Introduction

The true extend of prokaryotic diversity, encompassing the spectrum of variability among bacteria, remains unknown. Early discussions on prokaryotic diversity were frequently devoted to sterile arguments about how much? Or how many? Increasingly, however, the focus is turning towards trying to understand why prokaryotic diversification occurs, its underlying mechanisms and its impact. The significance of such studies has a broad appeal and the popular scientific press highlights such topics as the emergence of new diseases, the attribution of existing diseases to unrealized actions of prokaryotes and the activities of prokaryotes in key environmental processes. The dynamic nature of prokaryotic world and continuing advances in the technological tools available to this field of study ensures that the latest story illustrating prokaryotic diversity is never away (**Logan et al., 2006**).

The organism now known as *E. coli* was first isolated and described by Theodr Escherich in 1885 under the name *bacterium coli commune* from infant faces.

It belongs to family *Enterobacteriacae*.There are several other species in the genus *Escherichia*, the best known is *E.coli*, which has a variety of serotypes. Although *E.coli* is a normal inhabitant of the intestinal tract of man and animals, specific pathogenic forms are associated with different types of intestinal and extraintestinal infections. It can be also used as an indicator of fecal pollution (**Holt et al., 1994 and Azam, 2010**).

It grows readily on simple culture and synthetic media with glycerol or glucose as the sole source of carbon and energy. When growing anaerobically, there is absolute requirement for fermentable carbohydrates. Glucose and other carbohydrates are fermented to produce acid and gas. Lactose is fermented by most strains. It is catalase positive, oxidase negative, indole positive, methyl red positive, voges-proskauer negative, usually citrate negative, negative for H_2S, urea hydrolysis and lipase.

In liquid culture, smooth (S) strains of *E.coli* produce uniform turbidity, while rough strains (R) tend to leave a clear supernatant medium over a granular deposit of growth. The generation time of most strains is 20-30 min. On solid medium, colonies are circular and smooth and some strains produce mucoid colonies (**Macone et al., 1981**). On media containing washed erythrocytes, a soluble α-hemolysin can be demonstrated. There is a cell- associated β- hemolysin which is released from cells by lysis (**Smith, 1963**).

The primary habitat of *E. coli* is the lower intestinal tract of warm blooded animals. Colonization occurs shortly after birth and its source is to be found in the mother. At any one time, most normal individuals carry several strains of *E.coli* in their intestinal tract, including a small number of resident clones exhibiting a rate of replacement measured in weeks or months and much larger number of transient clones that are replaced in a matter of days or weeks. The overall density of *E.coli* in the mammalian colon is approximately 10^6 cells /gram of colon contents. The secondary habitats of *E.coli* are soil, sediment and water usually as a result of fecal

contamination, where its half-life is thought to be only a few days **(Hartl and Dykhuizen, 1984 and Sussman, 1985).**

E.coli is a gram negative, asporogenous, facultative anaerobic, non acid fast, Motile by peritrichous flagella or non motile, fimbriated and straight rods, 1.1-1.5 x 2.0-6.0µm, occurring singly or in pairs. A capsule or microcapsule is often present and few strains produce a polysaccharide slime layer (**Krieg and Holt, 1984 and Holt et al., 1994)**

Strains show a wide range of growth temperature, from 5-45°c with an optimum of 37°c.The minimum water activity value is 0 .94 in the pH range 3.3-10, with pH 6-8 as the average. The genome size is between 2.3 x 10^9 and 3.0 x10^9 daltons and G + C content is 48 to 52% **(Krieg and Holt, 1984; Sussman, 1985 and Holt et al., 1994).**

The differences in the ability of strains to cause disease and the diverse syndromes caused by different strains can be attributed to the existence in some of specific genes encoding virulence factors and the capacity of *E.coli* for genetic exchange (**Lawrence and Ochman, 1998**).

Genetic transfer via conjugation remains the important mode by which the species acquires new genes as evidenced by the large number of pathogenic factors present in various strains of *E.coli* on plasmids. In deed, a pathogenic serotype O157:H7 *E.coli* strain has hundreds of blocks of foreign DNA encoding virulence factors that are lacking in a non pathogenic *E.coli* K-12 strain that has many blocks missing from the pathogenic strain **(Perna et al.,2001)**. Thus the pathogenic potential of a particular *E.coli* strain depends on the specific virulence genes present within its genome and particular virulence gene combinations define

specific pathotypes of *E.coli*. *E. coli* which is the most abundant facultative anaerobe present in the intestine of humans and many other worm blooded animals, is here as a model for studying biodiversity.

Pathogenic types of *E. coli.* At least 12 different classes, or pathotypes, of *E. coli* have been identified based on distinct disease phenotypes (**Donnenberg and Whittman, 2001 and Kaper et al., 2004),** and often contain different pathogenicity islands encoding virulence factors responsible for conferring the different disease phenotypes. This allows grouping of strains based on similar disease causing ability, rather than serological similarities based on typing of the somatic O antigen and flagellar H antigens. So while two strains may share the same serotype (O103:H2, for example), they may not necessarily share the same virulence profile. Although the more closely related two strains are genetically, the more likely they will have the same array of virulence factors and cause the same disease.

Parallel evolution of virulence. Strains from each pathotype are often not evolutionarily related, but rather have acquired similar pathogenic ability through parallel acquisition of the same or similar virulence factors. Despite the harm caused to the host, selection favors increased virulence when it allows for better spread or transmission of the pathogen **(Ewald, 1996**). Often better transmission equals more severe disease symptoms.

Evolution by horizontal transfer of genetic material. The genome of *E. coli* is extremely heterogeneous, having undergone remarkable sequence divergence between the common pathotypes that

cause disease. Horizontal transfer of genetic material is common, with striking variability between 0157:H7 isolates and the commensal K12 strain. Nearly 1400 genes in 0157:H7 are unique compared to K12, including many that encode virulence functions. The genome is nearly 20% larger in 0157:H7 strains vs. K12 (5.5MB vs. 4.6MB) (**Perna et al., 2001).** Other pathotypes are also highly divergent. For example, the recent genome sequencing of the uropathogenic strain CFT073 revealed only 39.2% of the total protein content in CFT073, the laboratory K12 strain MG1655, and EHEC 0157:H7 strain EDL933 is conserved between all 3 strains. CFT073 is as different genetically from the commensal K12 strain as it is from EDL933 (**Welch et al., 2002)**.

In the present thesis, a novel visualization Crystal Image Software (CIS) (**El-Daly and Tolba, 2009**) based on combination of image processing and numerical analysis, this technique is used to image the internal structures of *E.coli.* This technique depends on the fact that the causes of color in many minerals are in response to structural irregularities(**Klein and Dutrow,2008**) such use of that property can be considered the key factor for mapping the way in which the electron beam of TEM interact with internal structure of nano-sized materials to produce the final digital image.however,the digital image consists of a square array of image elements or pixels; at each pixel,the image brightness was sensed and assigned with integer value (from 0 to 255 in the gray scale image).

Aim of the study:

1- Applying different strategies to reveal biodiversity of *E.coli* isolated from different water sites.

2- The use of Image Processing technique and numerical analysis for following up the nanoscale diversity at the organismal and organellar level.

3- Studying the ability of *E.coli* for synthesis of Gold nanoparticles (GNPs).

4- Correlating the biodiversity of *E.coli* to evolution.

5- Correlating the biodiversity of *E.coli* to taxonomic relationships within the same species.

6- To set an overall suggestion for interpreting *E.coli* diversity.

Literature review

The two habitats for *E.coli*, the primary (within the host) and the secondary (outside the host), these two habitats represented distinct ecosystems that differ in both the number and heterogeneity of harmful stimuli, for example, secondary habitat stimuli such as temperature, UV, radiation and predation decrease the density of individual strains to undetectable levels under controlled conditions **(Sinton et al., 2002; Arena et al., 2003 and Janes et al., 2004)**.

Based on such evidence, it was often concluded that secondary habitats do not actively support *E.coli* growth and therefore have little effect on the adaptive evolution of species **(Winfield and Groisman, 2003)**. One implication of this hypothesis was that natural selection in the primary habitat was a dominant influence on the genetic structure of populations sampled from secondary habitat (**Gordon et al., 2002**). However, there are strikingly few data with which to assess the contribution of selection and other genetic processes in the secondary habitat on the variability and organization of natural *E.coli* populations. Two observations suggested adaptive evolution in the secondary habitat can substantially influence population genetic structure of the *E.coli* species as a whole; a) First, the population size of *E.coli* in the secondary habitats may be very large, as it is estimated that half of all living cells are presently outside of a host (**Savageau , 1983**). b) Second, data from multiple studies in both tropical and temperate regions suggested that this organism can replicate and reach high densities under favorable conditions outside of mammalian hosts **(Alm et al., 2003; Whitman et al., 2004 and Anderson et al., 2005).**

It was suggested that Divergent bacterial lineages existed within the biochemically-defined species, *E.coli*, may represent novel species.

