000. Small, stable shuttle vectors based on the minimal pVS1 replicon for use in gram-negative, plant-associated bacteria. Mol Plant Microbe Interact; 13:232-7.

97. **Dombrecht, B., J. Vanderleyden, J. Michiels,** Stable RK2-derived cloning vectors for the analysis of gene expression and gene function in gram-negative bacteria. Mol Plant Microbe Interact;14:426-30

98. **Datsenko, K. A., B. L. Wanner,** 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci USA; 97:6640-5.

99. **Hepinstall, J.,** 1998. Isolation of total RNA from bacteria. In Methods in molecular biology vol 86: RNA isolation and characterization protocols. Totowa, NJ: Humana Press.

100. **Becker, P., W. Hufnagle, G. Peters, M. Herrmann,** Detection of differential gene expression in biofilm-forming versus planktonic populations of Staphylococcus aureus using micro-representational-difference analysis. Appl Environ Microbiol; 67:2958-65.

101. **Schultheiss D., D. Schuler,** 2003. Development of a genetic system for Magnetospirillum gryphiswaldense. Arch Microbiol; 179:89-94.

102. **Dean, A. J.,** 1999. Initial genomic characterization of magnetotactic bacteria. Master's thesis.

103. **Wolska, K. I.,** 2003. Horizontal DNA transfer between bacteria in the environment. Acta Microbiol Pol; 52:233-43.

104. **Bates, G., M. Brunori, G. Amiconi, E. Antonini, J. Wyman,** 1968. Studies on hemerythrin. I. Thermodynamic and kinetic aspects of oxygen binding. Biochemistry; 7:3016-20.

105. **Dunn, J. B., A. W. Addison, R. E. Bruce, J. S. Loehr, T. M. Loehr,** 1977. Comparison of hemerythrins from four species of sipunculids by optical absorption, circular dichroism, fluorescence emission, and resonance Raman spectroscopy. Biochemistry; 16:1743-9.

106. **Farmer, C. S., D. M. Jr. Kurtz, Z. J. Liu, B. C. Wang, J. Rose, J. Ai, J. Sanders-Loehr,** 2001. The crystal structures of Phascolopsis gouldii wild type and L98Y methemerythrins: structural and functional alterations of the O2 binding pocket. J Biol Inorg Chem; 6:418-29.

107. **Xiong, J., D. M. Jr. Kurtz, J. Ai, J. Sanders-Loehr,** 2000. A hemerythrin-like domain in a bacterial chemotaxis protein. Biochemistry; 39:5117-25.

108. **Jurica, M. S., B. L. Stoddard,** 1999. Homing endonucleases: structure, function and evolution. Cell Mol Life Sci; 55:1304-26.

109. **Mowbray, S. L., M. O. Sandgren,** 1998. Chemotaxis receptors: a progress report on structure and function. J Struct Biol; 124:257-75.

110. **West, A. H., A M Stock,** 2001. Histidine kinases and response regulator proteins in two-component signaling systems. Trends Biochem Sci; 26:369-76.

111. **Ponting, C. P., L. Aravind,** 1997. PAS: a multifunctional domain family comes to light. Curr Biol; 7:674-7.

112. **White, D,** 2000. The physiology and biochemistry of prokaryotes, Inorganic metabolism Oxford University Press Inc, New York, NY.

113. **Bhattacharya, P.,** 1999. Evidence for integration of IS1 and IS10 in recombinant plasmid containing rearrangement hot-spot(rhs) common-shared block(CSB) in Escherichia coli K-12. Indian J Exp Biol; 37:1022-4.

114. **Mantripragada, K. K., P. G. Buckley, T. D. de Stahl, J. P. Dumanski,** 2004. Genomic microarrays in the spotlight. Trends Genet; 20:87-94.

APPENDIX. DNA SEQUENCES OBATAINED FROM THIS STUDY

A) Homologous DNA sequences from cytochrome a_1-like subunits I and II from wt MV-1.

```
  1  TAAGGGAGGG ATAATACATG TCTATTACTC CTCCAGAACA AAAAANCGTN
     ATTCCCTCCC TATTATGTAC AGATAATGAG GAGGTCTTGT TTTTTNGCAN

 51  CTGAAGTGTT TGCAAACGTA CTGAAGAAGA TGGCTGCGAA ATATAAAGTT
     GACTTCACAA ACGTTTGCAT GACTTCTTCT ACCGACGCTT TATATTTCAA

101  GGCGAAGAAC GTGCACACAG GCGTGCCTGT GGTTCGTCCT CCTGCAGGCA
     CCGCTTCTTG CACGTGTGTC CGCACGGACA CCAAGCAGGA GGACGTCCGT

 +3                 Met  Leu  Gly  Arg  Leu  Trp  Glu  Trp  Trp  Pro  Val  Leu  Glu
151  GCGACGTTTA TATGCTTGGT CGTTTGTGGG AGTGGTGGCC GGTCCTGGAA
     CGCTGCAAAT ATACGAACCA GCAAACACCC TCACCACCGG CCAGGACCTT

 +3  Leu  Glu  Lys  Gly  Gln  Ser  Tyr  Arg  Leu  His  Leu  Ser  Ser  Met  Asp  Trp  Met
201  TTGGAAAAGG GACAATCCTA TCGCTTGCAC TTGTCATCCA TGGACTGGAT
     AACCTTTTCC CTGTTAGGAT AGCGAACGTG AACAGTAGGT ACCTGACCTA

 +3  Met  His  Gly  Phe  Ser  Leu  Gln  Pro  Glu  Asn  Ile  Asn  Leu  Glu  Ile  His  Pro  Gly
251  GCATGGCTTT TCCTTACAGC CGGAAAACAT CAATTTAGAG ATCCACCCTG
     CGTACCGAAA AGGAATGTCG GCCTTTTGTA GTTAAATCTC TAGGTGGGAC

 +3  Gly  Tyr  Asp  Met  Val  Leu  Thr  Val  Thr  Pro  Thr  Ser  Ser  Gly  Glu  Phe  Ser
301  GATACGACAT GGTGTTGACG GTAACGCCCA CGAGCTCTGG TGAGTTCAGT
     CTATGCTGTA CCACAACTGC CATTGCGGGT GCTCGAGACC ACTCAAGTCA

 +3  Val  Val  Cys  Asn  Glu  Phe  Cys  Gly  Ile  Gly  His  His  Thr  Met  Val  Gly  Lys
351  GTTGTTTGTA ACGAATTTTG CGGTATTGGC CATCATACGA TGGTCGGCAA
     CAACAAACAT TGCTTAAAAC GCCATAACCG GTAGTATGCT ACCAGCCGTT

 +3  Lys  Ile  His  Val  Val  Asp  Lys
 +2                                    Met  Ala  Gln  Phe  Arg  Thr
401  AATTCATGTG GTCGATAAGT GAGGAAACGG GTCATGGCAC AATTTAGAAC
     TTAAGTACAC CAGCTATTCA CTCCTTTGCC CAGTACCGTG TTAAATCTTG

 +2  Thr  Cys  Pro  Asp  Thr  Gly  Phe  Ala  Ile  Asp  Leu  Ala  Ala  Asp  Lys  Leu  Ile  Lys
451  TTGCCCGGAC ACCGGGTTCG CAATCGATCT TGCGGCTGAC AAATTGATCA
     AACGGGCCTG TGGCCCAAGC GTTAGCTAGA ACGCCGACTG TTTAACTAGT

 +2  Lys  Trp  Asn  Ala  Val  Thr  Ala  Val  Val  Phe  Leu  Phe  Val  Gly  Gly  Leu  Phe
501  AGTGGAACGC GGTTACCGCC GTTGTCTTCT TGTTCGTCGG CGGCTTGTTC
     TCACCTTGCG CCAATGGCGG CAACAGAAGA ACAAGCAGCC GCCGAACAAG

 +2  Gly  Ile  Leu  Val  Ala  Leu  Thr  Arg  Trp  Pro  Ala  Val  His  Leu  Leu  Asp  Ala
551  GGCATCCTGG TCGCGCTGAC GCGTTGGCCT GCGGTTCACT TATTGGACGC
     CCGTAGGACC AGCGCGACTG CGCAACCGGA CGCCAAGTGA ATAACCTGCG

 +2  Ala  Asp  Leu  Phe  Tyr  Leu  Ala  Leu  Thr  Ala  His  Gly  Leu  Asp  Val  Leu  Leu  Val
601  TGATTTATTT TACCTGGCGC TGACGGCGCA CGGTTTAGAC GTGTTGTTGG
     ACTAAATAAA ATGGACCGCG ACTGCCGCGT GCCAAATCTG CACAACAACC

 +2  Val  Trp  Ile  Ile  Phe  Phe  Glu  Met  Ala  Leu  Leu  Tyr  Phe  Ser  Ser  Ser  Ile
651  TTTGGATTAT CTTTTTCGAG ATGGCCTTGC TGTATTTCTC GTCGTCAATA
     AAACCTAATA GAAAAAGCTC TACCGGAACG ACATAAAGAG CAGCAGTTAT

 +2  Leu  Leu  Lys  Thr  Arg  Leu  Ala  Ala  Pro  Trp  Leu  Gly  Trp  Val  Gln  Tyr  Thr
701  TTGTTGAAAA CCAGACTGGC CGCGCCTTGG CTTGGATGGG TTCAGTATAC
     AACAACTTTT GGTCTGACCG GCGCGGAACC GAACCTACCC AAGTCATATG

 +2  Thr  Leu  Met  Ser  Val  Gly  Ala  Leu  Ile  Thr  Asn  Val  Ala  Val  Leu  Gln  Gly  Glu
751  GCTTATGTCG GTCGGCGCTT TGATCACCAA CGTTGCAGTG TTGCAGGGCG
     CGAATACAGC CAGCCGCGAA ACTAGTGGTT GCAACGTCAC AACGTCCCGC

 +2  Glu  Ser  Ser  Val  Met  Phe  Thr  Ser  Tyr  Val  Pro  Met  Met  Ala  Ala  Pro  His
801  AAAGCAGCGT GATGTTCACG TCTTACGTTC CGATGATGGC AGCACCGCAT
     TTTCGTCGCA CTACAAGTGC AGAATGCAAG GCTACTACCG TCGTGGCGTA
```

