Chapter 1

Literature Review

1 Introduction

Natural humic acids are polymeric brown-black organic acids that are ubiquitous in nature, being found in almost all soils and surface water (Aiken, 1985; Berzelius, 1893; Davies, 1996; Hoppe-Seyler, 1889; Kononova, 1966; Stevenson, 1982). Humic acids are of the allomelanin class of compounds and are derived mostly from plant, but also animal, material. The chemical characteristics and physical properties of humic acids vary depending on the source from which they have been extracted. Humic acids present in soils and most fresh waters, are acknowledged to be important for soil fertility (Vaughan & Malcolm, 1985; Visser, 1986), are fairly refractive to chemical and biological decomposition (Hedges & Oades, 1997) leading to their dominance as the organic matter found in soils. Humic acids are also precursors to, or inclusions in many of the abundant natural resources like peat, bitumen, coal and petroleum. They are claimed to possess some therapeutic medicinal activities (Janecek & Chalupa, 1969; Priegnitz, 1986; Reichert, 1966; Visser, 1973; Ziechmann, 1996), a characteristic known from ancient times. Despite the abundance, the known effects and uses and the many years of research, a probable but not universally accepted chemical structure was only proposed relatively recently (Schnitzer, 1985; Schulten *et al*, 1991). The reason for the long delay in determining the structure is straightforward: humic acids are complex heterogeneous mixtures of compounds (Wilson *et al*, 1987), often containing unrelated inclusions, and which are undergoing continuous chemical metamorphosis.

Humic substance research has seen slow progress as humic acid structure and chemistry is difficult to unravel, even with the technology available today. Many new researchers that recently joined the search for the elusive structure of humic acid have added to the impetus of this research.

The term humic acid is generally used to describe the brown-black, polymeric, alkali-soluble organic acid fraction of humus that is found in geological sediments, soils, wetlands, surface and underground water, and that more recently has even been identified in certain living fungi and plant materials (Kühnert *et al*, 1982).

Because humic acids are ubiquitous in nature and can be extracted in varying proportions by the many subtly different methods used by researchers in the past, many compounds with slight chemical differences and various inorganic inclusions and organic aggregates have been identified as humic acid. The variations in the origin and extraction methods utilized resulted in humic acids with different solubilities, colours, and textures which were the primary properties used to characterize compounds until structural chemical analysis became readily available in the latter 1900's.

2 Historical background

Humus was originally named as such in 1761 by the German scientist Wallerius (1761) using the Latin name for soil: Humus (the term humus is generally applied to all organic compounds of plant origin found in the soil). Twenty-five years later the first extraction of a soluble sub-fraction termed "humic acid" was reported by Achard (1786) who used peat, a rich source of humic acids, for his experiments. At that time it was recognized that humus and the alkali extractable portion named humic acids contributed to soil fertility but the mechanism of action was still unknown and some wild assumptions were made in this respect.

One of the first comprehensive studies on humic acids that covered the chemistry, physical properties and its uses was published by Sprengel (1826) and highlights how little progress has been made since then with regard to chemical structure.

Synthetic substances that looked like humic substances and referred to as artificial ulmin were made as early as 1819 by Braconnot (1819) who added acids to starch or sucrose each of which then formed a dark precipitate that looked like the humic acids extracted from soils. Glucose was found to give the same type of products, and Malguti (1835) published his views of the transformation of carbohydrates to synthetic humic acid. In 1839 Mulder (Mulder, 1839) published his work on the synthesis of humic substances from cellulose, resulting in a then widely accepted assumption that humic acids were derived from polysaccharides.

By the late 1800's humic substances had been formed from many other types of organic molecules, some very simple and the polysaccharide theory was essentially discarded in favour of the newly developing idea of a micro-organism mediated transformation of plant material into humic substances. In the mid 1800, microorganisms were recognized as widespread in soils and the possibility that humic substances were formed in the soil by their action was offered as the origin of humic substances. This hypothesis better explained the then recently described ethers, esters,

ketones, anhydrides, furans, aromatic compounds as well as the nitrogen and sulphur that had been reported to be present in humic substances.

Schreiner and Shorey (1908; 1910), found organic compounds such as simple hydrocarbons, fatty acids, glycerides, resin esters, chitin, cellulose, xylans, sugar alcohols, sterols, lecithins, pyridines, amides, amino acids, purine bases, vanillin, numerous aliphatic and aromatic acids and elemental carbon in humus. Some workers (Marcusson, 1920; Gortner, 1916) still favoured the polysaccharide theory as late as 1914 as they found a strong correlation for the furan structure in both coal and humic substances. Other coal researchers however disputed this theory again by demonstrating that microorganisms rapidly consume polysaccharides (Fischer & Schrader, 1921) and that there could not be sufficient polysaccharides remaining in the soil to give the quantities of humic acids found. These same authors proposed that lignin was the precursor to humic substances, which started a new school of thought around the origins of humic acids.

Marcusson (1925) and Hilpert and Littman (1934), opposed the new lignin theory but it still gained popularity from both a microorganism based synthesis (Waksman, 1938) as well as from a chemical synthesis point of view. The presence of aromatic compounds provided support for the idea that lignins must be involved in the formation of humic substances, as polysaccharides contained no aromatic functionality.

Shapiro (1957) published a study of humic acid where he used chromatography and liquid phase infrared spectrophotometry. He also introduced organic solvents into the fractionation of humic acids that had until then been described as insoluble in any organic solvents. An ethyl acetate soluble sub-fraction of the humic acid that he isolated apparently had no aromatic character or compounds.

Gas chromatography (GC) became a popular method of analysis in the late 1950's but humic acid samples needed derivatization before this type of analysis could be done due to the requirement that the sample be volatile for the technique to work. To overcome this volatility problem, the oxidation method of Bone *et al.* (1934) was used to split the humic acid into volatile sub-fractions. Most compounds identified by GC were aromatic, which gave support to the lignin derived humic acid theory (Wright & Schnitzer, 1958). Several reports do however indicate that oxidation can result in the formation of aromatic compounds from certain reactive aliphatic substances (Reuter *et al*, 1983). Despite this shortcoming, chemical oxidation became a common method to aid in the analysis of humic substances especially when the oxidation products were further analysed by gas chromatography.

Kononova (1961) published her book that supported the lignin derived humic acid theory as did Schnitzer (1985) in a report in Nature, and Stevenson (Stevenson, 1982; Stevenson, 1994) published two editions of a book several years apart, both supporting the lignin derived humic acid theory despite mounting evidence that there were many aliphatic compounds not found in lignins present in humic acids.

In 1982 the International Humic Substances Society (IHSS) was established in an attempt to coordinate humic acid research, and to collect and maintain a reference collection of humic acid samples to give researchers access to standardized samples.

The aliphatic nature of humic acid has always been a contentious issue as it contradicted the generally accepted origin of the humic acids. However a marine humic acid described in 1972 was shown to be essentially aliphatic (Nissenbaum & Kaplan, 1972), and Harvey *et al.*, (1984) later proposed that marine humic acids were derived from fatty acids. Farmer and Pisaniello (1985) found no aromatic compounds in the samples of humic acids that they analysed by NMR. A year after this report, Ikan *et al.*, (1986) reported that they had also found mostly aliphatic compounds in the humic acids that they were studying. Few studies have been published on chemically reduced humic acids but those that have, have also indicated an aliphatic nature for the humic acids analysed (Mendez & Stevenson, 1966; Martin *et al*, 1987). The reports of humic acids that have exhibited little or no aromaticity are relatively few compared to the bulk of the data that points to aromatic structure, but it should be accepted that humic acids from different sources can exhibit different structures due to different starting material and environmental conditions in which they have matured and accumulated.

3 Structural considerations

Due to the variation in solubility of the various humic substances from different sources, there are large differences in the reported composition of the humic substances isolated from different soils, sediments and other environments. Added to this is the fact that humic acids tend to bind to each other physically in a relatively tight but random fashion (Wershaw, 1993), often incorporating molecules that are not strictly speaking humic substances such as lipids, carbohydrates, nucleic acids and amino acids (Khairy *et al*, 1996a; Khairy *et al*, 1996b). These non-humic compounds enter the complex at a rather early stage of the humification process, whereas the aromatic compounds tend to become detectable at the later stages of the process (Ziechmann *et al*, 2000). Humic acids also bind tightly to a number of minerals (especially silica and alumina) (Stevenson, 1994) and easily complex heavy metals (especially iron, copper, chrome and the lanthanide series) (Griffith & Schnitzer,

1975), which in turn affect their solubility and which can catalyse chemical reactions within the complex of compounds making up the humic acid (Liu & Huang, 2000). Figure 1-1 illustrates some of the typical organic functionality that occurs within the humic acid structure. This includes aliphatic, aromatic, heterocyclic and polyphenolic functionality. The figure in no way attempts to assign a structure to any humic acids.

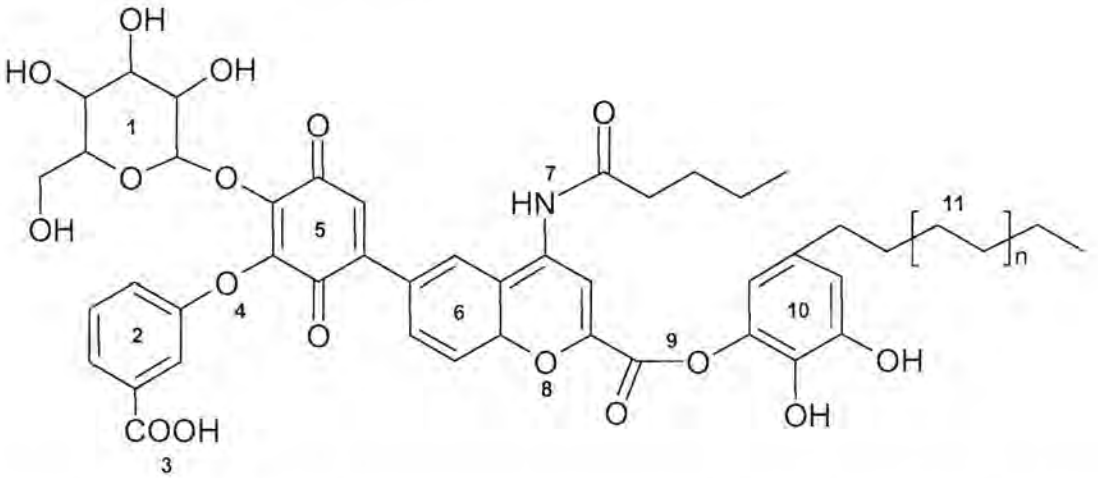

Figure 1-1: A diagram showing typical functional groups reported to occur in humic acids. This diagram in no way attempts to give a structure for a humic acid. 1, monosaccharides; 2, aromatic rings; 3, acid groups; 4, ether groups; 5, quinone groups; 6, fused aromatic rings; 7, amides; 8, heterocyclic rings; 9, esters; 10, polyphenols; 11, aliphatic carbon chains up to C34.

Coal is formed through a coalification process that involves a slow progression of reactions under distinct physical conditions that includes heat, pressure and submersion. The process takes place over many thousands to millions of years and is usually formed from humic acid rich deposits of rotting plant material (Ziechmann *et al*, 2000). This process of coalification can be reversed when coal undergoes weathering, usually when coal seams become exposed to oxygen and water. The medium to lower rank coals appear to spontaneously undergo this type of depolymerisation to form compounds that are humic substances, but which are probably not identical to the original compounds from which the coal formed (Rausa, 1994). Many of the compounds formed have typical chemical and physical properties of humic acids, although the more soluble fulvic acids and highly mineralized insoluble fractions are also present. The soluble products have high concentrations of carboxylic and aliphatic acids, phenolics and quinones (Wender *et al*, 1981).

4 Classification of humic substances

Humic substances were traditionally classified according to their solubility in aqueous medium at different pH values, but this was not a standardized method. Water-derived humic acids were precipitated at pH 2 whereas humic acids from soil were precipitated at pH 1. Generally, the weakest acids (highest pKa value) should be first to precipitate as the pH is lowered due to the suppression of ionisation. Unfortunately it is not this simple, as many other factors also influence the solubility of weak acids, such as the predominant counter ion, valence of the cations present, ionic strength, the presence of chaotropic agents, solvents or surfactants. These complications gave rise to different characteristics being reported for the isolated humic acids with respect to elemental analysis, ash content and solubility.

Humic substances are still generally divided into three classifications or sub-fractions based on the aqueous solubility of the humic substance (Odén, 1919). These are;

- fulvic Acids – the sub-fraction that is soluble in water at all pH values. They are generally reported as multivalent organic acids and have been found to have molecular mass varying from as low as 45 up to about 1000 Dalton.
- humic acids – the dark sub-fraction that is soluble in alkaline aqueous solutions but precipitates at pH of less than 2. They are not soluble in short chain alcohols. The reported molecular mass of humic acids have generally been found to be between 10 000 and 350 000 Dalton.
- humin – the water insoluble fraction. These compounds have very high apparent molecular mass and are more condensed molecules than the previous two fractions, being closer to peat and coal. They are often mineralized.