It was hypothesized that a similarly divergent lineage of *E.coli* represented the breadth species level diversity that existed some10-15 million years ago (**Wirth et al., 2006**).

An alternative or additional hypothesis is that multiple lineages of biochemically similar, but genetically distinct bacteria have evolved, so that multiple species currently exist within what is now considered to be *E.coli*.

For characterizing such diversity, *E.coli* was isolated from different water resources and the following approaches were used.

Part I

Physico-chemical analyses of water samples

The influence of environmental (physical and chemical) factors on growth and activity of *E.coli* was outlined. The density of *Bifidobacterium spp*., fecal coliforms, *E. coli*, and total anaerobic bacteria, percentages of total bacterial community activity and respiration, and 12 physical and chemical parameters were measured simultaneously at six sites for 12 months in the Mameyes River rain forest watershed, Puerto Rico. The densities of all bacteria were higher than those reported for uncontaminated temperate rivers, even though other water quality parameters would indicate that all uncontaminated sites were oligotrophic. The highest densities for all indicator bacteria were at the site receiving sewage effluent; however, the highest elevation site in the watershed had the next highest densities. Correlations between bacterial densities, nitrates, temperature, phosphates, and total phosphorus indicated that all viable counts were related to nutrient levels, regardless of the site sampled. *In situ* studies at two different sites indicated that *E. coli* could survive, remain physiologically active, and regrow at rates that were dependent on nutrient levels of the ambient waters. *Bifidobacterium adolescentis* did not survive at either site but did show different rates of decline and physiological activity at the two sites. Bifidobacteria showed promise as a better indicator of recent fecal contamination in tropical freshwaters than *E. coli* or fecal coliforms (**Carillo, 1985**).

A sampling program was set up within the Water Board (Sydney) drinking water distribution system to determine the presence and extent

of bacterial regrowth. System, physical, chemical and microbiological parameters, including the identification of bacteria isolated, was measured over a 14 month period and correlation analysis was carried out. Results showed that regrowth was present within the system and that certain parameters, such as turbidity and distance from the initial treatment point, correlated with the presence of high bacterial numbers. Identification of bacteria isolated on m-Endo broth indicated *Klebsiella oxytoca* as the principal coliform involved. High numbers of *Aeromonas hydrophila* were also isolated **(Kaye and laslo, 1998).** The physical and chemical parameters of water were studied by **(Hafez et al., 2008)** for El-Salam water. Seven locations were selected for measuring pH, Dissolved oxygen (DO), conductivity, Chemical Oxygen Demand (COD), Biological Oxygen Demand (BOD), Total Dissolved Salts (TDS) and selected anions were also included. The site was found to be polluted with high concentrations of organic compounds. From such studies, the physico-chemical parameters had affected bacterial growth and so their behavior.

PartII

Isolation and identification of *E.coli*

Methods for isolation and identification of *E.coli* depend on the type of the sample and the purpose of the study. The standard multiple tube test was described by (**Anon, 1984)** for enumeration of *E.coli* in water. He used mineral- modified glutamate medium at 37°c as a first selective step, with lauryl sulphate lactose broth as an alternative, subcultured the gas producing in brilliant green bile broth (BGLB) or lauryl sulphate tryptose broth at 44°c and confirmed by the indole test. A rapid

alternative is the single tube confirmation test based upon formation of gas from mannitol, and indole from tryptophan, according to (**Pugsley et al., 1973).** The American standard methods for the examination of water and waste water**(American Public Health Association,1992)** describe the standard most probable number, presence-absence test and membrane filter enumeration methods both for total and fecal coliforms,but not for *E.coli*. However, it is recommended that detection of *E.coli*; if wanted can be performed by the IMViC test **(Hofstra et al., 1988).**

Fluorogenic detection and enumeration methods for *E.coli* in water and other samples had been reported by **(Robison, 1984)**.

Usually large numbers of *E.coli* are expected in feces and sewage treatment plant effluents. In these cases, samples can be plated directly on macconkey agar or eosin methylene blue agar, while often a sample is enriched overnight in tryptic soy broth (TSB) or brain heart infusion (BHI) and samples from TSB or BHI are plated on macconkey or EMB agar the next day.

Identification can be performed by IMViC test or by using commercially available identification systems (API, Enterotube, etc) **(Hofstra et al., 1988).**

Assays for identification; in other cases were based on activity of β-glucuronidase which cleaves the substrate 4-methylumbelliferyl-β-d-glucuronide (MUG) into 4-methylumbelliferyl moiety which produce a blue fluorescence when irradiated with long wave ultraviolet (UV) light. Because β-glucuronidase activity is not fully specific for *E.coli*, it should be confirmed by biochemical tests. Another problem is that most entero

pathogenic hemorrhagic *E.coli* (EHEC) isolated are β-glucuronidase-negative **(Doyle and Schoeni, 1984)**, which interferes with their detection. Detection based on MUG identification should not, therefore, be the method of choice when the presence of EHEC strains is suspected. MUG can also be added to selective media, e.g. lauryl sulphate tryptone broth (LTB). After incubation, fluorescence can be examined by direct irradiation of the culture with long wave UV light. A new membrane filter agar medium (MI) containing a chromogen ,indoxyl-β-D-glucuronide, and a flurogen , 4- methyl umbelliferyl- β-D-galactopyronoside,was developed to detect and enumerate *E.coli* and total coliforms(TC) in water samples on the basis of enzymatic activity. TC produces β-galactosidase, which cleaves 4-methylumbelliferyl- β-D-galactopyronoside to form 4-methyl umbelliferone that fluoresced under long wave UVlight. *E.coli* produces β-glucuronidase, which cleaves indoxyl-β-D-glucuronide to form a blue color. Recently, significant progress has been made in rapid detection of *E.coli* using modern membrane filter techniques. These membrane filter methods, in principle, can be used without limitations in all displines of microbiology, both for enumeration of bacteria as marker organisms and for identification of bacteria with specific traits such as virulence factors. The introduction of new confirmation tests such as MUG hydrolysis has contributed in this respect. Furthermore, the combination of membrane filter methods with monoclonal antibody or DNA hybridization allow a rapid confirmation of pathogenic *E.coli* strains due to availability of monoclonal antibody and DNA probes specific for heat labile toxin (LT) and heat stable toxin (ST) **(Hofstra et al., 1988).**

The Biolog GN2 MicroPlate is designed for identification and characterization of a very wide range of aerobic gram-negative bacteria. The database for the GN2 MicroPlate is now over 500 species. It is by far the largest kit-based identification database available. Automated microbial identification systems, such as theMicroLog/Biolog identification system (Biolog, Hayward, Calif.), have become widely used in both food, clinical and research laboratories **(Odumeru et al., 1999)**. The analysis of carbon- source utilisation is a simple and rapid method to identify bacterial isolates from various sources. This system establishes identifications based on the exchange of electrons generated during bacterial respiration**.**

Previous studies showed that 16s rRNA gene sequencing based identification was more reliable as compared to other biochemical based systems. Over a period of 26 months, **(Bosshard et al., 2006**) have evaluated in a prospective fashion the use of 16S rRNA gene sequencing as a means of identifying clinically relevant isolates of nonfermenting gram-negative bacilli (*non-Pseudomonas aeruginosa*) in the microbiology laboratory. The study was designed to compare phenotypic with molecular identification. Results of molecular analyses were compared with two commercially available identification systems (API 20 NE, VITEK 2). By 16s rRNA gene sequence analyses, 92% of the isolates were assigned to species level and 8% to genus level. Using API 20 NE, 54% of the isolates were assigned to species and 7% to genus level, and 39% of the isolates could not be discriminated at any taxonomic level. The respective numbers for VITEK 2 were 53%, 1%, and 46%, respectively. Fifteen percent and 43% of the isolates

corresponded to species not included in the API 20 NE and VITEK 2 databases, respectively. 16S rRNA was concluded by them to be the gene to be sequenced as an effective means for the identification of clinically relevant gram-negative bacilli. Based on their experience, they proposed an algorithm for proper identification of gram-negative bacilli in the diagnostic laboratory.

