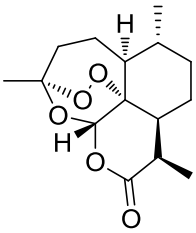
Research interests (GDB)

1. The Biosynthesis of Artemisinin

The unusual sesquiterpene artemisinin (qinghaosu) was first isolated from the Chinese plant *Artemisia annua* almost forty years ago as part of an initiative to discover novel natural products with antimalarial activity.



Artemisinin is now the most important drug for treating malaria. However, the high cost of producing artemisinin from the natural source often places this drug beyond the reach of those most in need; and, in the recent past, there have been shortages in artemisinin supply, which could recur as the world-wide demand for artemisinin continues to grow. One way to address this issue is to genetically engineer the pathway to artemisinin within a microbial host, in order to permit its production by fermentation. Obviously, a complete understanding of the biosynthetic route to artemisinin is a pre-requsite for the success of this approach. An overview of the biosynthetic route to artemisinin is given in Fig. 1:

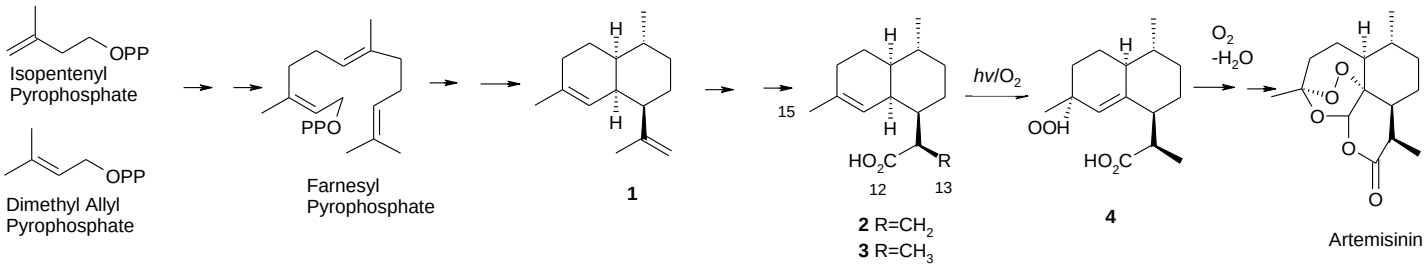


Figure 1.The biosynthesis of artemisinin.

Through feeding experiments with suitably isotopically-labeled precursors, we have been able to show that the biosynthesis of artemisinin proceeds from amorpha-4,11-diene (**1**) *via* dihydroartemisinic acid (**3**)

rather than artemisinic acid (**2**), as had been previously assumed. In the most recent extension of this work (BBSRC grant BB/G008744/1), we have now established that the transformation of amorpha-4,11-diene (**1**) to dihydroartemisinic acid (**3**) proceeds *via* the pathway shown in Fig. 2.

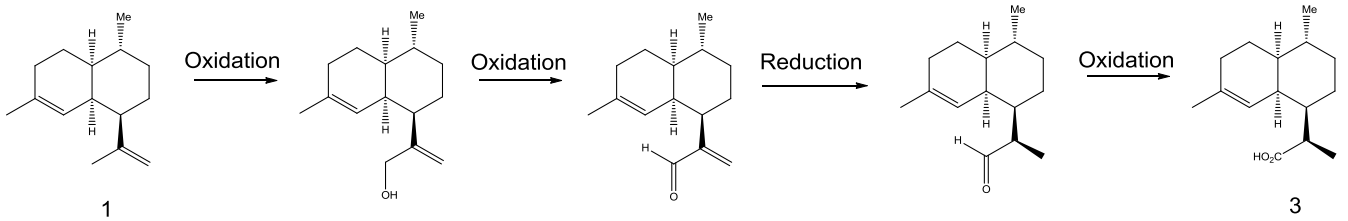


Figure 2. Pathway for conversion of amorpha-4,11-diene (**1**) to dihydroartemisinic acid (**3**)

During this same project, we have also completed a chemical synthesis the *enantiomer* of dihydroartemisinic acid (**3**) in labeled form (*i.e.* with ^{13}C and/or ^{2}H incorporated at the 15-position). This is being used in feeding experiments to establish whether the subsequent transformations of dihydroartemisinic acid (**3**) to artemisinin proceed *via* a spontaneous autoxidation process, involving a hydroperoxide intermediate (**4**), such as that depicted in Fig. 1, or by an enzymatic route. (If enzymes are involved in the further biotransformations of dihydroartemisinic acid, then, due to the highly specific nature of enzymatic catalysis, the unnatural enantiomer of dihydroartemisinic acid should not be accepted as substrate for further conversion to artemisinin; simpler chemical reactions, such as autoxidation, on the other hand, should not discriminate between the two enantiomers).

