USE OF SYNTHETIC DNA AS NEW TRACERS FOR TRACING GROUNDWATER FLOW AND MULTIPLE CONTAMINANTS

Iftikhar H. Sabir 1, Peter Aleström 1, Sylvi Haldorsen 2

1 Geology and Water Section, Department of Soil and Water Sciences, Agricultural University of Norway, P.O. Box 5028, N-1432 Ås, Norway
2 Department of Biochemistry, Physiology, and Nutrition, Norwegian College of Veterinary Medicine, P.O. Box 8146 Dep. N-0033 Oslo, Norway

Also at: ChemTAG AS, c/o Det Norske Veritas, N-1322 Høvik, Norway

Abstract: It has recently been possible to trace groundwater and contaminants flow paths by using synthetic DNA molecules as tracers. The DNA tracers injected in groundwater can be as distinctive and traceable as a person’s fingerprint. Alphanumeric information like names, dates, batch numbers, can be defined into the DNA sequence. This novel tracing technique can provide an unlimited number of uniquely labelled tracers with different sizes and valences. Using these tracers is a great advantage because they allow us to evaluate a number of origins of pollutants simultaneously but released over different times. The ability to figure out where water or a pollutant originated and where it has travelled, and such flow paths are a big concern for hydrologists trying to predict presence of multiple contaminants in the subsurface environments. Short pieces of single-stranded DNA produced by automated standard oligonucleotide synthesis have successfully been used in groundwater flow and contaminant transport studies. Polymerase Chain Reaction (PCR) is utilised to detect and read information in unique DNA tracers. PCR has extremely low detection limit and allows specific detection of one DNA sequence in a mixture of tracer signatures. Synthetic DNA tracer analyses could simplify interpretations of groundwater pollution problems and should be routinely performed in groundwater studies where there is a demand of simultaneous use of two or more tracers. In this chapter the DNA tracing technique is briefly reviewed.

Keywords: Synthetic DNA tracing technique.

Introduction
Groundwater constitutes a significant source of water all over the world. A problem of growing concern is the wide varieties of materials that have been identified as contaminants in groundwater predominantly as a result of domestic, agricultural, and industrial activities. Common examples of contaminants include synthetic organic chemicals, hydrocarbons, inorganic anions and cations, pathogens, viruses, bacteria, and radionuclides. Much of the emphasis in groundwater investigations has shifted from problems of locating and developing potable groundwater resources to problems pertaining to the occurrence and migration of contaminants in groundwater (Cherry, 1983). Sources of contamination may be either localized or spread over a large area. The origin, source geometry, and release time of numerous contaminants are usually the most elusive goals in studies of contaminant migration in groundwater.

Tracers (either artificial, or environmental) have gained widespread use in providing much needed information in areas such as surface water dynamics, atmospheric transport, hydrologic pathways, residence times, erosion and sediment transport, and solute transport Peters et al., (1993). Tracing techniques using natural tracers and isotopes, but also various artificial tracers exist which are contributing to provide temporal information in different conditions. Until now, none of the available tracing techniques provide the information on the real-time three-dimensional simultaneous detection of numerous contaminants in a system emitted at different times and their relative persistence in the time framework. It is only recently synthetic DNA tracing technique with information capacity is developed (Aleström, 1995) to characterizing temporal contaminant transport in groundwater.

Given the unique power of information storage and coding potential, the DNA tracers provide the means of gathering information about time and origin of individual contaminants and its distribution, which could thus be important in predicting the real-time location of contaminants. Such type of information is not possible to obtain with already available tracers. Hence, the DNA tracers have the potential to be an extremely useful tool in predicting the natural distribution of dissolved contaminants. Some special features are that DNA can be detected at extremely low concentration and no additional effort is required to detect several DNA tracers simultaneously. A few milligrams of DNA is sufficient for tagging of one billion m³ of water. Also, DNA's sorbed phase can be of interest for studying the transport of colloids and sediments in the subsurface environments, Mahler et al., (1998). More information on the features of DNA tracing technique is given in Sabir et al., (1999; 2000).