PartIII

Characterization of *Escherichia coli* biodiversity

1. Cultural characters:

The biochemical characteristics of 97 invasive *E. coli* strains of different Oserogroups were studied. Considered as a group, the behavior of the strains was quite variable. However, none of them decarboxylated lysine and all but seven strains, belonging to the O124 serogroup, were nonmotile. The growth of 25 strains obtained on Macconkey, salmonella-shigella, xylose-lysine-desoxycholate, and Hektoen enteric agars was compared. MacConkey and Hektoen enteric agars yielded the highest average growth for these strains, whereas salmonella-shigella agar had the lowest average counts (**Silva et al., 1980**).

MacConkey, eosine-methylene blue, deoxycholate-citrate, salmonella-shigella, and xylose-lysine-deoxycholate agars were compared for their ability to support the growth and to facilitate the recovery of enteroinvasive Escherichia coli strains from artificially contaminated as well as from clinical faecal samples. When grown as pure cultures, the 78 enteroinvasive Escherichia coli strains, as a group, exhibited the same growth characteristics as did Shigella isolates, i.e.

both organisms grew more weakly than did Salmonella strains on the various selective plates. Xylose-lysine-deoxycholate and deoxycholate-citrate plates were more effective in recovering enteroinvasive Escherichia coli from faecal samples than was salmonella-shigella agar. Likewise, xylose-lysine-deoxycholate agar, similar to the differentiating Macconkey and eosin-methylene blue agars, was less inhibitory for defined "sensitive" strains than were the selective media tested. Preincubating clinical faecal samples in selenite F did not influence the recovery of enteroinvasive Escherichia coli significantly. These data show that the use of xylose-lysine-deoxycholate, in combination with MacConkey or eosine-methylene blue agar, provides the best chance for recovery of enteroinvasive Escherichia coli when randomly selecting colonies from faecal cultures for subsequent molecular or immunological identification assays**(Szakal and pal, 2003).**

Escherichia coli cells were seen as coccobacillary, short rods, arranged singly and in pairs, stained negatively and non-motile. The cultural characteristics recorded by **(Fazlani et al., 2009)** were moist, gray, shiny and convex with entire margin colonies, non-haemolytic and easily dispersible in normal saline solution. Uniform turbidity was observed in the nutrient broth. It produced lactose fermentative colonies on MacConkey's agar. *Escherichia coli* as short rods of 0.5-3.0µ with a variable shapes from coccoid bipolar to long rods. Usually occurred singly but short chains were uncommon.

2. Biotyping:

A two- miniaturized scheme of eight tests was devised for biotyping strains of *Escherichia coli* in microwell plates. Primary biotypes were defined by positive and negative reactions in tests for fermentation of raffinose, sorbose, dulcitol and 2-deoxy-D-ribose and for decarboxylation of ornithine when read after specified periods of incubation; subtypes were identified within primary biotypes according to results in secondary tests for rhamnose fermentation, lysine decarboxylation and motility. The method gave reproducible results on different occasions of testing. Among 100 *E. coli* strains from various sources, 26 of the 32 possible primary biotypes and 56 full biotypes, as defined by results in both primary and secondary tests, were identified, thus demonstrating a high index of strain discrimination . The scheme is recommended as a simple, reliable, inexpensive and efficient method of differentiating strains of *E. coli* (**Crichton et al, 1993**)

Carbon source utilization (CSU) is a phenotypic library-based technique that relies upon proprietary technology developed by Biolog, Inc. (Haywood, CA). A standardized suspension of each bacterial isolate is inoculated into a 96-well plate prepared by 5 Biolog, Inc. This plate contains 1 blank well (water) and 95 wells lined with different substrates to be used for metabolism by the bacteria. The wells are also coated with a color change reagent, tetrazolium violet. As bacteria metabolize the substrate within the well, the tetrazolium violet changes from a colorless, oxidized form to a purple reduced form (Biolog, Hayward, Calif.). This color change is measured in terms of intensity, approximately 22-24 hours after incubation of the plates, utilizing the Biolog MicroStation

plate reader. The intensity readings from each well combine to form a profile for that isolate, which Biolog software compares with known biochemical profiles to provide an identification of the bacteria. The intensity readings from each well can also be used to compile a CSU profile for each isolate that can then be analyzed with discriminant analysis for classification**.** Thus, the biochemical profile determines the biotypic pattern of the tested organism.

API 20 was used to to characterize biotypic patterns of 190 *E.coli* strains isolated from different sources. Of 20 biochemical reactions analyzed, 4 were positive for all strains, 6 were negative across the 190 strains. The remaining 10 assays were variable with some strains either positive or negative for each test. The results of of the 10 variable tests were used to generate a strain specific biotype. There were a total of 76 distinct biotypes among the *E.coli* strains (**Walk, 2007**).

3. Serotyping:

It was reported that the first successful attempt to classify *E.coli* by serological methods was carried out by **(Kauffmann**, **1944)** who was able to sub divide *E.coli* strains into 20 O-groups. He used a boiled culture for O-antiserum production and agglutination tests. It was soon extended by **(knipschildt, 1945)** to contain 25 O-groups.

Since then, many O-antigens have been added, increasing the present antigenic scheme to over than 180. Little is known about the distribution of different serogroups.the best examined are strains from humans and common domestic animals in temperate climates. It is not yet possible to state whether there are significant differences in O-group

prevalence between different animal groups. However, some investigations reported differences in O-groups between humans and animals **(Hartly et al., 1975 and Orskov et al., 1977).**

Some authors have indicated that differences in O-serogroup distribution exist, e.g. between Europe and the United States **(Glynn and Howard, 1970),** even within different areas in London **(Gruneberg and Bettelheim, 1969)**. Others have suggested that prevalence rates of certain O-groups could vary with time in the same area **(Mabeck et al., 1971).**

Three surface structures corresponding to the O, K, H, antigens form the basis of the identification. From literature, it has been found that K-antigens are only present in capsulated *E.coli* and have served a minor role in Serotyping. Therefore, Serotyping has been mainly based on O and H antigens. H- antigen is present only in actively motile *E.coli*. O-antigen is a thermostable surface structure found in all *Enterobacteriaceae***.** It constitutes O-specific polysaccharide of cell wall lipopolysaccharide and is determined by chromosomal genes **(Orskov et al., 1977)**.

Serotype distribution of *E.coli* isolated from animal feces, human feces and meat was compared and found that there were marked differences in serotypes distribution in strains from man and animals. The meat strains resembled the animal strains. These results suggest either, animal strains of *E.coli* are not reaching the general human population to any great extend, or if they do so fail to colonize the bowel **(Bettelheim et al., 1976**).

It was reported that the most widely used and perhaps best known method of identifying individual strains of *E.coli* is based on an immunological reaction **(Paveen, 1997).**

Serotyping of O- and H-antigens is regarded as the gold standard in classification of *E. coli* for taxonomic and epidemiological purposes similar to the Kaufmann-White scheme for *Salmonella enterica*. Molecular methods to replace or to support the serotyping had been applied. Using the molecular polymorphism of the flagella (H-antigen) gene fliC, more than 220 *E. coli* strains from clinical origins and other sources had been characterized and a reproducible and clear cut classification with very good correlation to serotyping was found **(Prager et al.,2003).**

Serotyping was the foundation of pathogenic *Escherichia coli* diagnostics; however, few laboratories have this capacity. A molecular serotyping protocol was developed that targets, genetically, the same somatic and flagellar antigens of *E. coli* O26:H11 used in traditional serotyping. It correctly serotypes strains untypeable by traditional methods, affording primary laboratories serotyping capabilities **(Durso et al., 2005).**

FliC PCR-restriction fragment length polymorphism (RFLP) was performed to investigate whether this technique would be better than classic serotyping for the characterization of the H antigen in enterotoxigenic *Escherichia coli* (ETEC) strains. It was showed that the fliC genes from ETEC strains can be characterized by restriction analysis of their polymorphism. Only one allele of the fliC gene from ETEC

strains was found for each flagellar antigen, with the exception of H21. Nonmotile strains could also be characterized using this molecular technique. Moreover, determination of the somatic antigen was guided by the identification of the flagellar antigen from previously unknown serotypes of ETEC strains by PCR-RFLP. The PCR-RFLP technique proved to be faster than classic serotyping for the characterization of the *E. coli* H antigen, taking 2 days to complete instead of the 7 or more days using classic serotyping. The H molecular typing for *Enterobacteriaceae* members may become an important epidemiological tool for the characterization of the H antigen of E. coli pathotypes. The PCR-RFLP technique is capable of guiding the determination of the H antigen and could partially replace seroagglutination. With the determination of the molecular profiles of alleles from strains obtained in epidemiological studies, new patterns will be described for ETEC strains or other *E. coli* pathotypes **(Ramos et al., 2006**).

