

2.3 Data pre-treatment and pre-processing

To achieve optimal and robust chemical shift referencing, an internal reference was applied to the single protonated peak at 7.05 ppm of the benzene ring using Bruker Topspin software version 3.6.5. The single protonated peak of the benzene ring was chosen due to its single peak representation, its location (end of the aromatic region) and its clear resolution from other signals of interest. Phase correction was performed in Mestre Nova software. Auto phasing was selected which consisted of performing a zero-order phase correction on the whole spectrum by selection of the PH0 algorithm in the processing parameters of the software. Thereafter, all NMR spectra were manually inspected for any phase distortions. There were several baseline correction algorithms in the Mestre Nova software available from which to select. For this experiment, the automatic baseline correction function in Mestre Nova which applies a Bernstein polynomial fit algorithm to the frequency domain of the NMR data was selected. Drift correction was used to remove a baseline offset of the spectrum resulting from a non-zero integral for the fid and zero-frequency spikes in the spectrum. The Mestre Nova software applied this automatically, by default, using the common procedure of averaging the last 5% of the points in the fid and subtracting these from the rest of the fid.

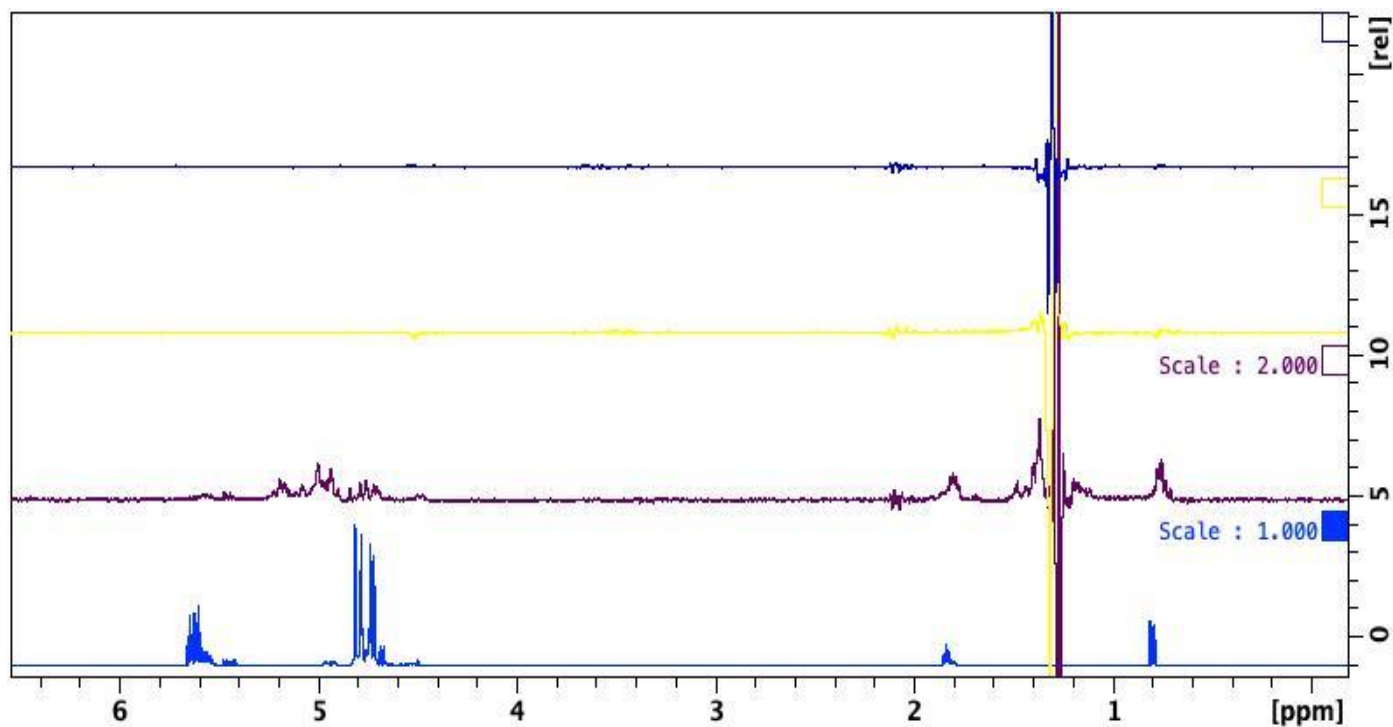
Subsequently, the ^1H NMR and ^1H TOCSY data was binned with a 0.01ppm bin size using Mestre Nova software for every petrol sample and the binned data was saved as CSV comma (*.csv) files. However, not all the variables in the binned data were relevant for discrimination. Visual inspection of the ^1H NMR and ^1H TOCSY spectra showed ranges of the spectrum that did not contain any information, which was considered as background noise that had to be filtered out of the binned data. The chemical shift ranges of the spectrum that did not contain any spectral information, i.e. the presence of no couplings, were 2.70ppm-3.10ppm, 3.50ppm-3.90ppm, 4.20ppm-4.50ppm and 5.70ppm-6.50ppm. The position of the relevant couplings in the binned data was significant for the investigation of the variables; for that reason, the spectral free regions of the data were not omitted or deleted but conditioned, so they were not considered for the classification model. To filter out the data noise and spectra-free regions, if the numerical value of the bin was <1 , the value in the bin was transformed to be equivalent to the numerical value of 0, else the binned data was the actual value.

The total area sum normalization consists of normalizing the intensity of each individual spectral bin to the total intensity of each spectrum. This type of normalization procedure is performed automatically in the MetaboAnalyst version 5.0 web platform or by applying a Log function in MS Excel. Single peak normalization is performed when each data point is divided by the amplitude of a specific peak of interest. Usually, the intensity of the internal standard peak is used as it is consistent for all samples. However, if an internal standard is not added to the sample mixture, a peak within the sample can be used for single peak normalization. This peak is usually the highest intensity peak in each sample. There are limitations of the single peak normalization method when there is no internal standard present; there is a possibility of skewing the relative abundance among samples because the abundance of the selected peak may not exhibit the same intensity in all samples. Therefore, total area sum normalization process was applied to the NMR data.

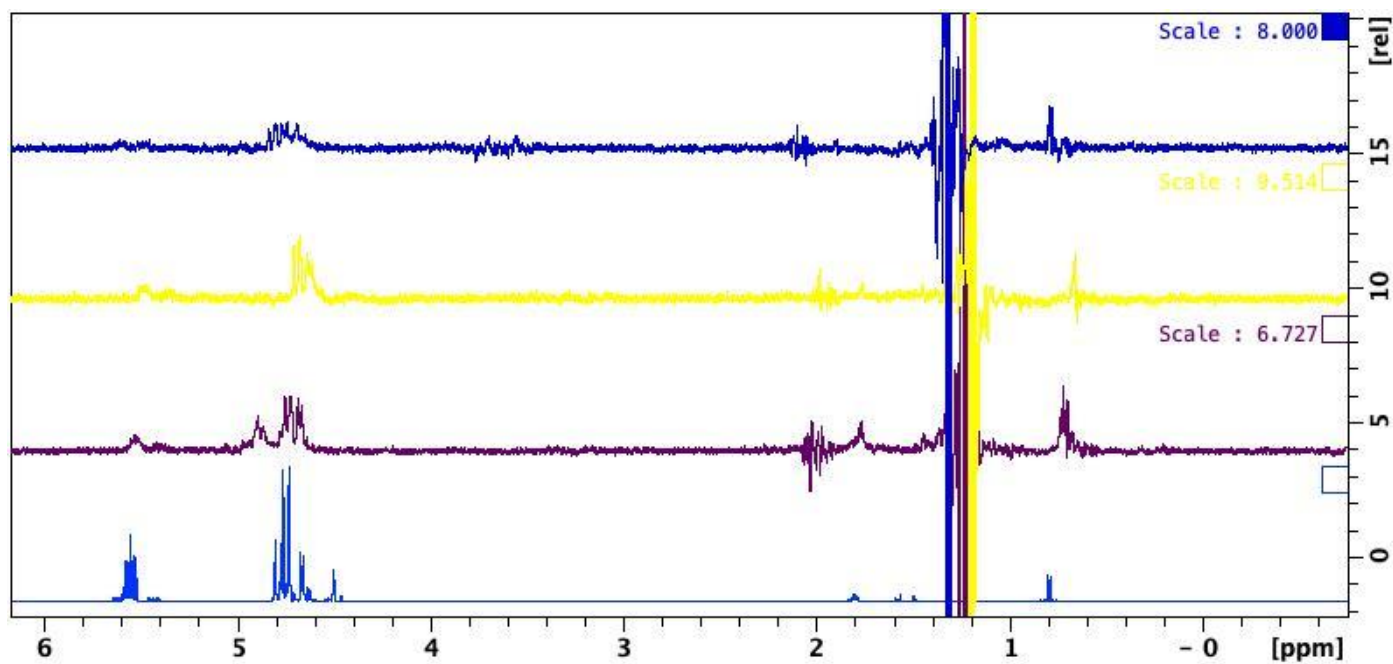
3.3 selTOCSY results for evaporated petrol samples.