Ziechmann, (1993) defined a fourth fraction of humic substances as the haematomelanic acids, the fraction that is water and short chain alcohol soluble and with an average molecular mass between 5000 and 10 000 Dalton.

More recent methods of isolating humic substance sub-fractions with macroreticular resins has again given raise to a different grouping of compounds within the older general classification of humic substances (Leenheer, 1985; Malcolm & MacCarthy, 1992; Thurman & Malcolm, 1981). A column packed with XAD-8 resin is used first to extract the hydrophobic acids, hydrophobic bases and the hydrophobic neutrals from an acidified solution. The polar pass-through fraction is then separated into hydrophilic acids, hydrophilic bases and hydrophilic neutrals on an XAD-4 resin column. These columns can be used in tandem (Malcolm & MacCarthy, 1992) to isolate humic acids from aqueous

samples. This newer method appears to be the method of choice for the IHSS, although many laboratories still use the classical separation using pH adjustment only. This variation in isolation methodology remains a problem as it introduces yet another area of uncertainty into the structure of the humic acids.

5 Chemical analysis of humic acids

Humic acid can be isolated from many different sources and this has resulted in many compounds with fairly different chemical and physical properties being characterized. A major problem encountered is that classical chemical analysis has not been successful in the complete elucidation of the composition or the structure of humic acids. This has resulted in many hypotheses (Hayes, 1998; Steelink, 1999) as to the chemical and physical structure of the humic acids isolated from soils (Ghabbour *et al*, 1998), peat bogs (Francioso *et al*, 2001), lake sediments (Ishiwatari, 1985), coals (Novac *et al*, 2001), compost, animal excreta (Khairy, 1989), as well as from numerous aqueous environments (Nissenbaum & Kaplan, 1972). Making things even more complicated is that although it is generally accepted that humic substances are derived from well decayed plant material, there are reports of humic substances being isolated from living plants and senescent leaves (Wershaw *et al*, 1998b; Wershaw *et al*, 1998a; Ghabbour *et al*, 1994). Some bacteria and fungi have also been reported to synthesize humic substances intracellularly and release these substances during the decomposition of these organisms (Kühnert *et al*, 1982).

Although there are many easily hydrolysable bonds in humic acids, most bonds are difficult to cleave making humic acids fairly refractory to microbial as well as chemical decomposition (Ghabbour *et al*, 1998). The use of decomposition analysis where the humic acid is broken down into smaller molecules has generally required relatively harsh chemical treatment that resulted in complete breakdown of susceptible parts of the complex. This approach enhances the possibility of finding artefacts of the decomposition method rather than actual compounds making up the original material (Farmer & Pisaniello, 1985; Hayes, 1998; Lehtonen *et al*, 2001; Saiz-Jimenez, 1994).
Solubility becomes a problem at lower pH values (a characteristic that has been used in the broad sense to define humic acids) and humic acids isolated from environments that have been in constant contact with water have tended to have fewer lower molecular weight and polar soluble fractions.

A fairly generally accepted structure for humic acids is one of a dynamic heterogeneous complex of many different molecules at various stages of degradation (Baldock *et al*, 1992; Kononova, 1961; Novac *et al*, 2001; Stevenson, 1994) that easily complexes further with diverse organic molecules such as sugars, amines, fatty acids, waxes and aromatic compounds (Khairy *et al*, 1996a; Khairy *et*

al, 1996b), to metal ions (Nifant'eva *et al*, 1999; Ruiz-Has *et al*, 1998; Stevenson & Chen, 1991) or minerals (Wershaw, 2000). Each of these types of complexed compounds alters the solubility of the humic acid differently, probably because they bind to different regions of the acid, form cross-links between different acid molecules, or alter the exposed surface of the complex. Other proposals of the structure include "Maillard" products (a condensation reaction between reducing sugars and amines), "aromatic core and periphery" models, aliphatically linked aromatic rings and polymeric chains with substituting molecules. Humic acids are also reported to easily form micelles, even though the critical micellar concentration (CMC) is high – between 0.4% and 0.9% (Engebretson & vonWandruszka, 1994; Guetzloff & Rice, 1994; Wershaw, 1993; Sato *et al*, 1987b; Wershaw, 1994), a characteristic that makes separation by chromatographic means unpredictable and irreproducible unless all conditions with respect to sample concentration, pH, counter ions, injection volume, column regeneration etc are rigorously adhered to (Preuse *et al*, 2000).

The classic methods of chemical analysis used for humic acid are elemental analysis, ash content, the formation of coloured condensation products (a technique hampered by the intense colour of the starting humic acids) and titration of the acid groups and reactive hydrogen atoms.

5.1 Solvent solubility

Strictly speaking humic acids are only soluble in alkaline aqueous medium, as this characteristic is used to define these compounds. However humic acids from different sources have been shown to have sub-fractions that are soluble in organic and even acidic medium (Hayes & Graham, 2000). Using XAD 8 resin it was possible to remove amino acids and neutral sugars from a solution of humic acid in DMSO containing 1% HCl. This result implied that non-humic compounds are co-precipitated or adsorbed to the humic acids during the isolation procedure and that if conditions are chosen that promote adsorption of the humic acid to a non-polar resin, then the polar non-humic compounds can be removed (Hausler & Hayes, 1996). An extraction procedure where the humic acid solution is passed through XAD 8 and XAD 4 resins used in tandem was originally proposed by Thurman and Malcolm (1981). Six sub-fractions were isolated with this procedure with the neutral hydrophobic and hydrophilic compounds (adsorbed onto the XAD 8 and XAD 4 columns respectively) requiring Soxhlet extraction using ethanol or acetonitrile to remove them from the resins.

Another anomaly is that all the reported mild oxidation and reduction reactions have released products of humic acids that are soluble in organic solvents and that have been found to be small

molecules (up to 35% of the original mass of humic acids) (Burges, 1963; Burges *et al*, 1964; Lehtonen *et al*, 2001; Lobbes *et al*, 1999).

Strictly speaking the organic solvent soluble fractions should not be classified as humic acid but it appears as though these fractions occur in most humic acid samples, even though hydrolysis or oxidative or reductive reactions are sometimes required to release them.

In this investigation, where active compounds are being sought, use is made of the classic acid precipitation method followed by organic solvent fractionation to release the low concentrations of compounds from within the humic acid complex and to facilitate bioassays for anti-inflammatory and cellular adhesion molecule expression.

5.2 Chromatography

5.2.1 Thin Layer Chromatography (TLC)

Not many reports of TLC separations of humic acids have been published, as these acids tend to remain on or very close to the origin and streak extensively in almost all mobile phase combinations. They also tend to separate into only three or four distinct bands of dark colour. There are no visualising agents reported for humic acids as such, therefore the natural dark colour is used to identify the position of the compounds that have moved. However the position of several bands separated from humic acid can be visualized with iodine vapour or their fluorescence under long wavelength ultra-violet light.

Khairy (1980) used various thin layer media and found silica gel 60 best as stationary phase combined with a very polar mobile phase consisting of 70% of a concentrated (25%) ammonium hydroxide solution and 30% n-propanol to achieved preparative TLC separations of humic acids obtained from faeces. Generally 3 bands were obtained. In the above method, the humic acid was first extensively and sequentially pre-extracted with diethyl ether, acetone, ethanol and dioxane in a Soxhlet apparatus to remove "non humic" substances. Water and dimethyl formamide were used further to extract polar non-humic compounds from the "defatted" samples and the small amounts of humic acid that dissolved in these two solvents were recovered by preparative TLC or acid precipitation. This extensive extraction procedure removed any lipophilic and polar substances that were loosely bound to the humic acids.

5.2.2 High Performance Liquid Chromatography (HPLC)

Several HPLC techniques or modes have been applied to the separation and quantification of humic acids. Due to the complexity of the humic acids that affects the spectroscopic characteristics and the low solubility of humic acids in organic solvents there has been limited success in the fractionation of humic acids with most of the common techniques used. As humic acids are only soluble in alkaline aqueous media or highly polar organic solvents, the use of unbonded normal phase stationary phases is precluded. The successful modes of separation used for humic acid separations have included size exclusion chromatography (SEC), reverse phase chromatography and ion exchange chromatography.

SEC on HPLC has been used in an attempt to separate humic acids into different molecular mass ranges so that further analysis and characterisation can be done on individual compounds (Chin *et al*, 1994; Piccolo *et al*, 2000; Rausa *et al*, 1991; Saito & Hayano, 1979). Even using HPLC SEC, the resolution of the compounds has not been as successful as would be expected from a sample consisting of a simple mixture of compounds, leading to the conclusion that humic acids must be non covalently-bound polymer-like associations of molecules (Piccolo & Conte, 2000). It is interesting to note that the same pattern of separation has been reported when using low-pressure gel filtration or size exclusion chromatography techniques using various types of media e.g. Sephadex, Sephacryl, Biogel P150. The reported molecular mass for humic acids has varied from as low as 1.7×10^3 D to as high as 1.36×10^6 D. The larger mass fractions can be reversibly separated into smaller fractions by various treatments, especially treatment with organic acids (Piccolo & Conte, 2000; Conte & Piccolo, 1999). Several publications have pointed out that humic acids are surface-active compounds and that they form micelles above a critical concentration that could easily be reached on a column during separation (Engebretson & vonWandruszka, 1994; Guetzloff & Rice, 1994; Sato *et al*, 1987b; Wershaw, 1993; Wershaw, 1994). This characteristic alone would make molecular size dependent chromatographic separations very difficult and unpredictable.

Another characteristic that would affect chromatographic separation is that the solubility of humic acids appears to change in the presence of certain ions, especially the divalent and trivalent cations, and that the solubility is particularly sensitive to pH variations.

Another HPLC technique used is reverse phase chromatography. This mode of separation depends on an adsorption equilibration of lipophilic compounds onto the aliphatic 18-carbon chains bonded to the surface of the silica particles (Engelhardt, 1986). The silica particles have a fairly open and porous structure and the larger the effective pore size, the larger the molecules that can interact with

the inner aliphatic carbon groups of these particles. Several studies with fulvic and humic acids of different origin resulted in reported sample recoveries from reverse phase columns as low as 64% to as high as 300% (Woelki *et al*, 1997; Gremm *et al*, 1991; Frimmel *et al*, 1992).

A study done to determine the effect of the stationary phase pore size on the chromatograms was done and it was found that a pore size of 1000Å was ideal (Woelki *et al*, 1997). Pores of this dimension were found to exclude only molecules with a mass in excess of 2×10^6 D. Recovery of the injected sample was initially lower than with a smaller size pore stationary phase but by using a gradient and increased flow rate recoveries of ca. 90% could be achieved. An interesting observation in this study was that well resolved fractions that were collected, re-concentrated and re-chromatographed still exhibited most of the peaks found in the original sample separated, although the ratios of the peaks did vary in the new chromatograms.

It was also found that the concentration and volume of injection would affect the chromatogram (peak distribution), total recovery and the peak areas of eluted peaks (Preuse *et al*, 2000). The peaks were generally broad and unresolved, an indication that there are several very closely related compounds in the humic acid. Significant changes were seen in the chromatograms as the total amount of humic acid injected was increased. If the injected amount of humic acid was kept constant by varying the concentration and the volume (to accommodate the dilution) of the injected sample there were still major changes in the chromatogram.

Susic and Boto (1989) presented a method for quantitative analysis of humic acids using reverse phase chromatography and the fluorescent characteristic of the humic acids. In the development of their method, various bonded normal and reverse phase stationary phases were investigated as media for the separation. Amine and anion exchange media retained the humic acids completely. However bonded diol, cyano (normal phases) and phenyl (reverse phase) phases all resulted in only one early eluting peak with the peak width and shape dependant on the mobile phase used to elute the humic acid. In their method development, use of an octadecyl silyl (ODS) reverse phase stationary phase resulted in poor humic acid recovery from injected samples, even if the stationary phase was endcapped to prevent irreversible binding of the humic acids to exposed polar groups within the silica. Use of 0.002% or higher concentrations of ammonium hydroxide as well as the use of glass-lined columns was claimed to prevent this recovery problem. Sample cleanup or pre-concentration on ODS solid phase extraction cartridges resulted in more anomalies being observed. It appeared that there was a slow dissociation-equilibrium taking place within the ODS-SPE- retained humic acid, and that additional retained humic acid could be eluted from the SPE column after it had been

left long enough for the equilibration to be re-established, despite the fact that the column had already been exhaustively washed with the same eluting solvent.

A reverse phase HPLC separation of ether-extractable copper oxide oxidation-products of Swannee river and marine humic acids revealed the presence of 11 of the main lignic phenols at concentrations of 50 – 100 nmoles (Lobbes *et al*, 1999). Interestingly a comparison of the total lignic phenol concentrations as determined by GC/MS and HPLC showed a variation of less than 15% but the individual peak concentrations varied by much larger margins.