2. A New Strategy for Assigning the NMR Spectra of Carbohydrates

It is now well established that oligosaccharides are involved in a multitude of biological processes including cell-cell recognition, differentiation and adhesion. However, despite their importance, the determination of the chemical structure of oligosaccharides remains a difficult chemical problem.

NMR spectroscopy is the most powerful analytical technique currently available for determining the solution structure of a carbohydrate. Unfortunately, there are limits to the size of an oligosaccharide which is amenable to routine ^{1}H NMR spectroscopic analysis because the "C(H)$_n$" hydrogen signals of the carbohydrate backbone normally occupy only a small portion of the full range of all possible ^{1}H chemical shifts (δ_H 3 - 5.5 ppm from a full range of 0 - 10 ppm). We are attempting to overcome this limitation by additionally employing the -OH peaks of a sugar in the structure elucidation process. This has never been possible before because the –OH hydrogens are not normally observed in ^{1}H NMR, due to rapid exchange with the D_2O solvent. We have now been able to reduce the rate of exchange of the –OH groups sufficiently that these peaks become observable in a ^{1}H NMR spectrum (between 5.5 and 8.5 ppm), as shown in Fig. 3. This was achieved by preparing glucose as an aqueous solution in 1% D_2O/99% H_2O in a capillary tube, which could then be cooled below the freezing point of water.

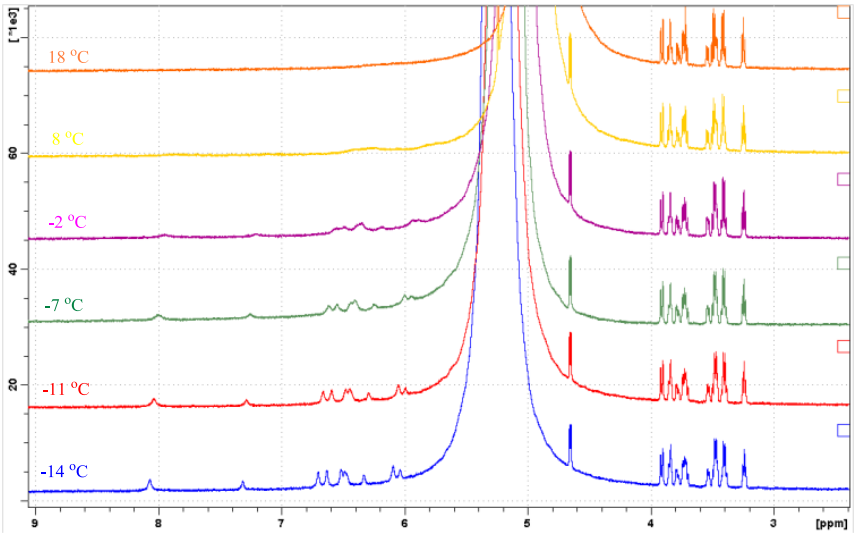


Figure 3. ^{1}H NMR spectra of glucose recorded in a capillary in 1% D_2O, as the temperature is lowered.

The 2D-NMR experiment, HSQC-TOCSY, was next used to "connect" these OH groups with their associated carbons, creating "jigsaw pieces" which correspond to small fragments of the molecule, such as that shown in Figure 4.