Figure S1. A comparison of ^1H selTOCSY of the mixture of 3-methyl-1-butene of (i) BP S (ii) Jet (iii) Esso (iv) Shell and (v) Texaco (vi) BP M native petrol sources with 25%, 50% and 75% evaporation rate

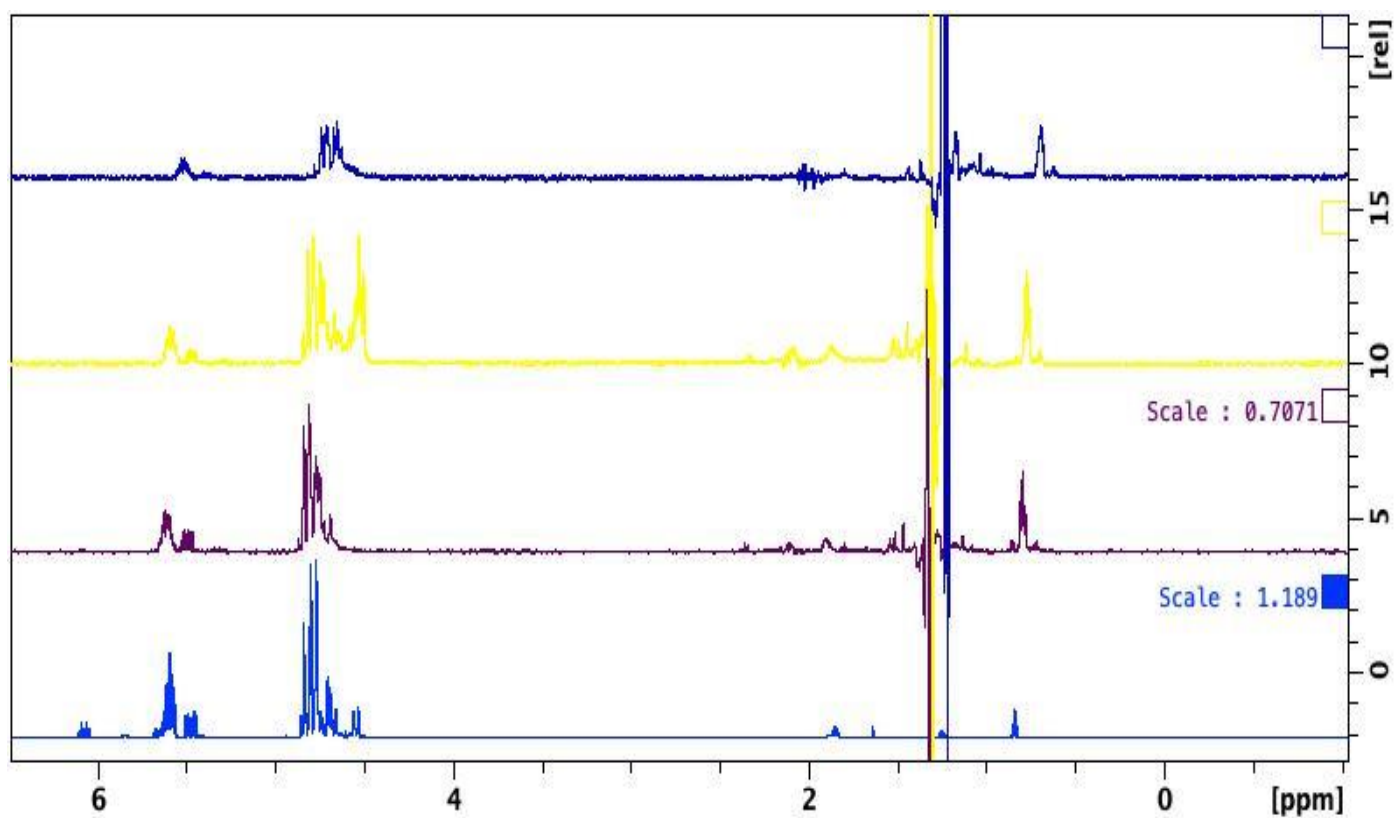
(i) BP S