The O-antigen gene clusters of *Escherichia coli* serogroups O117, O126, and O146 were sequenced, and 11, 10 and 11 open reading frames (ORFs) were identified, respectively. Genes required for O-antigen sugar biosynthesis, sugar transfer, and sugar processing were identified. Multiplex polymerase chain reaction (PCR) assays were developed targeting the wzx and wzy genes present in the O-antigen gene cluster of these serogroups. The assays were highly serogroup specific when tested against strains belonging to serogroups that were isolated from food, humans, animals, and environmental sources, as well as against representative strains belonging to165 different *E. coli* O serogroups and a number of non-*E. coli* bacteria. Thus, the results demonstrate that the

wzx and wzy gene sequences were specific to *E. coli* O117, O126, and O146 and can be used as diagnostic markers for rapid identification and detection of these serogroup (**Yanhong et al., 2007)**

4. Phage typing:

Phage typing was used to identify strains of *Escherichia coli* isolated from urinary and nonurinary sources. When eight phages isolated in Pennsylvania were used to type 717 cultures from Missouri, 50.3% of 624 urinary isolates and 34.4% of 93 nonurinary isolates were typable. Strains from nonurinary sources were not found commonly in urine. When five additional phages isolated in Missouri were added to the original set of eight phages, 80.4% of 331 urinary isolates were typable. When this set of phages was used to type 552 urinary cultures isolated in California, Minnesota, Ohio, Pennsylvania, Virginia, and West Virginia, 82.0% of the cultures were typable. Some common phage types were found in high incidence among cultures from the different regions. No correlation was found between phage type and the pattern of resistance to antibiotics. Phage typing data were presented also on the number of strains in individual urine specimens and the recurrences of strains in patients with chronic bacteriuria. Of 97 fecal isolates, 75.2% of the cultures were typable, and the most common phage type was observed in high incidence among the urinary isolates from this region. When 75 cultures from nine other genera of enteric bacteria were typed, only the shigellae were lysed. In view of the information obtained by phage typing and the ease and speed with which it can be done, it is suggested that phage typing be considered a new tool in epidemiological studies of urinary tract infections by *E. coli* (**Parisi et al., 1969).**

In Canada, the number of human isolates of verotoxigenic (VT + ve) *Escherichia coli* O157:H7 from diarrhoeal cases and haemolytic uraemic syndrome and haemorrhagic colitis has increased from 25 in 1982 to 2384 in 1989. A total of 3273 VT + ve *E. coli* O157:H7 strains (3255 strains isolated in Canada and 18 isolates from other countries) were phage typed. The phage typing scheme has been extended from 14 to 62 phage types. Of these, five types occurred exclusively in other countries (type 47 in Japan; and types 49, 50, 51 and 52 in the U.K.). Thirty-five different phage types were identified in Canada; only nine of these (1, 2, 4, 8, 14, 21, 23, 31 and 32), each accounted for more than 1% of the cases from human sources. The same nine types were the only ones observed among the isolates from non-human sources (meat and slaughter houses) suggesting a food-borne transmission in most of the human cases. Phage types 1 (30.5%); 4 (21%); 8 (13.5%); 31 (8.9%) and 14 (8%) were encountered in varying frequencies in most of the provinces; infrequently occurring phage types also showed regional variation. Thirteen different phage types were identified among 151 outbreaks representing 556 isolates of E. coli O157:H7. More than one phage types were encountered in 12 outbreaks whereas in 141 outbreaks, all strains in each had the same phage type (**Khakhria et al., 1990**).

The ability of a phage to infect a particular cell will be dependent on variation in primary surface receptors, which may comprise surface polysaccharides (lipopolysaccharide in Gram negative bacteria, the teichoic acids in Gram positive bacteria), (**Estrela et al., 1991**), the presence of surface structures such as flagellae or pili (**Merino et al., 1990**) , and the expression of a wide range of different types of Cell

surface–associated molecules including sugar uptake proteins (**Schwartz, 1983**), membrane proteins (**Heilpern and Waldor, 2000**), Slayer proteins (**Callegari et al., 1990**), and capsular polysaccharide (**Hung et al., 2002).** Even if phage successfully penetrate the cell envelope, replication leading to cell lysis and development of a plaque may be inhibited by a variety of mechanisms, such as the presence of prophage in the cell (**Harvey et al., 1993**), DNA restriction modification systems (**Frank, 1994**), and even specific phage inhibition genes (**Chopin et al., 2005**). These many factors which affect the efficiency of phage replication have lead to the development of phage typing schemes, where the ability of a phage to infect a cell is used as an indicator of biological variation at the cellular level. For phage typing, panels of phage characterized to have a limited host range are chosen and strains are infected with each phage at standard concentration (known as the Routine Test Dilution or RTD). Hence, a bacterial lawn is prepared and samples of the different phage at the RTD spotted onto the surface of the lawn. After incubation, infection is detected by the presence of plaques (zones of clearing) and patterns of susceptibility to individual phage are determined, leading to the characterization of a phage type. These panels of phages have successfully been used by epidemiologists to monitor changes in the predominant organism causing disease in the population and to identify the emergence of new dominant clones. For *Salmonella*, in particular, the phage typing system (**Callow**, **1959**) is successfully used to further differentiate strains of the same serovar, as the widespread nature of some serovars requires more discriminatory methods for subtyping than is afforded by serology alone **(McDonough et al., 1989; Threlfall et al., 1994**). The use of detailed subtyping has allowed sources of foodborne

disease to be identified. For instance, this technique has been successfully used to identify the emergence of clones that come to predominate in food industries so that new risks and intervention strategies can be established. Similarly, when rare phage types emerge (often through importation of foods), phage typing can greatly assist in the identification of outbreak sources. In 2003 *Salmonella enteritidis* PT 56, rare in the United Kingdom, caused an outbreak in the northeast of England, which was then traced back to a particular restaurant and workers in that restaurant were found to be carrying the strain responsible for the outbreak.

All bacteriophages are, to a greater or lesser extent, specific in their host range.Generally, they are specific for members of a certain species (or sometimes species that are very closely related). But in some cases the specificity may be more restricted than that – they may be able to infect some strains but not others. So, it can be used to distinguish between strains of the same species. This is known as phage typing. The best example of the use of phage typing is in distinguishing strains of *Staphylococcus aureus*. This bacterium, which is a common cause of infection in hospitals, is a common skin organism and is also able to persist for many weeks in dust. Phage typing provides a means of distinguishing between all the possibilities by testing each of the isolates for sensitivity to a battery of bacteriophages. If there is a single source of infection, all strains from infected patients will be the same and will show the same pattern of phage sensitivity.

Bacteriophages also have potential in a diagnostic laboratory for the detection and identification of pathogenic bacteria. This makes use of

the amplification that arises from the multiplication of the bacteriophage – commonly each infected bacterium will yield 100 or more phage particles, so giving rise to a 100-fold increase in the sensitivity of detection – and also the specificity of the bacteriophage, thus enabling a specific pathogenic species and not closely related nonpathogens to be detected. The sensitivity and ease of detection can be enhanced by using a reporter phage, i.e. one that has been engineered to carry a gene coding for an easily detected enzyme. The firefly luciferase gene is often used for this purpose as luciferase can be easily and specifically detected by means of the light emitted. The potential of this technique is greatest for those bacteria that are not easily detected in clinical samples – such as *Mycobacterium tuberculosis* which takes several weeks to form colonies – but is yet to achieve widespread routine use (**Dale and Park, 2004)**.