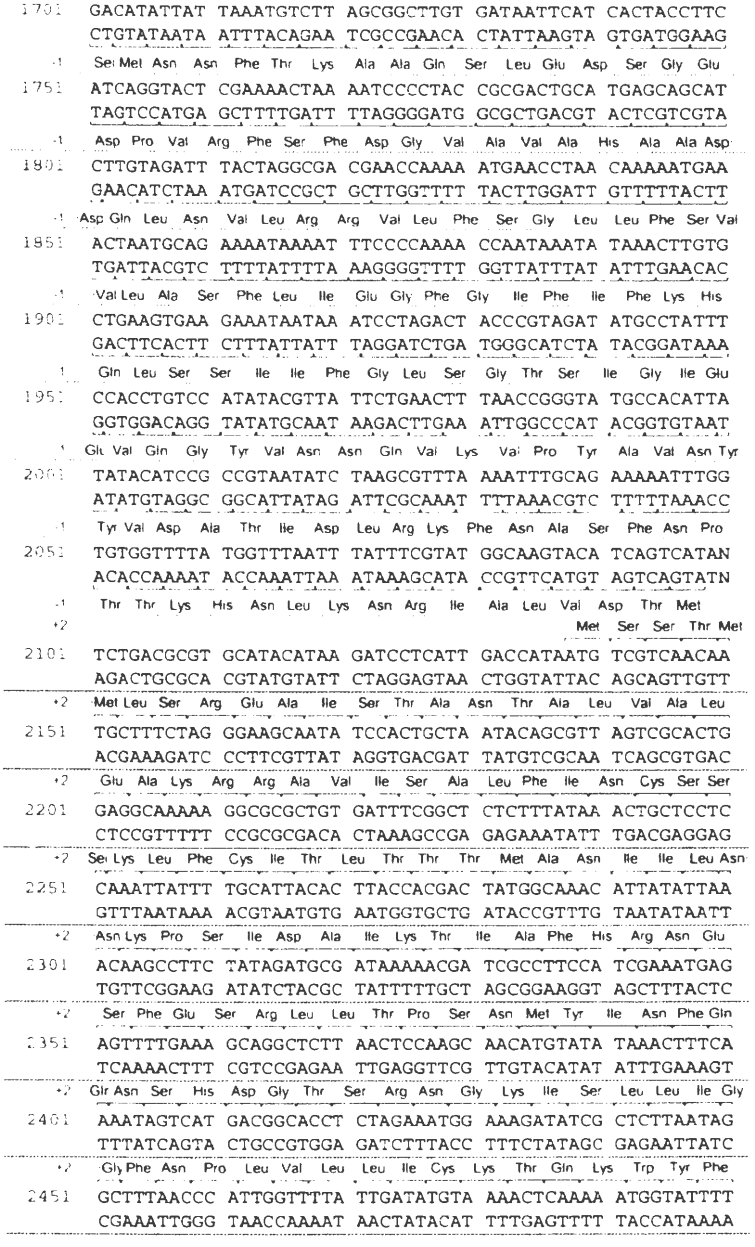
```
     +2  Phe Tyr Leu  Gly  Val  Ile  Leu  Phe Ala  Val  Gly  Ala Leu  Met  Gly  Cys Val
 801     TTTTACCTCG GCGTGATTTT GTTTGCCGTG GGTGCGTTGA TGGGCTGCGT
         AAAATGGAGC CGCACTAAAA CAAACGGCAC CCACGCAACT ACCCGACGCA
     +2  Va Val  Phe  Phe  Ser  Thr  Leu  Tyr  Ile  Ala  Lys  Arg  Asp  Lys  Thr  Tyr  Val Gly
 901     GGTCTTTTTC TCGACCTTGT ATATCGCCAA ACGCGATAAG ACCTATGTAG
         CCAGAAAAAG AGCTGGAACA TATAGCGGTT TGCGCTATTC TGGATACATC
     +2  Gly Phe Asn  Pro  Ala Gly  His  Val Trp Cys  Asp  Val Arg  Arg  Arg Leu Leu
 951     GGTTCAATCC CGCTGGTCAC GTTTGGTGCG ATGTACGGCG GCGATTATTG
         CCAAGTTAGG GCGACCAGTG CAAACCACGC TACATGCCGC CGCTAATAAC
     +2  Pro Cys Ser  Pro  Ser Trp  Arg  Ala Gln  Trp  Ile Leu  Ile  Pro  Thr Phe Phe
1001     CCGTGTTCAC CATCTTGGCG GGCGCAGTGG ATTTTGATCC CAACATTCTT
         GGCACAAGTG GTAGAACCGC CCGCGTCACC TAAAACTAGG GTTGTAAGAA
     +2  Phe Trp Ser  Leu  Gly Phe  Ile  Gly  Asn  Ile  Asp  Thr Leu  Met  Tyr  Arg Val Ile
1051     CTGGTCGTTG GGTTTTATCG GCAACATCGA TACGTTGATG TATCGTGTCA
         GACCAGCAAC CCAAAATAGC CGTTGTAGCT ATGCAACTAC ATAGCACAGT
     +2  Ile Trp Trp  Gly  Leu Gly  His  Ser Ser Gln  Gln  Ile Asn  Val  Ala Ala Gln
1101     TCTGGTGGGG CTTGGGACAC TCGTCTCAGC AGATCAACGT GGCAGCGCAG
         AGACCACCCC GAACCCTGTG AGCAGAGTCG TCTAGTTGCA CCGTCGCGTC
     +2  Ile  Ala Cys  Trp  Tyr  Ala  Ile  Ala Ala  Ile  Leu Phe Gly  Ala  Lys Pro Leu
1151     ATTGCCTGCT GGTACGCTAT TGCGGCCATC TTGTTCGGTG CCAAACCGCT
         TAACGGACGA CCATGCGATA ACGCCGGTAG AACAAGCCAC GGTTTGGCGA
     +2  Leu Ser Glu  Lys  Val Ser  Arg  Thr  Ala Phe  Leu  Phe Tyr  Ile  Leu Phe Leu Gln
1201     GTCCGAAAAA GTCAGCCGCA CCGCCTTTTT GTTCTACATT CTGTTCTTGC
         CAGGCTTTTT CAGTCGGCGT GGCGGAAAAA CAAGATGTAA GACAAGAACG
     +2  Gln Leu Ala  Ser  Ala His  His  Ile Leu  Val  Asp  Pro Gly  Ile  Ser  Ser  Glu
1251     AACTGGCATC CGCTCACCAC ATCTTGGTCG ATCCGGGCAT CAGTTCCGAA
         TTGACCGTAG GCGAGTGGTG TAGAACCAGC TAGGCCCGTA GTCAAGGCTT
     +2  Trp Lys  Ile  Phe Asn  Thr  Ser  Tyr Ala  Leu  Tyr Leu Ala  Val  Leu Gly Ser
1301     TGGAAAATCT TCAACACGTC TTACGCTCTG TACTTGGCTG TTTTGGGTTC
         ACCTTTTAGA AGTTGTGCAG AATGCGAGAC ATGAACCGAC AAAACCCAAG
     +2  Ser Met Leu  His  Gly Leu  Thr  Val  Pro Gly  Ala  Val Glu  Ala  Ala Leu Arg Lys
1351     CATGCTTCAT GGCCTGACGG TTCCGGGTGC GGTCGAAGCC GCGTTGCGCA
         GTACGAAGTA CCGGACTGCC AAGGCCCACG CCAGCTTCGG CGCAACGCGT
     +2  Lys Lys Gly  Tyr  Asn Arg  Gly  Leu Phe Glu  Trp  Leu Arg  Lys  Ala Pro Trp
1401     AAAAAGGCTA CAACCGCGGT CTGTTCGAGT GGCTCCGCAA AGCGCCTTGG
         TTTTTCCGAT GTTGGCGCCA GACAAGCTCA CCGAGGCGTT TCGCGGAACC
     +2  Gly Asn Pro  Val  Phe Ala  Gly  Phe Ala  Leu  Ser Val  Val  Leu  Phe Gly Phe
1451     GGTAATCCGG TATTTGCTGG ATTTGCTCTG TCNGTCGTGT TGTTCGGCTT
         CCATTAGGCC ATAAACGACC TAAACGAGAC AGNCAGCACA ACAAGCCGAA
     +2  Phe Leu Gly  Gly  Ile Thr  Gly  Val Met Met  Gly  Thr Glu  Gln  Leu Asn  Ile Ile
1501     CTTGGGCGGT ATCACGGGCG TTATGATGGG CACAGAACAG CTCAACATCA
         GAACCCGCCA TAGTGCCCGC AATACTACCC GTGTCTTGTC GAGTTGTAGT
     +2  Ile Ile His  Asn  Thr Leu  Tyr  Val Pro Gly  His  Phe His  Gly  Thr Val Val
1551     TCATTCACAA CACCTTGTAT GTTCCGGGTC ACTTCCATGG CACGGTCGTG
         AGTAAGTGTT GTGGAACATA CAAGGCCCAG TGAAGGTACC GTGCCAGCAC
     +2  Val Gly Thr  Thr  Leu Ala  Phe Met Ala  Met  Ala Tyr Trp  Leu  Val Pro Val
1601     GTCGGTACGA CCTTGGCCTT CATGGCGATG GCATACTGGT TGGTTCCGGT
         CAGCCATGCT GGAACCGGAA GTACCGCTAC CGTATGACCA ACCAAGGCCA
```

```
     +2  Val Leu Trp  Gln  Lys  Glu  Leu  Val  Phe  Pro  Gly  Leu  Ala  Lys  Trp  Gln  Pro  Tyr
1651  GTTGTGGCAG AAAGAACTGG TGTTCCCGGG TCTTGCCAAA TGGCAACCGT
      CAACACCGTC TTTCTTGACC ACAAGGGCCC AGAACGGTTT ACCGTTGGCA
     +2  Tyr Leu Phe  Gly  Leu  Gly  Met  Gly  Leu  Ala  Ser  Ile  Phe  Met  Met  Gly  Ala
1701  ACCTGTTCGG TCTCGGTATG GGTCTGGCAT CAATCTTCAT GATGGGTGCG
      TGGACAAGCC AGAGCCATAC CCAGACCGTA GTTAGAAGTA CTACCCACGC
     +2  Gly  Thr  Leu  Gly  Val  Pro  Arg  Arg  His  Trp  Asp  Ile  Ser  Phe  Ala  Asp  Ala
1751  GGTACGCTTG GTGTTCCTCG TCGTCACTGG GATATCTCGT TCGCCGACGC
      CCATGCGAAC CACAAGGAGC AGCAGTGACC CTATAGAGCA AGCGGCTGCG
     +2  Ala Met  Met  Val  His  Glu  Phe  Pro  Ala  Ala  Ala  Tyr  Met  Met  Met  Gly  Leu  Ala
1801  CATGATGGTG CATGAGTTCC CCGCCGCCGC CTATATGATG ATGGGCTTGG
      GTACTACCAC GTACTCAAGG GGCGGCGGCG GATATACTAC TACCCGAACC
     +2  Ala Gly  Ile  Ser  Gly  Val  Ile  Gly  Gly  Ile  Gly  Gly  Gly  Ile  Phe  Ile  Leu
1851  CAGGGATTTC GGGTGTAATC GGCGGCATTG GTGGGGGTAT CTTTATCCTC
      GTCCCTAAAG CCCACATTAG CCGCCGTAAC CACCCCCATA GAAATAGGAG
     +2  Asn  Met  Val  Ser  Thr  Ile  Leu  Phe  Gly  Lys  Ala  Lys  Val  Pro  Ala  Pro  Ala
1901  AACATGGTCA GCACGATTTT GTTTGGTAAG GCTAAAGTGC CTGCGCCTGC
      TTGTACCAGT CGTGCTAAAA CAAACCATTC CGATTTCACG GACGCGGACG
     +2  Ala Ala  Glu  Glu  Ala  Pro  Ala  Ala  Pro  Ile  Gly  Gly  Ala  Ala  Thr  Glu  Tyr  Gly
1951  AGCCGAAGAA GCTCCGGCCG CACCGATCGG TGGTGCCGCA ACTGAATACG
      TCGGCTTCTT CGAGGCCGGC GTGGCTAGCC ACCACGGCGT TGACTTATGC
     +2  Gly  Ser  Ala  Ser  Thr  Leu  Glu  Val  Pro  Gly  Thr  Phe  Val  Leu  Ala  Leu  Val
2001  GCAGTGCCAG CACCTTGGAG GTTCCTGGAA CCTTCGTGTT GGCTCTGGTG
      CGTCACGGTC GTGGAACCTC CAAGGACCTT GGAAGCACAA CCGAGACCAC
     +2  Phe  Leu  Thr  Ala  Phe  Val  Leu  Tyr  Tyr  Phe  Val  Asn  Trp  Lys  Tyr  Leu  Ala
2051  TTCCTAACAG CTTTTGTCCT TTACTACTTC GTCAACTGGA AATACTTGGC
      AAGGATTGTC GAAAACAGGA AATGATGAAG CAGTTGACCT TTATGAACCG
     +2  Ala  Ser  Val  Trp  Gly  Met  Ser
2101  GAGCGTTTGG GGTATGAGCT AACCGCACGC ACCCTTTTAG AAGGGTGCTC
      CTCGCAAACC CCATACTCGA TTGGCGTGCG TGGGAAAATC TTCCCACGAG
2151  CGGGGGGGCG GATACTACGA TGCCCACAGA ATAGTATCCG CCCTTTCTTA
      GCCCCCCCGC CTATGATGCT ACGGGTGTCT TATCATAGGC GGGAAAGAAT
2201  ATCTTGCAGG CTGGCTTGTG GGCAAGCCAC ACTCAAGGCT TTGTGGGTGA
      TAGAACGTCC GACCGAACAC CCGTTCGGTG TGAGTTCCGA AACACCCACT
2251  GCCGAATGAA CAAAAAAATC ATGAAAAAGC GAATGATGAC AATAATGCTC
      CGGCTTACTT GTTTTTTTAG TACTTTTTCG CTTACTACTG TTATTACGAG
2301  GCCGCGACGC TCTGTGCAGC GGTCGTGGCG GTGTCATCTG ACGCTCACGC
      CGGCGCTGCG AGACACGTCG CCAGCACCGC CACAGTAGAC TGCGAGTGCG
2351  AGGCCAGCAA GCGCCCGGCA CGATTTCCAA TTCGTACACC GGTGCCTTGG
      TCCGGTCGTT CGCGGGCCGT GCTAAAGGTT AAGCATGTGG CCACGGAACC
2401  ATGAAGAAAA AGCATTGGCG CTGTCTCAGG CGGCCATTGG TCGGGCCATC
      TACTTCTTTT TCGTAACCGC GACAGAGTCC GCCGGTAACC AGCCCGGTAG
2451  GGTGATTACA GTTTTCGCGA CACCCAAGGC AATCGCATTT CTCTTTCAGA
      CCACTAATGT CAAAAGCGCT GTGGGTTCCG TTAGCGTAAA GAGAAAGTCT
2501  TTTTCGCGGT AAGCCACTGG TGATTTCGTT GATCTATTCC AGTTGCGCGG
      AAAAGCGCCA TTCGGTGACC ACTAAAGCAA CTAGATAAGG TCAACGCGCC
2551  ATGTTTGTCC GGTGATTACG GCGACGTTGC AAAACGTTGA TGAAATGGCC
      TACAAACAGG CCACTAATGC CGCTGCAACG TTTTGCAACT ACTTTACCGG
```

```
2601  CGCGACGCAC TGGGCGACGA TAGCTATTCG ATCGTAACCA TAGGCTTTGA
      GCGCTGCGTG ACCCGCTGCT ATCGATAAGC TAGCATTGGT ATCCGAAACT
2651  CATTGGGGCG GACAACCCTG TGCAAATGAA AAGTTTTGCG CGCAAGTACG
      GTAACCCCGC CTGTTGGGAC ACGTTTACTT TTCAAAACGC GCGTTCATGC
2701  GTGTAAAGAT TAATGACCAC TGGAAATTTC TCAGTGGAGA TTTATTGGCC
      CACATTTCTA ATTACTGGTG ACCTTTAAAG AGTCACCTCT AAATAACCGG
2751  GTAACAGCGT TAAGCGAAGA CTTAGGATTT CAATTTTTCG AATCCCCCCA
      CATTGTCGCA ATTCGCTTCT GAATCCTAAA GTTAAAAAGC TTAGGGGGGT
2801  AAGGATTTGA CCATTTAACG CAGACCACGA TTTTGGACAA AACCGGGGTT
      TTCCTAAACT GGTAAATTGC GTCTGGTGCT AAAACCTGTT TTGGCCCCAA
2851  GTGTTTCGCC AGATTTACGG CGAGCAGTTT GAGGGCTCCA CATTTTGTCN
      CACAAAGCGG TCTAAATGCC GCTCGTCAAA CTCCCGAGGT GTAAAACAGN
2901  AGCCNTTAAA AAGTTTGGTT TTTGGNACGG CGTCNCCNTT TTTCANCCTT
      TCGGNAATTT TTCAAACCAA AAACCNTGCC GCAGNGGNAA AAAGTNGGAA
2951  TANCGATCTG ATCAACAAAG GTACGGTTGA
      ATNGCTAGAC TAGTTGTTTC CATGCCAACT
```

B) DNA sequences obtained from the subtractive library present in wt MV-1 and absent in MV-1nm1.

```
901  TACCACCGTG TCGATTTATT AANCCTGCGG CCAATGGCAG ACCCAAGCCT
     ATGGTGGCAC AGCTAAATAA TTNGGACGCC GGTTACCGTC TGGGTTCGGA
 -1  Lys Gly Gly His Arg Asn Ile Leu Gly Ala Ala Leu Pro Leu Gly Leu Gly Thr
951  GTCCCCCCAA TTGGGCGTTT CATGCTATCA TCAACTTGCT CAAATGCTTT
     CAGGGGGGTT AACCCGCAAA GTACGATAGT AGTTGAACGA GTTTACGAAA
 -1  Thr Gly Gly Ile Pro Arg Lys Met Ser Asp Asp Val Gln Glu Phe Ala Lys
1001 AAATATAACA TCCAGTTTAT CATCCGGAAT CCCGCATCCA GTATCGCTAA
     TTTATATTGT AGGTCAAATA GTAGGCCTTA GGGCGTAGGT CATAGCGATT
 -1  Phe Ile Val Asp Leu Lys Asp Asp Pro Ile Gly Cys Gly Thr Asp Ser Ile
1051 TGACAAAGTG CACAACATTT TTTTCTTGTT CAACAGATAC GCTAACGTGC
     ACTGTTTCAC GTGTTGTAAA AAAAGAACAA GTTGTCTATG CGATTGCACG
 -1  Ile Val Phe His Val Val Asn Lys Glu Gln Glu Val Ser Val Ser Val His Gly
1101 CCTTGATCAC AAAATTTGAT GGCATTTCCG AGCAGATTAA TAAGGATTTG
     GGAACTAGTG TTTTAAACTA CCGTAAAGGC TCGTCTAATT ATTCCTAAAC
 -1  Gly Gln Asp Cys Phe Lys Ile Ala Asn Gly Leu Leu Asn Ile Leu Ile Gln
1151 ATTTAAGCAT ACAGGATCCG CCCAAACCCA TAAATTATCA GAAGCAGTCA
     TAAATTCGTA TGTCCTAGGC GGGTTTGGGT ATTTAATAGT CTTCGTCAGT
 -1  Asn Leu Cys Val Pro Asp Ala Trp Val Trp Leu Asn Asp Ser Ala Thr Val
1201 CTTTTAATTT TAGATTTTTT TTTGTGATTG TAGAGCCGAC AATATTGATA
     GAAAATTAAA ATCTAAAAAA AAACACTAAC ATCTCGGCTG TTATAACTAT
 -1  Val Lys Leu Lys Leu Asn Lys Lys Thr Ile Thr Ser Gly Val Ile Asn Ile Thr
1251 GTGGTTTTAA TTGCAGCTGC AATATTAACT AACTTTGGAG TGATGGATAA
     CACCAAAATT AACGTCGACG TTATAATTGA TTGAAACCTC ACTACCTATT
 -1  Thr Thr Lys Ile Ala Ala Ala Ile Asn Val Leu Lys Pro Thr Ile Ser Leu
1301 TTTACCAGAA TCCATTTTTG CAAAATCCAA AACTTGGTTT ATTAAACAAA
     AAATGGTCTT AGGTAAAAAC GTTTTAGGTT TTGAACCAAA TAATTTGTTT
 -1  Lys Gly Ser Asp Met Lys Ala Phe Asp Leu Val Gln Asn Ile Leu Cys Leu
1351 GCAAGTGATT GGCAGATTCA AGTCCTCCAC TAGCGATTTC AACGACTTCT
     CGTTCACTAA CCGTCTAAGT TCAGGAGGTG ATCGCTAAAG TTGCTGAAGA
 -1  Leu Leu His Asn Ala Ser Glu Leu Gly Gly Ser Ala Ile Glu Val Val Glu Thr
1401 GTCACAAGGT CATTTAATGA TTCCTGAAGC GTCCCACGGC ACTCTAAAGG
     CAGTGTTCCA GTAAATTACT AAGGACTTCG CAGGGTGCCG TGAGATTTCC
 -1  Thr Val Leu Asp Asn Leu Ser Glu Gln Leu Thr Gly Arg Cys Glu Leu Pro
1451 AAGTGACATC TTCATTTGTT GTAAGCCTTG CACTCCTTTA AACAAATATA
     TTCACTGTAG AAGTAAACAA CATTCGGAAC GTGAGGAAAT TTGTTTATAT
  1  Leu Ser Met Lys Met Gln Gln Leu Gly Gln Val Gly Lys Phe Leu Tyr Ile
1501 TTTGCTTTGC TAATTCTTCA AATGGCCCCT TTAAACCCTG TGTTGGTGTA
     AAACGAAACG ATTAAGAAGT TTACCGGGGA AATTTGGGAC ACAACCACAT
 -1  Ile Gln Lys Ala Leu Glu Glu Phe Pro Gly Lys Leu Gly Gln Thr Pro Thr Arg
1551 CGCAATTCAT GGCTTATCAT ATTAATAAAA TCTGTTTTAG CTTTACTTGC
     GCGTTAAGTA CCGAATAGTA TAATTATTTT AGACAAAATC GAAATGAACG
 -1  Arg Leu Glu His Ser Ile Met Asn Ile Phe Asp Thr Lys Ala Lys Ser Ala
1601 ATGTTCAGCT ATATTGCGTG ACTCTTCGGA GACCTTATAT TGTTTTTTAA
     TACAAGTCGA TATAACGCAC TGAGAAGCCT CTGGAATATA ACAAAAAATT
 -1  His Glu Ala Ile Asn Arg Ser Glu Glu Ser Val Lys Tyr Gln Lys Lys Leu
1651 GTTCTTGGTT AGCTTCCTCT AAGCTTGTAA CTAATAATTC TAGCGAATAT
     CAAGAACCAA TCGAAGGAGA TTCGAACATT GATTATTAAG ATCGCTTATA
  1  Leu Glu Gln Asn Ala Glu Glu Leu Ser Thr Val Leu Leu Glu Leu Ser Tyr Ser
```