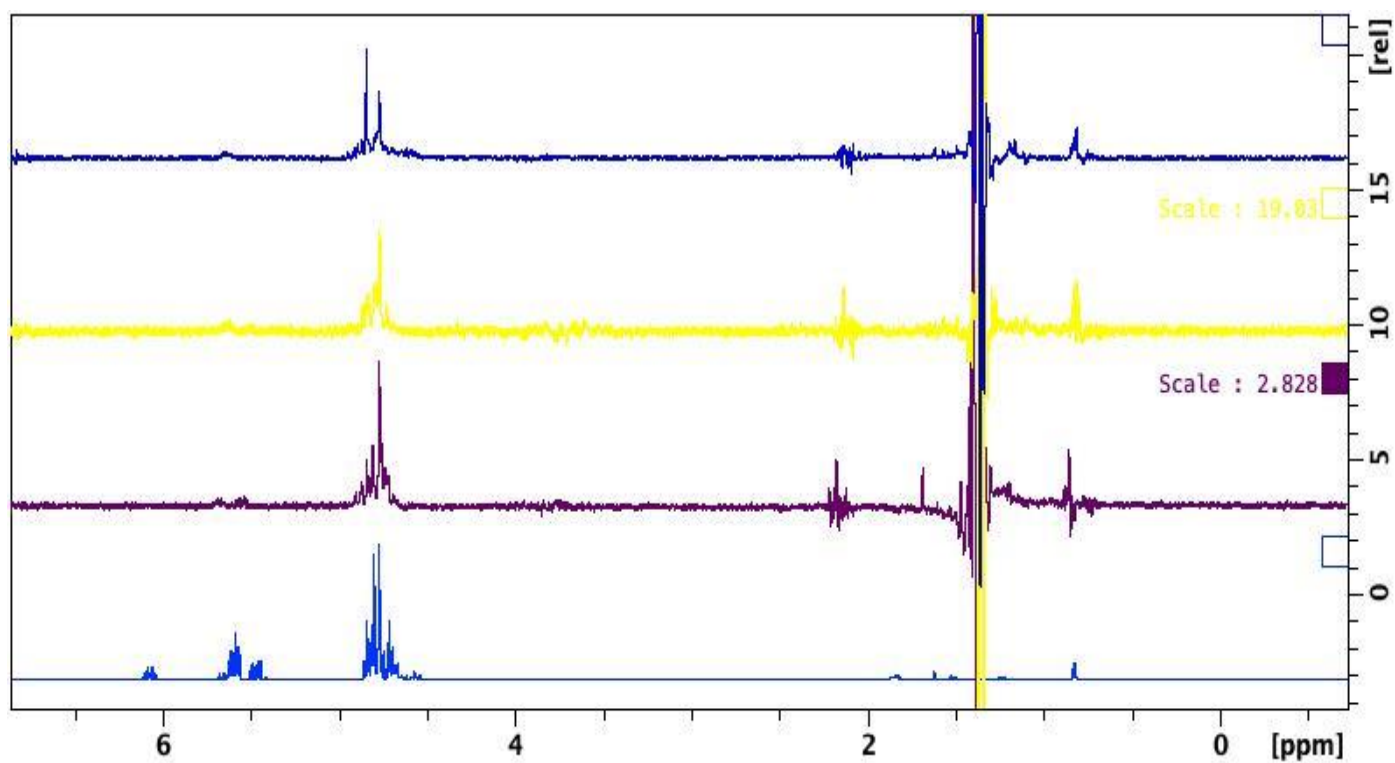
(ii) Jet



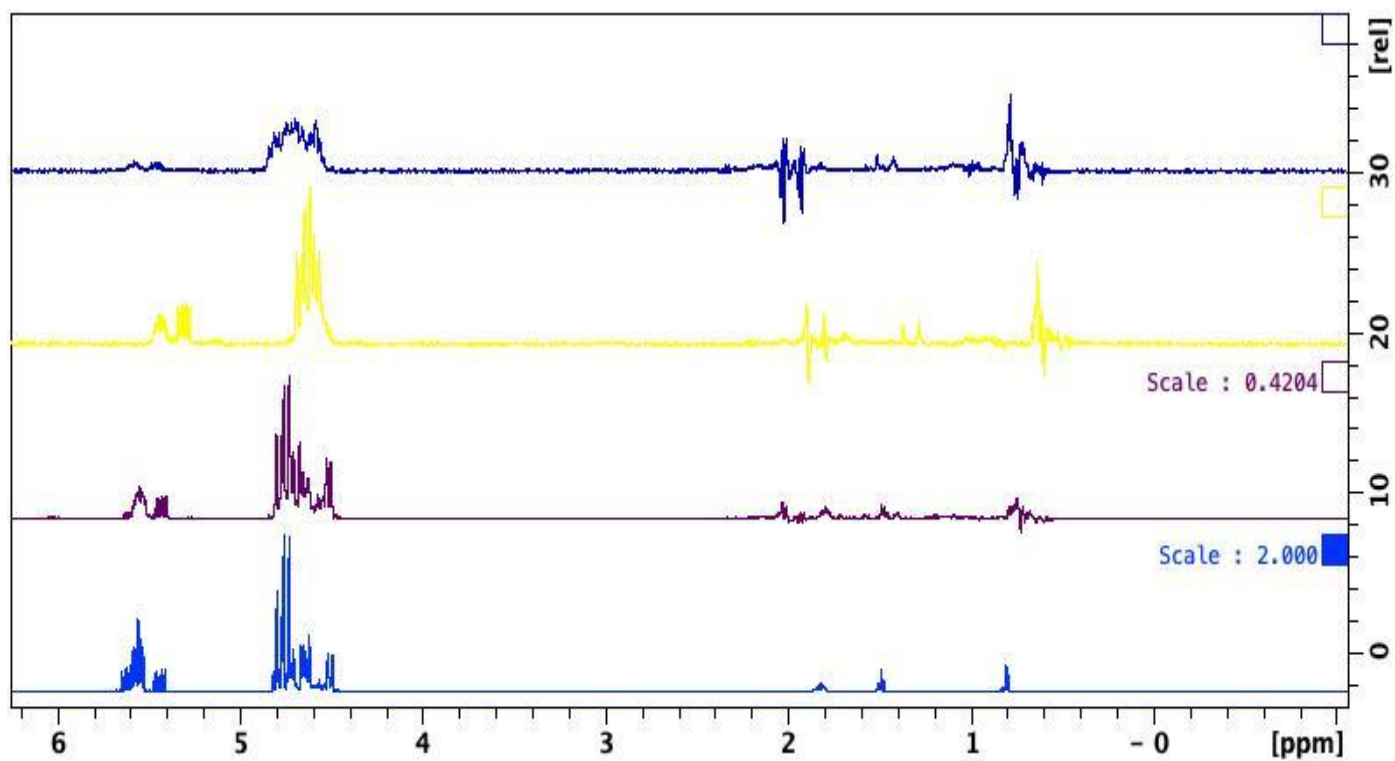
(iii) Esso



(iv) Shell



(v) Texaco



(vi) BP M

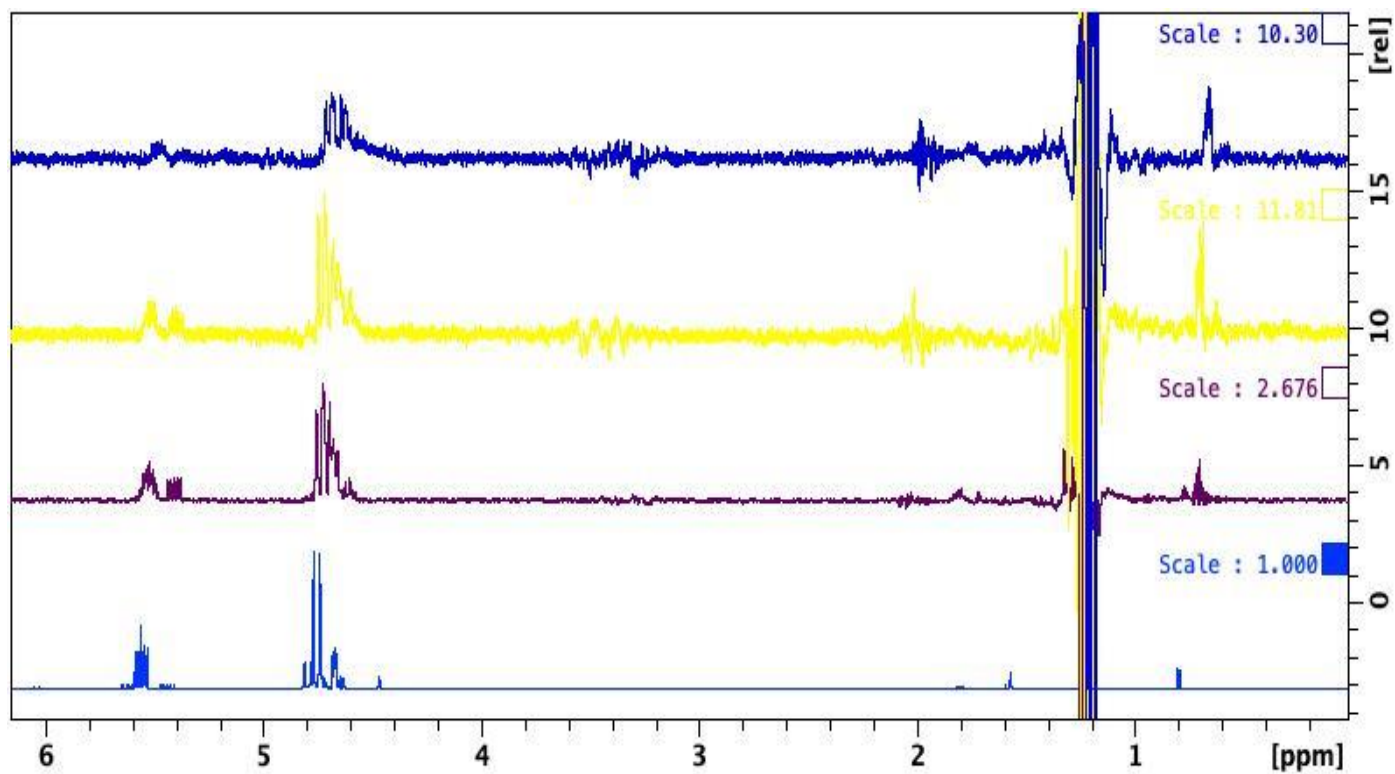
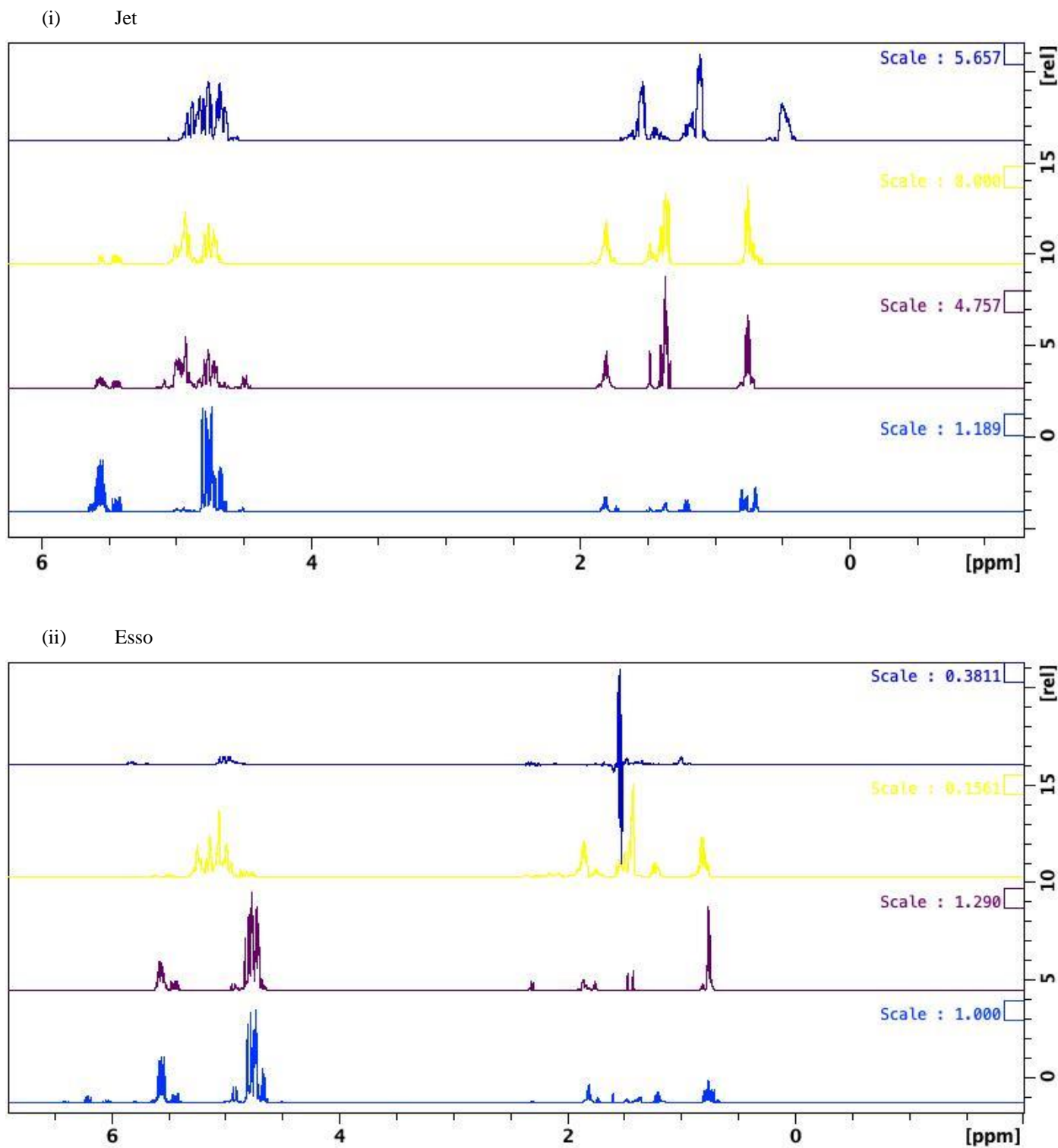
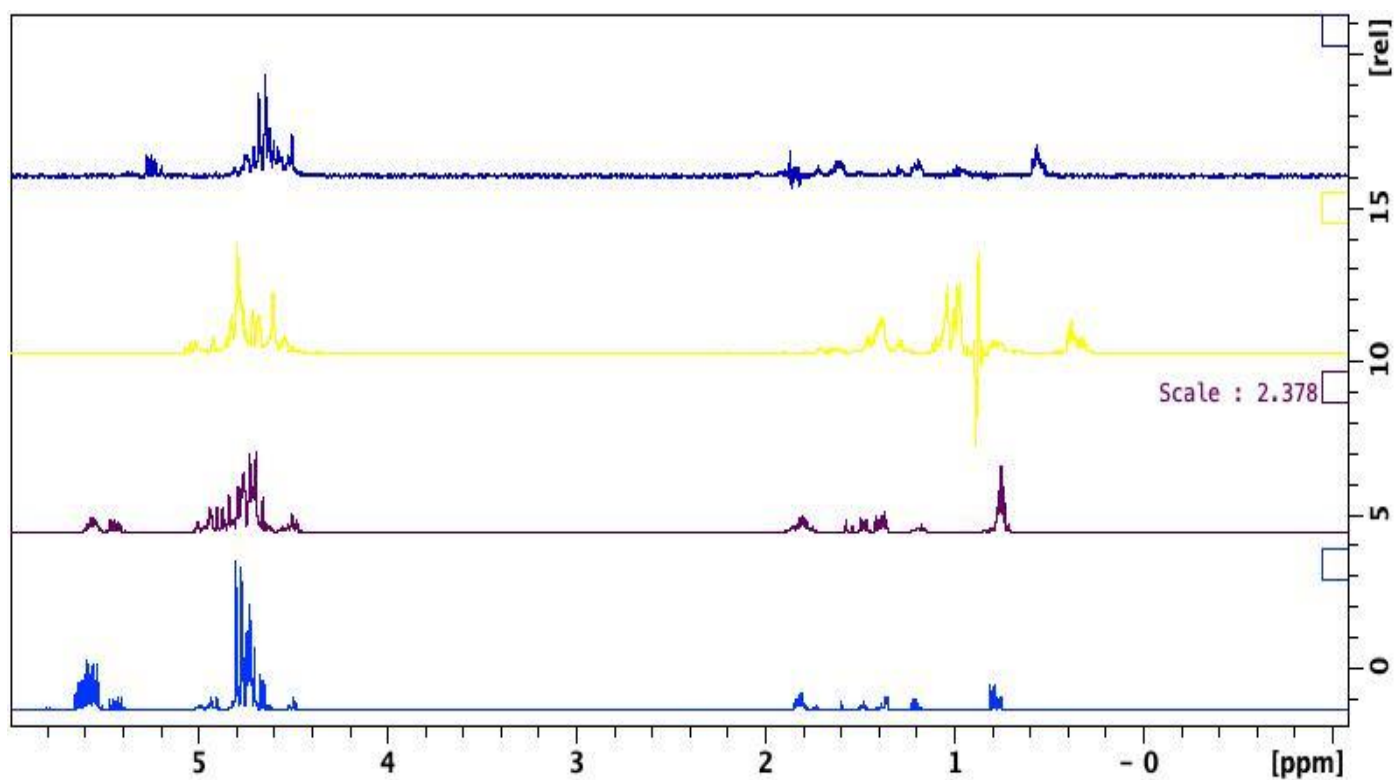