5. Congo Red dye agar (CR) test for detecting *E.coli* invasiness:

Congo red dye agar test (CR test) has been used to differentiate invasive and non invasive *E. coli* in poultry. This simple test was used to detect enteroinvasive *E. coli* of bovine origin, isolated from cases of neonatal calf diarrhoea. Out of 97 isolates tested 46 showed CR-positive reactions, while 51 were CR negative. CR test was found 100% specific and 58.89% sensitive. This test can be used for primary screening of non invasive *E. coli* from potentially invasive *E. coli* **(Sharma et al., 2006).**

6. 16s rRNA sequencing and analysis:

16s rDNA gene sequencing is a powerful tool that has been used to trace phylogenetic relationships between bacteria, and to identify bacteria from various sources, such as environmental or clinical specimens.This technology is used today in clinical laboratories for routine identifications, especially for slow-growing, unusual or fastidious bacteria, but also for bacteria that are poorly differentiated by conventional methods; however, it provides no information about antibiotic resistance. Phenotypic methods present some inherent problems: there can be a substantial amount of variability among strains belonging to the same species, the corresponding database may not yet include newly described species and the test may rely on an individual and subjective interpretation. Identification based on the 16s rDNA sequence is of interest because ribosomal SSU exists universally among bacteria and includes regions with species specific variability, which makes it possible to identify bacteria to the genus or species level by comparison with databases in the public domain (**Vandamme et al., 1996**). The molecular approach has been used for bacterial phylogeny and is of major importance for species definition and identification (**Fredricks and Relman**, **1996; Rossello-Mora and Amann, 2001; Raoult et al., 2004 and Clarridge, 2004**). For these reasons, 16s rDNA sequencing is also used as a method of detecting pathogens in normally sterile clinical specimens, or for detecting species that cannot be cultured. Other authors have also reported its use as a tool for bacterial identification. They have usually compared this molecular identification tool to classical phenotypic methods. The great potential of the method

has been reported for Gram-positive rods and coryneiform bacterial identification (**Tang et al., 2000 and Bosshard et al., 2003)**. For Gram-positive, catalase-negative cocci, Bossard obtained better identification with this method than with conventional methods despite the low resolving power for some organisms (**Bosshard et al., 2004**). For Gram-negative rods, Tang et al. concluded that it provided some improvement, with 89.2% of strains identified to the species level **(Tang et al., 1998**). Of particular interest was reported by Ferroni et al. with regard to Gram-negative rods isolated from patients with cystic fibrosis where phenotypic methods are not suitable (**Ferroni et al., 2002; Coenye et al., 2002; and Drancourt et al., 2004).** It also seems to be of interest for bacteria that remained unidentified or misidentified by phenotypic methods (**Drancourt et al., 2000 and Petti et al., 2005**). Their study was to assess the use of 16S rDNA sequencing for the identifying bacteria (excluding all mycobacteria) that are difficult to identify using phenotypic methods under strictly routine conditions in a clinical microbiology laboratory.

Comparative sequence analysis of small subunit rRNA was documented to be currently one of the most important methods for the elucidation of bacterial phylogeny as well as bacterial identification. Phylogenetic investigations targeting alternative phylogenetic markers such as large subunit rRNA, elongation factors, and ATPases have shown that 16s rRNA-based trees reflect the history of the corresponding organisms globally. However, in comparison with three to four billion years of evolution the phylogenetic information content of these markers is limited. Consequently, the limited resolution power of the marker molecules allows only a spot check of the evolutionary history of

microorganisms. This is often indicated by locally different topologies of trees based on different markers, data sets or the application of different treeing approaches. Sequence peculiarities as well as methods and parameters for data analysis were studied with respect to their effects on the results of phylogenetic investigations. It is shown that only careful data analysis starting with a proper alignment, followed by the analysis of positional variability, rates and character of change, applying alternative treeing methods and, finally, performing confidence tests, allows reasonable utilization of the limited phylogenetic information **(Ludwig et al., 1998)**.

Sequence heterogeneities of variable positions located at regions V1 and V6 of 56 cloned 16s rRNA genes were determined from six *Escherichia coli* strains by **(Martinez-Murcia et al., 1999)**. These nucleotides were involved in secondary structure base-pairing of stem-loops. Single mutations have occurred but secondary structure was conserved. Eight different sequences were found in the stem at region V1 indicating that in these sites mutation rates are higher than those of homogenization processes. Region V6 showed two different structures (V6-I and V6-ll) although heterogeneities were determined in nine sites. Strains ECOR52 and ECOR56 only showed the V6-I sequence, ECOR35 showed V6-ll, whereas clones from ECOR42 and ECOR49 showed both types of V6 structures. Results were confirmed by PCR using V6 sequence-specific probes. Stem V6-ll was also found in 16s rRNA sequences deposited in the RDP (Ribosomal Database Project), belonging to distantly related taxa; ancestral sequence V6-II seems to be homogenized in all *rm* operons of the multigene family of strain ECOR35

producing effects of distortion in the molecular clock. V6 sequence-specific probes were applied to the 72 ECOR strains: half showed both V6-I and V6-ll, and the rest had one or another. Only strain ECOR24 did not yield products in the PCR test and sequencing of 12 cloned 16s rRNA genes revealed a third form, V6-Ill, also found in the RDP.

For many years, sequencing of the 16s ribosomal RNA (rRNA) gene has served as an important tool for determining phylogenetic relationships between bacteria. The features of this molecular target that make it a useful phylogenetic tool also make it useful for bacterial detection and identification in the clinical laboratory. Sequence analysis of the 16s rRNA gene is a powerful mechanism for identifying new pathogens in patients with suspected bacterial disease, and more recently this technology is being applied in the clinical laboratory for routine identification of bacterial isolates. Several studies have shown that sequence identification is useful for slow-growing, unusual, and fastidious bacteria as well as for bacteria that are poorly differentiated by conventional methods. The technical resources necessary for sequence identification are significant. This method requires reagents and instrumentation for amplification and sequencing, a database of known sequences, and software for sequence editing and database comparison. Commercial reagents are available, and laboratory-developed assays for amplification and sequencing have been reported. Likewise, there are an increasing number of commercial and public databases. Despite the availability of resources, sequence-based identification is still relatively expensive. The cost is significantly reduced only by the introduction of

more automated methods. As the cost decreases, this technology is likely to be more widely applied in the clinical setting **(Patel, 2001).**

Due to the inadequate automation in the amplification and sequencing procedures, the use of 16s rRNA gene sequence-based methods in clinical microbiology laboratories is largely limited to identification of strains that are difficult to identify by phenotypic methods. Using conventional full-sequence 16s rRNA gene sequencing as the "gold standard," the usefulness of the MicroSeq 500 16s ribosomal DNA (rDNA)-based bacterial identification system, which involves amplification and sequencing of the first 527-bp fragment of the 16s rRNA genes of bacterial strains and analysis of the sequences using the database of the system, for identification of clinically significant bacterial isolates with ambiguous biochemical profiles was evaluated. Among 37 clinically significant bacterial strains that showed ambiguous biochemical profiles, representing 37 nonduplicating aerobic gram-positive and gram-negative, anaerobic, and *Mycobacterium species*, the MicroSeq 500 16s rDNA-based bacterial identification system was successful in identifying 30 (81.1%) of them. Five (13.5%) isolates were misidentified at the genus level (*Granulicatella adiacens* was misidentified as *Abiotrophia defectiva*, *Helcococcus kunzii* was misidentified as *Clostridium hastiforme*, *Olsenella uli* was misidentified as *Atopobium rimae*, *Leptotrichia buccalis* was misidentified as *Fusobacterium mortiferum*, and *Bergeyella zoohelcum* was misidentified as *Rimerella anatipestifer*), and two (5.4%) were misidentified at the species level (*Actinomyces odontolyticus* was misidentified as *Actinomyces meyeri* and *Arcobacter cryaerophilus* was misidentified as *Arcobacter butzleri*). When the same

527-bp DNA sequences of these seven isolates were compared to the known 16s rRNA gene sequences in the GenBank, five yielded the correct identity, with good discrimination between the best and second best match sequences, meaning that the reason for misidentification in these five isolates was due to a lack of the 16s rRNA gene sequences of these bacteria in the database of the MicroSeq 500 16s rDNA-based bacterial identification system. In conclusion, the MicroSeq 500 16s rDNA-based bacterial identification system is useful for identification of most clinically important bacterial strains with ambiguous biochemical profiles, but the database of the MicroSeq 500 16s rDNA-based bacterial identification system has to be expanded in order to encompass the rarely encountered bacterial species and achieve better accuracy in bacterial identification **(Woo et al., 2003)**.