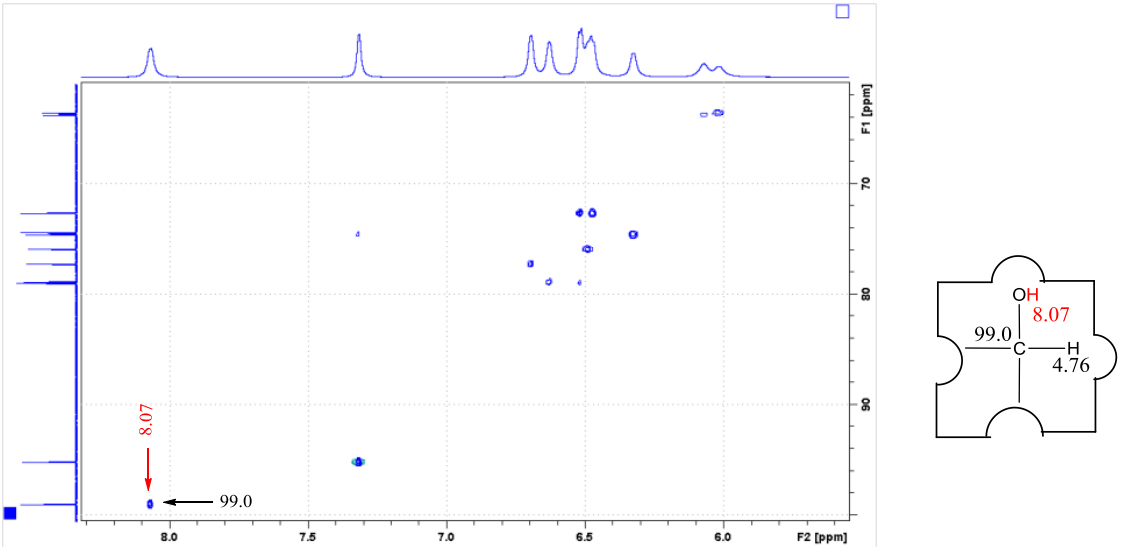


Figure 4. 2D-NMR HSQC-TOCSY spectrum of "supercooled" glucose.

Finally, application of HSQC-TCOSY spectra recorded with a variety of mixing times allowed these fragments to be joined into successively larger pieces, eventually leading to the elucidation of the structure of the entire carbohydrate. Thus, in Fig. 5, all the ^{13}C and ^{1}H assignments of glucose (both CH and –OH) have been deduced by application of this technique.

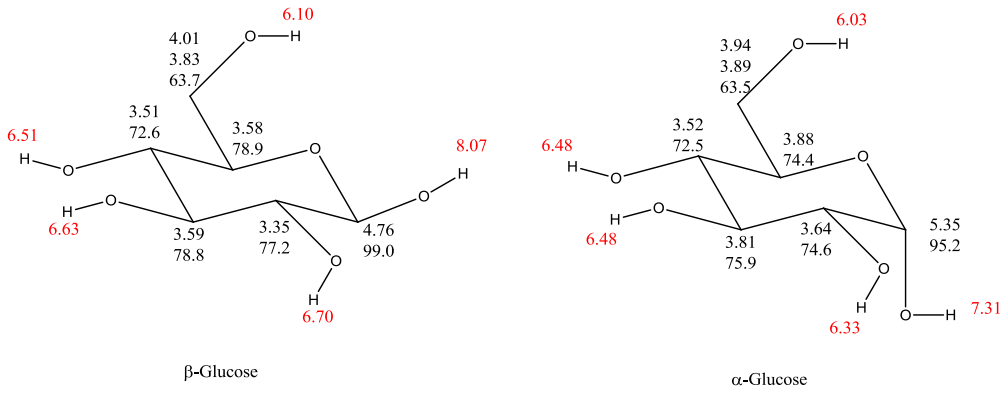


Figure 5. Complete CH and -OH assignments for glucose established by using HSQC-TOCSY experiments with short and long mixing times.

The range of –OH chemical shifts in a large number of carbohydrate studied to date have been found to be quite wide and, furthermore, it does not overlap with the CH chemical shift range. Investigations are currently underway to establish whether application of this novel technique will permit the structure

determination of oligosaccharides larger than those which are currently possible, when using only the CH resonances.

3. NMR of Peptides and Proteins

In collaboration with Prof. Ian Hamley and Dr David Nutt, we have studied the conformation of a model peptide AAKLVFF, which is based on a fragment of the amyloid peptide A 16-20. (The self-assembly of such peptides into amyloid fibrils - so-called amyloid formation - has been implicated in a considerable number of disease states, including Alzheimer's). Using 2D-NMR experiments such as COSY and TOCSY it has been possible to fully assign all the protons in the 1D- ^{1}H NMR spectrum of this heptapeptide (Figure 6).

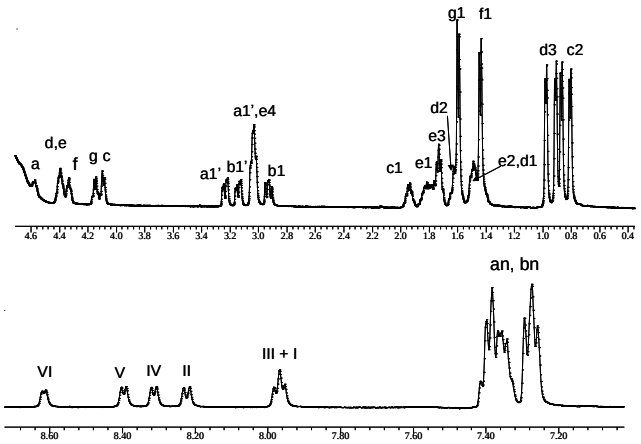


Figure 6. Assigned ^{1}H NMR spectrum of AAKLVFF in water

Correlations observed in 2D- ROESY spectra were then used to generate a set of distance constraints between these assigned protons. Molecular dynamics simulations based on these distance constraints suggest that in aqueous solution this peptide fragment is able to stack into antiparallel β-sheets (Fig. 7), which may have relevance to the mechanism of amyloid formation.