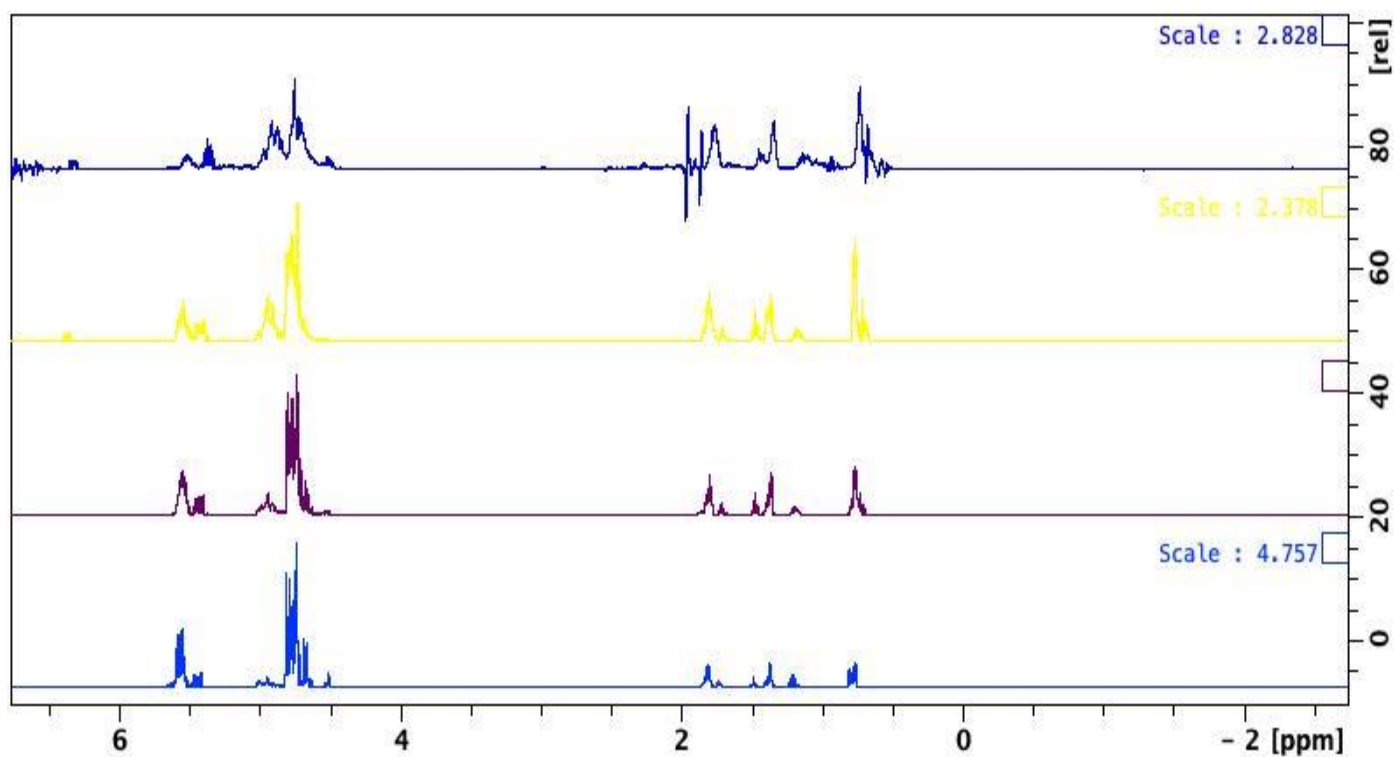
Figure S2. A comparison of ^1H selTOCSY of the mixture of 3-methyl-1-butene and 1-pentene of (i) Jet (ii) Esso (iii) Shell and (iv) Texaco (v) BP M native petrol sources with 25%, 50% and 75% evaporation rate