Reliable automated identification and susceptibility testing of clinically relevant bacteria is an essential routine for microbiology laboratories, thus improving patient care. Examples of automated identification systems include the Phoenix (Becton Dickinson) and the VITEK 2 (bioMerieux). However, more and more frequently, microbiologists must isolate "difficult" strains that automated systems often fail to identify. An alternative approach could be the genetic identification of isolates; this is based on 16s rRNA gene sequencing and analysis. To evaluate the possible use of MicroSeq 500 (Applera) for sequencing the 16s rRNA gene to identify isolates whose identification is unobtainable by conventional systems **(Fontana et al., 2005**) analyzed 83 "difficult" clinical isolates: 25 gram-positive and 58 gram-negative strains that were contemporaneously identified by both systems--VITEK 2 and Phoenix--while genetic identification was performed by using the

MicroSeq 500 system. The results showed that phenotypic identifications by VITEK 2 and Phoenix were remarkably similar: 74% for gram-negative strains (43 of 58) and 80% for gram-positive strains were concordant by both systems and also concordant with genetic characterization. The exceptions were the 15 gram-negative and 9 gram-positive isolates whose phenotypic identifications were contrasting or inconclusive. For these, the use of MicroSeq 500 was fundamental to achieving species identification. In clinical microbiology the use of MicroSeq 500, particularly for strains with ambiguous biochemical profiles (including slow-growing strains), identifies strains more easily than do conventional systems. Moreover, MicroSeq 500 is easy to use and cost-effective, making it applicable also in the clinical laboratory.

It was found that accurate identification of bacterial isolates is an essential task in clinical microbiology. Phenotypic methods are time consuming and either fails to identify some bacteria such as Gram-positive rods entirely or at least fail to do so in some clinical situations. 16s rDNA sequencing is a recent method of identification which offers a useful alternative. The usefulness of this method for identifying a range of bacteria in a clinical laboratory under routine conditions was investigated. Over a period of 30 months, 683 isolates were obtained from clinical specimens, sequenced and analyzed. For 568 of these isolates (83.1%), the sequence provided species level identification. For 108 isolates (15.8%), the identification was limited to the genus level, and for 7 isolates (1%), the sequence remained unidentifiable by 16s rDNA sequence analysis. For the isolates identified only to the genus level, the 16s rDNA approach failed to identify bacteria to the taxonomic level for

3 reasons: failure to differentiate between species in 72 isolates (66%), the lack of any closely related sequence in the database for 15 isolates (13.8%) and the presence of more than 1% of undetermined position in the sequence for 13 isolates (12%) **(Mignard and Flandrois**, **2006)**.

Identification of mycobacteria to the species level is of therapeutic significance. Conventional methods are laborious and time consuming so **(Therese et al., 2009)** performed 16s rRNA sequencing using a commercial MicroSeq sequencing kit, which includes DNA sequencing with software package for identification and phylogenetic analysis of clinical mycobacterial isolates. A total of 47 mycobacteria were tested by both conventional and genotypic method using commercially available MicroSeq 500 amplification kit assay. The identification was determined by comparing the 500 bp amplified product of 16s rDNA sequence to the MicroSeq database. The phenotypic identification was concordant with genotypic identification in 33 (70.2%) isolates of 14 *Mycobacterium tuberculosis*, 11 *M. fortuitum,* 7 *M. abscessus* and 1 *M. duvalii*. For the discrepant isolates, identification was possible only by DNA sequencing in 14 (29.7%) isolates. The 14 discrepant isolates were 5 *M. farcinogenes*, 3 *M. genavense*, 2 *M. species. nov* and 1 each of *M. fortuitum, M. immuogenum*, *M. simiae* and *M. wolinskyi*. Of these, five were uncommon species that were difficult to identify by phenotypic method. The MicroSeq DNA sequencing is an excellent tool for species identification of mycobacteria, which reduces the turn around time, makes repeat analysis easy as compared to phenotypic identification especially for mycobacterial isolates with ambiguous biochemical profiling.

It was assured that molecular phylogeny increasingly supports the understanding of organismal relationships and provides the basis for the classification of microorganisms according to their natural affiliations. Comparative sequence analysis of ribosomal RNAs or the corresponding genes currently is the most widely used approach for the reconstruction of microbial phylogeny. The highly and less conserved primary and higher order structure elements of rRNAs document the history of microbial evolution and are informative for definite phylogenetic levels. An optimal alignment of the primary structures and a careful data selection are prerequisites for reliable phylogenetic conclusions.rRNA based phylogenetic trees can be reconstructed and the significance of their topologies evaluated by applying distance, maximum parsimony and maximum likelihood methods of phylogeny inference in comparison. Phylogenetic trees based on almost equivalent data sets of bacterial 23S and 16s rRNAs are in good agreement and their overall topologies are supported by alternative phylogenetic markers such as elongation factors and ATPase subunits. Besides their phylogenetic information content, the differently conserved primary structure regions of rRNAs provide target sites for specific hybridization probes which have been proven to be powerful tools for the identification of microbes on the basis of their phylogenetic relationships **(Therese et al., 2009).**

Partial sequences of 16s rRNA gene of 19 strains of *Mycobacterium synoviae* were sequenced and analyzed in order to obtain molecular characterization and evaluation of the genetic variability of strains from distinct Brazilian areas of poultry production. Different polymorphic patterns were observed. The number of polymorphic

alterations in the studied strains ranged from 0 to 6. The nucleotide variations, including deletion, insertion and substitutions, ranged from3 to 5. The genotypic diversity observed may be explained by spontaneous mutations that may occur when a lineage remains in the same flock for long periods **(Marcos et al., 2008)**.

Sequence analysis of the 16s ribosomal RNA (rRNA) gene has been widely used to identify bacterial species and perform taxonomic studies **(Choi et al., 1996; Schmalenberger et al., 2001; Clarridge, 2004; Munson et al., 2004 and Petti et al., 2005**). Bacterial 16s rRNA genes generally contain nine "hypervariable regions" that demonstrate considerable sequence diversity among different bacterial species and can be used for species identification **(Van de Peer et al., 1996**). Hypervariable regions are flanked by conserved stretches in most bacteria, enabling PCR amplification of target sequences using universal primers **(Lu et al., 2000**; **Baker et al., 2003 and Munson et al., 2004).** Numerous studies have identified 16s rRNA hypervariable region sequences that identify a single bacterial species or differentiate among a limited number of different species or genera **(Lu et al., 2000; Bertilsson et al., 2002; Rothman et al., 2002; Becker et al., 2004; Clarridge, 2004; Maynard et al., 2005**). Rapid approaches that detect certain species-specific sequences within a single hypervariable region are also in common use **(Varma-Basil et al., 2004 and Stohr al., 2005**). Unfortunately, 16s rRNA hypervariable regions exhibit different degrees of sequence diversity, and no single hypervariable region is able to distinguish among all bacteria. Molecular diagnostic methods, such as real-time PCR **(Varma-Basil et al., 2004 and Selim et al., 2005**) or

melting temperature analysis **(Skow et al., 2005**) generally use fluorescent probes that hybridize to relatively short amplicons. Given the increasing importance of real-time PCR to medical diagnostics, it is surprising that few studies have focused on matching short segments of hypervariable 16s rRNA gene regions with the common pathogenic or environmental bacteria that can be differentiated by each segment.

7. Nanoscale evaluation of biodiversity

7.1. Using image processing:

The atomic force microscope is a useful tool for imaging native biological structures at high resolution. In analogy to conventional immunolabeling techniques, antibodies were used and directed against the C-terminus of bacteriorhodopsin to distinguish the cytoplasmic and extracellular surface of purple membrane while imaging in buffer solution. The extracellular surface of purple membrane was imaged at 0.7 nm resolution, exhibiting a major and a minor protrusion per bacteriorhodopsin monomer. As confirmed by immuno-dot blot analysis and sodium dodecyl sulfate-gel electrophoresis, labeling of the purple membrane was not observed if the C-terminus of bacteriorhodopsin was cleaved off by papain (**Muller et al., 1997**)