5.3 Ultraviolet/visible absorbance spectroscopy

Spectrophotometric and colorimetric analysis has often been carried out on humic acids as they are non-destructive and non-intrusive technique and can easily be repeated on the same sample under numerous slightly different conditions where pH or metal ion or the effect of added organic marker compound concentrations are varied. The use of dilute humic acid solutions for analysis by UV/visible spectrometry gave a general tendency, irrespective of the source of the humic acids: the absorption is intense at low UV wavelengths and decreases logarithmically with increasing wavelength. This effect, where a extended decreasing wavelength-dependant "absorption tail" exists can be described as an Urbach phenomenon (Urbach, 1953), caused by a structurally or thermally disordered system comprising numerous different wavelength absorption sites. Some humic acids do show broad low intensity absorption bands superimposed on the general absorption trend described above. These bands have been ascribed to contaminants and "immature humic acids" (Sato & Kumada, 1967).

Several absorption ratios have been used to give an indication of the degree of aromaticity of the humic acid under investigation although it must be pointed out that there has been no definite proof of this assumption. One advantage of using absorbance ratios is that it gives a characteristic that is independent of the concentration of the humic acid under investigation. The E4/E6 ratio (the ratio of absorbance at 465nm divided by the absorbance measured at 665 nm) (Welte, 1957) has become the most commonly used ratio in this regard and values of between 2.8 and 4.7 are fairly common for soil humic acids. Unfortunately no structural information can be obtained from these ratios.

5.4 Infrared spectroscopy

Fourier transform infrared spectroscopy is a rapid and convenient method of gaining information about the chemical structure of compounds. The chemical bonds connecting any two atoms in a

molecule are flexible and subject to various movements; stretch, twist, distortion and deformation when irradiated with energy that matches the energy required to allow the electrons involved in the bond to be excited to a high energy state (Gunzler & Gremlich, 2002). Each type of chemical bond has a characteristic combination of infrared absorption frequencies. By interpreting the absorption (or more commonly the transmission) spectrum it can be concluded that certain chemical bonds exist in a molecule being analysed e.g. aliphatic chain, aromatic rings, carbonyl groups and hydroxyl groups.

Due to the chemical complexity of humic acids and the high carbon content the spectra obtained are not as sharp and discrete as would be expected from clean chemicals. However infrared spectra of humic acids extracted from various sources have all shown most of the following chemical functional groups (with the typical IR absorption bands indicating their presence in brackets): free hydroxyl groups (3420 cm^{-1}), conjugated double bonds or substituted aromatic rings (3060 - 3050 cm^{-1}), aliphatic groups (CH stretch bands for CH_3: 2962 & 2872 cm^{-1}, and for CH_2: 2926 & 2853 cm^{-1}), carbonyl groups (1720 – 1710 cm^{-1}), aromatic rings (1630 cm^{-1}), carboxylate salts (1400 cm^{-1}) aliphatic asymmetric deformation (CH_2: 1465 - 1450, CH_3: 1380 - 1370 cm^{-1}) νC-O of acids, ethers or alcohols (1100 – 1000 cm^{-1}) and the presence of minerals or metal ions (1100 – 1000 & 525, 465, 415 cm^{-1}). This data was collected from Theng *et al.* (Theng *et al*, 1966; Theng & Posner, 1967), Khairy (1980), Tacacz *et al.* (Tacacz & Alberts, 1999) and (Pouchert, 1981).

5.5 Fluorescence spectrophotometry

Fluorescence is a phenomenon generally exhibited by humic acids and was originally believed to be due to the humic acid structure itself. Using this fluorescence characteristic, quantitation of humic acids has been determined using HPLC to separate the humic acid from other contaminants (Susic & Boto, 1989). This technique for quantitation of humic acids is however dubious as it has subsequently been proved that the fluorescence is caused by a small isolatable fraction of the humic acids and that this fraction can easily be separated from the bulk of the humic acids by size exclusion chromatography (Aoyama, 1999). Several workers have used fluorescence spectrometry to characterize humic acids (Alberts *et al*, 2000; Aoyama *et al*, 2000; Mobed *et al*, 1996; Spark & Swift, 1994; Tacacz & Alberts, 1999; Senesi *et al*, 1991) including the generation of total luminescence (3-D fluorescence) spectra. Determination of binding-coefficients of various fluorescent quenching compounds (Coolidge & Ryan, 2000; Engebretson & von Wandruszka, 1994) has also been used to determine the binding capacity of humic acids for various compounds. Humic acids have often been reported to have two discrete fluorescent bands and these vary depending on the source of the humic acid. A soil humic acid exhibited λ_{ex} 380 and 460nm which resulted in λ_{em} of

460 – 470nm and 510 – 530nm (Aoyama, 1999) while peat humic acids had a λ_{ex} 335nm resulting in λ_{em} 466 – 480nm with a small band at 512nm (Tacacz & Alberts, 1999). Humic substances from a Norwegian lake also revealed two fluorescent peaks; one with λ_{ex} 330 – 335nm and λ_{em} 437 – 480nm and the other with λ_{ex} 225nm and λ_{em} 426 – 428nm, a result claimed to be in common with humic acids from oceans, estuaries and rivers (Alberts *et al*, 2000). Suwannee River humic substances (a marsh drainage river) had λ_{ex} at 366nm and λ_{em} 450 – 465nm. These wavelength combinations however, appear to be dependant on the apparent molecular size of the humic acids, with the smallest molecules contributing most to the fluorescence.

5.6 Molecular mass determination

Indirect molecular mass measurement using size exclusion chromatography (von Wandruszka *et al*, 1999; Obenaus *et al*, 1966; Ouatmane *et al*, 2000; Piccolo & Conte, 2000; Rausa, 1994), ultra-filtration through membranes with select nominal molecular mass cut-off ranges (Trubetskoj *et al*, 1997), various electrophoretic methods (Duxbury, 1989) including poly-acrylamide gel electrophoresis (PAGE) (Trubetskoj *et al*, 1997; Trubetskoj *et al*, 1994; Trubetskoj *et al*, 1998) and capillary electrophoresis (DeNobili *et al*, 1998; Pokorna *et al*, 1999), and analytical ultra-centrifugation (Cameron *et al*, 1972; Flaig & Beutelspacher, 1968; Stevenson *et al*, 1953) has resulted in large variations of molecular mass being reported, with each mass range apparently having slightly different properties. When isolated, the properties of humic acids from the different mass ranges have generally been observed to revert back to the properties of the original humic acid with respect to molecular mass distribution if given time to re-equilibrate at high pH (Piccolo & Conte, 2000). This would indicate that there is an apparent dynamic complexation process or "supramolecular association" at work which could explain why it has been difficult to analyse humic acids by the classical chemical analysis techniques such as elemental analysis, colour analysis, solubility, total acidity, melting points, formation of specific derivatives and chemical degradation studies.

Mass spectrometry (MS) is a powerful technique used to determine the structure of compounds by identifying the mass of the whole molecule (the mother ion) and several sub-units ("daughter ions") of the compound, each dependant on the chemical functionality within the molecule. Pyrolysis-MS, where a sample to be analysed is heated until decomposition begins and the volatile products of the decomposition are directed into the MS for mass determination, was the first technique used to analyse humic acids by MS. The problem with this type of analysis is that although structural information is obtained for the compounds entering the MS, the pre-analytical thermal

decomposition often gave rise to new compounds that did not exist in the original sample and only a fingerprint of the original compound could be obtained.

Although structural detail of a molecule can be identified for single compounds when using electron impact ionisation, use of "soft" ionisation techniques has allowed determination of the masses of several compounds in mixtures without the requirement for separating the individual compounds from each other, but these mixtures have generally been fairly well defined or the compound of interest was known to be in the mixture. Many researchers have attempted direct molecular mass determinations of the compounds making up the humic acids using mass spectroscopy. Different "soft" ionisation techniques used include fast atom bombardment (FAB) on various matrixes (Brown *et al*, 1998), electrospray ionisation (ESI) (Brown & Rice, 2000) and matrix assisted laser desorption ionisation time-of-flight (MALDI-TOF), (Pokorna *et al*, 1999; Haberhauer *et al*, 1999; Brown *et al*, 1998). Most of these studies revealed that the molecular mass of the majority of components in humic acid appear to have molecular masses of less than 1000 Dalton. Some researchers have found high mass compounds in humic acids from various sources but these results are not always reproducible, even with the same sample and same sample preparation (Haberhauer *et al*, 1999) and may indicate unpredictable ionisation or secondary effects in the separation chamber. The high occurrence of low molecular mass molecules could be due to impurities that reduce the ion yields or intramolecular dispersion involving transfer of the ionic charge in larger molecules causing them to loose their apparent charge. It may also be that humic acids are in fact loose complexes made up from only small molecules.

A major complication with regard to the molecular mass data collected so far is the different sources and pre-treatments of the humic acid samples used. This makes comparing the data from different laboratories very difficult.

5.7 Nuclear magnetic resonance spectroscopy

Nuclear magnetic resonance spectroscopy (both proton 1H and carbon 13C) (Wershaw *et al*, 1998b; Wershaw *et al*, 1998a; Wershaw, 2000; Preston, 1996) has been fairly successful in determining structural detail, bearing in mind that the humic acids are complex combinations of different molecules in different ratios.

Solid state (i.e. no solvent added) ^{1}H- and ^{13}C-NMR of humic acids using the cross polarization magic angle spin (CP-MAS) technique has been used extensively and is claimed to give an indication of the ratio of aliphatic, anomeric and aromatic carbon atoms in the structure (Wershaw *et*

al, 1998b; Wershaw *et al*, 1998a; Preston, 1996; Mao *et al*, 1998). Liquid phase NMR where the sample is dissolved in a deuterated solvent reveals much more detail due to narrow radio frequency absorption bands but the baseline noise increases and the effect of the "inactive" carbons with long relaxation times is more obvious. It should be remembered that humic acid may be a complex mixture of many compounds with similar chemical structures and that there is a high ratio of inactive nuclei, so the overlap of peaks and distorted peak area ratios are to be expected which results in broader than expected peaks. Typical NMR spectra of humic acids look like distant mountain peaks rather than the sharp peaks of pure synthesized compounds. One problem of liquid phase NMR of humic acids is that humic acid is soluble in very few solvents other than deuterated water and that there are many active hydrogen sites within the humic acid structure that exchange for deuterium in the solvent thereby creating a large water or hydroxyl peak in the spectrum with simultaneous loss of these active hydrogens from the humic acid molecules.

5.8 Other spectroscopic techniques used to analyse humic acids

Raman spectroscopy, which is complementary to infrared spectroscopy, has often revealed graphite-like or amorphous carbon as major constituents of humic acids. Unfortunately not much useful information has been gleaned from Raman spectra. The intrinsic fluorescence of humic acids has also been found to interfere with Raman spectroscopy when using lasers for excitation.

Electron spin resonance (ESR) spectroscopy gives information with respect to the occurrence of free radicals in the humic acid molecules or in humic acid metal complexes and was apparently used in the original confirmation of the presence of quinone-hydroquinone radicals in humic acids.

X-ray absorption spectroscopy (Frenkel & Korshin, 1999; 2000) has been used to determine the type of complexation that takes place between humic acids and various metal ions or salts. Information with respect to electron structure, bonding geometry and neighbouring atoms can be gained from extremely small samples making this a powerful technique where very small areas of the humic acid are to be analysed.

6 Role of humic acid in soil

Even though the humic acid content of mineral soils is generally less than 5% it is an essential component of fertile soil and required for crop growth and agricultural success. For more than 150 years humic acid has been recognized as a plant growth-promoting component in soil and this was

the reason for much of the early research into humic acids, especially their chemical composition and physical properties (Sprengel, 1826).

Humic acids show benefits in the agricultural food production (Faust, 1996) with several presentations and papers documenting the growth stimulating and crop enhancement of up to 30% by plants grown in humic acid enriched soils. Many articles published on the subject of plant growth stimulation by humic acid was reviewed by Vaughan and Malcolm (1985), which points out the positive effect of humic substances on plant growth in general, with an increase in the growth at almost every stage of plant development evident. The growth parameter that appears to be most affected is an increase of the wet and dry mass of the plants (Vaughan & Linehan, 1976; Sladky & Tichy, 1959) attributed to an increase in the length of both the roots and the stems (Vaughan & Linehan, 1976; Tan & Nogamornbodi, 1979; Sladky & Tichy, 1959; Rauthan & Schnitzer, 1981; Malik & Azam, 1985; Furter *et al*, 1996; Furter *et al*, 1997) as well as an increase in number of leaves and flowers (Rauthan & Schnitzer, 1981), all of which manifest as a positive effect on crop yields (Varshney & Gaur, 1974). The mechanism of action has been attributed to many different characteristics and wild assumptions have sometimes been made.

Humic acids are reported to alter soil properties by making them more friable, by buffering soil pH, by allowing soils to swell with moisture retention, by complexing trace metals and making them more readily available to plants, and by releasing bound nutrients such as phosphates from clays (Visser, 1986; Day *et al*, 2000; Clapp *et al*, 1998; Chen *et al*, 1999).