(iii) Shell



(iv) Texaco



(v) BP M

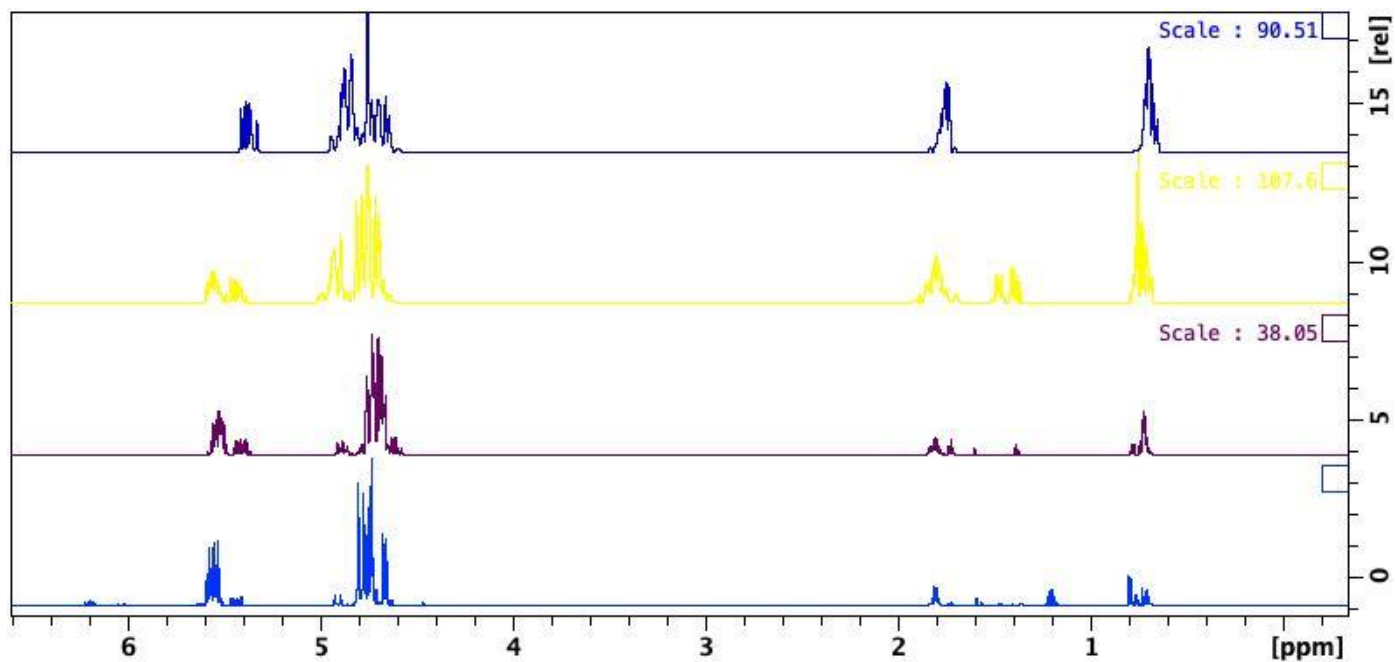
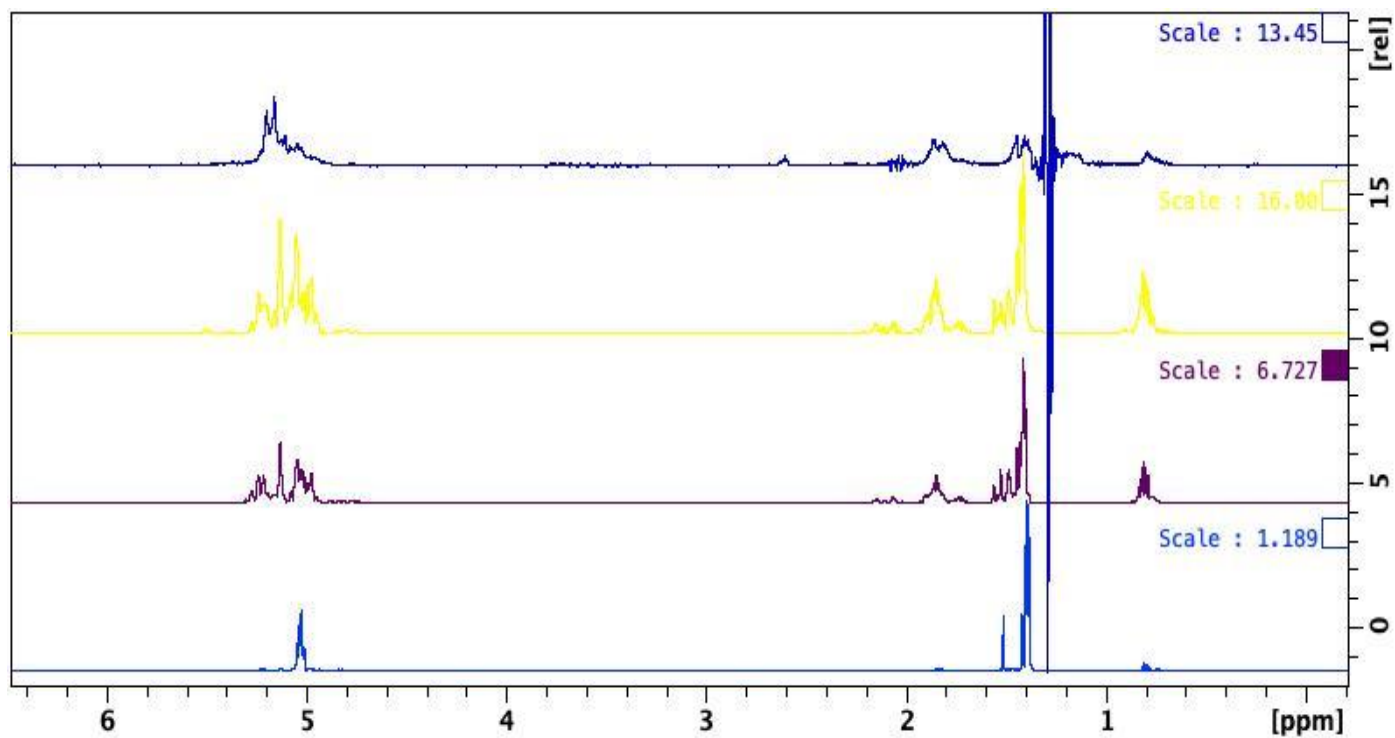
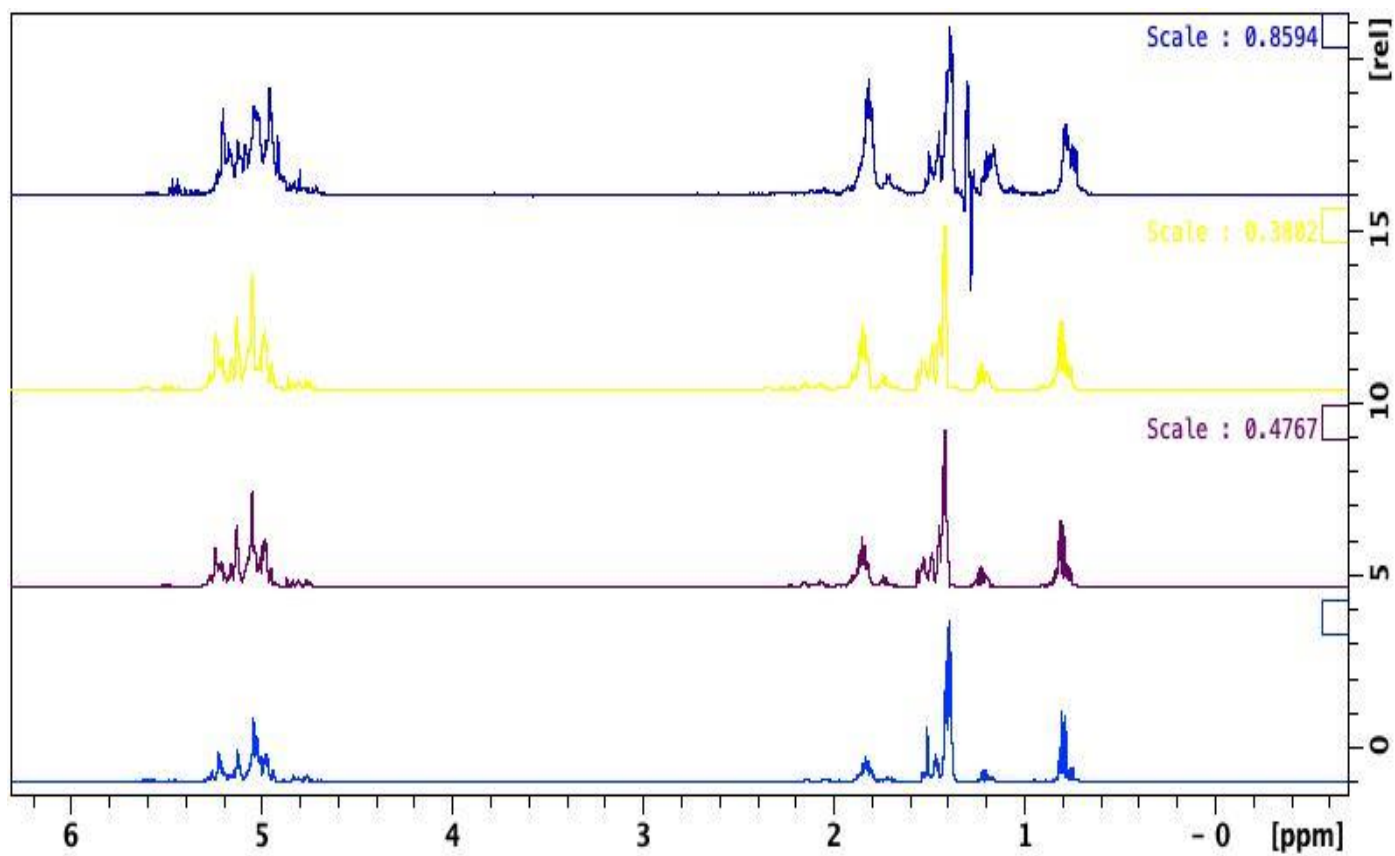


Figure S3. A comparison of ^1H seltocsy of the 2-methyl-2-butene of (i) Jet (ii) Esso (iii) Shell and (iv) Texaco (v) BP M native petrol sources with 25%, 50% and 75% evaporation rate