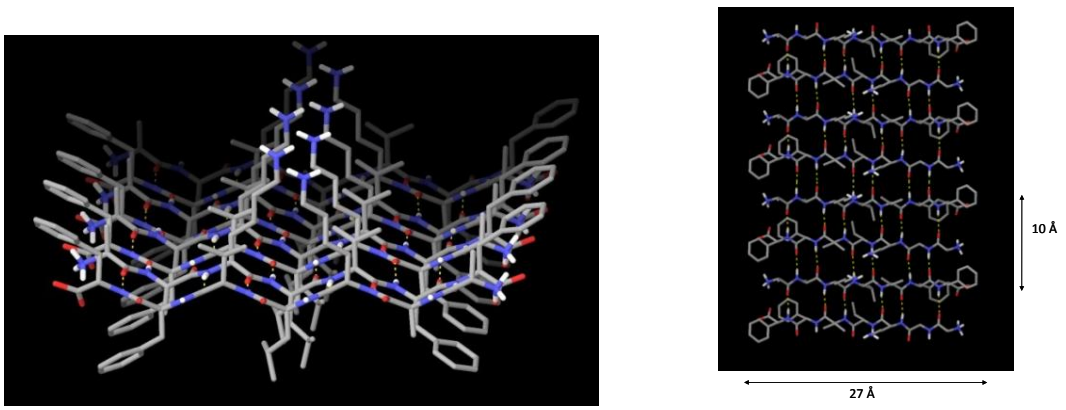


Figure 7. Proposed antiparallel β-sheet structure of AAKLVFF

More recently, we have been attempting to determine the structure of a protein which is a component of EfeUOB, a new type of bacterial iron transporter discovered by Prof. Simon Andrews at the University of Reading. This cupredoxin domain protein is more than ten times the size of the amyloid peptide fragment previously investigated, and its study has required that protein to be over-expressed in doubly-labelled form (both ^{13}C and ^{15}N) in order that 3D-NMR spectra, such as that shown in Fig. 8, can be acquired.

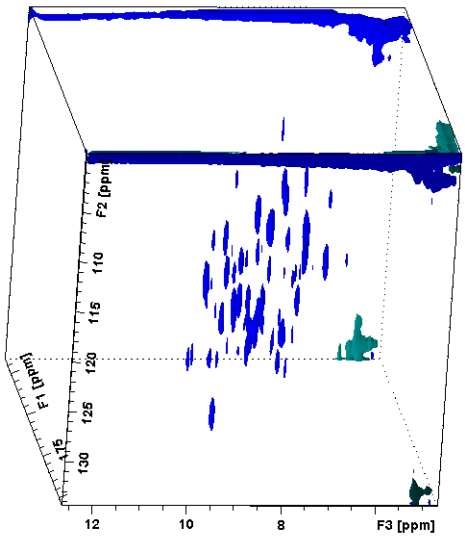


Figure 8. 3D- HNCO spectrum of the cupredoxin domain protein.

4.Chemical Approaches to the Origin of Life – Did Life Begin as a Disproportionation Reaction?

Over the past 50 years, there have been a number of ideas concerning the chemical processes which must have been involved in the very early stages of the origin of life on Earth. The first serious study was made by Oparin who observed the spontaneous formation of coacervates from modern-day biological polymers, and suggested this as a model for the origin of the very first proto-cell. In the 1950's and 60's, Miller and Urey were able to show that many of the small molecules involved in present-day metabolism could be formed by applying electric discharges to various mixtures of reducing gases. Fox later discovered that thermal polymerization of amino acids could yield "proteinoids" which bore some striking resemblances to modern-day biological cells. In the 1980's, Cairns-Smith led a movement for an inorganic origin (from clay minerals) and the most recent suggestion has been that of the "RNA world", proposed by Cech, following on from the discovery that RNA can behave as a catalyst, as well as fulfilling its more usual role of storing information.