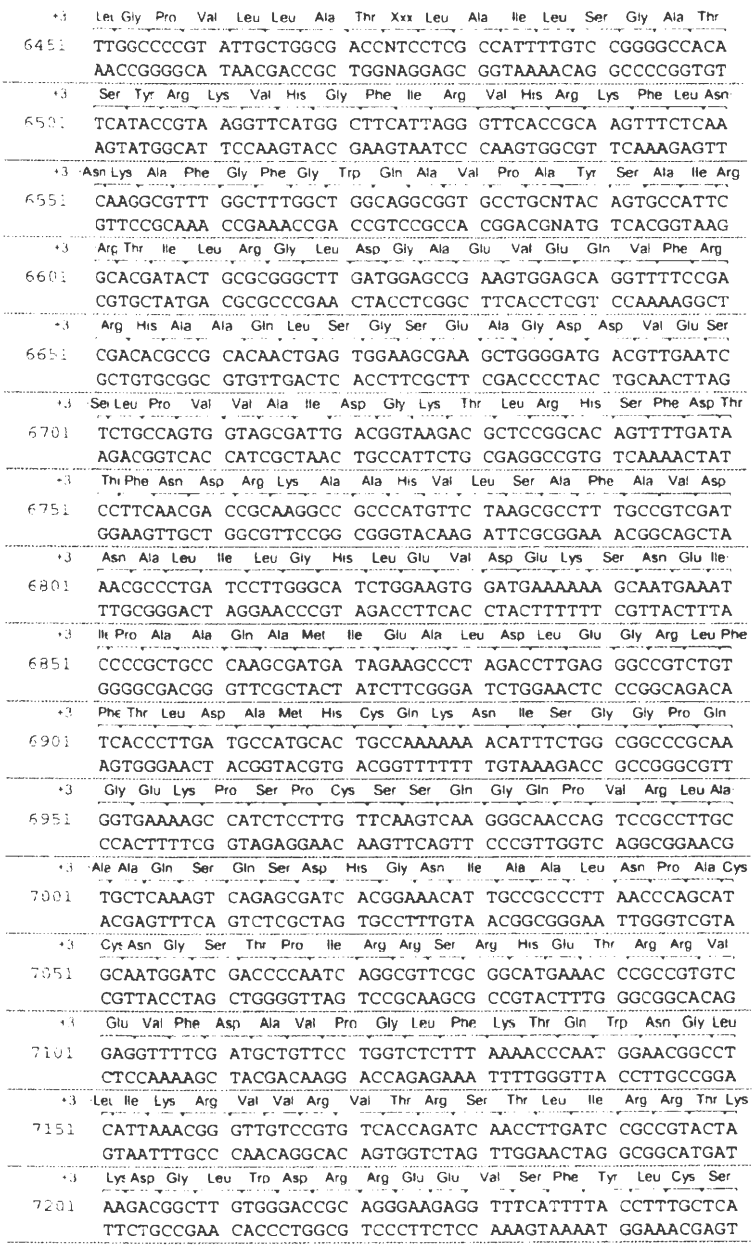

1701 GACATATTAT TAAATGTCTT AGCGGCTTGT GATAATTCAT CACTACCTTC
CTGTATAATA ATTTACAGAA TCGCCGAACA CTATTAAGTA GTGATGGAAG
-1 Ser Met Asn Asn Phe Thr Lys Ala Ala Gln Ser Leu Glu Asp Ser Gly Glu
1751 ATCAGGTACT CGAAAACTAA AATCCCCTAC CGCGACTGCA TGAGCAGCAT
TAGTCCATGA GCTTTTGATT TTAGGGGATG GCGCTGACGT ACTCGTCGTA
-1 Asp Pro Val Arg Phe Ser Phe Asp Gly Val Ala Val Ala His Ala Ala Asp
1801 CTTGTAGATT TACTAGGCGA CGAACCAAAA ATGAACCTAA CAAAAATGAA
GAACATCTAA ATGATCCGCT GCTTGGTTTT TACTTGGATT GTTTTTACTT
-1 Asp Gln Leu Asn Val Leu Arg Arg Val Leu Phe Ser Gly Leu Leu Phe Ser Val
1851 ACTAATGCAG AAAATAAAAT TTCCCCAAAA CCAATAAATA TAAACTTGTG
TGATTACGTC TTTTATTTTA AAGGGGTTTT GGTTATTTAT ATTTGAACAC
-1 Val Leu Ala Ser Phe Leu Ile Glu Gly Phe Gly Ile Phe Ile Phe Lys His
1901 CTGAAGTGAA GAAATAATAA ATCCTAGACT ACCCGTAGAT ATGCCTATTT
GACTTCACTT CTTTATTATT TAGGATCTGA TGGGCATCTA TACGGATAAA
1 Gln Leu Ser Ser Ile Ile Phe Gly Leu Ser Gly Thr Ser Ile Gly Ile Glu
1951 CCACCTGTCC ATATACGTTA TTCTGAACTT TAACCGGGTA TGCCACATTA
GGTGGACAGG TATATGCAAT AAGACTTGAA ATTGGCCCAT ACGGTGTAAT
1 Glu Val Gln Gly Tyr Val Asn Asn Gln Val Lys Val Pro Tyr Ala Val Asn Tyr
2001 TATACATCCG CCGTAATATC TAAGCGTTTA AAATTTGCAG AAAAATTTGG
ATATGTAGGC GGCATTATAG ATTCGCAAAT TTTAAACGTC TTTTTAAACC
-1 Tyr Val Asp Ala Thr Ile Asp Leu Arg Lys Phe Asn Ala Ser Phe Asn Pro
2051 TGTGGTTTTA TGGTTTAATT TATTTCGTAT GGCAAGTACA TCAGTCATAN
ACACCAAAAT ACCAAATTAA ATAAAGCATA CCGTTCATGT AGTCAGTATN
-1 Thr Thr Lys His Asn Leu Lys Asn Arg Ile Ala Leu Val Asp Thr Met
+2 Met Ser Ser Thr Met
2101 TCTGACGCGT GCATACATAA GATCCTCATT GACCATAATG TCGTCAACAA
AGACTGCGCA CGTATGTATT CTAGGAGTAA CTGGTATTAC AGCAGTTGTT
+2 Met Leu Ser Arg Glu Ala Ile Ser Thr Ala Asn Thr Ala Leu Val Ala Leu
2151 TGCTTTCTAG GGAAGCAATA TCCACTGCTA ATACAGCGTT AGTCGCACTG
ACGAAAGATC CCTTCGTTAT AGGTGACGAT TATGTCGCAA TCAGCGTGAC
+2 Glu Ala Lys Arg Arg Ala Val Ile Ser Ala Leu Phe Ile Asn Cys Ser Ser
2201 GAGGCAAAAA GGCGCGCTGT GATTTCGGCT CTCTTTATAA ACTGCTCCTC
CTCCGTTTTT CCGCGCGACA CTAAAGCCGA GAGAAATATT TGACGAGGAG
+2 Ser Lys Leu Phe Cys Ile Thr Leu Thr Thr Thr Met Ala Asn Ile Ile Leu Asn
2251 CAAATTATTT TGCATTACAC TTACCACGAC TATGGCAAAC ATTATATTAA
GTTTAATAAA ACGTAATGTG AATGGTGCTG ATACCGTTTG TAATATAATT
+2 Asn Lys Pro Ser Ile Asp Ala Ile Lys Thr Ile Ala Phe His Arg Asn Glu
2301 ACAAGCCTTC TATAGATGCG ATAAAAACGA TCGCCTTCCA TCGAAATGAG
TGTTCGGAAG ATATCTACGC TATTTTTGCT AGCGGAAGGT AGCTTTACTC
+2 Ser Phe Glu Ser Arg Leu Leu Thr Pro Ser Asn Met Tyr Ile Asn Phe Gln
2351 AGTTTTGAAA GCAGGCTCTT AACTCCAAGC AACATGTATA TAAACTTTCA
TCAAAACTTT CGTCCGAGAA TTGAGGTTCG TTGTACATAT ATTTGAAAGT
+2 Gln Asn Ser His Asp Gly Thr Ser Arg Asn Gly Lys Ile Ser Leu Leu Ile Gly
2401 AAATAGTCAT GACGGCACCT CTAGAAATGG AAAGATATCG CTCTTAATAG
TTTATCAGTA CTGCCGTGGA GATCTTTACC TTTCTATAGC GAGAATTATC
+2 Gly Phe Asn Pro Leu Val Leu Leu Ile Cys Lys Thr Gln Lys Trp Tyr Phe
2451 GCTTTAACCC ATTGGTTTTA TTGATATGTA AAACTCAAAA ATGGTATTTT
CGAAATTGGG TAACCAAAAT AACTATACAT TTTGAGTTTT TACCATAAAA

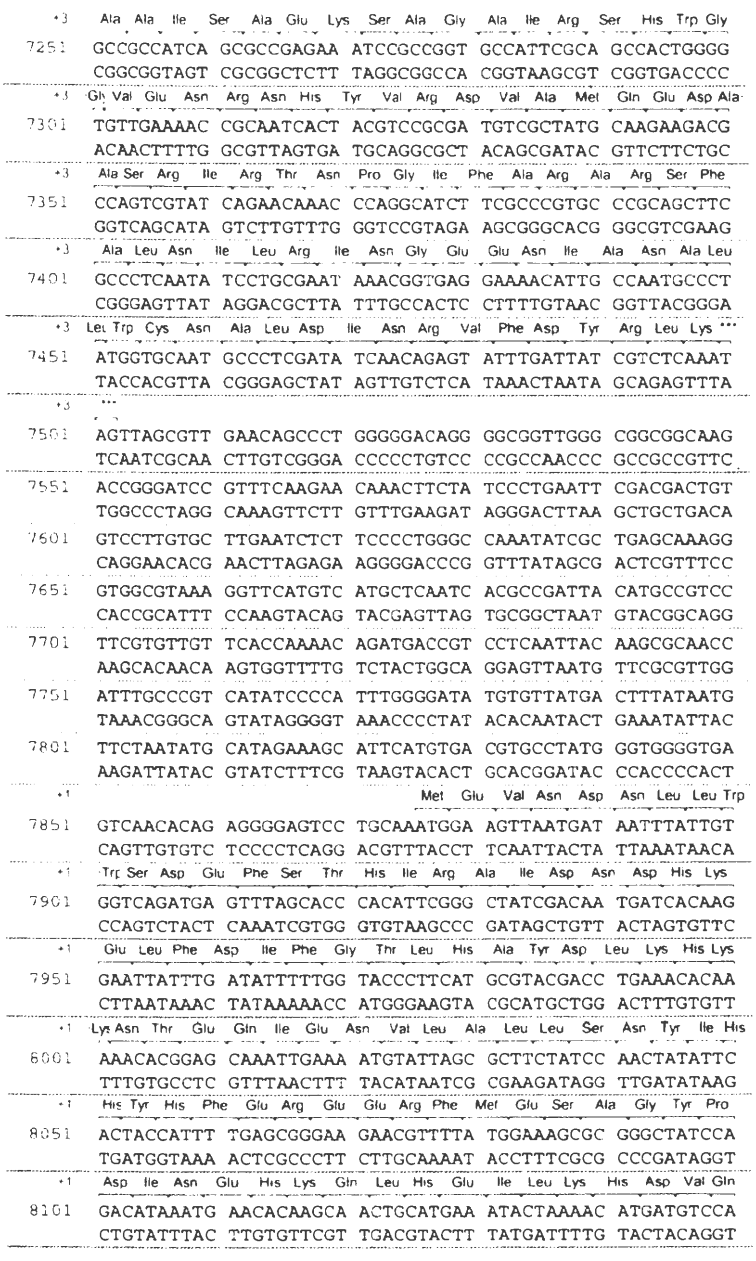
```
 +2   ***
2501  TAGAGGTGCC TATGAAAGTG TTATTTGGGA AGAATATCTA TCTGAAGTTC
      ATCTCCACGG ATACTTTCAC AATAAACCCT TCTTATAGAT AGACTTCAAG
 -2                          *** Lys Pro Leu Ile Asp Ile Gln Leu Glu
2551  CCTTACGTCA TTCCAGTCAG AATCAGTGCC GCGTTCTATT CCTTGTGTGA
      GGAATGCAGT AAGGTCAGTC TTAGTCACGG CGCAAGATAA GGAACACACT
 -2   Arg Val Asp Asn Trp Asp Ser Asp Thr Gly Arg Glu Ile Gly Gln Thr Phe
2601  ACCCAAGTTC ATTTAACAGT TTTTTCCCAT CTTTATTATC AAACATTTTA
      TGGGTTCAAG TAAATTGTCA AAAAAGGGTA GAAATAATAG TTTGTAAAAT
 -2  Phe Gly Leu Glu Asn Leu Leu Lys Lys Gly Asp Lys Asn Asp Phe Met Lys Leu
2651  AGTAAAACTC TTTGTACAGT ATCAACAACA TTTGTTGGCA CTCTCGGGTG
      TCATTTTGAG AAACATGTCA TAGTTGTTGT AAACAACCGT GAGAGCCCAC
 -2   Leu Leu Val Arg Gln Val Thr Asp Val Val Asn Thr Pro Val Arg Pro His
2701  TGCTGCAAAT GCGTGAGGCG TATACCCTTC AGAGACCCAA AAAATATTTA
      ACGACGTTTA CGCACTCCGC ATATGGGAAG TCTCTGGGTT TTTTATAAAT
 -2   Ala Ala Phe Ala His Pro Thr Tyr Gly Glu Ser Val Trp Phe Ile Asn Leu
2751  GTTTTTCTTT GACTTTTGGG TCGACCGTTT CCAACGTTCG TTGGATGCCT
      CAAAAAGAAA CTGAAAACCC AGCTGGCAAA GGTTGCAAGC AACCTACGGA
 -2  Leu Lys Glu Lys Val Lys Pro Asp Val Thr Glu Leu Thr Arg Gln Ile Gly Gly
2801  CCTCCTGCAA CAAACATCCC CCTCGCGACA CCAATGTAAA CTGAATCGTG
      GGAGGACGTT GTTTGTAGGG GGAGCGCTGT GGTTACATTT GACTTAGCAC
 -2   Gly Gly Ala Val Phe Met Gly Arg Ala Val Gly Ile Tyr Val Ser Asp His
2851  ACTTCCTACA TATTTAGAAG TGAATTTAAC ATCAATTTTC GACAAATTAG
      TGAAGGATGT ATAAATCTTC ACTTAAATTG TAGTTAAAAG CTGTTTAATC
 -2   Ser Gly Val Tyr Lys Ser Thr Phe Lys Val Asp Ile Lys Ser Leu Asn Ser
2901  ATTGTGGCAA TATAGTTGCT GCAAATGCAG CAGGCGCAGG AAATGCTATT
      TAACACCGTT ATATCAACGA CGTTTACGTC GTCCGCGTCC TTTACGATAA
 -2  Ser Gln Pro Leu Ile Thr Ala Ala Phe Ala Ala Pro Ala Pro Phe Ala Ile Glu
2951  TCTTTCCCAT CAAGTTCTTT CAAATTTTTT ATTAGGGTAT CCTTTTGAGT
      AGAAAGGGTA GTTCAAGAAA GTTTAAAAAA TAATCCCATA GGAAAACTCA
 -2   Glu Lys Gly Asp Leu Glu Lys Leu Asn Lys Ile Leu Thr Asp Lys Gln Thr
3001  AACCAATATA CCCTTAATAT TTTTTCCTTT CTGTTTGCCA AATGCACGAT
      TTGGTTATAT GGGAATTATA AAAAAGGAAA GACAAACGGT TTACGTGCTA
 -2   Val Leu Ile Gly Lys Ile Asn Lys Gly Lys Gln Lys Gly Phe Ala Arg Tyr
3051  ATCCTGGATT CTGAGAAAAA ACCGTATAAT GTAACGGGTT CATATAGGCA
      TAGGACCTAA GACTCTTTTT TGGCATATTA CATTGCCCAA GTATATCCGT
 -2  Tyr Gly Pro Asn Gln Ser Phe Val Thr Tyr His Leu Pro Asn Met Tyr Ala Phe
3101  AAATCATATT CTCCAGCAGC TAATCGAGCT TCAAATGTAG GAATGTCAGG
      TTTAGTATAA GAGGTCGTCG ATTAGCTCGA AGTTTACATC CTTACAGTCC
 -2   Phe Asp Tyr Glu Gly Ala Ala Leu Arg Ala Glu Phe Thr Pro Ile Asp Pro
3151  AGCCGTGCTA AAATTTATAG TTAAACCAGT ATCATGGCTT AGTTTTTTCA
      TCGGCACGAT TTTAAATATC AATTTGGTCA TAGTACCGAA TCAAAAAAGT
 -2   Ala Thr Ser Phe Asn Ile Thr Leu Gly Thr Asp His Ser Leu Lys Lys Leu
3201  GAATTGGCCC CCAAAGTTTT ACGAGAACTG ATGGAGACTG TTGAGGAACG
      CTTAACCGGG GGTTTCAAAA TGCTCTTGAC TACCTCTGAC AACTCCTTGC
 -2  Leu Ile Pro Gly Trp Leu Lys Val Leu Val Ser Pro Ser Gln Gln Pro Val Phe
3251  AAGCCAAATG AAAAAGAGTT TTCAACAGCG TAAACTGGTT TAACCGCTAA
      TTCGGTTTAC TTTTTCTCAA AAGTTGTCGC ATTTGACCAA ATTGGCGATT
 -2   Phe Gly Phe Ser Phe Ser Asn Glu Val Ala Tyr Val Pro Lys Val Ala Leu
```

```
3301  TACCAATATC GTGGAAATCA AAACAAAAAA CTTACACACA TCAATTCCTT
      ATGGTTATAG CACCTTTAGT TTTGTTTTTT GAATGTGTGT AGTTAAGGAA
 -2    Val Leu Ile Thr Ser Ile Leu Val Phe Phe Lys Cys Val Asp Ile Gly Lys
3351  TCATACAAAT AGGGCACATC GAAAGGGTGT TCACTCGCGA TGGCTATACT
      AGTATGTTTA TCCCGTGTAG CTTTCCCACA AGTGAGCGCT ACCGATATGA
 -2  Lys Met Cys Ile Pro Cys Met
3401  GGTCACCCAA ATTTCCAAAA AAATAATAAT ATACTTCAAT AAATTGTCAT
      CCAGTGGGTT TAAAGGTTTT TTTATTATTA TATGAAGTTA TTTAACAGTA
3451  ATGCTGGCGC TATAGCGGCT TATCACGTTG CCCCATTATC TCCCTTTTGG
      TACGACCGCG ATATCGCCGA ATAGTGCAAC GGGGTAATAG AGGGAAAACC
3501  GGCGGCTTGA CCCGTTTCCT CTGCCCGATA CCGTCAAACT GATCACCGTC
      CCGCCGAACT GGGCAAAGGA GACGGGCTAT GGCAGTTTGA CTAGTGGCAG
3551  CTGGGTCTTG CATCGGCGGC GGACTTTCGG TTCACCGGCT GATTTTCGGG
      GACCCAGAAC GTAGCCGCCG CCTGAAAGCC AAGTGGCCGA CTAAAAGCCC
3601  GGCTGCCGCG CTTGATGATG TCCCAATACG ACATAAGCGC GTGGTCGGTG
      CCGACGGCGC GAACTACTAC AGGGTTATGC TGTATTCGCG CACCAGCCAC
3651  GACAGAGGCA TGTCACGTTG ATACCCATAA TGCGGCTCGA TCTCCTGATT
      CTGTCTCCGT ACAGTGCAAC TATGGGTATT ACGCCGAGCT AGAGGACTAA
3701  TGATCAAGTT GCCATGGGTC AACGTGATAA GCCGTGCGCT ATAGCCTCAA
      ACTAGTTCAA CGGTACCCAG TTGCACTATT CGGCACGCGA TATCGGAGTT
3751  TAGGCCTATC AAAAACATAA ATCACTCATT TTTCACCATC GCAAGTATTA
      ATCCGGATAG TTTTTGTATT TAGTGAGTAA AAAGTGGTAG CGTTCATAAT
3801  ATAATCAGTC ACTAAAGACT TATCCCGTTG TCCCAGGCAT CCGGGAGCAC
      TATTAGTCAG TGATTTCTGA ATAGGGCAAC AGGGTCCGTA GGCCCTCGTG
3851  CTTGCAATAG ATTGGAGAGA GATGAGCATT ATTTCGTATA ATCACCATCA
      GAACGTTATC TAACCTCTCT CTACTCGTAA TAAAGCATAT TAGTGGTAGT
3901  TCCGGCTCAT TTTGTTCCTC AGAGAAGTAA AATCAGCACC AAACCTTCCA
      AGGCCGAGTA AAACAAGGAG TCTCTTCATT TTAGTCGTGG TTTGGAAGGT
3951  GACCTTTGTC CCTTGCCAGT CATTGCGTTT ATGACCTGGG ACAACGGGAT
      CTGGAAACAG GGAACGGTCA GTAACGCAAA TACTGGACCC TGTTGCCCTA
4001  AAGCCACAGT CACTATTGCC TTTTATTATT TCAATAGGTT ATTTTGAGGC
      TTCGGTGTCA GTGATAACGG AAAATAATAA AGTTATCCAA TAAAACTCCG
4051  ACTCAGCACA GTACCCGACT CTTCTCATAA AGCAAGGCCG TACAACGGTT
      TGAGTCGTGT CATGGGCTGA GAAGAGTATT TCGTTCCGGC ATGTTGCCAA
4101  TTCGAGGTAT ATTTGGCTTG AACAACCCGG TCATATAATT GCGCAATTCC
      AAGCTCCATA TAAACCGAAC TTGTTGGGCC AGTATATTAA CGCGTTAAGG
4151  AATTCAGACA TTCTGACTGA GGCTTTTCTA AACACATTGA TGCCGAGCGT
      TTAAGTCTGT AAGACTGACT CCGAAAAGAT TTGTGTAACT ACGGCTCGCA
4201  GAAGATGGAT TGCTGCAATC GGCCATGGGC TTTGCGTCGT GGTGGTTTAG
      CTTCTACCTA ACGACGTTAG CCGGTACCCG AAACGCAGCA CCACCAAATC
4251  TTCTCGCCAC CCAAAGTCCT GTTTTGTAGG CGATGCAAAA AGCAATCGCC
      AAGAGCGGTG GGTTTCAGGA CAAAACATCC GCTACGTTTT TCGTTAGCGG
4301  AATACGCTCA TCAACGTTGC TAATTTATGC GGATTTGTTA TATGCGTATC
      TTATGCGAGT AGTTGCAACG ATTAAATACG CCTAAACAAT ATACGCATAG
4351  TTCCAATCCT AAACCTCGTG TTTTCAAGCA GGAAAATAAC GTTTCAATGC
      AAGGTTAGGA TTTGGAGCAC AAAAGTTCGT CCTTTTATTG CAAAGTTACG
 -1                                                *** His Gly
```