The use of DNA as new tracers in groundwater is of interest for at least four main reasons: 1) the transport of DNA as tracers for tracing and investigation of groundwater flow in porous medium and in fractured rocks, 2) the distribution of dissolved, i.e. the mobile phase contaminants, 3) DNA tracer's sorbed phase associated with aquifer solids, and 4) to understand the transport behaviour of other materials, when introduced intentionally to the environment or naturally occurring in groundwater. Recently, evidence has been accumulated that shows DNA tracers can successfully move through the porous medium and fractured systems, and discriminates between the sources of contaminants (Sabir et al., 1999; 2000).
An important but yet poorly understood aspect of DNA behaviour as tracers is related to the possible interactions with the dissolved phase chemicals and solid matrix; adsorption, desorption, and persistence processes, and the effects of different hydrological and physical changes which may occur during the course of DNA transport along hydrological pathways. With these aims in mind, in this paper, first, the physical and chemical properties of DNA molecules are discussed in a general fashion from the literature. Second, a tabular summary on the few potential applications of the DNA tracing technique in various fields is presented. Finally, a conceptual model is presented that can be used for insights into the understanding of mechanisms that control DNA migration.

**DNA molecule structure:** The cell is a basic unit of any living organism. It is a small compartment filled with proteins and other chemicals dissolved in water. The main type of chemical inside the nucleus of a cell is deoxyribonucleic acid (DNA). The DNA functions as the "hard disc" of the cell and carries through the combinations of sequence codes all genetic instructions needed by the cell. DNA is built up from subunits known as nucleotides. Each nucleotide consists of a nitrogenous base, which may be adenine (A), guanine (G), cytosine (C) or thymine (T), and the sugar-2'-deoxyribose interlinked by phosphodiester bonds. The shapes of A and T and of C and G are "complementary" through a hydrogen bond binding pattern. Each of the two pairs fit together neatly like pieces of a jigsaw puzzle. Adenine (A) pairs with thymine (T), while cytosine (C) pairs with guanine (G). Chemical analysis of double-stranded DNA has shown that the content of adenine (A) is equal to that of thymine (T) and that the content of guanine (G) is equal to that of cytosine (C), Fig. 1. illustrates different conformations of base-paired DNA, Smellie (1969).

DNA can be thought as a long-chain polymer built up of nucleotide units. A huge number of DNA molecules exist in nature, differing from one another in size and in sequence of nucleotides. Most naturally occurring molecules of DNA are double-stranded. Single-stranded DNA molecules occur in some viruses. In order to specify the precise nature of any given molecule of DNA, it would be necessary to establish the sequence of nucleotides in one of its two complementary strands. The two strands of connected nucleotides run in opposite directions (opposite polarity) to one another to comprise the DNA double helix and are held together by hydrogen and other non-covalent bonds between the bases. The ends of the nucleic acid polymers are named 3' and 5' because of the numbering of carbon atoms in the sugar moiety.

The four nitrogenous bases of DNA (G, C, A and T) are arranged along the sugar-phosphate backbone in a particular order (the DNA sequence), encoding all genetic instructions. The space between one base pair and the next is 3.4 Å (3.4x10⁻10 m), and the double helix makes one complete turn every 34 Å, i.e. after ca 10 base pairs (Fig. 2). By virtue of its structure and size, a molecule of DNA normally is very long (millions of base pairs) and thin (20 Å in diameter). The molecular weight of a single nucleotide is ca 330 D in average and will be as much as 2x10⁶ for a plasmid with 3x10⁶ base pairs. Such a structure would be approximately 1 mm in length but only 20 Å in diameter (translated in terms of a piece of wool 0.1 inch in diameter, the length would be close to approx. 1.6 Km). It will be understood, therefore, that the very long DNA molecules found in nature are rather fragile and easily broken into smaller pieces. The non-covalent association between the two strands of a DNA molecule can be broken by several mild procedures such as heating (100 °C) or exposure to alkaline conditions (pH 12). The ease with which this occurs is partly determined by the proportion adenine/thymine and guanine/cytosine base pairs in the molecule, and this property has proved important in the characterisation of DNA molecules from different sources.

DNA molecules range in density from about 1.68 to 1.74 and this is related to the content of GC (3 hydrogen bonds) and AT (two hydrogen bonds) base pairs. The higher the proportion of GC pairs in the DNA the higher is the density of the molecule. Thus an expression has been devised relating the buoyant density (p) of DNA to its content of GC base pairs:

\[ p = 1.660 + 0.098 \times [\text{GC}/2.44] \]

When DNA is heated in solution the double helical structure eventually collapses giving rise to single strands coiled in a random manner. This process is referred to as denaturation or melting of DNA, and the temperature at which this occurs is characteristic of different species of DNA. The melting temperature, like the density, is related to the GC content of DNA and is also influenced by the salt concentration of the solution. Since both buoyant and melting temperature of DNA can be measured experimentally, these expressions provide convenient methods for determining the GC content of DNA samples. The following expression has been devised to relate the melting temperature \( T_m \) to the GC content of DNA under standard conditions of ionic concentration:

\[ T_m = 69.3 + [\text{GC}/2.44] \]