Humic acids contribute to the formation and stabilization of soil aggregates, as well as adding to the nutritional value of the soil. They increase the water retention capacity, decrease soil density, control retention and release of micro and macro-nutrients and are involved in the carbon, nitrogen, phosphorus and sulphur cycles (Kononova, 1966). They act as buffers that regulate pH and bind many metal ions (Guminski *et al*, 1983; Vaughan & McDonald, 1971) in such a way that they are easily available for uptake by the roots of plants.

Within the plant tissues, metabolic cycles appear to be modified by humic acid and were demonstrated to occur both *in vitro* and *in vivo* (Vaughan & Malcolm, 1985; Vaughan *et al*, 1985; Visser, 1987). Samson and Vaughan (Samson & Vaughan, 1989) indicated that these metabolic changes may have been due to changes in the membrane porosity although an earlier study by the same group had implied that enhancement of the iron transport mechanism could have been the reason for the metabolic changes seen (Vaughan & McDonald, 1971).

Some researchers have found hormone like activity, although this is small (±100 times less) when compared to that of known plant hormones such as indoleacetic acid (Cacco & DellÁgnola, 1984; Chen *et al*, 1999).

At present there is a tendency for the agricultural sector to use synthetic fertilizers that are rich in nitrogen and phosphorus but public pressure is growing to use organic fertilizers such as humic acid (Faust, 1996).

7 Therapeutic and medicinal use of humic substances

Therapeutic and medicinal use of humic substances was already known in the Babylonian times, and during the Roman Empire mud baths were used to treat a number of ailments (Priegnitz, 1986). Humic acids were also used as folk remedies for a wide variety of ailments (Lotosh, 1991). "Shilajit", "asphaltum", "vegetable asphalt" or "mumie" are names used for humic substance found in the Himalayan and Caucasus mountains and has been used for centuries to treat numerous ailments and improve the immune system. A recent review article describes the many applications for which these compounds have been used (Schepetkin *et al*, 2002). Among the many indications for which these humic acid type compounds are used include: burns, injuries, bone fractures, dislocations, disease of the skin, neuralgia, arthritis, poisoning and as an anti-inflammatory, antibacterial, anticancer, a diuretic and an immune stimulating agent. Other diverse conditions claimed to respond to these compounds are diabetes, cholesterolemia, eczema, amnesia, epilepsy, asthma, dysmenorrhoea and digestive disorders - including ulcers.

Extranit®, Kalumat®, Kalumin®, Salhumin® and Sulumin® are examples of humic acid products that have been used for either humic acid baths or as topically applied gels. TPP® (Torf peat preparation), Torfot®, Humisol®, FiBS®, Shilagen®, Cystone®, Rumalaya® and Geriforte® are humic or fulvic acid containing medications manufactured for oral administration. Many other medications that are combinations of humic and herbal extracts are sold worldwide to treat many of the ailments mentioned above. Germany, Hungary, Poland, Russia and India appear to be the countries where most of these humic acid-containing medications are formulated.

In Europe, in the 19th century, many of the health spas extensively used peat mud baths especially for rheumatic and gynaecological conditions (Baatz, 1988; Kleinschmidt, 1988; Kovarik, 1988; Lent, 1988). Extracts, tinctures and infusions of peat were also given as a tonic for liver and gastric ailments (Kallus, 1964). This implies that humic acids were deemed safe for external as well as

internal applications as long as 150 years ago, which correlates with traditional uses of mumie in the East.

Peat from the Belgian town of Spa was used during the First World War (before penicillin was discovered) to treat wounds and amputations in field hospitals (van Beneden, 1971) The peat was applied directly to the wounds to prevent infections, relieve pain and facilitate healing (Haanel, 1924). In an experiment with rabbits, Biber and Bogolyubova (1952) reported faster wound healing in the rabbits injected with humic acid. Tazhimamentov *et al.* (1987) reported decreased suppurative wound healing times when humic acids were administered. Salz (1974) offered some thoughts on why humic acid could accelerate wound-healing rates. This author highlighted the bacteriostatic action, anti-inflammatory nature and steroid concentration altering characteristics of humic acids as well as an increased blood flow to the skin. Ghosal *et al.* (1988) and Rajic *et al.* (2001) eluded to the strong anti-inflammatory properties of 4'-methoxy-6-carbomethoxybiphenyl and the tirucallane type triterpenoids respectively, both of which are found in Shilajit, the Himalayan humic acid used traditionally as an anti-inflammatory remedy.

Experimental bone fractures demonstrated accelerated osteoid formation and mineralization if the humic acids were administered during the first week after the fracture, however if treatment was delayed until the second week, osteoid mineralization was reduced significantly (Tkachenco *et al*, 1979). Mumie extracts have also been reported to show positive effects on bone regeneration of fractures in children (Kelginbaev *et al*, 1973).

Adjuvant-induced arthritis in rats was suppressed by doses of crude mumie extract at concentrations of 50mg/kg (Goel *et al*, 1990). Carrageenan-induced foot oedema was also inhibited at the same concentration (Goel *et al*, 1990). Arthritis patients that were treated with humic acid baths showed decreased serum albumin and increased amino acid, globulin and properdin concentrations. The urinary corticosteroid concentrations were also found to be elevated (Reichert, 1966; Miehlke & Thürigen, 1961; Hiller, 1952; Hiller, 1953b) after humic acid bath treatments. It was noted that these elevated urinary corticosteroid concentrations returned to normal after the treatment regimen ended, implying that the changes were directly due to the treatment (Hartman, 1967). Eichelsdörfer (1976) pointed out that the mechanism of humic acid baths is still unknown and that several workers doubt that the benefits are due to the humic acid. However several commercially available humic acid based medications have been successfully and widely used for the treatment of arthritis and rheumatism.

Other balneotherapy combinations that have shown positive effects on rheumatoid arthritis patients consisted of mixtures of salicylic acid and humic acid in various ratios (2.5% humic acid and 47% salicylic acid) (Brandt, 1964; Miehlke & Thürigen, 1961). Some researchers therefore claimed that the positive effect has been due to the salicylic acid rather than the humic acid. The humic acid bath treatment is usually in the order of a 20-minute exposure to a bath at 38°C three times a week for up to 6 weeks.

Gastrohumit®, a bismuth salt and Huminit®, a calcium salt of humic acid are commercially available and used for the treatment of heartburn, gastric ulcers and acute gastroenteritis (Brandt, 1964; Kinzlmeier, 1954; Reichert, 1966; Schlepper, 1960; Weithaler, 1954). The calcium in the Huminit® had an immediate alkylating effect in the stomach but systemic alkylosis was avoided due to the buffering capacity of the humic acid. The humic acid was reputedly bound to the mucus in the stomach, leaching out slowly to give long lasting protection as opposed to the newer generation ion exchange type medication that gave only a few hours protection.

Gastrointestinal applications of humic acids in the veterinary discipline were very successful and included: treatment of enteritis in calves (Kühnert *et al*, 1980) and in piglets (Golbs & Kühnert, 1986; Gramsch, 1961), while successful treatment of diarrhoea, gastroenteritis, colitis and parvovirus was reported in dogs and cats (Bartels, 1986). None of these authors indicated any severe side effects, even after prolonged use, and found the humic acids to be as effective as commonly prescribed drugs. These effects have been ascribed to the toxin absorbing ability, anti-inflammatory, bacteriostatic and immune stimulating effects of humic acid. Kühnert (1989) also pointed out that there is no detectable residual drug after treatment and as there is negligible allergy incidence and no observed symptoms of overdosing, humic acids can easily be incorporated into the animal feed. This would be especially beneficial in meat producing animals where residual drugs are a problem.

Salz (1974), the developer of the topically applied Salhumin®-gel, successfully treated the following conditions with his product: osteochondritis, torn ligaments, lumbago, hip pain, chronic rheumatoid arthritis, arthrose, vertebrae pain, bruises, muscle spasms, oedema and phlebitis. These conditions are all related to inflammation, therefore it appears as though humic acid has anti-inflammatory properties and that the active components can be absorbed through the skin.
Tolpa Peat Preparation (TPP) a peat-derived product is sold in Eastern Europe as an immune enhancing agent. Studies have shown that TPP inhibits IgE induced anaphylaxis in mice (Wyczolkowska *et al*, 1993), enhances the humoral response to sheep red blood cells (SRBC) in mice in a dose dependant manner (Obminska-Domoradzka *et al*, 1993) and that TPP is effective in

the treatment of recurrent respiratory infections in human volunteers, reportedly due to enhanced granulocyte phagocytosis (Jankowski *et al*, 1993).

Taugner (1963) observed an impressive reduction in oedema in rats after intravenous injection of sodium humate. Klöcking *et al*, (1968) induced oedema experimentally in rats by two different methods and found that a peritoneal injection of humic acid reduced the oedema better than some proven medications to which the humate was being compared.

Salz (1974), Motohisha *et al*, (1974) and Iubitskaia and Ivanov (1999) reported analgesic and anti-inflammatory effects as well as serum lipid modulation and promotion of metabolic processes when using sodium humate in balneotherapy on osteoarthritic patients.

Anti-inflammatory properties for humic acids were reported (Kühnert *et al*, 1982; Ye *et al*, 1985) and gynaecological inflammatory conditions have been successfully treated by a group in Poland (Woyton *et al*, 1993).

Kühnert *et al*, (1982) described the anti-inflammatory action of humic acids as being as effective as that of dimethyl sulfoxide (DMSO), a well-known anti-inflammatory agent and drug carrier. Klöcking *et al*. (1968) ascribed the anti-inflammatory action of humic acid to the polyphenolic structure (based on the lignin derived humic acid theory) while Salz (1974) attributed it to the fact that it stimulated higher blood flow.

Klöcking (1994b) provided a feasible biochemical explanation of the anti-inflammatory properties of humic acids, basing their findings on earlier work by the same group where synthetic humic acids inhibited the lipoxygenase enzyme from rabbit reticulocytes strongly, while prostaglandin H synthase from sheep vesicular glands was only weakly inhibited (Schewe *et al*, 1991). Salts of humic acid and some synthetic humic acids were shown to inhibit the lipoxygenase and prostaglandin H synthase pathways involved in the conversion of arachidonic acid, released after membrane damage, into leukotrienes or prostaglandin H respectively. The lipoxygenase pathway resulting in the production of leukotrienes, lipoxins and other hydroxylated but unsaturated fatty acids, was blocked by low concentrations of the synthetic humic acids but required relatively high concentrations of the natural products to achieve the same inhibition effect. Leukotrienes act as mediators of the inflammatory processes affecting bronchial dilation, vascular permeability, chemotactic leukocyte infiltration, oedema and hypersensitivity reactions (Penrose *et al*, 1999). The prostaglandins are antagonists of the leukotrienes that show almost exactly the opposite effects

despite the fact that the origin of both these classes of potent biologically active C_{20} lipids is arachidonic acid. Natural humic acid appears to have a smaller effect on the prostaglandin synthesis pathway than the synthetic humic acids as approximately 1100µg/ml of natural humic acid was required to elicit a 50% inhibition of prostaglandin synthesis by a kidney homogenate (Breng *et al*, 1981).

It has also been found that the expression of protein type inflammatory mediators is inhibited (Gau *et al*, 2000). Intercellular adhesion molecule-1 (ICAM-1), vascular adhesion molecule-1 (VCAM-1) and E-selectin expression by cultured HUVEC were inhibited after treatment of the cells with humic acid and then stimulated by addition of a lipopolysaccharide (LPS). An immuno-fluorescence technique using fluorescent-labelled monoclonal antibodies against each of the above adhesion molecules was used and quantitation performed by measuring the fluorescent intensity of the cell-bound monoclonal antibody by flow cytometery. The LPS induced expression of all three of these adhesion molecules was inhibited in a dose and contact time dependant manner by the added humic acid. As the nuclear transcription factor kappa B (NF-κB) plays a central regulatory role for all three of these adhesion molecules, the effect of humic acid on the NF-κB activity was determined and found to have a similar dose and contact time dependant response as that of the adhesion molecules.

Inhibition of hyaluronidase activity that is associated with the anti-inflammatory response to humic acid baths is possibly due to an increase in the oestrogen concentrations in the blood (Sprunt *et al*, 1938; Vasterling, 1958; Wattenberg & Glick, 1949). This could either be from "contaminating" oestrogens in the humic acid (Lotmar, 1960; Taugner, 1963) or from stimulation of endogenous oestradiol and oestrone production by the adrenal cortex. These increased hormone concentrations are identified by the increase of up to 100% in urinary oestrogen excretion within 24 hours of balneology in both male and female patients (Hiller, 1952; Hiller, 1953b; Hiller, 1953a).

Human neutrophil functions, especially the respiratory burst that produces reactive oxygen species like hydrogen peroxide, were stimulated by two natural and one synthetic humic acid (Riede *et al*, 1991). There was however no stimulation of chemotaxis or chemokinesis.