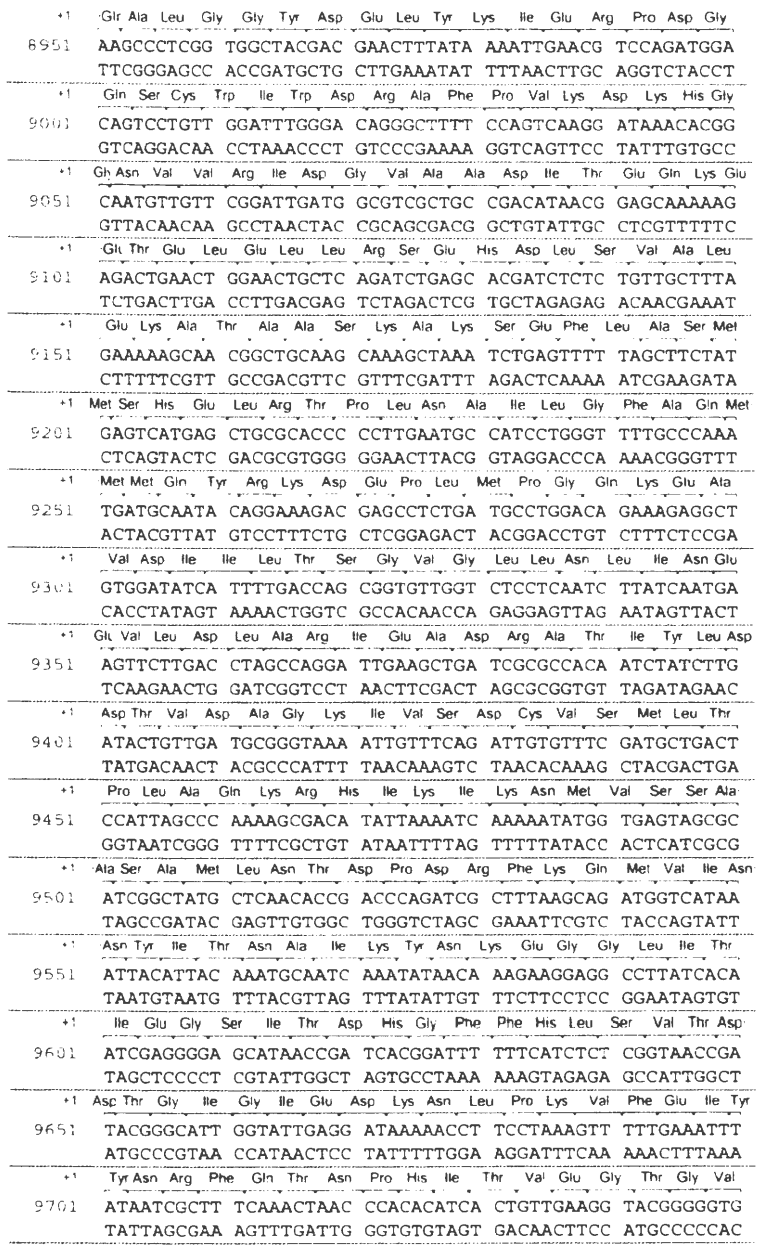
```
4401  CCCCATCGGT TGCGATAGAT CGGGAGTGCC GTTTGGATGC GGATAAGACT
      GGGGTAGCCA ACGCTATCTA GCCCTCACGG CAAACCTACG CCTATTCTGA
 -1     Gly Trp Arg  Asn  Arg Tyr  Ile  Pro Leu Ala  Thr  Gln  Ile  Arg  Ile  Leu Ser
4451  TGTGGCGACA ATAAGCATTT CTCCGGTTTT CAAACGCATC ATGGAAATTT
      ACACCGCTGT TATTCGTAAA GAGGCCAAAA GTTTGCGTAG TACCTTTAAA
 -1     Thr  Ala  Val  Ile  Leu Met  Glu  Gly Thr  Lys  Leu Arg  Met  Met  Ser  Ile Lys
4501  TTATGGGCGT TGTTGTTTTG GATGGGTCTC GACCTAAATA CCACGTCCCT
      AATACCCGCA ACAACAAAAC CTACCCAGAG CTGGATTTAT GGTGCAGGGA
 -1  Lys Ile  Pro  Thr  Thr  Thr  Lys  Ser  Pro Asp  Arg  Gly  Leu  Tyr  Trp  Thr  Gly Lys
4551  TTTAAAATGC ATTTTTGACC GTTTTTAAGA TGACGTGCAT GGGCCGATAG
      AAATTTTACG TAAAAACTGG CAAAAATTCT ACTGCACGTA CCCGGCTATC
  1   Lys Leu  Ile  Cys  Lys  Gln  Gly  Asn Lys  Leu  His  Arg  Ala  His  Ala  Ser  Leu
4601  TGGAGCCGGA GCATATCCAT CCTGCCAAAT GAACATATTT TCTTTCAGGC
      ACCTCGGCCT CGTATAGGTA GGACGGTTTA CTTGTATAAA AGAAAGTCCG
 -1     Pro  Ala  Pro  Ala  Tyr  Gly  Asp  Gln  Trp  Ile  Phe  Met  Asn  Glu  Lys  Leu Arg
4651  GCAAAACAAA AGGAATGCCT TGGTTTTGAA GCCAATTAAA CCATGCCTCC
      CGTTTTGTTT TCCTTACGGA ACCAAAACTT CGGTTAATTT GGTACGGAGG
  1  Arg Leu  Val  Phe  Pro  Ile  Gly  Gln  Asn Gln  Leu  Trp  Asn  Phe  Trp  Ala  Glu Gly
4701  CCAATAAATT CACGATCCCC AGACAAAGAT GCTATGGGTT GATTTGGGAA
      GGTTATTTAA GTGCTAGGGG TCTGTTTCTA CGATACCCAA CTAAACCCTT
 -1     Gly  Ile  Phe  Glu  Arg Asp  Gly  Ser  Leu Ser  Ala  Ile  Pro  Gln  Asn  Pro Phe
4751  GACTTTGTTG AGTTTTGGCA TGAGGTCTGT GCGCTCTTTG GCATTGGAAT
      CTGAAACAAC TCAAAACCGT ACTCCAGACA CGCGAGAAAC CGTAACCTTA
 -1     Val  Lys  Asn  Leu  Lys  Pro  Met  Leu Asp  Thr  Arg  Glu  Lys  Ala  Asn  Ser Asn
4801  TCCCAGCCTT GCCTAACAAG ACCCAAAAAA GGGGAATGCC GACTTTCTCA
      AGGGTCGGAA CGGATTGTTC TGGGTTTTTT CCCCTTACGG CTGAAAGAGT
 -1  Asn Gly  Ala  Lys  Gly  Leu  Leu  Val  Trp Phe  Leu  Pro  Ile  Gly  Val  Lys  Glu His
4851  TGGATAATGG TCAGCATTAG AATGTTGATG TGGCATCGCC CAAACTTCCA
      ACCTATTACC AGTCGTAATC TTACAACTAC ACCGTAGCGG GTTTGAAGGT
 -1     His  Ile  Ile  Thr  Leu Met  Leu  Ile Asn  Ile  His  Cys  Arg  Gly  Phe  Lys  Trp
4901  GTTTGTTCGA TCAATGGCCA AGTGCCAGGG CTTTCCAGAC AGTCCCATAA
      CAAACAAGCT AGTTACCGGT TCACGGTCCC GAAAGGTCTG TCAGGGTATT
 -1     Asn  Thr  Arg  Asp  Ile  Ala  Leu  His Trp  Pro  Lys  Gly  Ser  Leu  Gly  Met
4951  TGTGCACAAC CAGACGGGCA ACACAGCCTT CATCAAGACG AATATGTTGA
      ACACGTGTTG GTCTGCCCGT TGTGTCGGAA GTAGTTCTGC TTATACAACT
5001  AAAAAACGTT CAAAACGACG ATGTACAGAT AACGTTTTTG ACCGTGTATC
      TTTTTTGCAA GTTTTGCTGC TACATGTCTA TTGCAAAAAC TGGCACATAG
5051  AACATGACCC GCAAGCCGCC AAAGACTGAC CGTTCCGGTT TGGATGATCA
      TTGTACTGGG CGTTCGGCGG TTTCTGACTG GCAAGGCCAA ACCTACTAGT
5101  GCGTAACAAG CCAAACAAGC GTTTCTTGGC TTGTGCGCCA GAGAGGGATG
      CGCATTGTTC GGTTTGTTCG CAAAGAACCG AACACGCGGT CTCTCCCTAC
5151  TGGTTTCCTA ACTTCAGGGA TAAATACTTA GTNCAGCGTT CTGTTCATAT
      ACCAAAGGAT TGAAGTCCCT ATTTATGAAT CANGTCGCAA GACAAGTATA
5201  TCAAAGACTT TNCAGTGGTT TGAGGATCTT AAAGTATGAA TTGCTTCAAG
      AGTTTCTGAA ANGTCACCAA ACTCCTAGAA TTTCATACTT AACGAAGTTC
5251  TGTTTATCCC GTTAAGTCGG GTGCTGTGGG CACTCAGACA AGTCCAGGCA
      ACAAATAGGG CAATTCAGCC CACGACACCC GTGAGTCTGT TCAGGTCCGT
```