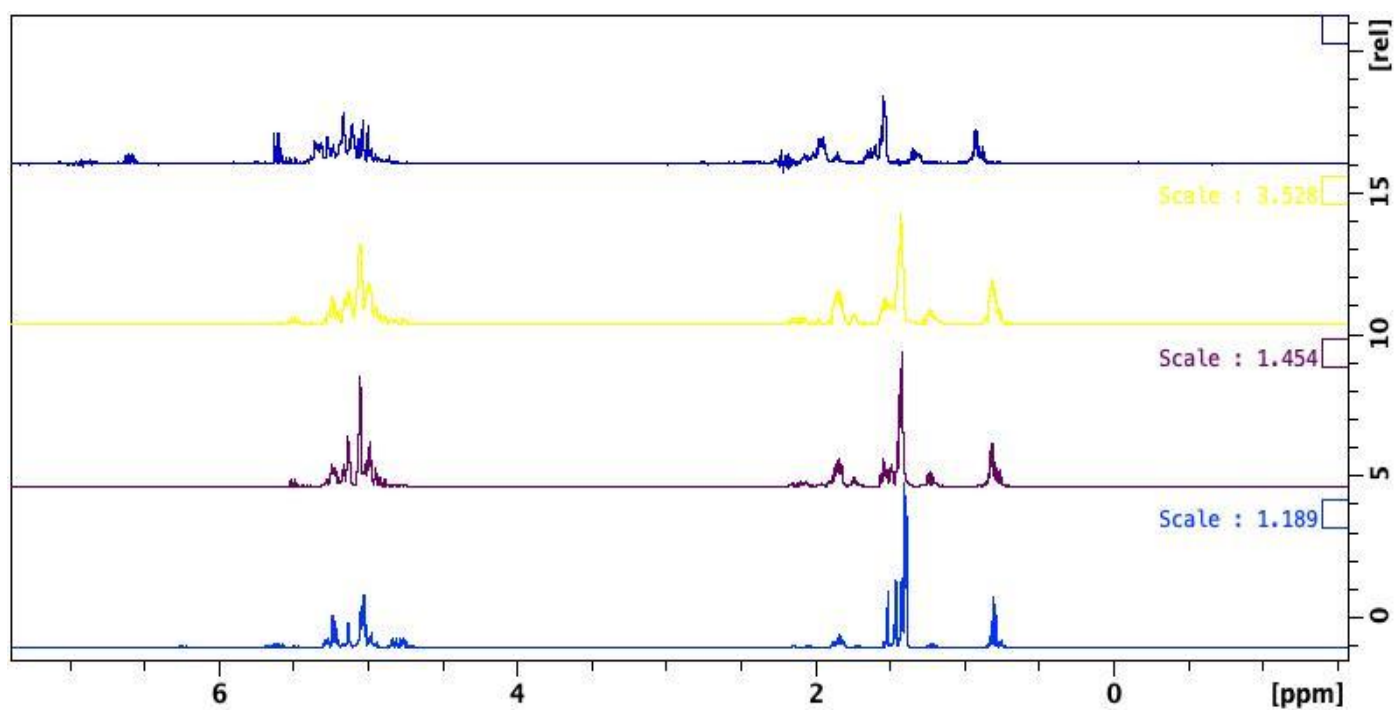
(i) Jet



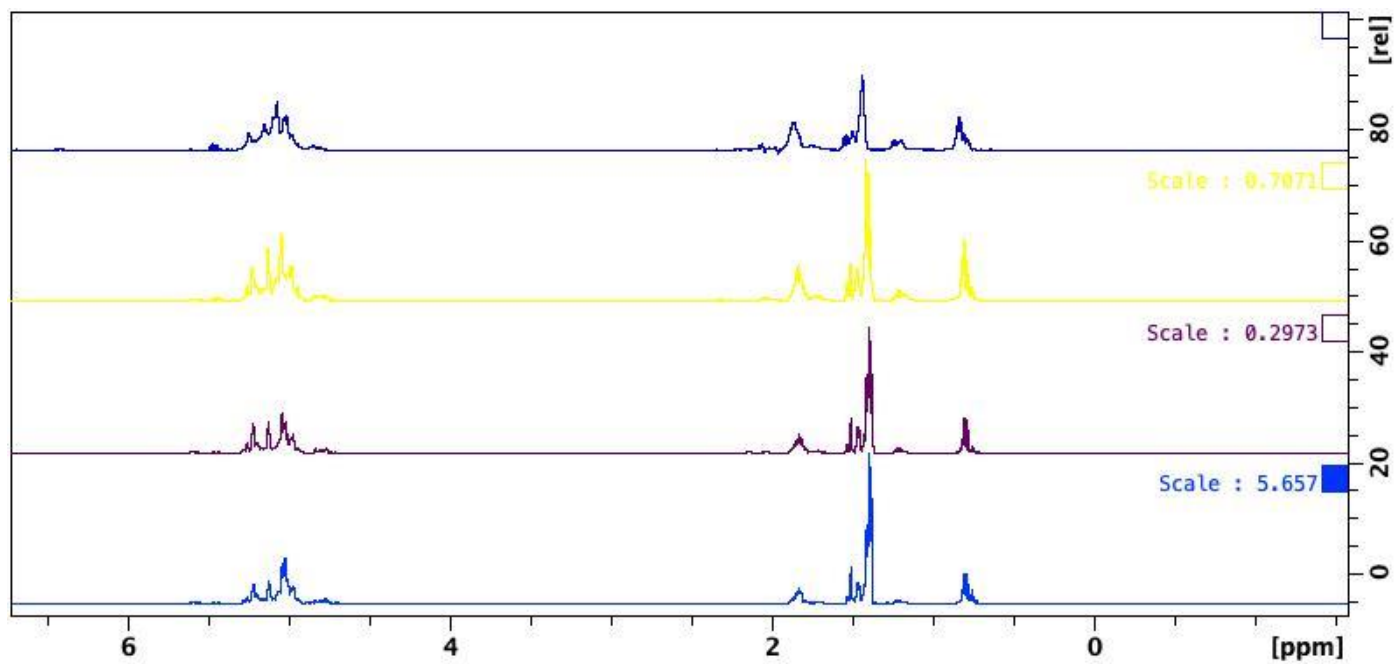
(ii) Esso



(iii) Shell



(iv) Texaco



(v) BP M

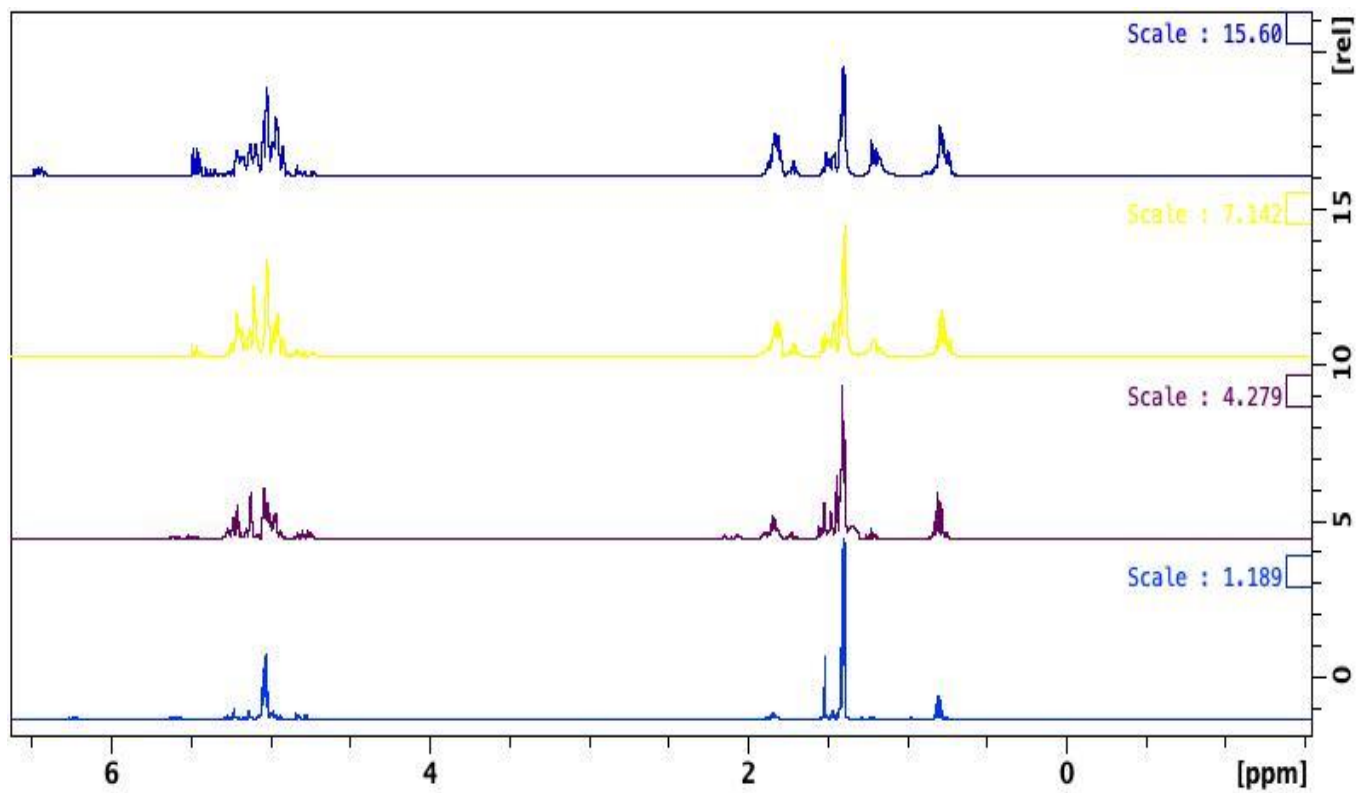
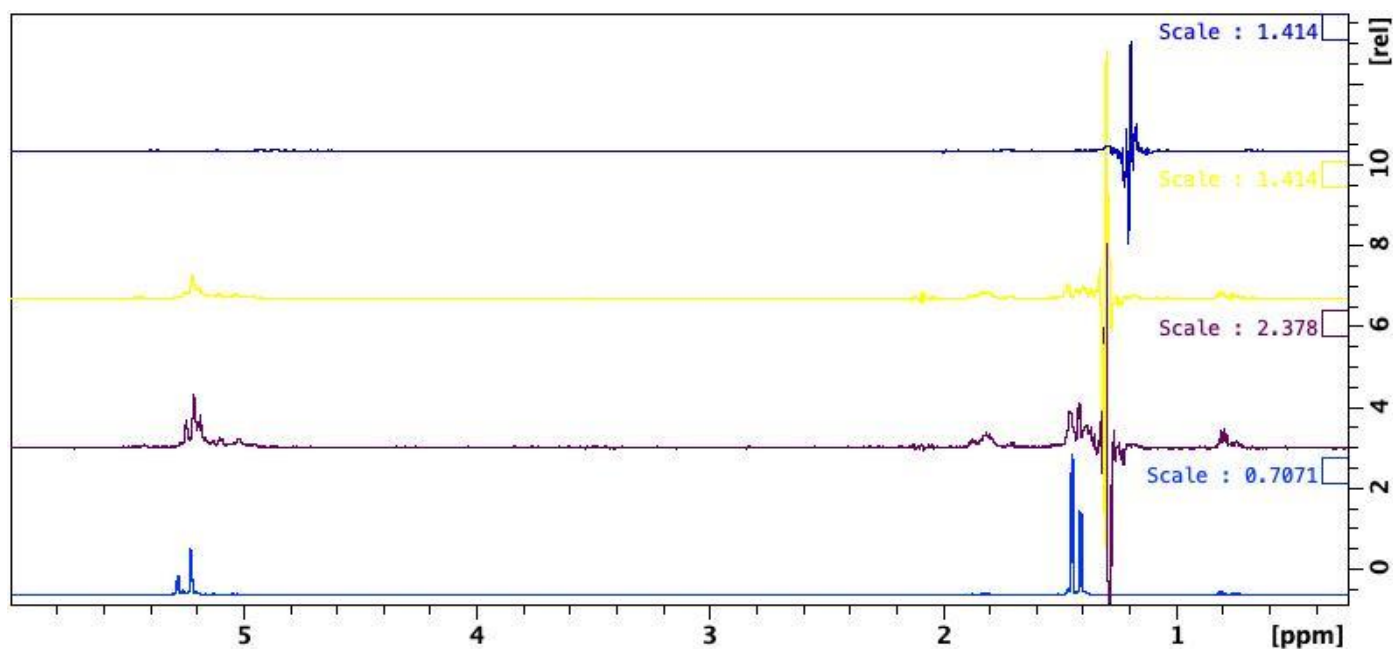