Lange *et al.* (1985) found increased leukocyte counts and globulin concentrations in rats after oral administration of a combination of pesticides and humic acids. The increase in globulins was attributed to increased antibody production, a direct effect on the immune system, in response to the humic acid-pesticide complex. In another study on white rats by the same author (Lange *et al*, 1987), different humic acids alone were given as single intra-gastric doses, which resulted in changes in the

blood leukocyte profile. The changes were interpreted as indicating an effect on the immune system. Plasma protein concentrations and phagocytic activity were also analysed.

Soloveyva and Lotosh (1984) treated anaemia and hypercholesterolemia with humic acids and concluded that the action of humic acids could be attributed to a positive effect on liver detoxification, several enzyme systems and the immune system.

The beneficial effects of humic substances have been ascribed to the direct effect of the humic acids due to their surface effects and ability to bind or chelate compounds or metals (Marx & Heumann, 1999; Nifant'eva *et al*, 1999), adsorption of xenobiotics (Nielsen *et al*, 1997; Prosen & Zupancic-Krajl, 2000; Schulze *et al*, 1999) and to its polyacidic and polyphenolic nature (Woelki *et al*, 1997). Interaction with biologically important molecules like proteins, including peptides and enzymes, polysaccharides and lipids, including the steroids and lipophilic hormones, has been proposed as mechanisms of the humic acid action. The effect of humic acid has also been ascribed to the "contaminating" compounds that are bound to, or complexed within, the structure of humic acids and not to the humic acids themselves (van Beneden, 1971).

Studies on the anti-viral properties of humic acid and several synthetic humic acids formed by oxidation of phenolic precursors have revealed some interesting results. Natural humic acid was shown *in vitro* to have antiviral effects against both herpes simplex types 1 and 2 viruses (Klöcking & Sprössig, 1975; Klöcking & Helbig, 1991; Thiel *et al*, 1977; Thiel *et al*, 1981) and against Coxsackie virus (Klöcking & Sprössig, 1972). Synthetic humic acids formed by oxidatitive condensation of chlorogenic or caffeic acids showed even better antiviral properties than the natural humic acids (Eichhorn *et al*, 1984; Hils *et al*, 1986; Klöcking *et al*, 1983; Thiel *et al*, 1984). Other viruses tested for sensitivity to natural and synthetic humic acids are influenza type A and B (with selective sensitivities), cytomegalovirus, vaccinia, adenovirus which all showed sensitivity but poliovirus type 1, parainfluenza virus type 3, retrovirus type 1 and sindbis virus showed no sensitivity to any of the tested humic acids (Neyts *et al*, 1992). Inhibition of viral replication was evident at humic acid concentrations of between 20 and 100ug/ml. A topical application of a 1% ammonium-humate solution was effective against herpes virus infections in 90% of cases.
Several workers have tested HIV-1 and HIV-2 for natural and synthetic humic acid sensitivity. Schols *et al.* (1991) and Schneider *et al.* (1996), both found inhibition that appears to be due to inhibition of viral entry into the target cells, a process that involves the V-3 loop of the virus protein. It is possible that direct binding of the poly-anionic humic acid to critical regions the protein sheath

of the virus prevents cell attachment. Other mechanisms may be that humic acid binds to the viral attachment site on the host cells.

Studies done both *in vitro* and *in vivo* where the ability of humic acid to remove or counteract toxic compounds were studied (Fuchs *et al*, 1982; Fuchs *et al*, 1986; Golbs *et al*, 1984; Lange *et al*, 1985; Rottinghaus, 2000; Solovyva & Lotosh, 1984). Heavy metal, alkaloid, chemical and pesticide toxicity could be counteracted in varying degrees by the oral administration of humic acids. It was found that serum immunoglobulin concentrations increased if humic acids were administered together with the toxic compound (Lange *et al*, 1985) and that the detoxification activity of the liver appeared to have been stimulated (Solovyva & Lotosh, 1984). These authors attributed the *in vivo* antitoxic and prophylactic characteristic of humic acid to the combined effects that they have on the immunoglobulin concentrations and the liver activity.

The mutagenic effect of some well known carcinogenic compounds (benzo[a]pyrene and 3-aminoanthracene) was counteracted by humic acid, but this was found to be a desmutagenic effect where the mutagen was adsorbed onto the humic acid and therefore prevented from reaching the target tissue where it could exert its mutagenic effect. Larger molecular mass fractions of humic acid were found to be more effective in prevention of the mutagenesis in Ames test systems (Sato *et al*, 1986; Sato *et al*, 1987a; Sato *et al*, 1987b; Bernacchi *et al*, 1996).

Mao *et al* (1998) proposed that a possible mechanism of action of the humic acid compounds may have been due to their chemically reducing properties, thereby preventing oxidative radicals from exerting their documented mutagenic and carcinogenic effects.

8 Negative aspects of humic acid exposure

As far back as 1950 humic acid was implicated as a causative agent of goitre of the thyroid gland (Woodward, 1963; Hettche, 1955; Hettche, 1956; Galcenko, 1950; Burkat, 1965). These authors also implicate urochrome (a natural metabolic product of haemoglobin) as being a cooperative compound. The similarity between these compounds is that both humic acid and urochrome can complex copper ions ($Cu^{2}+$) and iodide, which in turn affects thyroxin synthesis and hence the thyroid gland is affected. However, Schierbaum (1966) and Janeček & Chalupa (1969) could not detect goitrogenic activity in thyroid glands of rats kept on a diet high in humic acid over extended periods.

Humic acids have also been implicated in the development of Black-foot Disease, an endemic disease in rural parts of south-western Taiwan where wells are the main source of drinking water. Black-foot disease is a vascular abnormality affecting the peripheral circulatory system to such an extent that the reduced blood flow to the extremities of the sufferer's results in small localized gangrenous areas and ulceration. The condition presents as cold feet and dark discolouring of the skin, resulting in "black feet" (Tseng *et al*, 1961). The well water, from which the implicated humic acids were extracted, also had relatively high arsenic concentrations. The metal complexing/chelating ability of humic acid results in an arsenic-humic acid complex that appears to result in a synergistic deleterious effect by the arsenic and humic acid on the vascular epithelial cells of the capillary walls (Lu *et al*, 1988; Gau *et al*, 2000).

In another study on Black-foot disease, it was demonstrated that tissue plasminogen activator (t-PA) and plasminogen activator inhibitor 1 (PAI-1) production by human umbilical cord arterial epithelial cells (HUVEC) in culture, were both induced by humic acids of natural and synthetic origin. t-PA is a proteolytic enzyme involved in conversion of plasminogen into plasmin, which in turn activates fibrinolysis and has a central role in plasma fluidity and plasma leakage. PAI-1 is normally released from activated platelets and is thought to be important in controlling fibrinolytic activity within thrombi. HUVEC cell cultures also showed changes on the cell surface and inhibition of growth with a simultaneous increase in mRNA concentrations after brief exposure to humic acids (Yang *et al*, 1996). Later studies on the same HUVEC model showed increased intracellular chelatable iron concentrations after humic acid exposure and this could be directly associated with an increased reactive oxygen radical formation. Formation of the oxygen radicals could be inhibited by intracellular antioxidants and intracellular iron chelators, but not by enzyme inhibitors or calcium chelators (Gau *et al*, 2001).

Hseu (2000) investigated the effect of humic acid on non-immune system blood cells in relation to Blackfoot disease. His study found that humic acid induced oxidation of normal membrane-proteins of human erythrocytes and that these were then exchanged for high molecular mass proteins resulting in echinocyte formation. Simultaneously haemoglobin inside the erythrocytes was oxidized increasing the oxidative stress on the cells.

Klöcking (Klöcking, 1994a) investigated the *in vivo* fibrinolytic and coagulation effects of a number of synthetic and of two natural humic acids in rats at doses of between 5 and 10 mg/kg. It is interesting to note that he found a decrease in PAI-1 activity in contrast to the results of Yang *et al*

(1996). However *in vitro* results using the same compounds indicated that some of the synthetic humic acid compounds had strong anticoagulation properties while others showed significantly shortened coagulation times.

Humic acids in water react with active chlorine to form carcinogenic compounds and are therefore seen as a problem in water purification plants throughout the world. Various chlorinated humic acid derivatives have been demonstrated to have mutagenic properties in the Ames test (Bernacchi *et al*, 1996; van Duuren, 1986).

9 Pharmacokinetics of humic acid

For any pharmaceutical compound to have an effect on a target organ or cell within a living organism, it must be absorbed from the point of administration and should be distributed effectively so that the target region, organ or cells become exposed to a therapeutically effective concentration of the pharmaceutical compound (Harvey, 1990).

In humans, the most acceptable and therefore the most common method of introducing a drug is via the enteral route involving oral administration of a dose at defined intervals. An effective uptake or absorption of the compound from the gastrointestinal tract (GIT) must occur before the compound would elicit any effect.

This absorption is dependent on several factors (Sommers, 2000);

- solubility
- lipophilic nature
- molecular size
- degree of ionisation
- pKa of the compound if it carries an ionisable group
- the rate of transit through the gastrointestinal tract
- enzymatic or chemical changes that may occur in the GIT

Once absorbed from the GIT, the compounds need to be effectively distributed to the target organ or cells and this is usually via the blood circulatory system but may initially be transported via the lymphatic system.

Blood is a complex suspension of numerous cell types and biologically active soluble components, including hormones, enzymes and lipids in a protein rich aqueous electrolyte solution. The primary function of blood is that of transport for nutrients, metabolites and respiratory gasses to and from all the cells in the body. Secondary functions are homeostasis (whereby pH, temperature and electrolyte concentrations are kept within approximately constant limits) and defence against foreign compounds and agents (Weiss & Jelkmann, 1989). Any compound that is introduced into the blood (indirectly by absorption from the GIT, subcutaneous or intramuscular injection or directly by intravenous injection) can complex or bind to the plasma proteins, enzymes or cells in the blood or may be transported in the serum as a soluble compound without interaction with any of the blood components (Harvey, 1990).

Many compounds are absorbed from the lumen of the gut by passive permeability, making use of the paracellular transport route, where the absorbed compound enters the interstitial space via the so called tight junctions between the mucosal enterocytes. This passive absorption is facilitated by osmotic, hydrostatic, chemical and electrical gradients but has a "pore" size limit of about 0.8nm in the jejunum reducing continuously down the length of the gut to only 0.25nm in the colon. This passive permeability is selective to some extent with cationic compounds being favoured above anions due to a negative charge of the surface of epithelial cells (Ewe & Karbach, 1989). Both diffusion and convection (i.e. "going with the flow") mechanisms are involved in this type of passive transport.

Some compounds are absorbed by an active transcellular transport mechanism where the compound is initially absorbed from the lumen either by diffusion across the cell membrane or actively sequestered and transferred by carrier proteins or in vesicles.

Some general characteristics of active transport are (Ewe & Karbach, 1989; Harvey, 1990);

- that metabolic energy is utilized in the process
- that inhibitors of metabolism stop the process
- that the process can be saturated, meaning that it can only transfer a limited amount of compound per unit time.
- that it is specific for particular compounds
- that compounds having similar chemical structure to the preferred transport compound can competitively inhibit the process.

- that maximal efficiency is only achieved when a coupled transport mechanism is in operation, meaning that there is an exchange of compounds from either side of the membrane.

Facilitated diffusion is a process where diffusion takes place faster than would be expected from either the concentration or electrical gradient that predominates. It is a form of active transport where only some of the above mentioned active transport criteria are met, such as a carrier protein being used but not metabolic energy.

Mixed transport occurs when there is a combination of mechanisms operating such as when an active transport process occurs at low concentrations of compound but when the process is saturated due to a high concentrations, passive diffusion rate exceeds the active uptake rate.

Pinocytosis is a process of absorption where the compound, or even a particle, is enclosed by an invagination in the cell membrane that then envelops the compound to form a small internalized vesicle. The vesicle then is transported across the cell to the basolateral membrane where it is expelled in the reverse manner to the initial engulfment. Pinocytosis is the mechanism by which intact particles, macromolecules and some proteins are absorbed.

Lipophilic compounds can enter the cell from the lumen by a diffusion mechanism whereas polar and charged (especially anionic) molecules require carrier proteins to assist the transfer. The compounds enter the epithelial cell cytosol (Harvey, 1990) from where they must again be transported across the basolateral membrane and deposited into the interstitial space and finally into the subepithelial capillaries from where they are rapidly transported further. The capillary venules surrounding the G.I.T. all drain into the hepatic portal system consisting of the superior mesenteric vein, inferior mesenteric vein, the gastric veins, the splenic vein and the portal vein (Netter, 2003). Compounds absorbed from the G.I.T. therefore pass through the liver prior to reaching the heart from where they are circulated to the rest of the body.

Absorbed compounds can complex with blood components or may be enzymatically altered even before they reach the liver (Harvey, 1990). One of the common interactions that occur is the binding of ionized and lipophilic compounds to the serum albumin proteins. This effectively reduces the apparent concentration of the compound in the blood, as the compound is no longer free to leave the blood and interact with the target organ. Antibodies and blood cells may also interact with the

absorbed compounds, which is equally deleterious for the therapeutic effect of the compounds for the same reason (Harvey, 1990).