From ancient times until the invention of the optical microscope by Van Leeuwenhoek in the seventeenth century, humankind could only study those biological objects that were visible to the naked eye, which means to say, bigger than one millimetre across. Optical microscopy took observation down to the micron, an improvement by a factor of a thousand. Then, in the last century, the invention of the electron

microscope opened up the world on a scale of ten nanometers. Today, the nanometer and even the angstrom unit (0.1 nm) mark the limits of resolution in a whole range of modern molecular imaging techniques. Indeed, since the middle of the twentieth century, crystallographic studies of DNA and proteins have gradually revealed the structure of biological entities with a resolution slightly greater than 0.1 nm, which is the size of a hydrogen atom. Extremely complex protein molecules, built up from several smaller proteins, have now been successfully characterized and their three-dimensional coordinates can be obtained from data bases. Molecular imaging has made considerable progress since the heroic work carried out by crystallographers at the Cavendish Laboratory in Cambridge. It is possible today to obtain the complete structure of a single molecule, whatever its size. The most difficult thing when working with a single molecule, if it is complex, is to be able to visualise its motion within a protein or, an even more delicate task, within a multiprotein structure. Electron microscopy and crystallographic techniques can usually only be applied to preparations in which there is no molecular motion, although over the past twenty years or so, NMR techniques have become available to follow the internal motions of small proteins with a size of a few nanometers **(Boisseau et al., 2007**). There are two approaches applied basically in nanobiotechnology.The so-called top–down approach consists in miniaturizing the investigative or analytic tools we possess in order to move from subjects of centimetric or millimetric dimensions to ones with the same function but much smaller. In a sense it is as though one were climbing down the length scale. The opposite approach aims to climb up this same scale by arranging nanoscopic elements like's atoms or clusters of atoms into novel

structures, assembled in some appropriate way. It is clear that today most nanobiotechnological activity subscribes to the first approach, for while the second, bottom–up approach remains extremely attractive, it still faces many difficulties. One of these is that we are unable to predict the properties of elements conceived in this way, working solely from the knowledge of the individual properties of their components. In association with these two approaches, new imaging and measurement techniques are being developed to observe phenomena that have now become accessible to us **(Boisseau et al., 2007).**

It is well known that biological objects are investigated by various microscopes such as transmission electron microscope (TEM), high resolution transmission electron microscope (HRTEM), scanning tunnel microscope (STM) and even atomic force microscopy (AFM). These techniques however have some advantages and disadvantages, in terms of complexity of methods and the obtained data. It is well documented that the images from these tools are the first step for predicting the shape and size of nanostructures for novel properties. These images stored a lot of data about the real internal structure of these materials, which has not yet been discovered till now. So, great advances to image and manipulate these structures give birth to revolution in nanotechnology **(sally, 2009**)

One of the first applications of AFM imaging in biology was to the imaging of nucleic acids. There has been growing interest in this technique in the past few years, with more than 160 publications in the PubMed data base over the last months. AFM can be used to study the structure of DNA and RNA (Size and secondary structure), and also the condensation or bending of DNA (**Hansma et al., 2004**). AFM has also

proven itself in the study of DNA– protein interactions. The binding of the RNA polymerase of *Escherichia coli* with its target DNA and the binding of p53, a tumour-suppressing protein, to a specific target have also been visualised. In the latter case, AFM also revealed the dynamics of this interaction. The dynamics of DNA–protein complexes has been revealed using AFM when visualizing the degradation of DNA by the endonuclease DNAseI **(Abdelhady et al., 2003)**.

The first soluble proteins for which AFM was able to obtain high-resolution images (about 1 nm) were GroEL and GroES, two chaperones of the bacterium *Escherichia coli* . However, this image quality was obtained with AFM, which is a highly sophisticated experimental setup developed by Zhifeng Shao and coworkers at the University of Virginia (USA). This same type of equipment has been used to investigate changes in position of myosin heads, thiophosphorylated or otherwise **(Sheng et al., 2003**), and the structure of actins filaments. Other elements of the cytoskeleton such as microtubules have been observed in a liquid medium with a resolution of a few nanometers, using amore standard AFM.

This is the case for the membrane fractions of the photosynthetic apparatus in certain bacteria. AFM is then a unique tool for gathering information about the organisation of supramolecular complexes in physiological conditions, and this at a resolution close to 10˚A **(Scheuring et al.,,2004**) . Some membrane proteins, such as those of the communicating junctions in eukaryotes or bacteriorhodopsin in *Halobacterium halobium,* self-organise spontaneously into a 2D lattice in order to carry out their function. The trimeric organization of bacteriorhodopsin, a proton pump with 7 transmembrane helices, is

revealed with a lateral resolution as good as 5˚A and a vertical resolution in the range 1–2˚A. The dimeric organization of rhodopsin in rod membranes has also been demonstrated recently. The structural characterization of molecules was revealed by AFM. In analogous to that TEM and image processing were adopted in our study.

7.2. Biosynthesis of Gold NanoParticles (GNPs):

Many of the properties of gold nanometal colloids make them particularly amenable to biological applications, namely: they are highly customizeable in size, shape, and surface chemistry, they are stable in a wide variety of environments, they are inert, non-toxic, and they have controllable optical-electronic properties. The use of these materials for biological sensing and medical therapeutic applications becomes now a challenge. Gold NPs demonstrate massive dipole interactions with light rays; their (surface plasmon Resonance) SPR bands are very sensitive to their environment. Plasmon resonance is the basis behind a popular analytical technique for biomolecular detection, where the index of refraction changes upon biomolecular binding (e.g., protein) to thin gold films. Mirkin and coworkers used such a technique to monitor nucleic acid interactions through shifts in colloidal gold SPR bands **(Taton et al., 2000)**.

The aggregation of gold nanoparticles may be characterized by a drastic shift in color. The color shift is a result of the interacting electric fields of neighboring particles and has a tendency to lower the resonant frequency of plasmon oscillations (lower energy SPR absorption band.) The closer particles get to one another, the greater the magnitude of this

shift. Shifting SPR energies therefore give the basis for colorimetric tests using capture-target bimolecular binding schemes. Fig. (1) (Left) demonstrates this colorimetric scheme for the detection of specific DNA strands. When gold nanoparticles containing a specific recognition strand of DNA are introduced to a solution in which the complementary strand is also bound to gold nanoparticles, the proper hybridization will draw the nanoparticles together and shift the SPR frequency to lower energy. A change in the color of the solution leads to the detection of the DNA strands. The use of gold nanoparticles in biomolecular detection schemes have advanced rapidly by Mirkin's group to include not only shifting in the SPR bands for colorimetric detection but to incorporate silver ion reduction into the detection platform for enhanced detection sensitivity via darker color intensity or surface-enhanced Raman spectroscopy (SERS) **(Cao et al., 2002)** . Using a surface-based capture assay, a gold NP containing a specific sequence of single-stranded DNA will bind to a patterned surface of different complementary sequences. After washing the slide to remove non-specifically bound NPs, only the patterned regions of interest contain gold NPs via DNA hybridization. The NPs can then be detected either colorimetrically by reducing silver ions onto the gold NPs for an enhanced color change, or using SERS. The use of many spots combined with the unique DNA sequence on each spot can be used for high-throughput analysis and detection of genes or for profiling hereditary disease or genomic-based pathogen detection. In addition to Mirkin's group, others demonstrated the use of gold nanoparticles to detect single basepair mismatches of complementary DNA by using the electronic properties of gold particles not in an absorption based colorimetric assay, but as the quenching species of a fluorescent molecule

(**Dubertret et al., 2001**). Fig.(1) (right) demonstrates the experiment performed by Libchaber et al. in which a hairpin loop of single-stranded oligonucleotide sequence was synthesized to contain a small gold NP at one end and a fluorescent dye at the other. When the oligonucleotide is wrapped up in the hairpin loop, the fluorescent dye and the NPs are very close together and the dye is incapable of emitting photons of light. However, if a matching oligonucleotide sequence, or sequence to be detected, with a perfect complementary match is introduced into solution, the hairpin structure will unravel to form a linear double-stranded structure that is more stable than the hairpin conformation. As a result, the dye and NP are spatially separated and this leads to the dye emitting photons of light. The output correlates to the concentration of the sequence to be detected, which can be from cells or tissues associated with a disease. Although the ability for a gold NP to quench the fluorescence of a dye molecule is clear, the mechanism responsible for this observation requires further elucidation. Aside from colorimetric assays and fluorescence quenching, the SPR of spheres, rods, and even gold shells is being hotly pursued for the potential use in photothermal therapy.

The SPR band, as described earlier, is a gigantic dipole, with extinction coefficients ranging into the billions (M-1cm-1) if the particles are large enough.