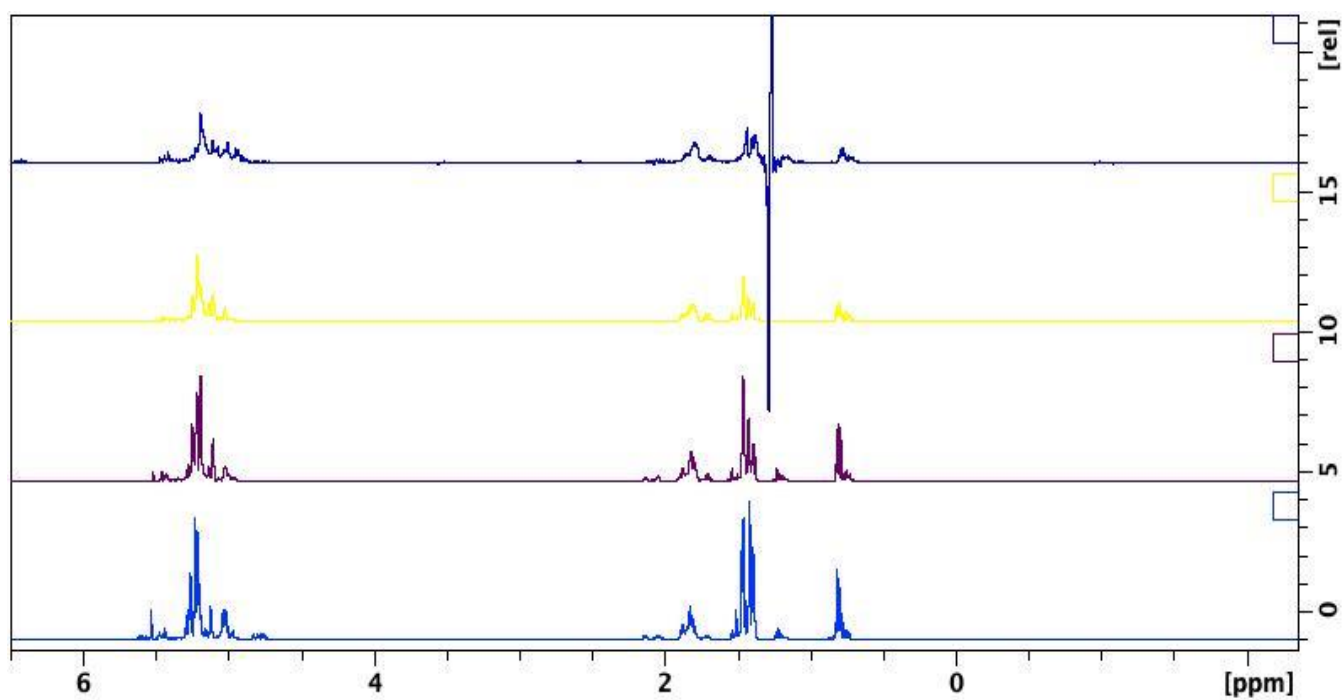


Figure S4. A comparison of ^1H seITOCY of the mixture of cis and trans-2-pentene of (i) BP S (ii) BP M (iii) Jet (iv) Esso (v) Shell and (vi) Texaco native petrol sources with 25%, 50% and 75% evaporation rate