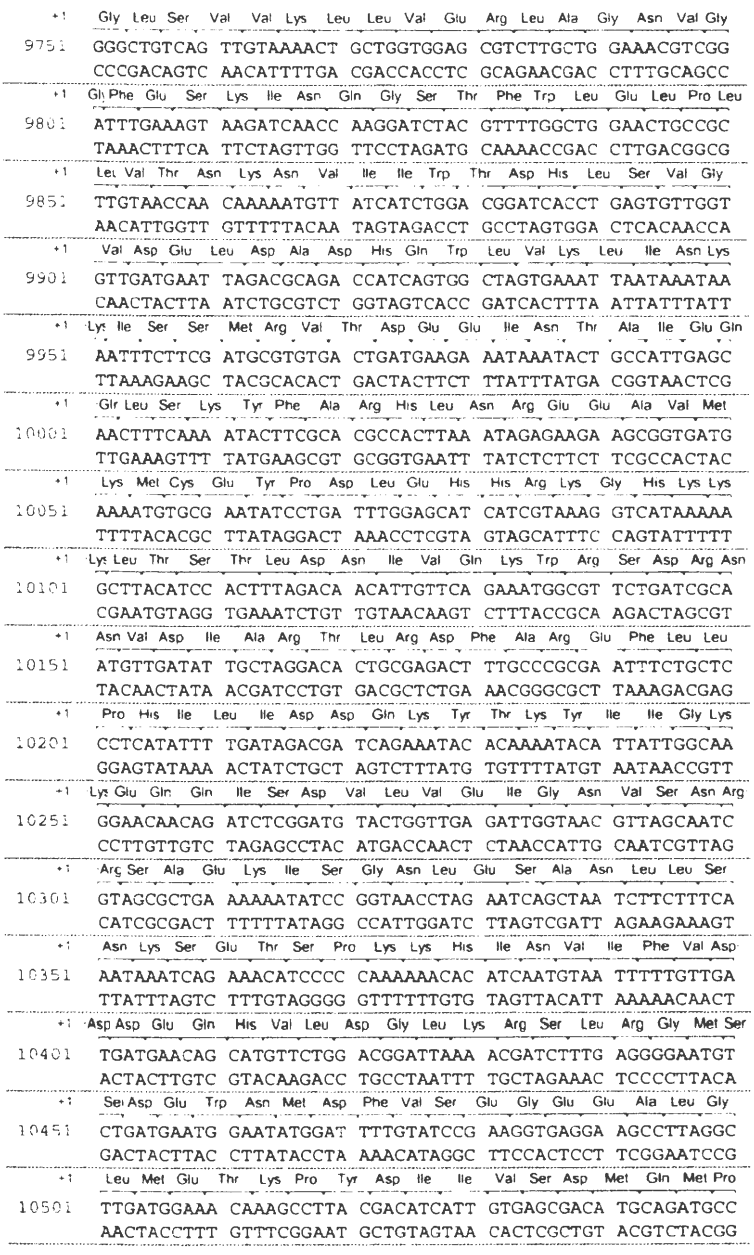
```
5301  TTTGGTGTCT GGGTCTTGAA GTGACGACGC TGATGGCGAA GAAACGGGCT
      AAACCACAGA CCCAGAACTT CACTGCTGCG ACTACCGCTT CTTTGCCCGA
5351  CAGGTGTTTG CATGTTTCGG GAGGACAGAA TACGAATGGC TGCTGCGCCT
      GTCCACAAAC GTACAAAGCC CTCCTGTCTT ATGCTTACCG ACGACGCGGA
5401  TTGTTGTAAG CGTGTTCCAT AACACCCTCC TGTTTTATGT AGCTTTTTGC
      AACAACATTC GCACAAGGTA TTGTGGGAGG ACAAAATACA TCGAAAAACG
5451  ATCAAGCAAC ATGCCCAAAA GTGTGGCCTC GTCGAGATTG TCCATTCCGA
      TAGTTCGTTG TACGGGTTTT CACACCGGAG CAGCTCTAAC AGGTAAGGCT
5501  CAACCCGCAA CACTCTTAAG GGATCATCAC ATGCGGGATG ACCTTGAAAG
      GTTGGGCGTT GTGAGAATTC CCTAGTAGTG TACGCCCTAC TGGAACTTTC
5551  CCGAGCTTAT TGGTTTCTCT TTTGAGCATG CAACACCTCA ATTAAATGTG
      GGCTCGAATA ACCAAAGAGA AAACTCGTAC GTTGTGGAGT TAATTTACAC
5601  TTGTTGGTGA GGTCATCCCT AAATTTTACA ACCCGTTCAC GCAACTTCAA
      AACAACCACT CCAGTAGGGA TTTAAAATGT TGGGCAAGTG CGTTGAAGTT
5651  AATATCTATG CCACTCTTGG CTAAGGCGTA TGCTTCTGCG TTCAGCGTGG
      TTATAGATAC GGTGAGAACC GATTCCGCAT ACGAAGACGC AAGTCGCACC
5701  AAATTGCATC ATCAACAGCA AACTTGAGAT CCAAAAGGAC CGCACAAGAT
      TTTAACGTAG TAGTTGTCGT TTGAACTCTA GGTTTTCCTG GCGTGTTCTA
5751  TTCAACGAAA TATATTTATC TTTTGGATTG AGAGAGATGT GCTGTTTCGG
      AAGTTGCTTT ATATAAATAG AAAACCTAAC TCTCTCTACA CGACAAAGCC
5801  ATCCTTGTTC AGGAGAAACT CTATAGCCTC CGAAGAGAGA TATTTTTTTA
      TAGGAACAAG TCCTCTTTGA GATATCGGAG GCTTCTCTCT ATAAAAAAAT
5851  TAAATTCAAT ATCTAAGCAG TGCATGTCGG TTCCATTTTC ATCTTCGGAC
      ATTTAAGTTA TAGATTCGTC ACGTACAGCC AAGGTAAAAG TAGAAGCCTG
5901  TGATTAAATT CAACATTATG TTGGGTCATT ATACGGGACC TTTTACAGAA
      ACTAATTTAA GTTGTAATAC AACCCAGTAA TATGCCCTGG AAAATGTCTT
5951  TAAAATATCC ATATAACTAT TGCTATATAC AATGCTTACT CCATTTTATG
      ATTTTATAGG TATATTGATA ACGATATATG TTACGAATGA GGTAAAATAC
6001  ACTTTAACAA TCCCCATATG GGGATAGGTA ATGACGTGAT GGTGTGGACG
      TGAAATTGTT AGGGGTATAC CCCTATCCAT TACTGCACTA CCACACCTGC
6051  GTTTATTCCC TTGCTGTTCT GCCATCTGGC GCGAGCGGAT AGTCTTCGTG
      CAAATAAGGG AACGACAAGA CGGTAGACCG CGCTCGCCTA TCAGAAGCAC
6101  AGATTTGCGG CGGCCCTTCG TGTCAACGGC GTAGTAAAAA GGGGCCGCTG
      TCTAAACGCC GCCGGGAAGC ACAGTTGCCG CATCATTTTT CCCCGGCGAC
6151  GGCGGAGCAA TAGTGGTCCC CTTTTCTTGA TGTTGCGCAC GGCGGTGCGG
      CCGCCTCGTT ATCACCAGGG GAAAAGAACT ACAACGCGTG CCGCCACGCC
6201  GTCGTCCCGG TAGCCAACGA GCACCGCCCC GCCGTGCATC TGTTTTTGAC
      CAGCAGGGCC ATCGGTTGCT CGTGGCGGGG CGGCACGTAG ACAAAAACTG
6251  AGCCACGGAG TGCTTTAGGG GACAGGGCAG GGCTGTTCAA CGCTAATTAC
      TCGGTGCCTC ACGAAATCCC CTGTCCCGTC CCGACAAGTT GCGATTAATG
6301  TTTAAAGTGA AGTAAGGGAT TCCCAAAGCG GGTTTTTTCT GAATCTATGT
      AAATTTCACT TCATTCCCTA AGGGTTTCGC CCAAAAAAGA CTTAGATACA
  +3                                  Met  Ser  Leu  Ala  Phe  Leu  Glu Leu
6351  GAGTGCAGAT GAATTACAGG AGCACTCATG TCGTTAGCAT TTTTGGAACT
      CTCACGTCTA CTTAATGTCC TCGTGAGTAC AGCAATCGTA AAAACCTTGA
  +3  Leu Leu  Gly  Gln  Ile  Pro  Asp  Pro  Arg  Arg  Ser  Gln  Gly  Lys  Lys  Trp  Gln Leu
6401  GTTGGGGCAA ATCCCCGACC CGCGCCGTTC GCAGGGTAAG AAATGGCAGC
      CAACCCCGTT TAGGGGCTGG GCGCGGCAAG CGTCCCATTC TTTACCGTCG
```


+3 Leu Gly Pro Val Leu Leu Ala Thr Xxx Leu Ala Ile Leu Ser Gly Ala Thr
6451 TTGGCCCCGT ATTGCTGGCG ACCNTCCTCG CCATTTTGTC CGGGGCCACA
AACCGGGGCA TAACGACCGC TGGNAGGAGC GGTAAAACAG GCCCCGGTGT
+3 Ser Tyr Arg Lys Val His Gly Phe Ile Arg Val His Arg Lys Phe Leu Asn
6501 TCATACCGTA AGGTTCATGG CTTCATTAGG GTTCACCGCA AGTTTCTCAA
AGTATGGCAT TCCAAGTACC GAAGTAATCC CAAGTGGCGT TCAAAGAGTT
+3 Asn Lys Ala Phe Gly Phe Gly Trp Gln Ala Val Pro Ala Tyr Ser Ala Ile Arg
6551 CAAGGCGTTT GGCTTTGGCT GGCAGGCGGT GCCTGCNTAC AGTGCCATTC
GTTCCGCAAA CCGAAACCGA CCGTCCGCCA CGGACGNATG TCACGGTAAG
+3 Arg Thr Ile Leu Arg Gly Leu Asp Gly Ala Glu Val Glu Gln Val Phe Arg
6601 GCACGATACT GCGCGGGCTT GATGGAGCCG AAGTGGAGCA GGTTTTCCGA
CGTGCTATGA CGCGCCCGAA CTACCTCGGC TTCACCTCGT CCAAAAGGCT
+3 Arg His Ala Ala Gln Leu Ser Gly Ser Glu Ala Gly Asp Asp Val Glu Ser
6651 CGACACGCCG CACAACTGAG TGGAAGCGAA GCTGGGGATG ACGTTGAATC
GCTGTGCGGC GTGTTGACTC ACCTTCGCTT CGACCCCTAC TGCAACTTAG
+3 Ser Leu Pro Val Val Ala Ile Asp Gly Lys Thr Leu Arg His Ser Phe Asp Thr
6701 TCTGCCAGTG GTAGCGATTG ACGGTAAGAC GCTCCGGCAC AGTTTTGATA
AGACGGTCAC CATCGCTAAC TGCCATTCTG CGAGGCCGTG TCAAAACTAT
+3 Thr Phe Asn Asp Arg Lys Ala Ala His Val Leu Ser Ala Phe Ala Val Asp
6751 CCTTCAACGA CCGCAAGGCC GCCCATGTTC TAAGCGCCTT TGCCGTCGAT
GGAAGTTGCT GGCGTTCCGG CGGGTACAAG ATTCGCGGAA ACGGCAGCTA
+3 Asn Ala Leu Ile Leu Gly His Leu Glu Val Asp Glu Lys Ser Asn Glu Ile
6801 AACGCCCTGA TCCTTGGGCA TCTGGAAGTG GATGAAAAAA GCAATGAAAT
TTGCGGGACT AGGAACCCGT AGACCTTCAC CTACTTTTTT CGTTACTTTA
+3 Ile Pro Ala Ala Gln Ala Met Ile Glu Ala Leu Asp Leu Glu Gly Arg Leu Phe
6851 CCCCGCTGCC CAAGCGATGA TAGAAGCCCT AGACCTTGAG GGCCGTCTGT
GGGGCGACGG GTTCGCTACT ATCTTCGGGA TCTGGAACTC CCGGCAGACA
+3 Phe Thr Leu Asp Ala Met His Cys Gln Lys Asn Ile Ser Gly Gly Pro Gln
6901 TCACCCTTGA TGCCATGCAC TGCCAAAAAA ACATTTCTGG CGGCCCGCAA
AGTGGGAACT ACGGTACGTG ACGGTTTTTT TGTAAAGACC GCCGGGCGTT
+3 Gly Glu Lys Pro Ser Pro Cys Ser Ser Gln Gly Gln Pro Val Arg Leu Ala
6951 GGTGAAAAGC CATCTCCTTG TTCAAGTCAA GGGCAACCAG TCCGCCTTGC
CCACTTTTCG GTAGAGGAAC AAGTTCAGTT CCCGTTGGTC AGGCGGAACG
+3 Ala Ala Gln Ser Gln Ser Asp His Gly Asn Ile Ala Ala Leu Asn Pro Ala Cys
7001 TGCTCAAAGT CAGAGCGATC ACGGAAACAT TGCCGCCCTT AACCCAGCAT
ACGAGTTTCA GTCTCGCTAG TGCCTTTGTA ACGGCGGGAA TTGGGTCGTA
+3 Cys Asn Gly Ser Thr Pro Ile Arg Arg Ser Arg His Glu Thr Arg Arg Val
7051 GCAATGGATC GACCCCAATC AGGCGTTCGC GGCATGAAAC CCGCCGTGTC
CGTTACCTAG CTGGGGTTAG TCCGCAAGCG CCGTACTTTG GGCGGCACAG
+3 Glu Val Phe Asp Ala Val Pro Gly Leu Phe Lys Thr Gln Trp Asn Gly Leu
7101 GAGGTTTTCG ATGCTGTTCC TGGTCTCTTT AAAACCCAAT GGAACGGCCT
CTCCAAAAGC TACGACAAGG ACCAGAGAAA TTTTGGGTTA CCTTGCCGGA
+3 Leu Ile Lys Arg Val Val Arg Val Thr Arg Ser Thr Leu Ile Arg Arg Thr Lys
7151 CATTAAACGG GTTGTCCGTG TCACCAGATC AACCTTGATC CGCCGTACTA
GTAATTTGCC CAACAGGCAC AGTGGTCTAG TTGGAACTAG GCGGCATGAT
+3 Lys Asp Gly Leu Trp Asp Arg Arg Glu Glu Val Ser Phe Tyr Leu Cys Ser
7201 AAGACGGCTT GTGGGACCGC AGGGAAGAGG TTTCATTTTA CCTTTGCTCA
TTCTGCCGAA CACCCTGGCG TCCCTTCTCC AAAGTAAAAT GGAAACGAGT