**Hydrophobic forces in the DNA molecule:** The DNA consists of two components, the bases and the sugar-phosphate backbone. The bases have quite low solubility in water, owing to the aromatic ring, whereas the sugar-phosphate backbone is hydrophilic. High solubility in water indicates the ability to interact with water molecules; a substance of low solubility cannot interact significantly with water and tends to interact with itself (a so-called hydrophobic force). In double-stranded DNA the hydrophobic interactions between the bases help stabilise the double helix structure generated by the hydrogen bond base pairing. A single-stranded DNA nucleotide incapable of forming any hydrogen bonds thus will expose both hydrophilic phosphodiester bonds and hydrophobic (water-repelling) groups. The material presented in this section is adapted from Freifelder (1978).

**Electric properties of DNA molecules:** DNA molecule is, because of the phosphodiester linkages, negatively charged polymer ion (anion). In addition to the masses, shape and sizes, its migration may take place in response to its electric charge. The changes in the transporting environment (e.g. in heterogeneous porous media with variable cation-exchange capacities of coarse and fine fractions) may raise special problems. The characteristics of DNA tracers can be chemically modified to overcome such problems.

**Obtaining single-stranded synthetic DNA tracers:** A total 72-nucleotides long fragments of single
stranded DNA with two oligonucleotide primers designed for amplification with PCR was used. Both the longer oligonucleotides (50-100 nucleotides) to be used as tracers and the shorter (15-20 nucleotides) to be used in the PCR analyses were produced by chemical synthesis. The chemical composition is identical to single stranded DNA molecules found in nature, although natural DNA molecules normally are orders of magnitude longer Sabir et. al. (1999 and 2000).

**Polymerase Chain Reaction (PCR):** PCR is repeated cycles of DNA replication carried out in vitro using a thermostable DNA polymerase enzyme and a thermocycler melting the two strands of DNA apart between every cycle of replication. The oligonucleotide primers act as the starting point (the 'primer') for synthesis of a new polynucleotide chain, complementary to the existing ('template') strand. The first step in obtaining a single stranded DNA molecule is to attach ('anneal') a short oligonucleotide on to each single stranded molecule and delineate the region to be amplified. The chain elongation reaction is catalysed by a DNA polymerase enzyme. To carry out PCR one must know at least some of the sequence of the target DNA, the binding sites for the primers. The code of the PCR amplified tracer molecule can be verified by direct sequence analysis (for details of PCR and DNA sequencing see any modern textbook in biochemistry).

The advantage of DNA tracers over other existing tracers is largely due to the capability of analysis via the polymerase chain reaction (PCR) developed for biochemical analysis of DNA molecules and commonly used in forensic medicine. A repetitive series of cycles involving template denaturation, primer annealing, and the extension of the annealed primers by DNA polymerase results in the exponential accumulation of a specific fragment whose termini are defined by the 5' ends of the primers. Because the primer extension products synthesized in one cycle can serve as a template in the next, the number of target DNA copies approximately doubles at every cycle. Thus 20 cycles of PCR yields about a million-fold (20) amplification. PCR is an extremely sensitive method theoretically requiring only a single molecule as substrate for amplification. It is extremely specific, as the primers can be designed to anneal to a single, unique target sequence in a mixture of other DNA sequences, and a virtually infinite number of differently tagged molecules can be synthesized and used simultaneously. All such characteristics put DNA tracers in a special niche among the other available tracers. Effective use of the DNA tracer will require quantification of the amount of tracer present in each sample. With the latest PCR technology (TaqMan PCR instruments), it is possible to quantify the number of tracer molecules in the sample (Applied Biosystems, Roche). Precautions when evaluating the results from DNA tracer studies by quantitative measures must be taken because of possible losses of tracer DNA through adsorption to clay minerals and chemical degradation. Using quantitative PCR, the time of the arrival, concentrations, and shape of the resulting breakthrough curves for the tracer can be determined.

**Gel electrophoresis and sequence analysis:** The DNA fragment pattern from PCR analysis was visualised by agarose gel electrophoresis and ethidium bromide staining. PCR amplified tracer DNA was used as template in DNA sequencing reactions. Oligonucleotide primers were used in cycle sequencing reactions, using ABI PRISM Dye terminator cycle sequencing kit (Perkin Elmer Corporation) and a Perkin Elmer thermo cycler. The sequencing products were finally analysed on a Perkin Elmer ABI 377 automatic DNA sequencer (Sabir et. al., 1999).