Many non-nutritious compounds are initially altered by the liver where they are enzymatically metabolized by one of the cytochrome P-450 enzymes, an inducible mono-oxygenase enzyme group found in high concentrations in the liver, to form compounds with more polar functionality (Waterman *et al*, 1986). These enzymatically-altered compounds are often altered further by conjugation to highly polar molecules that enhance aqueous solubility and the excretion of these compounds by the kidneys.

The anionic nature of humic acids, the reported large molecular mass (Obenaus *et al*, 1966; Ouatmane *et al*, 2000; Rausa *et al*, 1991) and the fact that humic acids precipitate at low pH (Aiken, 1985; Hayes, 1998; Odén, 1919) are all factors which indicate that the absorption of humic acids would be slow if it occurs at all. There is doubt whether these compounds are in fact absorbed after oral administration and that any absorbed compounds are probably breakdown products of the administered humic acids.

If an *in vivo* therapeutic effect is to be achieved there must be sufficient absorption within a reasonably short time and the elimination half life must be long enough for the compound to reach a therapeutically active concentration in the blood and be able to reach the target organs. It follows that if the cellular immune system is the target of the active compound it is likely that contact with the target cells takes place as soon as the humic acid enters the venous capillaries following absorption from the lumen of the GIT and that a fairly rapid elimination half-life would not completely eliminate the therapeutic effects. Binding of the humic acid to serum protein would however render the concentration of the free active compound(s) very low and a therapeutic concentration of these compounds may then never be achieved.

The fact that physiologically measurable changes in hormone concentrations (Hiller, 1953a; Reichert, 1966) and immune system functions (Lange *et al*, 1985) together with the fact that there is apparent relief of the symptoms of rheumatoid arthritis after sufferers have taken humic acid orally, implies that humic acids or at least some of the physiologically active compounds in the humic acids are in fact absorbed from the gastrointestinal tract. Visser (1986) argued that there is no reason why humic acids could not be absorbed even though they are reported to have such a high molecular mass. Humic acids have been shown to be absorbed by animals after subcutaneous injection, to

circulate in the blood after intravenous administration with some of the humic acids bound to serum proteins (Klöcking *et al*, 1978) and appear to be metabolized in the liver (Visser, 1973).

Obenaus et al., (1965) demonstrated that humic acids bind to serum albumin *in vitro* and that metals like lead and iron increase this binding affinity. Immuno-electrophoretically it was shown that there were no changes in the immune effect of the humic acid bound albumin. Klöcking et al., (1967) demonstrated that this binding takes place *in vivo* in rats with 69% of an intra-cardially injected dose of humic acid bound within 20 minutes of administration.

Although two studies have been done using ^{14}C labelled humic acids, [a synthetic phenolic oxidation product by Lange *et al.*(1996) and a humic acid extracted from a fungus grown using ^{14}C glucose as the only carbon source by Visser (1973)], in an attempt to determine skin penetration depths or organ distribution, there have been very few studies on the pharmacokinetics of humic acids. Very few publications referring to experiments where the absorption of orally administered humic acids was determined, could be found in an extensive literature search. One reference compared different routes of administration of multi layer liposome-encapsulated humic acids on chickens. In these experiments liposomes containing the sodium salt of humic acid were injected subcutaneously, intra-cardially or given orally (Hampl *et al*, 1994). The sodium humate was labelled with ^{125}I prior to encapsulation and the liposomes were deemed to be stable, as 35 days of *in vitro* incubation in chicken serum did not show any liposome degradation (Hampl *et al*, 1992). Bioavailability of sodium humate was calculated as being highest after subcutaneous injection of the free unencapsulated sodium humate. The oral route of dosing resulted in approximately 30% of the dose being detected in the blood after liposome encapsulated humic acid was used. Unencapsulated sodium humate absorption was lower implying that the encapsulation did increase the absorption (Hampl *et al*, 1994). No data were given with regard to areas of accumulation of activity in the birds during the study. Bioavailability of sodium humate was calculated as being highest after subcutaneous injection of the free unencapsulated sodium humate.

Another study determined the penetration of humic acid through human skin and was performed *ex vivo* using synthetic ^{14}C-humic acids (Wohlrab *et al*, 1984). In this study it was found that almost 60% of a 1% humic acid ointment penetrated the keratin layer of the dermis while about 3% penetrated as deep as the epidermis within 30 minutes. It was concluded that therapeutic concentrations could be achieved by topical applications.

In vitro studies of ill defined mixtures are often not realistic as models that can be extrapolated directly to an identical *in vivo* effect due to over- or underestimation of the absorption of the active compounds or because a bias for absorbing only certain compounds exists, or due to selective binding and inactivation of only part of a mixture of compounds. The formulation of an administered dose can also play an important role in the rate and ability to cross membranes, an event that is not usually included in *in vitro* studies. The ability of humic acid to be rapidly absorbed from the gut is very important if it is to be used as an anti-inflammatory or for immuno-modulating therapy.

A baboon study using ^{123}I labelled oxihumate (Dormehl, 1998) and oral dosing revealed that approximately 11% of the activity was absorbed and that distribution appeared to be via the circulatory system. The uptake appeared faster when the dose was delivered directly into the duodenum using gastroscopic delivery. The areas that were reported to show accumulation of radioactivity after administering Iodide-123labelled oxihumate included the thyroid gland (possibly indicating free iodide), salivary glands, and the septum between the nostrils, the liver and the kidneys. Excretion was essentially via the urine and was persistent for up to 48 hours after which time the residual isotope activity became too low to detect. The liver and kidneys would be involved with the metabolism and excretion of the humic acid respectively, and except for the nasal septum these are areas where it could be safely speculated that free iodide would normally accumulate. Apparently none of the lymph nodes or organs of the immune system showed any uptake of activity.

No data could be found in the literature with respect to the normal distribution of free iodide in the baboon but discussions with several pharmacologists and physiologists (personal communications) pointed out that it is generally accepted that any drug that has covalently bound iodide within its structure, would be metabolized so that the iodide atom would be released. The thyroid gland and any other areas where iodide is known to accumulate, such as the salivary glands, eyes, testes etc would then scavenge the released iodide. Support for this theory can be found in several publications where *in vivo* administration of iodide labelled compounds have resulted in thyroid accumulation of iodide (Ercan & Senekowitsch, 1991; Sinn *et al*, 1990; Klett *et al*, 2003; Press *et al*, 1996)

The target organs and rate of elimination can give some idea of which organs and therefore an indication of the mode of action.

All potential drugs must be tested on animal models before human clinical trials may be initiated. The pharmacological parameters that are most often determined using animal models are the

toxicity, pharmacokinetics and the pharmacodynamics. The most common animal model used in the initial phase of testing are the mouse or rat models. Medical and pharmacological research makes extensive use of inbred or genetically deficient strains of mouse models (Festing, 1979), but the results from these models are often not directly applicable to humans. Despite this shortcoming murine models are popular due to availability, known susceptibilities, fast breeding, similar responses to the experimental procedures, ease of transferring techniques between laboratories and the low cost of maintenance compared to most other animal models. Extrapolation of data from rodent models to humans is not without problems due to anatomical, physiological, immunological and genetic differences. As rodent models are often from closely inbred or specific hybrid models they show no or very little variation to induced responses. This "homogenous" response is not always representative of the effect that would be observed in a general human population where heterogeneous responses can be expected.

In contrast the primates are the animals that are closest to humans with respect to physiology (including similarity in the nervous, respiratory, circulatory, digestive, immune and endocrine systems) and gross anatomy. Anatomical similarity is evident although the normal quadruped movement of the primates as opposed to the bipedal motion in the human results in some unique variations in gross anatomy (Swindler & Wood, 1982).

Despite the anatomical differences the baboon model remains the animal model that is closest matched to the human and which is fairly accessible for pharmacokinetic studies.

Some differences in the anatomy of the baboon digestive system exist when comparing the human and baboon. The most obvious difference is that the baboon has a well-developed caecum, whereas the human has an insignificant appendix as the equivalent organ. As the caecum is situated distal to the small intestine where most nutrient and drug absorption is expected to take place, this anatomical difference should have little effect on the pharmacokinetics of most drugs.

9.1 Iodide-123

Iodide is the most widely used radioactive label for proteins and phenolic type compounds that are used as tracers in radioimmunoassays and ligand derivatives for receptor binding studies. The most common iodide labelling reaction is a two-step process that involves *in situ* oxidation of a radioisotopic iodide salt to form molecular iodine that in turn reacts with phenolic groups (e.g. tyrosine and histidine residues in proteins) or other conjugated double bonds. This oxidation can be done chemically using chloramine-T or enzymatically using lactoperoxidase and hydrogen peroxide.

The latter technique has been used to label cell surface markers and due to the relatively mild reaction conditions, biological activity of the labelled surface markers was apparently not affected (Morrison, 1980). The labelled products are generally chemically stable. The oxidation of thiols and thio-ethers that can take place during the oxidative conversion of the starting iodide to iodine can be avoided by the use of a very mild biologically friendly acylation reaction that avoids this problem (Bolton & Hunter, 1973).

Iodine-123 is a cyclotron-generated gamma-emitting radioisotope of iodine with a half-life of 13.2 hours. This isotope has a single gamma radiation emission with an energy of 159 keV resulting from an electron capture event. These characteristics make ^{123}I a nearly ideal medical and pharmacokinetic radiotracer isotope. ^{123}I is commonly used for thyroid function and imaging studies in the free iodide form (Martindale, 1993), for renal imaging and perfusion studies as iodohippurate (Jorgensen & Ladefoged, 1987) as well as for cerebral imaging as iofetamine (Cohen *et al*, 1988).

Oxihumate (see below) has many available phenolic and activated aromatic rings, permitting successful electrophilic substitution by molecular iodine under mild conditions. The iodide labelled oxihumate prepared from iodide-123 and chloramine-T (Hunter & Greenwood, 1962) results in a relatively efficient labelling of numerous compounds in the oxihumate, but these labelled compounds cannot be rapidly isolated as individual labelled compounds by HPLC because of its complex chemical composition, its micelle forming characteristics and its low solubility in organic solvents. The alkaline salts are reasonably soluble in water although they precipitate quickly when lowering the pH of the solution to below pH 2.0 by addition of acids.

An earlier baboon study using labelled oxihumate did not identify any organs that could be associated with the immune system and only showed up the organs that would have been expected to accumulate free iodide. In spite of the typical iodide distribution being seen in the previous study (Dormehl, 1998), no attempt was made to analyse the labelled humic acid for free iodide but assumed that any residual free iodide would have been removed from the reaction mixture by the extensive washing procedure used.

One of the objectives of the present study was to repeat an earlier study where radiolabelled oxihumate was administered to baboons per os. The prime objective was to ascertain whether absorption of humic acid compounds did in fact take place from the gastro-intestinal tract. Additional data that was to be collected was kinetic data on the uptake and elimination of the

labelled compounds, to find possible immune system related target organs and to analyse the urine to find excreted labelled compounds.

10 The immune response

10.1 Overview

Immunity is the ability to ward off or protect the host against invading microorganisms or pathologic agents. The immune system is a complex combination of numerous cell types and bio-molecules, distributed throughout the host, especially in areas where contact with pathogens is high. Each component of the immune system, cellular or molecular, has a specialized role in this host defence. The immune system can be divided into two broad classes: an innate immunity involving immediate response and inflammatory processes and an acquired adaptive response that is antigen specific and involves immunological memory against a particular target antigen.

The innate response is the first line defence against exogenous invasions, especially against microbial assault, but it is also forms an integral part in the initiation of the adaptive response that follows a particular assault. The innate response is generally a localized response to an assault by microorganisms or foreign particles and involves the complement system of the blood, target independent phagocytic cells like the neutrophils also known as the polymorphonuclear leukocytes (PMNL), the macrophages (MØ) and their precursors the monocytes (MO) as well as non-phagocytising natural killer cells (NK) and the mast cells. All these cells can release inflammatory mediators and cytokines during the phagocytosis or when otherwise activated.

The adaptive response takes longer to be initiated and makes use of the T- and B- lymphocytes, cells which need to mature and proliferate in response to specialized antigen presenting cells (APC). The APC presents a processed antigen, from the invading agent, attached to the major histocompatibility complex class II molecules (MHC II) together with other stimulatory signalling molecules to immunocompetent Th 0 lymphocytes, which in turn proceeds through a series of differentiation steps depending on the cytokines released by the APC and accessory cells. This T-lymphocyte maturation results in either a cell mediated immunity (occasionally accompanied by a delayed type hypersensitivity) as a result of a Th 1 mediated response or a humoral immunity resulting from Th 2 mediated stimulation of B-lymphocytes to secrete immunoglobulins (Ig) which then bind to the exogenous microorganism causing disruption of cellular activity with suppression of microorganism growth and increases the recognition of the Ig-microorganism complex as being a foreign threat by the innate immune system.