+3 Ala Ala Ile Ser Ala Glu Lys Ser Ala Gly Ala Ile Arg Ser His Trp Gly
7251 GCCGCCATCA GCGCCGAGAA ATCCGCCGGT GCCATTCGCA GCCACTGGGG
CGGCGGTAGT CGCGGCTCTT TAGGCGGCCA CGGTAAGCGT CGGTGACCCC
+3 Gly Val Glu Asn Arg Asn His Tyr Val Arg Asp Val Ala Met Gln Glu Asp Ala
7301 TGTTGAAAAC CGCAATCACT ACGTCCGCGA TGTCGCTATG CAAGAAGACG
ACAACTTTTG GCGTTAGTGA TGCAGGCGCT ACAGCGATAC GTTCTTCTGC
+3 Ala Ser Arg Ile Arg Thr Asn Pro Gly Ile Phe Ala Arg Ala Arg Ser Phe
7351 CCAGTCGTAT CAGAACAAAC CCAGGCATCT TCGCCCGTGC CCGCAGCTTC
GGTCAGCATA GTCTTGTTTG GGTCCGTAGA AGCGGGCACG GGCGTCGAAG
+3 Ala Leu Asn Ile Leu Arg Ile Asn Gly Glu Glu Asn Ile Ala Asn Ala Leu
7401 GCCCTCAATA TCCTGCGAAT AAACGGTGAG GAAAACATTG CCAATGCCCT
CGGGAGTTAT AGGACGCTTA TTTGCCACTC CTTTTGTAAC GGTTACGGGA
+3 Leu Trp Cys Asn Ala Leu Asp Ile Asn Arg Val Phe Asp Tyr Arg Leu Lys ***
7451 ATGGTGCAAT GCCCTCGATA TCAACAGAGT ATTTGATTAT CGTCTCAAAT
TACCACGTTA CGGGAGCTAT AGTTGTCTCA TAAACTAATA GCAGAGTTTA
+3 ***
7501 AGTTAGCGTT GAACAGCCCT GGGGGACAGG GGCGGTTGGG CGGCGGCAAG
TCAATCGCAA CTTGTCGGGA CCCCCTGTCC CCGCCAACCC GCCGCCGTTC
7551 ACCGGGATCC GTTTCAAGAA CAAACTTCTA TCCCTGAATT CGACGACTGT
TGGCCCTAGG CAAAGTTCTT GTTTGAAGAT AGGGACTTAA GCTGCTGACA
7601 GTCCTTGTGC TTGAATCTCT TCCCCTGGGC CAAATATCGC TGAGCAAAGG
CAGGAACACG AACTTAGAGA AGGGGACCCG GTTTATAGCG ACTCGTTTCC
7651 GTGGCGTAAA GGTTCATGTC ATGCTCAATC ACGCCGATTA CATGCCGTCC
CACCGCATTT CCAAGTACAG TACGAGTTAG TGCGGCTAAT GTACGGCAGG
7701 TTCGTGTTGT TCACCAAAAC AGATGACCGT CCTCAATTAC AAGCGCAACC
AAGCACAACA AGTGGTTTTG TCTACTGGCA GGAGTTAATG TTCGCGTTGG
7751 ATTTGCCCGT CATATCCCCA TTTGGGGATA TGTGTTATGA CTTTATAATG
TAAACGGGCA GTATAGGGGT AAACCCCTAT ACACAATACT GAAATATTAC
7801 TTCTAATATG CATAGAAAGC ATTCATGTGA CGTGCCTATG GGTGGGGTGA
AAGATTATAC GTATCTTTCG TAAGTACACT GCACGGATAC CCACCCCACT
+1 Met Glu Val Asn Asp Asn Leu Leu Trp
7851 GTCAACACAG AGGGGAGTCC TGCAAATGGA AGTTAATGAT AATTTATTGT
CAGTTGTGTC TCCCCTCAGG ACGTTTACCT TCAATTACTA TTAAATAACA
+1 Trp Ser Asp Glu Phe Ser Thr His Ile Arg Ala Ile Asp Asn Asp His Lys
7901 GGTCAGATGA GTTTAGCACC CACATTCGGG CTATCGACAA TGATCACAAG
CCAGTCTACT CAAATCGTGG GTGTAAGCCC GATAGCTGTT ACTAGTGTTC
+1 Glu Leu Phe Asp Ile Phe Gly Thr Leu His Ala Tyr Asp Leu Lys His Lys
7951 GAATTATTTG ATATTTTTGG TACCCTTCAT GCGTACGACC TGAAACACAA
CTTAATAAAC TATAAAAACC ATGGGAAGTA CGCATGCTGG ACTTTGTGTT
+1 Lys Asn Thr Glu Gln Ile Glu Asn Val Leu Ala Leu Leu Ser Asn Tyr Ile His
8001 AAACACGGAG CAAATTGAAA ATGTATTAGC GCTTCTATCC AACTATATTC
TTTGTGCCTC GTTTAACTTT TACATAATCG CGAAGATAGG TTGATATAAG
+1 His Tyr His Phe Glu Arg Glu Glu Arg Phe Met Glu Ser Ala Gly Tyr Pro
8051 ACTACCATTT TGAGCGGGAA GAACGTTTTA TGGAAAGCGC GGGCTATCCA
TGATGGTAAA ACTCGCCCTT CTTGCAAAAT ACCTTTCGCG CCCGATAGGT
+1 Asp Ile Asn Glu His Lys Gln Leu His Glu Ile Leu Lys His Asp Val Gln
8101 GACATAAATG AACACAAGCA ACTGCATGAA ATACTAAAAC ATGATGTCCA
CTGTATTTAC TTGTGTTCGT TGACGTACTT TATGATTTTG TACTACAGGT


+1 Glu Gly Leu Arg Ser Leu Phe Thr Thr Ser Pro Glu Glu Val Lys Ile Thr Arg
8151 AGGCTTACGG TCTCTTTTTA CAACCTCTCC AGAGGAAGTA AAAATTACGA
TCCGAATGCC AGAGAAAAAT GTTGGAGAGG TCTCCTTCAT TTTTAATGCT
+1 Arg Ala Val Ala Phe Leu Phe Lys Trp Leu Gln Lys His Ile Leu Lys Glu
8201 GGGCGGTGGC TTTTCTCTTT AAATGGCTGC AAAAGCATAT TCTTAAAGAG
CCCGCCACCG AAAAGAGAAA TTTACCGACG TTTTCGTATA AGAATTTCTC
+1 Asp Met Lys Tyr Val Pro Tyr Met Arg Gly Glu Lys Asn Val Gln Lys Ser
8251 GATATGAAAT ACGTCCCCTA TATGCGTGGT GAAAAAAATG TACAGAAGTC
CTATACTTTA TGCAGGGGAT ATACGCACCA CTTTTTTTAC ATGTCTTCAG
+1 Ser Ala Gln Glu Lys Ala Asp Thr Cys Ile Lys Ile Ser Val Cys Val Pro Arg
8301 AGCACAAGAA AAGGCCGACA CATGCATTAA AATATCTGTC TGTGTACCTC
TCGTGTTCTT TTCCGGCTGT GTACGTAATT TTATAGACAG ACACATGGAG
+1 Arg Ala Lys Ala Ala Leu Ile Glu Asn Cys Ile Glu Val Ile Leu Glu His
8351 GTGCAAAGGC TGCCCTCATT GAGAACTGTA TTGAAGTGAT TTTAGAACAT
CACGTTTCCG ACGGGAGTAA CTCTTGACAT AACTTCACTA AAATCTTGTA
+1 Gly Leu Thr Gly Glu Leu Glu Gln Leu Met Asn Glu Leu Gln Lys Lys Lys
8401 GGCTTGACGG GAGAGTTAGA GCAGTTAATG AATGAACTTC AGAAAAAGAA
CCGAACTGCC CTCTCAATCT CGTCAATTAC TTACTTGAAG TCTTTTTCTT
+1 Lys Leu Lys Arg Ala Arg Phe Leu Phe Thr ***
8451 ACTCAAAAGG GCACGGTTTC TTTTTACATA ATGTCTGTGG TCTGCGGCGT
TGAGTTTTCC CGTGCCAAAG AAAAATGTAT TACAGACACC AGACGCCGCA
+1 Met Ser Glu Phe Leu Asn Ile Asn Gln Gln
8501 GTTGCTTAGG GGATACATTG ATGTCGGAAT TCTTGAATAT TAACCAACAA
CAACGAATCC CCTATGTAAC TACAGCCTTA AGAACTTATA ATTGGTTGTT
+1 Arg Met Leu Ser Met Leu Leu Arg Val Thr Leu Ile Phe Met Gly Cys Glu
8551 CGCATGTTAT CTATGCTTCT CAGAGTGACT CTTATTTTTA TGGGCTGCGA
GCGTACAATA GATACGAAGA GTCTCACTGA GAATAAAAAT ACCCGACGCT
+1 Glu Leu Phe Val Met Phe Val Leu Asp Val Phe Ile Lys Pro Leu Ser Leu Ser
8601 ACTTTTTGTC ATGTTTGTAC TGGACGTTTT TATAAAACCT CTAAGCCTTA
TGAAAAACAG TACAAACATG ACCTGCAAAA ATATTTTGGA GATTCGGAAT
+1 Ser Phe Ser Ile Glu Pro Ile Leu Asp Val Val Ile Leu Gly Ile Ile Ala
8651 GTTTTTCCAT AGAGCCTATT TTAGATGTTG TAATACTGGG CATAATTGCA
CAAAAAGGTA TCTCGGATAA AATCTACAAC ATTATGACCC GTATTAACGT
+1 Ser Pro Leu Val Tyr Phe Phe Ala Ile Arg Pro Phe Met Phe Lys Gln Ser
8701 TCACCTCTTG TGTATTTTTT TGCAATCCGC CCCTTCATGT TTAAGCAGTC
AGTGGAGAAC ACATAAAAAA ACGTTAGGCG GGGAAGTACA AATTCGTCAG
+1 Ser Leu Phe Arg Gln Val Thr Lys Lys Ile Gly Glu Val Phe Trp Met Ser Asp
8751 CCTTTTTCGT CAGGTGACAA AAAAAATAGG CGAAGTATTT TGGATGAGTG
GGAAAAAGCA GTCCACTGTT TTTTTTATCC GCTTCATAAA ACCTACTCAC
+1 Asp Pro Gly Lys Ser Lys Ile Leu Tyr Ile Ser Pro Ala Tyr Glu Thr Ile
8801 ATCCAGGTAA AAGTAAAATA CTTTATATTA GCCCCGCCTA TGAAACCATT
TAGGTCCATT TTCATTTTAT GAAATATAAT CGGGGCGGAT ACTTTGGTAA
+1 Trp Gly Arg Thr Cys Lys Ser Leu Tyr Gln Asn Pro Gln Ser Phe Ile Glu
8851 TGGGGGAGAA CGTGCAAAAG CCTATACCAA AACCCGCAAA GTTTTATAGA
ACCCCCTCTT GCACGTTTTC GGATATGGTT TTGGGCGTTT CAAAATATCT
+1 Glu Ala Ile Cys Gln Glu His Arg Glu Arg Val Val Ala Ala Met Gly Lys Gln
8901 AGCGATATGC CAGGAACATC GGGAACGGGT TGTTGCCGCA ATGGGGAAAC
TCGCTATACG GTCCTTGTAG CCCTTGCCCA ACAACGGCGT TACCCCTTTG


+1 Glr Ala Leu Gly Gly Tyr Asp Glu Leu Tyr Lys Ile Glu Arg Pro Asp Gly
8951 AAGCCCTCGG TGGCTACGAC GAACTTTATA AAATTGAACG TCCAGATGGA
TTCGGGAGCC ACCGATGCTG CTTGAAATAT TTTAACTTGC AGGTCTACCT
+1 Gln Ser Cys Trp Ile Trp Asp Arg Ala Phe Pro Val Lys Asp Lys His Gly
9001 CAGTCCTGTT GGATTTGGGA CAGGGCTTTT CCAGTCAAGG ATAAACACGG
GTCAGGACAA CCTAAACCCT GTCCCGAAAA GGTCAGTTCC TATTTGTGCC
+1 Gln Asn Val Val Arg Ile Asp Gly Val Ala Ala Asp Ile Thr Glu Gln Lys Glu
9051 CAATGTTGTT CGGATTGATG GCGTCGCTGC CGACATAACG GAGCAAAAAG
GTTACAACAA GCCTAACTAC CGCAGCGACG GCTGTATTGC CTCGTTTTTC
+1 Glu Thr Glu Leu Glu Leu Leu Arg Ser Glu His Asp Leu Ser Val Ala Leu
9101 AGACTGAACT GGAACTGCTC AGATCTGAGC ACGATCTCTC TGTTGCTTTA
TCTGACTTGA CCTTGACGAG TCTAGACTCG TGCTAGAGAG ACAACGAAAT
+1 Glu Lys Ala Thr Ala Ala Ser Lys Ala Lys Ser Glu Phe Leu Ala Ser Met
9151 GAAAAAGCAA CGGCTGCAAG CAAAGCTAAA TCTGAGTTTT TAGCTTCTAT
CTTTTTCGTT GCCGACGTTC GTTTCGATTT AGACTCAAAA ATCGAAGATA
+1 Met Ser His Glu Leu Arg Thr Pro Leu Asn Ala Ile Leu Gly Phe Ala Gln Met
9201 GAGTCATGAG CTGCGCACCC CCTTGAATGC CATCCTGGGT TTTGCCCAAA
CTCAGTACTC GACGCGTGGG GGAACTTACG GTAGGACCCA AAACGGGTTT
+1 Met Met Gln Tyr Arg Lys Asp Glu Pro Leu Met Pro Gly Gln Lys Glu Ala
9251 TGATGCAATA CAGGAAAGAC GAGCCTCTGA TGCCTGGACA GAAAGAGGCT
ACTACGTTAT GTCCTTTCTG CTCGGAGACT ACGGACCTGT CTTTCTCCGA
+1 Val Asp Ile Ile Leu Thr Ser Gly Val Gly Leu Leu Asn Leu Ile Asn Glu
9301 GTGGATATCA TTTTGACCAG CGGTGTTGGT CTCCTCAATC TTATCAATGA
CACCTATAGT AAAACTGGTC GCCACAACCA GAGGAGTTAG AATAGTTACT
+1 Glu Val Leu Asp Leu Ala Arg Ile Glu Ala Asp Arg Ala Thr Ile Tyr Leu Asp
9351 AGTTCTTGAC CTAGCCAGGA TTGAAGCTGA TCGCGCCACA ATCTATCTTG
TCAAGAACTG GATCGGTCCT AACTTCGACT AGCGCGGTGT TAGATAGAAC
+1 Asp Thr Val Asp Ala Gly Lys Ile Val Ser Asp Cys Val Ser Met Leu Thr
9401 ATACTGTTGA TGCGGGTAAA ATTGTTTCAG ATTGTGTTTC GATGCTGACT
TATGACAACT ACGCCCATTT TAACAAAGTC TAACACAAAG CTACGACTGA
+1 Pro Leu Ala Gln Lys Arg His Ile Lys Ile Lys Asn Met Val Ser Ser Ala
9451 CCATTAGCCC AAAAGCGACA TATTAAAATC AAAAATATGG TGAGTAGCGC
GGTAATCGGG TTTTCGCTGT ATAATTTTAG TTTTTATACC ACTCATCGCG
+1 Ala Ser Ala Met Leu Asn Thr Asp Pro Asp Arg Phe Lys Gln Met Val Ile Asn
9501 ATCGGCTATG CTCAACACCG ACCCAGATCG CTTTAAGCAG ATGGTCATAA
TAGCCGATAC GAGTTGTGGC TGGGTCTAGC GAAATTCGTC TACCAGTATT
+1 Asn Tyr Ile Thr Asn Ala Ile Lys Tyr Asn Lys Glu Gly Gly Leu Ile Thr
9551 ATTACATTAC AAATGCAATC AAATATAACA AAGAAGGAGG CCTTATCACA
TAATGTAATG TTTACGTTAG TTTATATTGT TTCTTCCTCC GGAATAGTGT
+1 Ile Glu Gly Ser Ile Thr Asp His Gly Phe Phe His Leu Ser Val Thr Asp
9601 ATCGAGGGGA GCATAACCGA TCACGGATTT TTTCATCTCT CGGTAACCGA
TAGCTCCCCT CGTATTGGCT AGTGCCTAAA AAAGTAGAGA GCCATTGGCT
+1 Asp Thr Gly Ile Gly Ile Glu Asp Lys Asn Leu Pro Lys Val Phe Glu Ile Tyr
9651 TACGGGCATT GGTATTGAGG ATAAAAACCT TCCTAAAGTT TTTGAAATTT
ATGCCCGTAA CCATAACTCC TATTTTTGGA AGGATTTCAA AAACTTTAAA
+1 Tyr Asn Arg Phe Gln Thr Asn Pro His Ile Thr Val Glu Gly Thr Gly Val
9701 ATAATCGCTT TCAAACTAAC CCACACATCA CTGTTGAAGG TACGGGGGTG
TATTAGCGAA AGTTTGATTG GGTGTGTAGT GACAACTTCC ATGCCCCCAC