**DNA tracers and interfacial interactions:** Being a macromolecule, DNA exhibits a range of chemical properties. For dealing with DNA tracers, the first challenge is to be able to well predict the spatial distribution and possible speciation of DNA during their reactive transport. The second challenge is estimating parameters mainly controlling their movement in the natural subsurface environments. The subsurface is a very complex system, which is heterogeneous with respect to physical and chemical properties. On the one hand subsurface plays the role of the skeleton through which transport of DNA occurs, and on the other hand it can react with DNA tracers. Its geometry is of great interest because of DNA tracer size and shape.

**Adsorption:** DNA is known to adsorb to clay minerals, Goring and Bartholomew (1952); Greaves and Wilson, (1969). Adsorption of DNA to sand was rapid and dependent on salt concentration, salt valency (ionic conditions), and pH, Lorenz and Wackernagel (1987). The adsorption of DNA increases with increasing diveral cations (e.g. Mg$^{2+}$ or Ca$^{2+}$). A considerably less DNA is adsorbed to the sand in presence of Na$^+$, and DNA probably adsorbs to sand by means of physical attraction forces, such as van der Waals forces, and hydrophobic interaction between DNA and sand plays minor role. The pH of the solution has a marked effect on the binding of DNA to sand and DNA adsorption is a charge-dependent process. In the presence of Mg$^{2+}$, DNA absorbed more efficiently at high pH (>7.5-9) than at low pH (<6), and by a minimum at neutral pH. Adsorption of DNA to sand by cation bridging has been proposed as a mechanism for nucleic acid adsorption to sand. This is in accordance with the results by Greaves and Wilson, (1969). In another study by Aardema et. al., (1983) more adsorption of DNA to sand was observed at a high salt concentration or low pH. The adsorption to clay is three to four orders of magnitude greater than that to sand. Adsorption of DNA may also be dependent by the original colonization by bacteria and the biofilm created by the attached bacteria. No correlation between temperature the extent of adsorption is reported. The DNA desorption process is much slower than adsorption and the DNA-sand complex is a relatively stable association. Lorenz and Wackernagel (1987).

**Degradation rate:** The use of DNA molecules in the aquatic environment raises the question of the survival of these because of DNA's susceptibility to microbial degradation. Studies have shown that the adsorption and binding of DNA to clay minerals and other particulates may protect DNA against degradation, and thus enhance its persistence, Stotzky (1989; Stotzky and Babich (1986); Greaves and Wilson (1970). Free DNA is more rapidly degraded then salt-adsorbed DNA and with variable kinetics, depending on the Mg$^{2+}$ concentration at which the DNA-sand complex is formed. The concentration of Mg$^{2+}$ ions (and presumably other divalent ions) is a critical factor of DNA degradation and adsorbed transforming DNA is protected from rapid degradation. Lorenz and Wackernagel (1987). The ability of DNA to bind on
soil particles and resist against degradation could be a useful traits and the basis of further developing the DNA tracing technique. Changes of the ionic strength in presence of some salts alter the type of DNA-sand association, which results into activation or inactivation of nuclease enzyme activities and consequently the DNA degradation rate. Additional studies are required to determine the types of interaction between DNA and mineral surfaces, and the ionic conditions that may be important in natural environments for adsorption of DNA and its protection against enzymatic degradation.

**DNA sorbed/labelled colloids**: The association of DNA molecules with colloidal matter in groundwater has important ramifications to quantify the mobile DNA tracers. A colloid is a particle ranging in size from 1 nm to 1000 nm, e.g., bacteria, viruses, large macromolecules of dissolved organic carbon (for example humic substances), small droplets of non-aqueous phase liquids, inorganic rock, or mineral fragments. Binding of DNA tracers to natural colloids can significantly alter their transport behaviour in porous media. The transport properties of colloids may differ from solute transport properties due to their size and charge. In experiments, the average velocity of inert colloids is found to be equal or larger than the average water velocity, Enfield and Bengtsson (1988): Pulse and Powell (1992) and Small (1974). Also, a fraction of the colloids may be immobile due to attachment to the solid matrix. Consequently, colloids may travel faster or slower than inert solutes. Dependent on the interaction between DNA, colloids, and the solid matrix, DNA transport may be enhanced or retarded. Table 1.