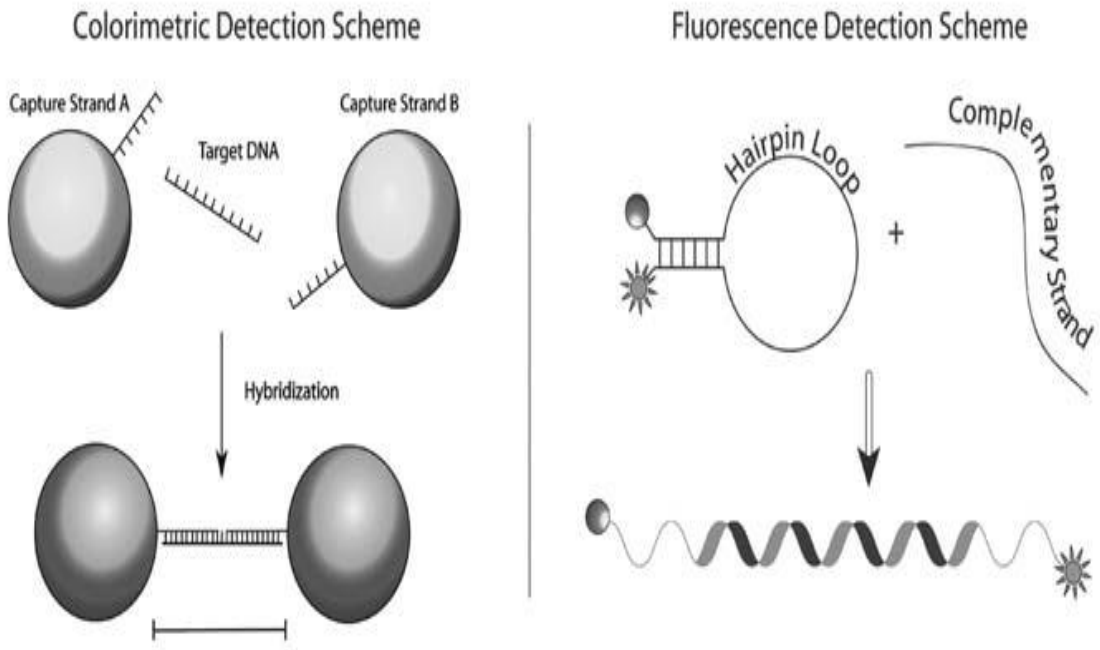

Fig.(1). Colorimetric detection scheme as developed by Mirkin et al. The hybridization of complementary surface-bound DNA strands brings gold nanoparticles together, which affects SPR behavior and will change the visible color of the solution.

It was realized that this massive absorption feature could extend metallic nanoparticles toward therapeutic purposes. Gold NPs are (mostly) non-photoluminescent structures where the absorbed energy by the NP were mostly dissipated to the surrounding environment as heat. If the NP exists in an environment such as a tumor cell, then the heat dissipated by the NP through light absorption will locally destroy its environment, leaving the healthy neighboring cells alone. To accomplish this goal, of course, the gold NPs must first target specific cancerous cells and not internalize into the healthy ones. The use of specific targeting antibodies, such as Herceptin's affinity for breast cancer, may be employed toward this end. Studies to understand the cell's uptake dependence on shape and size are being performed **(Chithrani et al., 2006)**, as well as the efficiency of photothermal therapy to destroy

specific cells while leaving healthy cells in tact are being performed in-vitro.28-30 Rod-shaped NPs may prove to be particularly useful for photothermal therapy because the longitudinal plasmon mode in the near IR wavelength range will be more accessible for external illumination through skin tissue or blood than is the transverse plasmon mode.

In comparison to most inorganic nanomaterials, gold nanoparticles can be synthesized in the largest variety of shapes and sizes. The most common shape, and arguably one of the easiest to synthesize, is the simple sphere. In the synthesis, gold chloride is added to aqueous solvents in the presence of a reducing agent. Gold nanoparticles are quickly formed, as observed simply by the change in the color of the solution (light yellow goes to red). Growth is rapid and no crystallographic face is favored over others when a neutral stabilizing agent such as citrate is used to bind the surface. In most solution-phase synthetic methods, stabilizing agents such as phosphine, thiol, carboxylate, or amine containing molecules are often used. These molecules are important primarily in mediating the interactions between the NP surface and the surrounding bath. Without stabilizing agents, particles in solution are more susceptible to either oxidation or the attractive interparticle Van der Waal's forces which cause them to aggregate and precipitate from solution. The most successful approaches to controlling nanoparticle size and dispersity are therefore often dependent upon the choice of ligand, surfactant, or stabilizing agent (**Warren, 2007)**.

The nanostructured materials have been receiving considerable attention, due to their unique physical and chemical properties and

important applications in optics, electronics, biomedicine, magnetics, mechanics, catalysis, energy science, etc. It is well known that biological systems can provide many examples of specifically tailored nanostructures with highly optimized properties and characteristics including magnetotactic bacteria synthesizing intracellularly magnetic nanocrystals in magnetosomes , diatom cell walls regarded as a paradigm for controlled production of nanostructured silica , bacterial S-layer as templates for the formation of regularly arranged nanoparticles and nanobacteria also known as calcifying nanoparticles (**Kajander, 2006)** . Synthesis and assembly of nanoscale materials using biological systems is a relatively clean, nontoxic and environmentally friendly, namely green chemistry procedure comparing with conventional chemical synthesis techniques. There has been great and increasing interest in looking at biological systems for inspiration and using microorganisms as workers in the living factory for the production of new functional nanomaterials. Many microorganisms including living plants, extract of plants, bacteria **(He et al., 2006)**, fungi, actinomycetes and human cells, have been reported to produce either intracellular or extracellular bio-nanocomposites. Microorganisms as possible eco-friendly nanofactories can exert control over size, morphology, composition, and crystallographic orientation of the as-prepared metal nanostructured particles. Biosynthesis of noble metal nanoparticles with well-defined shapes and morphologies has been observed including spheres, hexagons, triangles, rods, flat sheets, decahedrons, icosahedrons, hierarchical tube assembly and nodous ribbons and so on, especially gold nanotriangles with potentially exciting applications in hyperthermia of tumors, optical coatings and scanning tunneling microscopes as conductive tips .

Biosynthesis of gold nanoparticles assisted by *E. coli* DH5a and its application on direct electrochemistry of hemoglobin are reported. The gold nanoparticles formed on the bacteria surface are mostly spherical **(Liangwei et al., 2007).**

The increasing number of bacterial strains that are resistant to available pharmaceutical compounds is a vital issue for public health. Innovative approaches will be required to improve the methods for both diagnosis and destruction of these organisms. It was considered that the possible role that can be played by technologies based on gold nanoparticles. Gold nanoparticles generally are considered to be biologically inert but can be engineered to possess chemical or photothermal functionality. A growing body of research is devoted to the potential use of these nanoparticles in the diagnosis and treatment of bacterial infections. The results are both promising and intriguing, and suggest a range of new strategies to identify, target or destroy pathogenic organisms **(Pissuwan et al., 2007**).

Biological synthesis of gold and silver nanoparticles was carried out using the bacteria *Bacillus subtilis*. The reduction processes of chloroaurate and silver ions by *B. subtilis* were found to be different. Gold nanoparticles were synthesized both intra- and extracellularly, while silver nanoparticles were exclusively formed extracellularly. The gold nanoparticles were formed after 1 day of addition of chloroaurate ions, while the silver nanoparticles were formed after 7 days. The nanoparticles were characterized by X-ray diffraction, UV-vis spectra and transmission electron spectroscopy. X-ray diffraction revealed the formation of face-centered cubic (fcc) crystalline gold nanoparticles in the supernatant,

broth solution and bacterial pellet. Silver nanoparticles also exhibited diffraction peaks corresponding to fcc metallic silver. UV-vis spectra showed surface plasmon vibrations for gold and silver nanoparticles centered at 530 and 456 nm, respectively. TEM micrographs depicted the formation of gold nanoparticles intra- and extracellularly, which had an average size of 7.6 +/- 1.8 and 7.3 +/- 2.3 nm, respectively, while silver nanoparticles were exclusively formed extracellularly, with an average size of 6.1 +/- 1.6 nm. The bacterial proteins were analyzed by sodium dodecyl sulfonate-polyacrylamide electrophoresis (SDS-PAGE) before and after the addition of metal ion solutions. We believe that proteins of a molecular weight between 25 and 66 kDa could be responsible for chloroaurate ions reduction, while the formation of silver nanoparticles can be attributed to proteins of a molecular weight between 66 and 116 kDa. We also believe that the nanoparticles were stabilized by the surface-active molecules i.e., surfactin or other biomolecules released into the solution by *B. subtilis* (**Reddy et al., 2010)**.