+1 Gly Leu Ser Val Val Lys Leu Leu Val Glu Arg Leu Ala Gly Asn Val Gly
9751 GGGCTGTCAG TTGTAAAACT GCTGGTGGAG CGTCTTGCTG GAAACGTCGG
CCCGACAGTC AACATTTTGA CGACCACCTC GCAGAACGAC CTTTGCAGCC
+1 Gly Phe Glu Ser Lys Ile Asn Gln Gly Ser Thr Phe Trp Leu Glu Leu Pro Leu
9801 ATTTGAAAGT AAGATCAACC AAGGATCTAC GTTTTGGCTG GAACTGCCGC
TAAACTTTCA TTCTAGTTGG TTCCTAGATG CAAAACCGAC CTTGACGGCG
+1 Leu Val Thr Asn Lys Asn Val Ile Ile Trp Thr Asp His Leu Ser Val Gly
9851 TTGTAACCAA CAAAAATGTT ATCATCTGGA CGGATCACCT GAGTGTTGGT
AACATTGGTT GTTTTTACAA TAGTAGACCT GCCTAGTGGA CTCACAACCA
+1 Val Asp Glu Leu Asp Ala Asp His Gln Trp Leu Val Lys Leu Ile Asn Lys
9901 GTTGATGAAT TAGACGCAGA CCATCAGTGG CTAGTGAAAT TAATAAATAA
CAACTACTTA ATCTGCGTCT GGTAGTCACC GATCACTTTA ATTATTTATT
+1 Lys Ile Ser Ser Met Arg Val Thr Asp Glu Glu Ile Asn Thr Ala Ile Glu Gln
9951 AATTTCTTCG ATGCGTGTGA CTGATGAAGA AATAAATACT GCCATTGAGC
TTAAAGAAGC TACGCACACT GACTACTTCT TTATTTATGA CGGTAACTCG
+1 Gln Leu Ser Lys Tyr Phe Ala Arg His Leu Asn Arg Glu Glu Ala Val Met
10001 AACTTTCAAA ATACTTCGCA CGCCACTTAA ATAGAGAAGA AGCGGTGATG
TTGAAAGTTT TATGAAGCGT GCGGTGAATT TATCTCTTCT TCGCCACTAC
+1 Lys Met Cys Glu Tyr Pro Asp Leu Glu His His Arg Lys Gly His Lys Lys
10051 AAAATGTGCG AATATCCTGA TTTGGAGCAT CATCGTAAAG GTCATAAAAA
TTTTACACGC TTATAGGACT AAACCTCGTA GTAGCATTTC CAGTATTTTT
+1 Lys Leu Thr Ser Thr Leu Asp Asn Ile Val Gln Lys Trp Arg Ser Asp Arg Asn
10101 GCTTACATCC ACTTTAGACA ACATTGTTCA GAAATGGCGT TCTGATCGCA
CGAATGTAGG TGAAATCTGT TGTAACAAGT CTTTACCGCA AGACTAGCGT
+1 Asn Val Asp Ile Ala Arg Thr Leu Arg Asp Phe Ala Arg Glu Phe Leu Leu
10151 ATGTTGATAT TGCTAGGACA CTGCGAGACT TTGCCCGCGA ATTTCTGCTC
TACAACTATA ACGATCCTGT GACGCTCTGA AACGGGCGCT TAAAGACGAG
+1 Pro His Ile Leu Ile Asp Asp Gln Lys Tyr Thr Lys Tyr Ile Ile Gly Lys
10201 CCTCATATTT TGATAGACGA TCAGAAATAC ACAAAATACA TTATTGGCAA
GGAGTATAAA ACTATCTGCT AGTCTTTATG TGTTTTATGT AATAACCGTT
+1 Lys Glu Gln Gln Ile Ser Asp Val Leu Val Glu Ile Gly Asn Val Ser Asn Arg
10251 GGAACAACAG ATCTCGGATG TACTGGTTGA GATTGGTAAC GTTAGCAATC
CCTTGTTGTC TAGAGCCTAC ATGACCAACT CTAACCATTG CAATCGTTAG
+1 Arg Ser Ala Glu Lys Ile Ser Gly Asn Leu Glu Ser Ala Asn Leu Leu Ser
10301 GTAGCGCTGA AAAAATATCC GGTAACCTAG AATCAGCTAA TCTTCTTTCA
CATCGCGACT TTTTTATAGG CCATTGGATC TTAGTCGATT AGAAGAAAGT
+1 Asn Lys Ser Glu Thr Ser Pro Lys Lys His Ile Asn Val Ile Phe Val Asp
10351 AATAAATCAG AAACATCCCC CAAAAAACAC ATCAATGTAA TTTTTGTTGA
TTATTTAGTC TTTGTAGGGG GTTTTTTGTG TAGTTACATT AAAAACAACT
+1 Asp Asp Glu Gln His Val Leu Asp Gly Leu Lys Arg Ser Leu Arg Gly Met Ser
10401 TGATGAACAG CATGTTCTGG ACGGATTAAA ACGATCTTTG AGGGGAATGT
ACTACTTGTC GTACAAGACC TGCCTAATTT TGCTAGAAAC TCCCCTTACA
+1 Ser Asp Glu Trp Asn Met Asp Phe Val Ser Glu Gly Glu Glu Ala Leu Gly
10451 CTGATGAATG GAATATGGAT TTTGTATCCG AAGGTGAGGA AGCCTTAGGC
GACTACTTAC CTTATACCTA AAACATAGGC TTCCACTCCT TCGGAATCCG
+1 Leu Met Glu Thr Lys Pro Tyr Asp Ile Ile Val Ser Asp Met Gln Met Pro
10501 TTGATGGAAA CAAAGCCTTA CGACATCATT GTGAGCGACA TGCAGATGCC
AACTACCTTT GTTTCGGAAT GCTGTAGTAA CACTCGCTGT ACGTCTACGG

```
      +1  Pro Thr  Met  Ser  Gly  Glu  Glu  Leu  Leu  Gly  Lys  Val  Glu  Lys  His  Tyr  Pro Ser
10551     GACAATGAGC GGAGAAGAGT TGTTGGGAAA AGTAGAAAAA CACTATCCTT
          CTGTTACTCG CCTCTTCTCA ACAACCCTTT TCATCTTTTT GTGATAGGAA
      +1  Ser Thr  Ala  Arg  Val  Val  Leu  Ser  Gly  His  Val  Asp  Gln  Asp  Ala  Thr  Tyr
10601     CCACAGCCCG TGTTGTGCTT TCTGGGCACG TTGATCAGGA TGCCACGTAT
          GGTGTCGGGC ACAACACGAA AGACCCGTGC AACTAGTCCT ACGGTGCATA
      +1  Arg  Leu  Val  Gly  Ser  Ser  His  Leu  Phe  Leu  Ser  Lys  Pro  Cys  Ser  Thr  Glu
10651     CGTCTGGTGG GTTCCAGCCA TCTGTTTCTT TCAAAACCTT GTTCCACTGA
          GCAGACCACC CAAGGTCGGT AGACAAAGAA AGTTTTGGAA CAAGGTGACT
      +1  Glu  Leu  Leu  Ile  Asp  Thr  Met  Arg  Lys  Ser  Ile  Ser  Leu  Ser  Asn  Thr  Gln  Gly
10701     GCTTCTCATT GACACAATGA GAAAGTCGAT TTCCTTGAGC AATACACAAG
          CGAAGAGTAA CTGTGTTACT CTTTCAGCTA AAGGAACTCG TTATGTGTTC
      +1  Gly  Leu  Ile  Ser  Asn  ***
10751     GCCTAATATC GAATTAGTTT TATTGCCCGG CAGGGCTTGA CTGATATACT
          CGGATTATAG CTTAATCAAA ATAACGGGCC GTCCCGAACT GACTATATGA
10801     CCGACTCATT GAATAAGCGA AAAATCCTTG CCAAATTCAT TTTAAAGGAT
          GGCTGAGTAA CTTATTCGCT TTTTAGGAAC GGTTTAAGTA AAATTTCCTA
      +1                    Met  Ile  Arg  Ala  Gly  Phe  Leu  Thr  Thr  Thr  Glu  Arg  Ala
10851     TCTTGTTGTC AGCATGATAC GCGCAGGATT TTTAACCACG ACGGAACGGG
          AGAACAACAG TCGTACTATG CGCGTCCTAA AAATTGGTGC TGCCTTGCCC
      +1  Ala  Glu  Leu  Ile  Asp  Leu  Ala  Arg  Asn  Gly  Leu  Val  Glu  His  Arg  Leu  Ala
10901     CAGAGTTGAT TGATTTGGCT CGGAATGGCT TGGTCGAGCA TCGTCTGGCG
          GTCTCAACTA ACTAAACCGA GCCTTACCGA ACCAGCTCGT AGCAGACCGC
      +1  Arg  Arg  Ala  Asn  Ala  Leu  Val  Leu  Leu  Asp  Arg  Gly  Met  Ser  Cys  Ala  Glu
10951     CGCCGCGCCA ACGCACTTGT GCTTCTGGAT CGAGGCATGA GTTGCGCGGA
          GCGGCGCGGT TGCGTGAACA CGAAGACCTA GCTCCGTACT CAACGCGCCT
      +1  Glu  Val  Ser  Ser  Val  Phe  Leu  Leu  Asp  Asp  Asp  Thr  Val  Arg  Thr  Trp  Phe  Arg
11001     GGTCAGTTCG GTATTTCTGC TTGATGATGA TACGGTCAGG ACTTGGTTTC
          CCAGTCAAGC CATAAAGACG AACTACTACT ATGCCAGTCC TGAACCAAAG
      +1  Arg  Leu  Tyr  Arg  Glu  Glu  Gly  Ile  Glu  Gly  Leu  Ala  Gly  Phe  Gly  Tyr  Glu
11051     GGCTCTATAG GGAAGAGGGG ATCGAGGGTT TGGCCGGTTT TGGCTACGAG
          CCGAGATATC CCTTCTCCCC TAGCTCCCAA ACCGGCCAAA ACCGATGCTC
      +1  Gly  Ser  Ala  Gly  His  Leu  Asn  Ala  Arg  Gln  Leu  Gln  Gln  Leu  Lys  Ser  Trp
11101     GGTAGCGCCG GACATCTGAA CGCCCGGCAG CTCCAGCAAC TGAAATCCTG
          CCATCGCGGC CTGTAGACTT GCGGGCCGTC GAGGTCGTTG ACTTTAGGAC
      +1  Trp  Ile  Thr  Asp  Thr  Leu  Pro  Arg  Thr  Thr  Arg  His  Val  Gly  Ala  Trp  Ile  Glu
11151     GATTACCGAT ACCTTGCCGC GCACGACGCG CCATGTCGGG GCGTGGATTG
          CTAATGGCTA TGGAACGGCG CGTGCTGCGC GGTACAGCCC CGCACCTAAC
      +1  Glu  Gln  Glu  Phe  Gly  Ile  Thr  Tyr  Gln  Ser  Arg  Ser  Gly  Leu  Val  Ala  Leu
11201     AGCAAGAGTT CGGGATCACC TATCAAAGCC GCTCGGGGTT GGTTGCGCTG
          TCGTTCTCAA GCCCTAGTGG ATAGTTTCGG CGAGCCCCAA CCAACGCGAC
      +1  Leu  His  Arg  Leu  Gly  Met  Glu  His  Arg  Lys  Pro  Gln  Ala  Val  Ala  Arg  Ile
11251     CTTCATCGTT TGGGGATGGA GCACCGCAAG CCGCAAGCGG TGGCGCGTAT
          GAAGTAGCAA ACCCCTACCT CGTGGCGTTC GGCGTTCGCC ACCGCGCATA
      +1  Ile  ***
11301     TTAGGCCTGA CATCAAGGCG ATATCAATCA GGTGAGATGG ATTATTCGGG
          AATCCGGACT GTAGTTCCGC TATAGTTAGT CCACTCTACC TAATAAGCCC
11351     TCGCATCTCA AAATTTGGTG ACGGCTTATT GCGCAGTCTT AATTTACGAA
          AGCGTAGAGT TTTAAACCAC TGCCGAATAA CGCGTCAGAA TTAAATGCTT
```

```
11401  GCCGCGAACA GCCTGTTGAC CATTGTTCGC AAGTCCCATC CGATGAAAGA
       CGGCGCTTGT CGGACAACTG GTAACAAGCG TTCAGGGTAG GCTACTTTCT
11451  TTGGGCCCGA CGCATTCGCA AACGTATGCA TTGTCGTTGA TAAAGGTTCA
       AACCCGGGCT GCGTAAGCGT TTGCATACGT AACAGCAACT ATTTCCAAGT
11501  TGGTAAGCGG TGCGATGATC ACACGGGGGT TATTTCTGCG GTATGGTTCC
       ACCATTCGCC ACGCTACTAG TGTGCCCCCA ATAAAGACGC CATACCAAGG
11551  AAGGTAGGAG CGACTCGATC TGGCTTTGCC TGTGCCCTGC TACGATGGCC
       TTCCATCCTC GCTGAGCTAG ACCGAAACGG ACACGGGACG ATGCTACCGG
11601  TTGAGAGTTG CCGTTATATA GAGCCCACTC ATTAAGTCTA AGTAATTGAA
       AACTCTCAAC GGCAATATAT CTCGGGTGAG TAATTCAGAT TCATTAACTT
11651  TTTAAATAAT ATATTTCATT ATATTTATGT TATAATGCAA GCTGGACACC
       AAATTTATTA TATAAAGTAA TATAAATACA ATATTACGTT CGACCTGTGG
11701  CGCATTAACC CAAGAATTTT AGTGCTTTGG GCTCGGATAT CCGGTTTTGG
       GCGTAATTGG GTTCTTAAAA TCACGAAACC CGAGCCTATA GGCCAAAACC
11751  CGGATATTGC TCGAGATTTT GGCTCTCCGA TCGCGTTGAT CGGGCTGTTT
       GCCTATAACG AGCTCTAAAA CCGAGAGGCT AGCGCAACTA GCCCGACAAA
11801  TGCTGTGAGT TTCCGCGCTT TGGGCGGACT CATCCCATGG CGGCGCGGCG
       ACGACACTCA AAGGCGCGAA ACCCGCCTGA GTAGGGTACC GCCGCGCCGC
  -2                                        ***  Gly Met Ala Ala Arg Arg
11851  TTCATGGAAG ATCAGGCGGT CCATGTTGTA AGCGAGGTTG GCGAGGGTCA
       AAGTACCTTC TAGTCCGCCA GGTACAACAT TCGCTCCAAC CGCTCCCAGT
  -2   Glu His Phe Ile Leu Arg Asp Met Asn Tyr Ala Leu Asn Ala Leu Thr Leu
11901  GTTTGGCCTC GGCCCGCTTG ATGCCGATGG TGCGGATGAA TAGACCGAAA
       CAAACCGGAG CCGGGCGAAC TACGGCTACC ACGCCTACTT ATCTGGCTTT
  -2   Leu Lys Ala Glu Ala Arg Lys Ile Gly Ile Thr Arg Ile Phe Leu Gly Phe Arg
11951  CGGTTTTTCT GGTGGGCGAA GACGTGCTCC ACATGGGCAC GGATTGATGA
       GCCAAAAAGA CCACCCGCTT CTGCACGAGG TGTACCCGTG CCTAACTACT
  -2   Arg Asn Lys Gln His Ala Phe Val His Glu Val His Ala Arg Ile Ser Ser
12001  TTTCTTGGCG TTGGCTTTAC GCGTGGCTTC GGGCATCGCT TTGCCCTTGG
       AAAGAACCGC AACCGAAATG CGCACCGAAG CCCGTAGCGA AACGGGAACC
  -2   Lys Lys Ala Asn Ala Lys Arg Thr Ala Glu Pro Met Ala Lys Gly Lys Pro
12051  GCTTCTTGCG ATGAATGCGG CTGACCAGCA TCTTCCCCGC AAGCCATTTT
       CGAAGAACGC TACTTACGCC GACTGGTCGT AGAAGGGGCG TTCGGTAAAA
  -2   Pro Lys Lys Arg His Ile Arg Ser Val Leu Met Lys Gly Ala Leu Trp Lys Glu
12101  TCGTTATTTT GCGAGCGGTA GGCGCTGTCG GCCCAGACCT CACCCGAGGT
       AGCAATAAAA CGCTCGCCAT CCGCGACAGC CGGGTCTGGA GTGGGCTCCA
  -2   Glu Asn Asn Gln Ser Arg Tyr Ala Ser Asp Ala Trp Val Glu Gly Ser Thr
12151  GTTGTCCGTG CCAACCACAT GCTTGAGTTG GCGACCATCG GGTGCGGAGG
       CAACAGGCAC GGTTGGTGTA CGAACTCAAC CGCTGGTAGC CCACGCCTCC
  -2   Asn Asp Thr Gly Val Val His Lys Leu Gln Arg Gly Asp Pro Ala Ser Ala
12201  CCGACGTCAC ATGTTGCCCG CGAATAAAGC CGAAACGGCG GTCAATGCTG
       GGCTGCAGTG TACAACGGGC GCTTATTTCG GCTTTGCCGC CAGTTACGAC
  -2   Ala Ser Thr Val His Gln Gly Arg Ile Phe Gly Phe Arg Arg Asp Ile Ser Ile
12251  ATGTGCGACT TGTAACCGAA CACCGGCAGC GCGATCATCG GCAGCGGCGT
       TACACGCTGA ACATTGGCTT GTGGCCGTCG CGCTAGTAGC CGTCGCCGCA
  -2   Ile His Ser Lys Tyr Gly Phe Val Pro Leu Ala Ile Met Pro Leu Pro Thr
```

```
12301  GCCGTCGGGC CGATAGCGCA CCTTGCCGCC GATCTTCAGC GTCCAGCGCG
       CGGCAGCCCG GCTATCGCGT GGAACGGCGG CTAGAAGTCG CAGGTCGCGC
  -2   Gly Asp Pro Arg Tyr Arg Val Lys Gly Gly Ile Lys Leu Thr Trp Arg Ala
12351  CATCCACGTC TTTCTGCGCT GCTTTGTTGG GTTCATCAGG CCAAATCTCG
       GTAGGTGCAG AAAGACGCGA CGAAACAACC CAAGTAGTCC GGTTTAGAGC
  -2   Ala Asp Val Asp Lys Gln Ala Ala Lys Asn Pro Glu Asp Pro Trp Ile Glu His
12401  TGAGCGCTCT TGCCCGCCTT GACGGCCTCC TTCTCCGGAT CGGTATTGCG
       ACTCGCGAGA ACGGGCGGAA CTGCCGGAGG AAGAGGCCTA GCCATAACGC
  -2   His Ala Ser Lys Gly Ala Lys Val Ala Glu Lys Glu Pro Asp Thr Asn Arg
12451  CTGCTTGGGC GCCGGAACAA GGCTGGCGTC GACGATTTGT CCCGACATCG
       GACGAACCCG CGGCCTTGTT CCGACCGCAG CTGCTAAACA GGGCTGTAGC
  -2   Gln Lys Pro Ala Pro Val Leu Ser Ala Asp Val Ile Gln Gly Ser Met Pro
12501  GAATGTAGCC TTTCTTGTGG AGTTGCCAGT CGAACGCCTT CATCATCTGG
       CTTACATCGG AAAGAACACC TCAACGGTCA GCTTGCGGAA GTAGTAGACC
  -2   Pro Ile Tyr Gly Lys Lys His Leu Gln Trp Asp Phe Ala Lys Met Met
12551  CACGATGGTC AGGGTGCCTG TTTGTTTACC AAGCGCCTTG AGCGAGGCCG
       GTGCTACCAG TCCCACGGAC AAACAAATGG TTCGCGGAAC TCGCTCCGGC
12601  TTTTTTGTGG CCTTCGGCCG CTGACGGGGT GGTGGCAATC ACCTCCGCGC
       AAAAAACACC GGAAGCCGGC GACTGCCCCA CCACCGTTAG TGGAGGCGCG
12651  AGTTGGGTTA TTTGCTCGAA GGGATCGATT GGCGCATGCC GCAAAAAACC
       TCAACCCAAT AAACGAGCTT CCCTAGCTAA CCGCGTACGG CGTTTTTTGG
12701  TGGCGTCCGA CATCCGTTGG ATGATGTTTT GACTTGGCTT TTGGGGCCGG
       ACCGCAGGCT GTAGGCAACC TACTACAAAA CTGAACCGAA AACCCCGGCC
  +1                                Met Glu Lys His Glu Asn Pro Thr Thr Glu Pro
12751  ACGTGATTCC ATAAGCTCAT GGAAAAGCAT GAGAATCCAA CGACTGAACC
       TGCACTAAGG TATTCGAGTA CCTTTTCGTA CTCTTAGGTT GCTGACTTGG
  +1   Pro Met Pro Ile Asp Leu Asp Gly Ala Leu Ser Met Leu Leu Cys Glu Arg Ala
12801  CATGCCCATC GACCTTGATG GCGCGCTTTC GATGCTGCTT TGCGAACGCG
       GTACGGGTAG CTGGAACTAC CGCGCGAAAG CTACGACGAA ACGCTTGCGC
  +1   Ala Ala Arg Val Lys Ala Glu Ala Glu Ala Ala His Val Gln Ala Val Leu
12851  CGGCACGGGT CAAGGCTGAG GCTGAGGCGG CCCATGTTCA GGCGGTTCTG
       GCCGTGCCCA GTTCCGACTC CGACTCCGCC GGGTACAAGT CCGCCAAGAC
  +1   Ser Ser Ser Glu Ala Leu Ile Met His Leu Arg Leu Glu Ile Glu Lys Leu
12901  TCGAGCAGCG AAGCCCTGAT TATGCATCTC AGGCTGGAGA TTGAAAAGCT
       AGCTCGTCGC TTCGGGACTA ATACGTAGAG TCCGACCTCT AACTTTTCGA
  +1   Leu Arg Arg Glu Leu Phe Gly Thr Arg Ser Glu Arg Lys Ala Arg Leu Leu Asp
12951  GCGGCGCGAA CTCTTTGGCA CGCGTTCGGA GCGCAAGGCC CGGCTGCTCG
       CGCCGCGCTT GAGAAACCGT GCGCAAGCCT CGCGTTCCGG GCCGACGAGC
  +1   Asp Gln Leu Glu Leu Gln Leu Glu Asp Leu Glu Ala Ala Ala Ser Glu Asp
13001  ATCAATTGGA ACTGCAACTC GAAGATCTCG AGGCTGCGGC GTCCGAAGAC
       TAGTTAACCT TGACGTTGAG CTTCTAGAGC TCCGACGCCG CAGGCTTCTG
  +1   Glu Arg Ala Ser Glu Gln Ala Ser Arg Leu Thr Asp Val Thr Pro His Arg
13051  GAGCGCGCCA GCGAACAGGC AAGCCGATTG ACGGATGTGA CACCACACCG
       CTCGCGCGGT CGCTTGTCCG TTCGGCTAAC TGCCTACACT GTGGTGTGGC
  +1   Arg Arg Lys His Pro Gly Arg Gln Ser Phe Pro Glu His Leu Pro Arg Glu Arg
13101  CCGCAAACAC CCCGGACGTC AGTCTTTTCC CGAGCACCTG CCACGCGAGC
       GGCGTTTGTG GGGCCTGCAG TCAGAAAAGG GCTCGTGGAC GGTGCGCTCG
```