The hypothesis under investigation in our group differs from all of the above in that we propose that life may have its origins in the oligomerization and disproportionation reactions of formaldehyde – a precursor molecule which would have been readily available on the primitive Earth. Our initial studies are directed towards obtaining a better understanding of the formose reaction, in which this one-carbon "building block" is known to be converted into a wide variety of complex organic compounds under very mild conditions. The following simplified mechanism (Figure 9) has been proposed previously by Breslow to account for the autocatalytic process by which formaldehyde undergoes oligomerization to a variety of polyols (including several of the modern-day sugars) in the presence of Ca^{2+} .

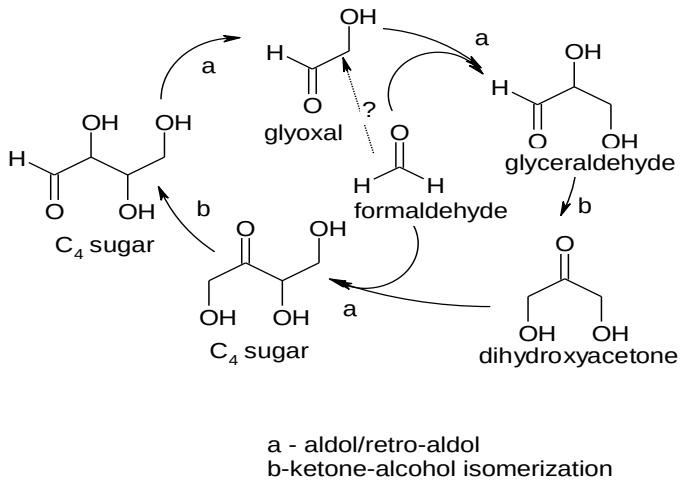


Figure 9. Breslow's mechanism to account for the autocatalytic nature of the formose reaction.

However, the formose reaction is, in reality, far more complex than the above scheme would suggest and we have also observed the formation of many other organic products, including hydrocarbons and carboxylic acids. These observations have suggested to us that the disproportionation of intermediates (at the aldehyde level of oxidation) may also be an important feature of the formose reaction, in addition to the more well established carbon-carbon bond formation reactions. If some of the lipids produced by the disproportionation of the products of formaldehyde oligomerization had the correct physical properties to spontaneously aggregate into micelles, then such micelles might serve as the very first proto-cells. In particular, if the oligomerization/disproportionation components of the formose reaction were to continue to operate within the micelle, then the micelle would continue to produce lipids, inevitably growing in size, until eventually a critical mass was attained, at which point it would sub-divide into "daughter" micelles.

We are currently attempting to demonstrate the possibility for the formation of such a "self-reproducing" vesicle in an aqueous system which is supplied with formaldehyde. To achieve this, we need a much better understanding of the formose reaction than is currently available, and, as in the other proposals, our understanding of the complex chemistry involved in the formose reaction, and our attempts to locate various components of the formose reaction within a micellar phase (if and when it is formed), are being undertaken primarily by the use of NMR spectroscopy.

3. Structure Elucidation of the Components of a Chemical Mixture by Diffusion Ordered Spectroscopy NMR (DOSY)

The application of NMR spectroscopy to the determination of chemical structure has seen rapid advances over the past twenty years, primarily as a result of the introduction of a variety of two-dimensional (2D-) NMR techniques. The structure elucidation of an "unknown" organic compound by NMR is now a fairly routine procedure and in most cases it can be achieved by the application of seven "classical" NMR experiments (^{1}H, ^{13}C, DEPT-135, HSQC, HMBC, ^{1}H-^{1}H COSY and NOESY). Such an unambiguous determination of chemical structure by the application of 2D-NMR generally requires a highly purified sample of a single compound. More recently, a novel NMR technique has been developed, which allows the various components of a chemical mixture to be analyzed simultaneously in terms of their differing diffusion coefficients. Diffusion Ordered Spectroscopy (DOSY) can be implemented on any NMR spectrometer which is equipped with a pulsed field gradient capability.

We are interested in developing two- and three-dimensional (3D-) NMR experiments, which will be hybrids between the "classical" 1D- and 2D-NMR experiments and the more recent DOSY technique, with the goal of being able to perform the structure elucidation of "unknown" organic compounds which are present as components of a mixture. The successful application of such hybrid 2D/3D-NMR techniques offers the possibility for the complete characterization of each and every component of a mixture of "unknown" organic compounds, without the need for any prior separation (e.g by chromatography). Such a capability should be of particular interest to the chemical and pharmaceutical industries, which currently invest a large amount of time and money in separating the metabolites of drugs and agrochemicals prior to their characterization.