10.2 Phagocytic cells of the innate immune system

10.2.1 Polymorphonuclear leukocytes (PMNL)

These are the most populous of the circulating leukocytes but also the shortest-lived cells in circulation. They are produced and mature in the marrow and are released into circulation where they spend as long as a week if they do not become stimulated. Stimulation results in the PMNL migrating by diapedesis (a sequence of attachment to the luminal vascular endothelium followed by a rapid spreading and flattening of the cells onto the endothelium followed by squeezing through the endothelium into the interstitial space) into the tissue where the stimulus originated. They are produced in the bone marrow in response to acute stress irrespective of where the stimulus originates, whether from infection, trauma, noxious stimuli, emotional stress, infarction or otherwise.

PMNL form the first line defence against any acute microorganism assault and is the first leukocyte to migrate (by means of amoeboid movement) into a lesion or region of infection. This migration is a well-coordinated chain of events involving chemotactic molecules (IL-8, PAF, LTB_4), adhesion molecules (β_2-integrins, ICAM-1, ICAM-2, the L-, P-, & E-selectins), the vascular endothelium, and several cytokines (IL-1, IL-8, TNF-α, GM-CSF). Many different receptors and molecules are involved, some with multiple functions e.g. L-selectin and CR3 on the PMNL and P- and E-selectins on the vascular endothelium (Witko-Sarsat *et al*, 2000). PMNL binding of endothelial P-selectin increases β_2-integrin mediated adhesion (via increased CR3 molecule expression) and also stimulates production of reactive oxygen species (ROS) (Ruchaud-Sparagano *et al*, 2000). Almost all the chemotactic molecules, phagocytic stimuli, activated complement molecules (C5a) and several cytokines (TNF-α and GM-CSF) stimulate expression of high levels of CR3 adhesion molecules by PMNL (Stewart & Hogg, 1996).

The functions of PMNL include phagocytosis of microorganisms, cell debris and denatured proteins, the release of proteolytic enzymes and highly reactive oxidant species (ROS) as well as the synthesis of cytokines, chemokines, and inflammation mediating lipid metabolites. Phagocytosis by PMNL always results in self-destruction, but this is a protective mechanism for the host. As the proteolytic enzymes and highly reactive oxygen species used by the PMNL to destroy the foreign particles cannot distinguish between target and host tissue, the PMNL engulfs the target particle and performs the destructive reaction in an intracellular phagosome thereby protecting the surrounding host tissue at its own expense.

Several cytokines are produced by PMNL, which appear to play an autocrine regulatory role on PMNLs. The main cytokines expressed by neutrophils are IL-8, a powerful PMNL chemotactic agent and inducer of degranulation, and TNF-α, an adhesion promoting and ROS inducing agent associated with inflammation (Cassatella, 1999). Exogenous inflammation suppressing cytokines (IL-4, IL-10 and IL-13) appear to inhibit the cytokine expression by neutrophils. This cytokine expression and control might open new avenues for drug targeting in chronic inflammatory conditions (Witko-Sarsat *et al*, 2000).

The arachidonic acid metabolites formed by PMNL are predominantly the leukotrienes, formed by the action of 5-lipooxygenase, with leukotriene B_4 (LTB_4) being the main metabolite (Alonso *et al*, 1998; Serhan, 1994). Recently it has been shown that prostaglandin E_2 (PGE_2) and thromboxane A_2 (TxA_2) are also synthesized by PMNLs via an inducible cyclooxygenase 2 (COX-2) in response to numerous known PMNL stimuli (Pouliot *et al*, 1998). The rate of upregulation of COX-2 in PMNL is dependant on the stimulant and differs greatly from that of the monocytes and macrophages that may have implications in the control of cell damage by PMNL in diseases like rheumatoid arthritis, sepsis and acute respiratory syndrome (Rocca & Fitzgerald, 2002). Like all inflammatory cells PMNL release platelet activating factor (PAF), which is a very powerful and omnipotent bioactive lipid. It potentiates the inflammatory response of PMNL and eosinophils, possibly through activation of specific G protein type receptors (Prescott, 1999).

10.2.2 Monocytes and macrophages

Monocytes and macrophages share several functions with other myeloid and lymphoid cells, illustrating the built in redundancy within the immune system, but are also the most adaptable cells. They are different stages of development of the same cell type, which are derived from the pluripotent stem cells in the bone marrow. These cells carry CD14 and CD68 cell surface markers. Initially newly formed monocytes circulate freely in blood and lymph (for 5 to 8 days) from where they migrate by diapedesis into almost all organs and body cavities where they differentiate to form resident macrophages that have varied function, properties and morphology depending on the tissue and conditions in their immediate environment (Hashimoto *et al*, 1999). Kupffer cells, Langerhans cells, microglia, dendritic cells and alveolar macrophages, although very different in morphology and function are all derived from monocytes. They are central to the host defence against pathogens, and like PMNLs are capable of phagocytosis, but without the self-destructive after-effects demonstrated by PMNL phagocytosis. They can also phagocytose many more microorganisms than PMNL.

Macrophages respond to circulating or tissue stimuli by secretion of cytokines, chemokines and lipid mediators of acute inflammation that results in the initiation of inflammatory processes, recruitment of leukocytes (especially PMNL). These cells can leave the circulatory system and return via the afferent lymph where they then interact with secondary lymphoid tissue (Gordon, 1999). They act as APCs to immunocompetent lymphocytes in the secondary lymphoid tissue but are not as effective as the dendritic cells (Bjercke & Gaudernack, 1985). They are found in elevated numbers in most areas of chronic inflammation or localized infection.

Bacterial lipopolysaccharides (LPS) and IL-1β induce COX-2 upregulation in macrophages but the Th 2 associated cytokines (IL-4, IL-10 and IL-13) suppress this induction (Berg *et al*, 2001).

Stimulated macrophages release pro-inflammatory cytokines and lipid inflammation mediators as well as express MHC class II molecules, all of which can condition dendritic cells or enable presentation of antigen directly.

10.3 Immune Mediators

Mediators play a central role in the immune response by linking the innate and acquired immune responses in a coordinated and controlled manner. They act as stimulators or amplifiers of specific responses and are central to immune cell trafficking as well as stimulation of cell proliferation and maturation in the myeloid and lymphoid tissue. Most mediators have small radii of action meaning that they only act on the cells in the immediate vicinity of the expressing cell, while a few diffuse throughout the system causing effect at distant organs or cells. Most mediators are associated with specific immunological stimuli and interact with specific high-affinity cell surface receptors. The immune mediators generally show pleiotropic activity (i.e. act on different types of cells and elicit several different responses) and exhibit redundancy (meaning that there are other mediators with the same or similar functions). They may have different effects on different cells or their effect may change depending on their concentration or the presence of other mediators. Mediators can be divided into the protein and lipid derived mediators.

10.3.1 Protein mediators

The protein mediators include amongst others the cytokines, chemokines, growth factors, interferons, adhesion molecules and tumour necrosis factors. The cytokines are expressed by myeloid and lymphoid cells and are ligands for receptors on other cells that are involved in the active immune response. This often results in a cytokine cascade where one cytokine initiates the

expression of another that again stimulates a third cytokine to be expressed. An example is the LPS induced TNF-α that causes the sequential cascade of IL-1 and IL-6 (Beutler & Cerami, 1987), which are all pro-inflammatory cytokines. The cytokines are generally stimulants of effects that take effect at gene expression level such as cell proliferation, differentiation, maturation or activation, but some cytokines inhibit these same effects. Cytokine production is regulated by inducing stimuli at gene transcription level with transient production and short radii of action. Their molecular mass is generally less than or equal to 30 kD (Vilcek & Le, 1994).

The interleukins are a subset of 18 cytokines that are produced mainly by the leukocytes and that act primarily on other leukocytes. Their functions are diverse, but mostly stimulatory for inflammatory processes (IL-2, IL-12, TNFα) or cell maturation (GM-CSF) although some are anti-inflammatory (IL-1β, IL-4, IL-6, IL-10, IL-13). They usually work in concert with other interleukins or cytokines.

The chemokines are a family of more than 40 polypeptides of less than 18kD that share many chemical and structural characteristics. They act as chemotactic attractants (at nanomolar concentrations) for the phagocytic cells of the innate immune response, enhance adhesion, stimulate lipid mediator synthesis, degranulation and the respiratory burst (Baggiolini *et al*, 1997). Chemokines fall into one of two classes; CC or CXC depending on the absence or presence of an amino acid between the two amino-terminal cysteine residues. The cell surface receptors for chemokines are involved in cell recognition; the CXCR4 chemokine receptor is used as an attachment site by the human immuno-deficiency virus (HIV-1) for entry into CD4+ T-lymphocyte cells (Feng *et al*, 1996).

The cell growth factors are stimulants of cell proliferation (especially of the myeloid cells in the bone marrow) and differentiation. GM-CSF is an example of a growth factor.

10.3.2 Lipid mediators

Lipid mediators, including prostaglandins, leukotrienes, thromboxanes, and platelet activating factor (PAF), are generally pro-inflammatory but can have a profound effect on cell maturation and expression of cytokines by cells in general. The first three types are all derivatives of arachidonic acid, which is released from membrane phospholipids by the action of the phospholipase A_2 enzymes. These mediators are active over short distances and have short half-lives. They exert their activity through surface membrane receptors.

Prostaglandins are formed from free arachidonic acid by the action of the cyclooxygenase enzyme 1 or 2. COX-1 is a constitutive enzyme while COX-2 is induced by numerous cytokines, endotoxins, phagocytic stimuli and growth factors (Smith & Langenbach, 2001). The main inflammatory related responses to prostaglandins are: fever, pain, oedema and regulation of leukocyte function. Pain and oedema are secondary effects due to hyperalgesia and increased blood flow respectively but the fever (pyrexia) and leukocyte regulation is a direct result of prostaglandins (Griffiths, 1999). There is an inhibitory effect on Th 1 helper cell formation by PGE_2 with the expressed cytokine profile being more inclined to the allergic response (Sergeeva *et al*, 1997). The main inflammation related prostaglandins are PGE_2 and PGI_2 and although PGD_2 and $PGF_{2\alpha}$ are also important they are more tissue specific (PGD is CNS and mast cell related while PGF is vascular smooth muscle related) (Griffiths, 1999).

Leukotrienes are formed by the action of 5-lipoxygenase on arachidonic acid (AA). This enzyme is found on the perinuclear membrane of most leukocytes and mast cells (Ford-Hutchinson *et al*, 1994). Two main classes of leukotrienes exist: the LTB_4 class of hydroxylated derivatives and the cysteinyl leukotrienes that include LTC_4, LTD_4 and LTE_4. LTB_4 enhances PMNL lysosome release and degranulation, mediates adhesion, migration and recruitment while B-lymphocytes are activated and enhanced production of immunoglobulins. It acts as an intracellular messenger acting on the peroxisome proliferator-activated receptor-α (PPAR-α) involved in fatty acid oxidation and adipogenesis. The cysteinyl leukotrienes have vasoconstrictive and smooth muscle constriction properties but augment vascular permeability (Penrose *et al*, 1999).

Two lipoxins (and two isomers that result from acetyl salicylic acid treatment) are also derived from AA by the action of 5-lipoxygenase and are lipid mediators, but in contrast have anti-inflammatory effects. They appear to be formed transcellularly i.e. synthesis occurs in two stages, each in a different cell. They are under IL-4 and IL-13 (i.e. inflammation suppressing) cytokine control and inhibit PMNL and eosinophil adhesion and migration but stimulates the same actions in monocytes and mediate vasodilation. They counteract the effect of prostaglandins and mediate nitric oxide generation in vascular smooth muscle (Bratt & Gyllenhammar, 1995; Serhan, 1999).

11 Oxihumate

11.1 Synthesis from coal

As mentioned above coal is formed slowly from humic substance rich material through a series of reactions. However, if conditions are right, coal can be converted back to a combination of fulvic

acids, humic acids, humic substances and highly mineralized compounds during the natural oxidative weathering of coal (Rausa, 1994). This natural process of oxidation of coal to form, amongst others, humic acids can be accelerated by chemical oxidation of the coal. By controlling the conditions used during the oxidation process, the yield of humic acids can be varied relative to the other products. Based on this fact a method for the production of a humic acid from a bituminous coal was developed (Bergh *et al*, 1997), patented (Cronjé, 1988; Cronjé *et al*, 1991) and used by a South African company – Enerkom (Pty) Ltd. - to reproducibly produce a formulated potassium salt of humic acid as a food supplement that was marketed as an immune stimulant (Dekker & Medlen, 1999b) under the name "Oximate". Another useful product formed during the process was fulvic acid, which could be applied in the agricultural and veterinary fields as well as formulated into a topical ointment for human use.

The patented process for the oxidation of bituminous coal was basically as follows;

- bituminous coal from a defined source was pulverized to a particle size of < 200μm.
- the powder was made into a slurry using water and pumped into a temperature- controlled high-pressure reaction vessel where the mixture was heated to 180°C under an atmosphere of oxygen maintained at a constant pressure of 4 MPa.
- the resulting exothermic reaction was then controlled to maintain the reaction temperature at 180°C while feeding oxygen at a high flow rate into the reactor for at least one hour. The reaction was stopped by shutting off the oxygen supply and cooling to ambient temperature.
- the reaction mixture was then filtered though an efficient high-pressure filter system to remove the insoluble humic acid and incompletely reacted coal fractions.
- the filtrate containing the fulvic acids was concentrated to a 24% solution in a reverse osmosis system.
- the retentate on the filter cakes contained incompletely reacted coal (termed oxicoal in the process) and humic acids that were resuspended in water and treated with an aqueous potassium hydroxide solution.
- this suspension was homogenized then dried in a spray-drying column to give the potassium humate powder.