+1 Arg Arg Val Ile Ala Ala Pro Ser Thr Cys Pro Cys Cys Gly Ser His Lys
13151 GTCGTGTCAT CGCGGCGCCT TCAACTTGCC CCTGCTGCGG TTCGCACAAA
CAGCACAGTA GCGCCGCGGA AGTTGAACGG GGACGACGCC AAGCGTGTTT
+1 Leu Ser Lys Leu Gly Glu Asp Ile Thr Glu Thr Leu Glu Ile Ile Pro Arg
13201 CTCTCGAAGT TGGGCGAGGA CATCACCGAG ACCTTGGAGA TCATTCCGCG
GAGAGCTTCA ACCCGCTCCT GTAGTGGCTC TGGAACCTCT AGTAAGGCGC
+1 Arg Gln Trp Lys Val Ile Gln Thr Val Arg Glu Lys Phe Ser Cys Arg Glu Cys
13251 CCAGTGGAAG GTGATCCAGA CCGTGCGCGA GAAGTTCTCA TGCCGTGAAT
GGTCACCTTC CACTAGGTCT GGCACGCGCT CTTCAAGAGT ACGGCACTTA
+1 Cys Glu Lys Ile Ala Gln Pro Pro Ala Pro Phe His Val Thr Pro Arg Gly
13301 GCGAGAAGAT TGCCCAACCG CCGGCACCGT TTCACGTCAC GCCGCGCGGC
CGCTCTTCTA ACGGGTTGGC GGCCGTGGCA AAGTGCAGTG CGGCGCGCCG
+1 Phe Leu Gly Pro Asn Leu Leu Ala Met Val Leu Phe Glu Lys Phe Gly Gln
13351 TTTCTGGGGC CGAACCTGCT GGCCATGGTG CTGTTCGAGA AGTTCGGGCA
AAAGACCCCG GCTTGGACGA CCGGTACCAC GACAAGCTCT TCAAGCCCGT
+1 Gln His Gln Pro Leu Asn Arg Gln Ser Ala Arg Phe Gly Arg Glu Gly Ile Asp
13401 GCATCAGCCC CTGAACCGCC AAAGCGCGCG TTTTGGCCGC GAGGGGATCG
CGTAGTCGGG GACTTGGCGG TTTCGCGCGC AAAACCGGCG CTCCCCTAGC
+1 Asp Leu Ser Val Ser Thr Leu Ala Asp Gln Val Gly Val Cys Thr Ser Leu
13451 ATCTGAGCGT CTCGACCCTG GCCGACCAGG TGGGGGTTTG CACGAGCCTA
TAGACTCGCA GAGCTGGGAC CGGCTGGTCC ACCCCCAAAC GTGCTCGGAT
+1 Leu Gln Pro Leu Glu Ala Leu Ile Glu Ala His Val Leu Ala Ala Glu Arg
13501 TTGCAACCCC TGGAAGCGTT GATCGAGGCC CATGTTTTGG CGGCGGAACG
AACGTTGGGG ACCTTCGCAA CTAGCTCCGG GTACAAAACC GCCGCCTTGC
+1 Arg Leu His Gly Asp Asp Thr Thr Val Pro Ile Leu Ala Arg Gly Lys Thr Val
13551 TCTGCATGGC GACGACACGA CCGTGCCGAT CCTGGCCCGT GGCAAGACGG
AGACGTACCG CTGCTGTGCT GGCACGGCTA GGACCGGGCA CCGTTCTGCC
+1 Val Thr Gly Arg Ile Trp Thr Tyr Val Arg Asp Asp Arg Pro Phe Gly Gly
13601 TCACGGGGCG GATATGGACC TATGTGCGCG ATGACCGACC TTTTGGCGGC
AGTGCCCCGC CTATACCTGG ATACACGCGC TACTGGCTGG AAAACCGCCG
+1 Lys Gly Pro Pro Ala Ala Leu Tyr Tyr Ala Ser Arg Asp Arg Arg Gly Ala
13651 AAGGGTCCGC CGGCGGCTCT CTATTATGCC TCGCGGGATC GGCGAGGCGC
TTCCCAGGCG GCCGCCGAGA GATAATACGG AGCGCCCTAG CCGCTCCGCG
+1 Ala His Pro Glu Glu His Leu Glu Gly Trp Ser Gly Ile Leu Gln Ala Asp Ala
13701 GCATCCCGAG GAGCACCTCG AGGGCTGGTC GGGCATTCTG CAGGCCGATG
CGTAGGGCTC CTCGTGGAGC TCCCGACCAG CCCGTAAGAC GTCCGGCTAC
+1 Ala Tyr Ser Gly Tyr Asn Gly Leu Phe Asn Pro Thr Arg Ala Gly Gly Ser
13751 CCTACAGCGG CTACAACGGG CTGTTCAACC CGACACGGGC CGGGGGCTCT
GGATGTCGCC GATGTTGCCC GACAAGTTGG GCTGTGCCCG GCCCCCGAGA
+1 Ala Thr Val Ala Leu Cys Trp Ala His Ala Arg Arg Gln Phe Phe Glu Leu
13801 GCAACGGTGG CGCTCTGCTG GGCGCATGCG CGGCGGCAGT TCTTCGAACT
CGTTGCCACC GCGAGACGAC CCGCGTACGC GCCGCCGTCA AGAAGCTTGA
+1 Leu Ala Asp Ile Ala Ala Asn Ala Arg Arg Gly Arg Thr Ala Gln Ala Ile Ser
13851 GGCCGATATT GCGGCCAACG CGCGGCGTGG TCGCACGGCG CAAGCGATCT
CCGGCTATAA CGCCGGTTGC GCGCCGCACC AGCGTGCCGC GTTCGCTAGA
+1 Ser Pro Val Ala Leu Glu Ala Val Lys Arg Ile Asp Ala Leu Phe Asp Ile
13901 CACCAGTGGC GCTCGAAGCG GTCAAACGTA TCGATGCCCT GTTCGACATC
GTGGTCACCG CGAGCTTCGC CAGTTTGCAT AGCTACGGGA CAAGCTGTAG

```
      +1  Glu Arg Thr  Ile  Ile Gly  Leu  Asn Asp  Ser  Glu Arg Leu  Arg  Val Arg Arg
13951 GAACGCACCA TCATCGGCCT GAACGATAGT GAGCGCCTGC GCGTCCGACG
      CTTGCGTGGT AGTAGCCGGA CTTGCTATCA CTCGCGGACG CGCAGGCTGC
      +1  Arg Glu  Arg  Ser  Ala Pro  Ile  Val  Ala Asp  Leu  Glu  Ala  Trp  Leu Arg  Glu Glu
14001 TGAGCGGAGC GCCCCCATTG TGGCGGATTT GGAAGCTTGG CTGCGTGAGG
      ACTCGCCTCG CGGGGGTAAC ACCGCCTAAA CCTTCGAACC GACGCACTCC
      +1  Glu Arg  Ala  Arg  Leu Ser  Arg  Ala Ser  Ser  Val  Ala Lys  Pro  Ile  Asp  Tyr
14051 AACGCGCGCG CCTATCGCGT GCGTCATCGG TCGCCAAGCC GATCGATTAC
      TTGCGCGCGC GGATAGCGCA CGCAGTAGCC AGCGGTTCGG CTAGCTAATG
      +1  Met Leu  Arg  Arg  Trp Glu  Ala  Phe Ala  Arg  Phe Leu Asp  Asp  Gly  His Ile
14101 ATGCTCAGGC GTTGGGAAGC CTTCGCTCGG TTCCTTGATG ACGGCCATAT
      TACGAGTCCG CAACCCTTCG GAAGCGAGCC AAGGAACTAC TGCCGGTATA
      +1  Ile Cys Leu  Thr  Asn Asn Ala  Ala  Glu Arg  Ala  Leu Arg  Gly  Phe Ala Leu Gly
14151 CTGTCTGACC AACAATGCCG CTGAGCGCGC GCTGCGAGGC TTCGCCTTGG
      GACAGACTGG TTGTTACGGC GACTCGCGCG CGACGCTCCG AAGCGGAACC
      +1  Gly Arg  Lys  Ser  Trp Leu  Phe  Cys Gly  Ser  Glu  Arg Gly  Ala  Glu  Arg  Ala
14201 GACGAAAATC TTGGCTATTT TGCGGCTCGG AACGAGGAGC CGAACGGGCG
      CTGCTTTTAG AACCGATAAA ACGCCGAGCC TTGCTCCTCG GCTTGCCCGC
      +1  Ala Ala  Met  Ala  Thr Leu  Ile  Glu Thr  Ala  Lys Leu Asn  Asp  Val Asp Pro
14251 GCAGCCATGG CAACGCTTAT CGAGACCGCG AAGCTGAACG ACGTCGATCC
      CGTCGGTACC GTTGCGAATA GCTCTGGCGC TTCGACTTGC TGCAGCTAGG
      +1  Pro His  Ala  Trp  Leu Ala  Asp  Val Leu  Gly  Arg  Ile Asn  Asp  Leu Pro  Gln Ser
14301 GCATGCTTGG CTGGCTGATG TGCTGGGGCG CATCAATGAT CTGCCGCAAA
      CGTACGAACC GACCGACTAC ACGACCCCGC GTAGTTACTA GACGGCGTTT
      +1  Ser Arg  Leu  Arg  Glu Leu  Leu  Pro Trp  Glu  Trp  Lys Thr  Lys  Arg  Asp  Val
14351 GCCGCCTGCG CGAACTGCTG CCCTGGGAAT GGAAAACGAA GCGCGATGTG
      CGGCGGACGC GCTTGACGAC GGGACCCTTA CCTTTTGCTT CGCGCTACAC
      +1  Leu Ser  Gly  Pro  Ala Ala  Ser  ***
14401 CTCTCCGGTC CAGCAGCCTC CTGAACCATG GCGGTGATTT CTCACGTATT
      GAGAGGCCAG GTCGTCGGAG GACTTGGTAC CGCCACTAAA GAGTGCATAA
14451 CACAATCCGA CGCGTATCGC AAATACTGGG GCGTGAGGAA GACCTGCTCC
      GTGTTAGGCT GCGCATAGCG TTTATGACCC CGCACTCCTT CTGGACGAGG
14501 GGGAAATCGC CAGCCAACTC GAGCCCGAGG NACGGCATGC TGTGGGTTTA
      CCCTTTAGCG GTCGGTTGAG CTCGGGCTCC NTGCCGTACG ACACCCAAAT
14551 TGATGTCGGC GACAGTCAAA TCCTTGCCTT CACCCAACAA GGGATCGAGA
      ACTACAGCCG CTGTCAGTTT AGGAACGGAA GTGGGTTGTT CCCTAGCTCT
14601 TGCTACGCGA CATCATTGAA GACCAAATCC AATGCGCCGA TTGATCGGTC
      ACGATGCGCT GTAGTAACTT CTGGTTTAGG TTACGCGGCT AACTAGCCAG
14651 GCGTCTGACA AACACGGAAC AAAAAACCGT GGTCTTCACC GGATGGTTAC
      CGCAGACTGT TTGTGCCTTG TTTTTTGGCA CCAGAAGTGG CCTACCAATG
```