**Table 1:** Overview of the Problems that might be studied by DNA tracers

| For tracing origins of groundwater and contaminants in porous medium and fractured rocks. |
| Identify the occurrence, residence time, stability, and mobility of multiple contaminants in groundwater. |
| Assess the evidence that contaminants bind to mobile contaminants in groundwater. |
| Identify mobile contaminants or suspended particles ( colloids facilitated transport). |
| Identify contaminants characterization process (particles themselves may be contaminants). |
| Determination of surface charge characteristics of selected trace colloids in groundwater (migration of radioactive waste). |
| Identify degradation patterns and stability kinetics of microorganisms. |
| Assess the occurrence, diversity, and enrichment of microorganisms in an environment and changes in community patterns. |

**Discussion**

Ground waters typically have a large range in chemical composition. This large range is attributed to two fundamental causes, the ability of groundwater to interact with the ambient environment and the systemised spatial distribution of its flow. The successful transport of DNA tracers in the groundwater requires an understanding of the processes, which control their movement. One of the most important controls is the sorption of the DNA species by interactions with the soil, which is poorly understood at the present time. Transport of DNA tracers through sandy aquifer sediments has been investigated in forced gradient
tracers test Sabir et al. (1999). In other field studies, differently coded DNA tracers have been shown to be competitive and intelligent tracers to distinguish the sources of pollution in porous media and origins of water sources in very complex fractured rock systems (Sabir et al. 2000). In all these studies, it was demonstrated only qualitatively, positive or negative. Based on the results of qualitative field studies, it is difficult to quantify the sorption process because of the lack of information on the relationship between DNA transport and its sorption with subsurface materials. In addition, DNA transport modeling has not been attempted to understand the processes that mainly govern DNA transport and fate.

In order to answer the preliminary questions about the sorption and particularly persistent of DNA tracers, these processes should be investigated involving an evaluation of sorption data with quantitative PCR. For this purpose, a substantially controlled laboratory column studies and application of simple transport models are required to assess the effect of different controlling factors responsible for DNA retention and transport behavior through porous media. Later on, the study can be extended by introducing advance-modelling routines to predict the coupled and kinetic interactions during transport of DNA tracers in presence of reactive solutes and colloids in the aquatic environments.

The transport processes of DNA labelled sediments and colloids are far more complex than the transport of DNA tracers with fluid flow only. To describe the distribution of DNA tracers between sediments and water, the difficulties lie not only in the complexity and diversity of aquatic sediments but also in the surface charge carried by the colloidal particles, which may influence the sorption significantly. A model should take into account the dependence of cations, anions, etc. to take into account all important mechanisms simultaneously for insight of the complete speciation of DNA tracers. Fig. 3 is a conceptual model showing speciation of DNA tracers during transport in typical aquatic porous media. As discussed before, a prominent fraction of DNA introduced into the aquatic environments may associate with resident sediments, suspended colloids or in the flowing water. Four different species of DNA are distinguished in the system: (1) free DNA transporting in the liquid phase (D1); (2) DNA sorbed on the solid matrix (D2); (3) DNA sorbed on mobile colloids (D3); and (4) DNA sorbed on immobile colloids (D4). The movement and adsorption of DNA tracers through all the four processes (D1 to D4) are of great interest especially to hydrogeologists working in a range of environments (Fig. 3).

Conclusions
This new technique offers a qualitative, and with certain reservations quantitative, detection of a large number of individual but chemically close to identical tracers. Different sequences of nucleotides will not be 100% identical chemical identity but for practical purposes behave as if they were identical. The DNA tracers are environmentally safe and being natural identical nucleic acid in very low concentration is fully non-toxic with respect also to human health. DNA tracers can be designed and altered in terms of its physical and chemical properties to suit any target media. Detection limit of the tracers can be as low as one single tracer molecule.

This new technique offers a cost-effective and versatile alternative to the use of bacterial and radioactive tracers. The technique has enormous potential for the monitoring of the real time colloidal transport, and sediment monitoring. There is neither an ideal tracer, nor a perfect isotope that is suitable, or can be recommended for use in all hydrogeological contexts and purposes, why the DNA tracers can be a valuable complement for several applications, and an environmental friendly alternative. The good choice of an adequate methodology involving the most suitable tracing products is of prime importance to the accuracy of any study.

The DNA tracers have been tested in field studies designed to understand their mobility in groundwater and to delineate fracture interconnectivity during pumping and injection. The intrinsic information-coding capacity of DNA tracers in quantifying real-time signals is of great significance with theoretical and practical viewpoints. However, sorption and degradation are crucial and yet poorly understood processes controlling the fate and transport of DNA tracers and need to be further characterized before quantification results can be used for accurate conclusions of simulated tracer movements.

Acknowledgments
Thanks to the Department of Soil and Water Sciences, Agricultural University of Norway for supporting the principal author.

References