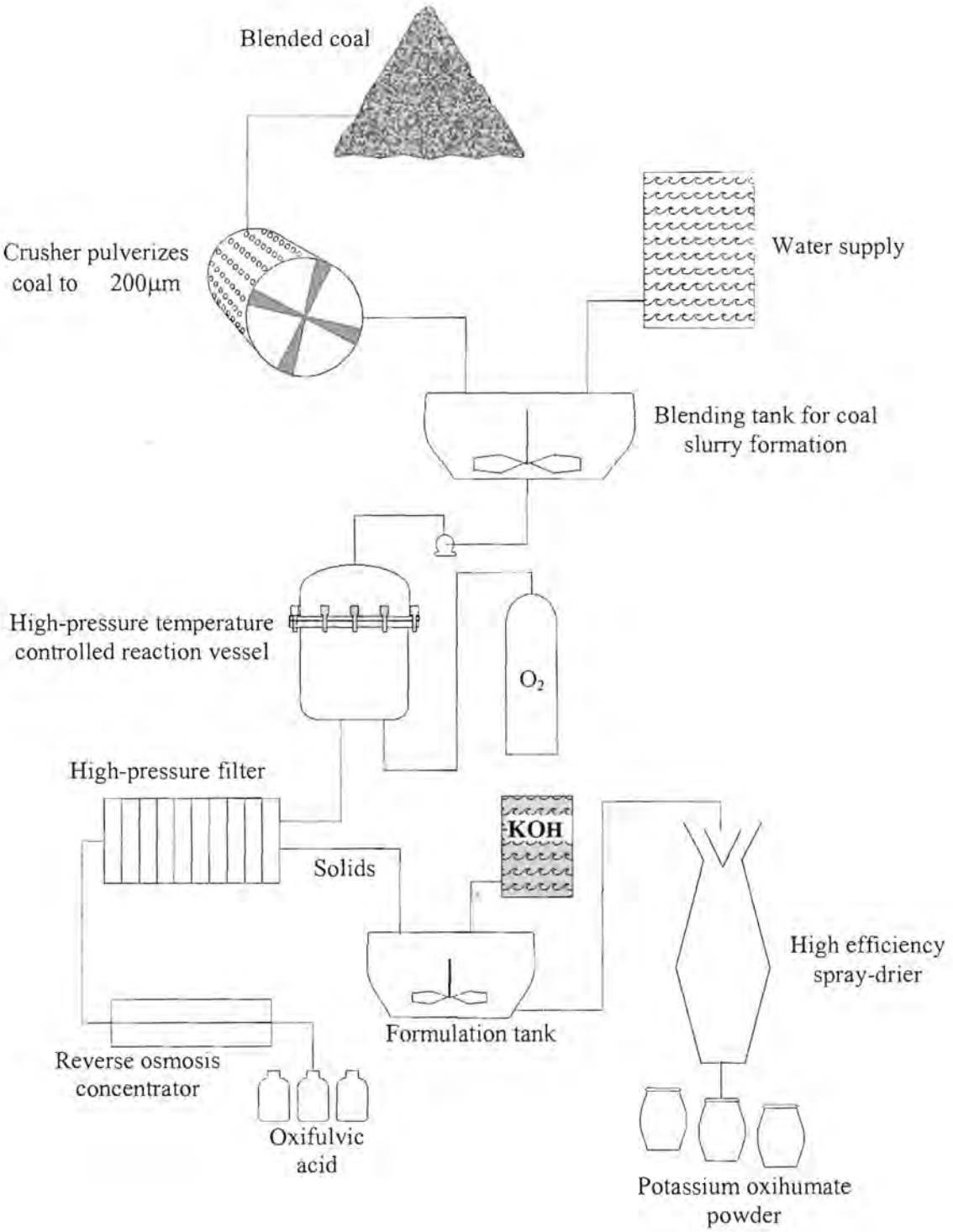

Figure 1-2: A diagrammatic representation of the patented process for manufacturing synthetic oxihumate from a bituminous coal (Cronjé, 1988; Cronjé *et al*, 1991)

To distinguish these synthetically produced products from the naturally formed equivalent compounds the fulvic acids and the potassium salt of humic acid were named oxifulvic acid and potassium oxihumate respectively.

Advantages that the above method has over other methods were that it was a rapid method that used no toxic compounds during the synthesis and that no toxic waste was generated. Due to the simplicity of the method it could easily be automated and the process could be adapted to batch or continuous flow processes. The wastage was also minimized as both the humic acid and fulvic acid fractions were recovered separately and could be used for different applications.

11.2 Therapeutic properties of oxihumate and oxifulvic acid

In vitro as well as *in vivo* toxicity tests have been performed on both oxifulvic acid and oxihumate using rat and dog models (Confidential Study Reports: Biocon (Pty) Ltd, Pretoria, South Africa, 1999 & 2001) where it was found that the sub-chronic long-term (90 day treatment) and the acute toxicity of oxihumate were low and that no toxic effects were exhibited up to oxihumate concentrations of 1000mg/kg/day during the test period. No drug related deaths occurred during any of the studies. Small changes in blood parameters (increased red blood cell size and haemoglobin content, increased serum globulin concentration, decreased circulating leukocytes, decreased serum inorganic phosphate) and a small but acceptable decrease of end body mass compared to control animals was found at concentrations greater than 100mg/kg/day. Under the conditions of the study the no-adverse-observed–effect-level (NAOEL) was concluded to be higher than 100mg/kg/day. The acute toxicity study using rats indicated that the oral LD_{50} was greater than 3456 mg/kg and the dermal LD_{50} greater than 4147mg/kg.

Oxifulvic acid studies indicated that it had antimicrobial activity against several known pathogens (van Rensburg *et al*, 2000) at concentrations of less than 15 g/l, was found to be effective as a topical cream (containing 5.3% oxifulvic acid) for the treatment of pyotraumatic dermatitis in dogs and cats as well as an inhibitor of contact dermatitis in animal studies using a mouse model (Dekker & Medlen, 1999a; van Rensburg *et al*, 2002). Another study evaluating the safety and efficacy of topically administered oxifulvic acid was done on atopic but allergic volunteers and was found not to cause sensitisation when applied topically nor did it alter any of a battery of safety parameters (Snyman *et al*, 2002). It was interesting to note that the formulation appeared to play a major role in the efficacy with the formulation containing the higher concentration of oxifulvic acid having less activity against the allergic response elicited by a challenge. This was ascribed to the lower pH of the cream with the higher oxifulvic acid concentration.

An *in vitro* study on the anti-viral activity of oxifulvic acid on Herpes Simplex Virus type-1 was conducted using BGM monkey kidney cells. These studies revealed that oxifulvic acid was toxic to the cells in culture at concentrations in excess of 1.25mg/ml. There was no virostatic effect on the HSV-1 virus in suspension but host cell entry was 75% inhibited and viral replication within the cells was completely inhibited at concentrations of 40µg/ml and 320µg/ml respectively (Williams *et al*, 2001).

The immunomodulating properties of oxihumate have been extensively investigated. The effect of oxihumate on the proliferation of lymphocytes stimulated with mitogens revealed that there was an increase in proliferation at oxihumate concentrations of more than 20µg/ml. This effect was seen *in vitro* and *ex vivo* in HIV positive individuals treated with oxihumate. It was determined that this increase in proliferation was probably due to increased IL-2 production and IL-2 receptor expression while IL-10 concentrations were suppressed (Jooné *et al*, 2003).

A two-week phase I clinical trial was performed using various orally administered oxihumate doses from 2 to 8 g per day. An increase in the weight of all the treated patients was found when compared to the control group receiving a placebo. There was however no significant improvement in either CD4+ counts or viral load (Botes *et al*, 2002) but this could have been due to the short duration of the study. No toxic effects were apparent during the treatment period or for one week after treatment had stopped. This study highlighted the non-toxic characteristic of oxihumate in humans. A later *in vitro* study on the anti-HIV properties of oxihumate (van Rensburg *et al*, 2001) demonstrated that replication of HIV was blocked through an inactivation of the virus particle that could occur after viral binding to the cell had taken place. This was identified as involving the binding of oxihumate to the V3 loop of the virus gp120 protein. No viral resistance developed against the oxihumate after as many as 18 passages.

In another study done using *in vitro* proliferating lymphocytes (Jooné, 2002), the concentrations of the following cytokines were determined in the culture medium: TNF-α, INF-γ, IL-2, IL-4, IL-6, IL-10 after exposing the cultures to various concentrations of oxihumate for 72 hours. By comparing the oxihumate induced cytokine profiles to equivalent controls, it was concluded that oxihumate induces mostly a Th 1 type of immune response. The expression of IL-2R – the cell surface IL-2 receptor (also referred to as CD25) was found to increase after oxihumate treatment.

Other immune related mediators, surface marker proteins and cell activities were also investigated for changes induced by oxihumate exposure. The effect of oxihumate on the concentrations of the

lipid immune mediators, prostaglandin E_2 and leukotriene B_4 were determined and showed a slight decrease in the former and a significant increase in the latter (Dekker & Medlen, 1999b). Prostaglandin E_2 is a pro-inflammatory mediator (stimulating pyresis, oedema and hyperalgesia) (Griffiths, 1999) that also has a suppressing effect on the Th 1 cytokine profile while LTB_4, which is also pro-inflammatory, especially with respect to neutrophils, stimulates the expression of IL-2, a Th 1 type cytokine (Penrose *et al*, 1999).

Due to phagocytic cell adhesion being one of the first stages in the inflammatory response, an *in vitro* investigation of the effect that oxihumate had on the expression of the CR3 adhesion molecules by PMNLs was included in the above study (Jooné, 2002). It was found that resting PMNLs were unaffected but that the expression of CR3 by activated PMNL (using either PMA or FMLP/cytochalasin B) was inhibited in a dose dependant manner. Additionally, the effect that oxihumate had on the binding of resting and stimulated PMNL to human ICAM-1 and human E-selectin molecules expressed on the surface of transfected hamster kidney cell lines were determined (Jooné *et al*, 2001). Again it was found that there was a dose dependant inhibitory response to oxihumate treatment.

To eliminate the possibility that the mechanism of action of oxihumate on the lymphocytes was due to disruption of the cell membrane integrity, the effect of oxihumate on sheep red blood cells, where the membrane is relatively fragile, was investigated and found to be minimal, which would exclude this membrane disrupting mechanism (Jooné, 2002).

12 Study design

Motivation

Various reports have been published on the lack of toxicity and possible anti-inflammatory properties of humic acid, but there is no conclusive evidence demonstrating that humic acid is absorbed when administered orally and that it is effective in the treatment of inflammation.

Unlike most other studies on humic acid, this study was done on oxihumate, the potassium salt of a semi-synthetic humic acid derived from a homogenous mixture of bituminous coal by a controlled process that ensures that the product is consistent from batch to batch. This product is easily available in large quantities, is relatively cheap and has already been registered as a food supplement.

Hypothesis

The immuno-modulating properties of oxihumate are attributable to only one or several discrete compounds that can be isolated *in vitro* and possibly also *in vivo* following oral administration of oxihumate.

Aims

The aims of this study are to determine whether;

- oxihumate can be sub-fractionated and that these sub-fractions can be chemically characterized
- biological activity of oxihumate can be assigned to one or more sub-fractions
- oxihumate can be absorbed from the gastro-intestinal tract
- oxihumate possesses anti-inflammatory activity *in vivo*

Objectives

The main objectives of the study are;

- to attempt to fractionate oxihumate using differential solubility of the semi-synthetic humic acid in aqueous organic solvent mixtures

- to chemically analyze the isolated sub-fractions using techniques such as thin layer chromatography, high-pressure liquid chromatography, fluorescence spectroscopy, infrared spectroscopy, UV/visible spectroscopy and the ash content.

- to assay oxihumate and any isolated sub-fractions for anti-inflammatory activity *in vitro* using two biological assays namely

    - expression of CR3, a pro-inflammatory adhesion molecule, on resting and stimulated human neutrophils
    - scavenging of pro-inflammatory oxidants formed by stimulated human neutrophils

- to perform an *in vitro* pharmacokinetic study using an isolated rat gut system to determine whether oxihumate can be absorbed from the gastro-intestinal tract

- to determine the anti-inflammatory activity effect of oxihumate and its sub-fractions in an *in vivo* contact hypersensitivity assay using a rodent model

- to perform an *in vivo* pharmacokinetic study in a baboon model to attempt to determine the target organs of absorbed humic acid components

Each phase of the research project depends on the outcome of the preceding phases and the *in vivo* studies will rely on the identification and isolation of an active fraction that can be confirmed *in vitro*.