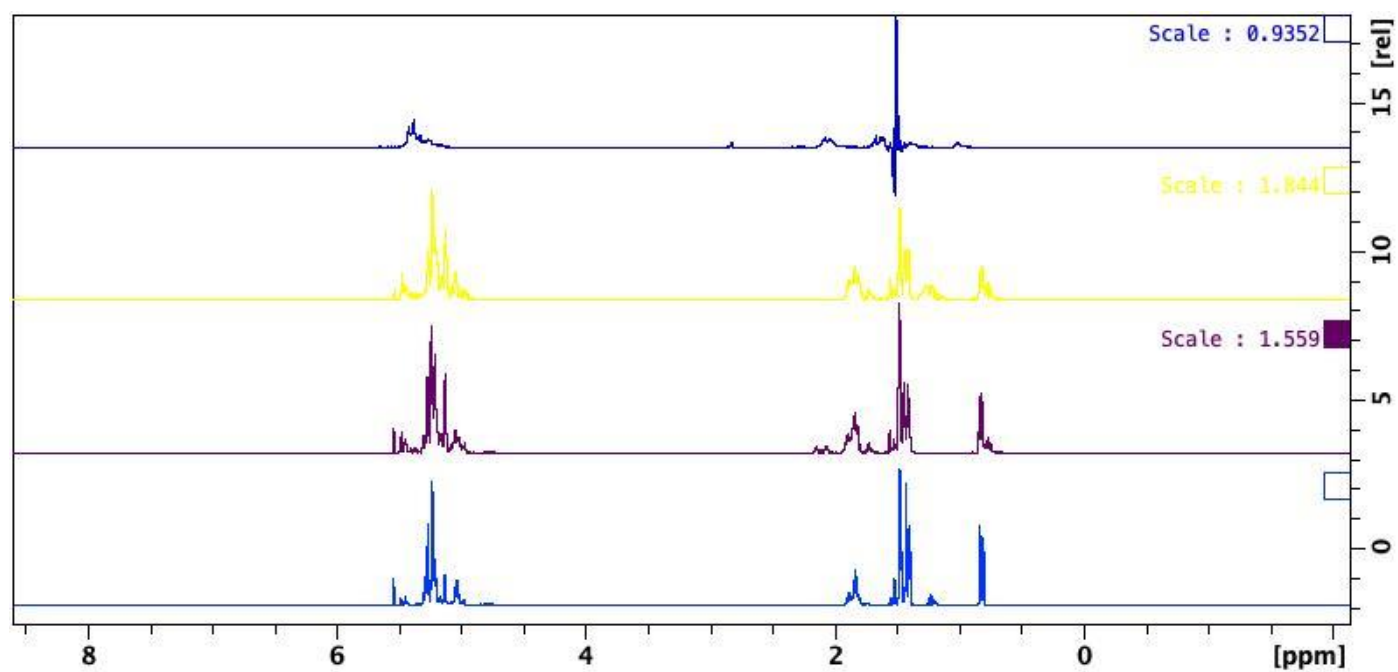
(i) BP S



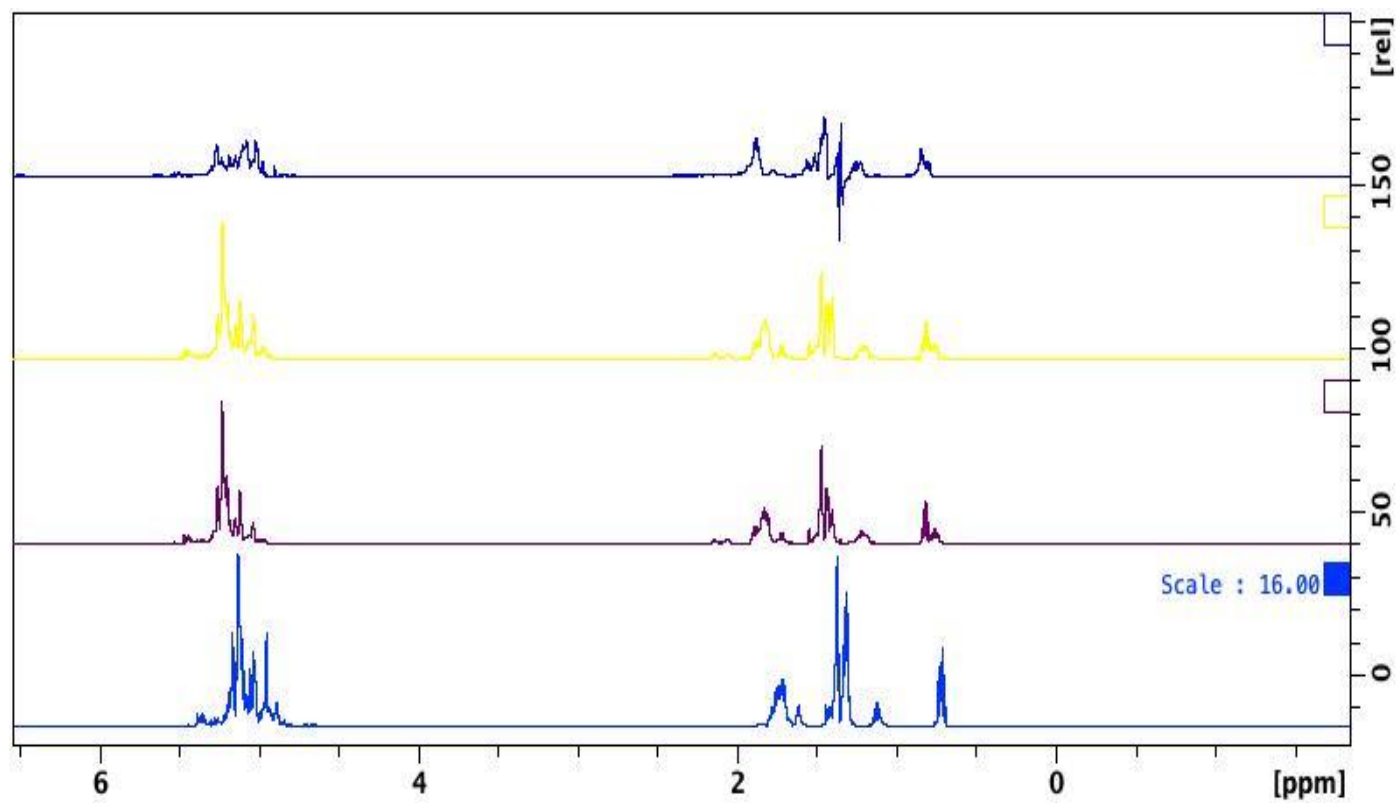
(ii) BP M



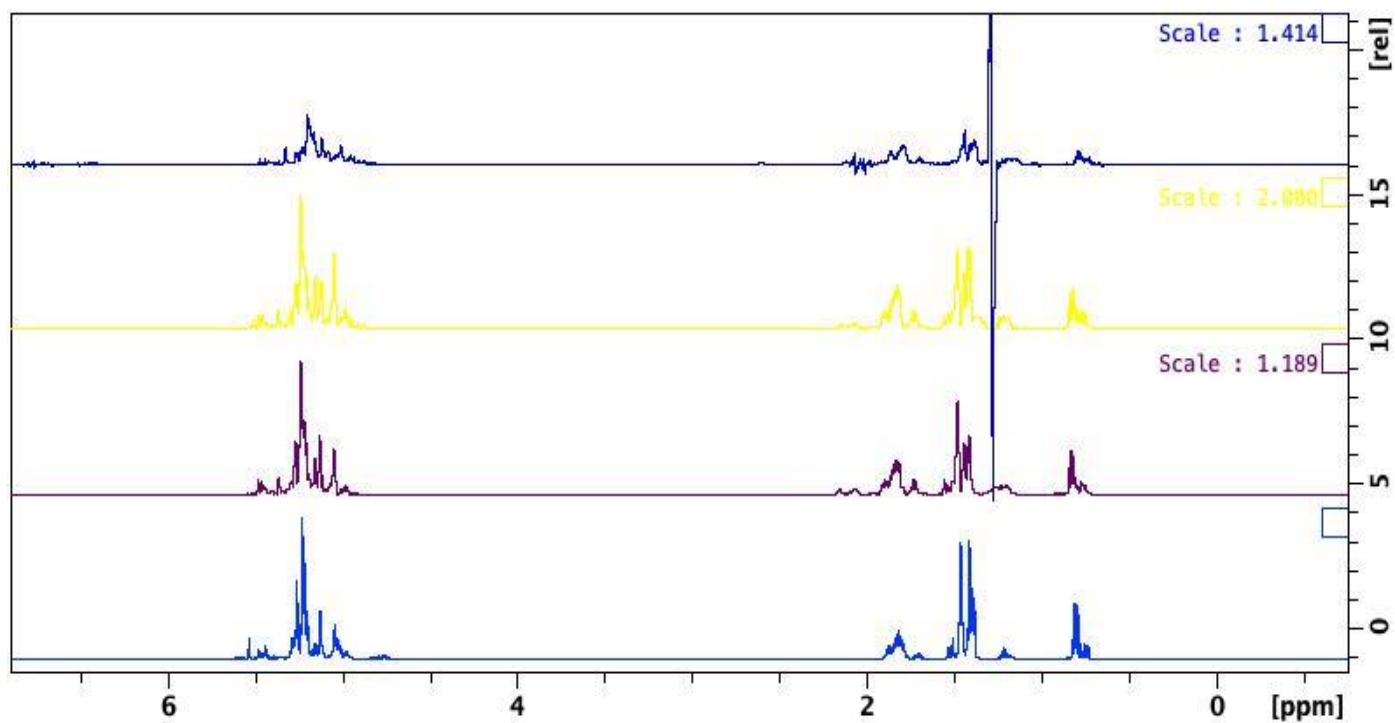
(iii) Jet



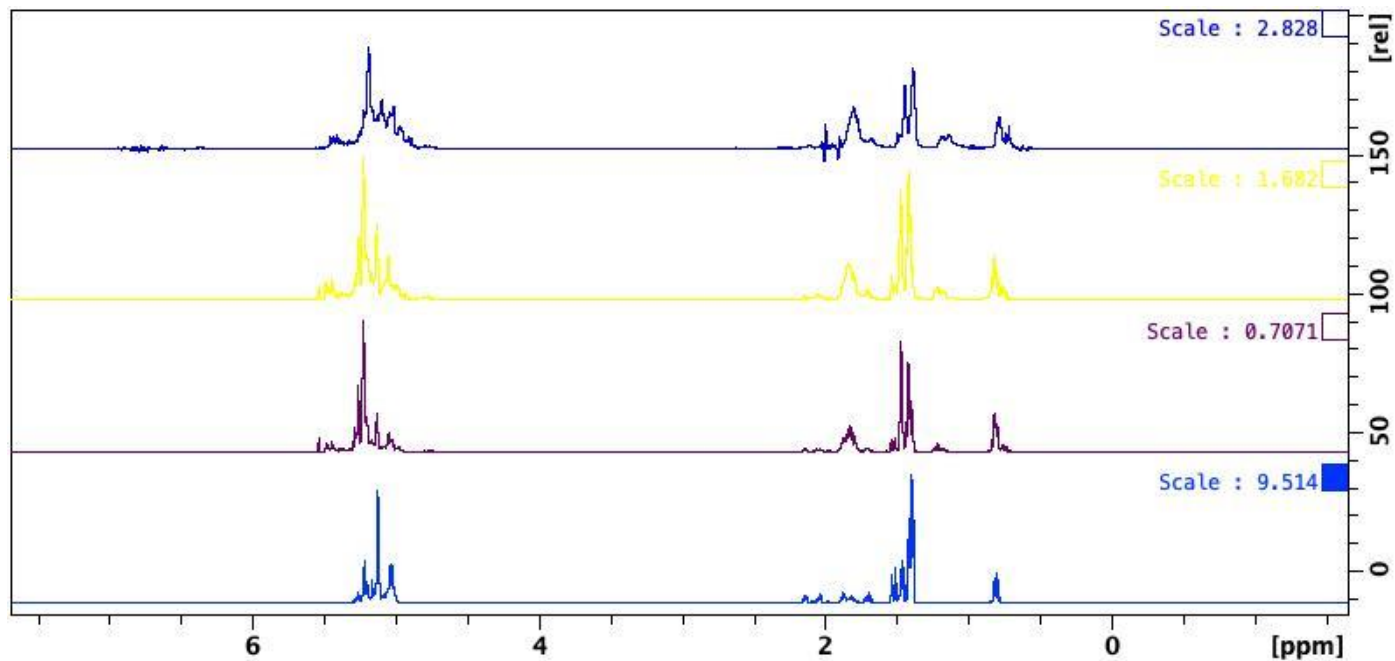
(iv) Esso



(v) Shell



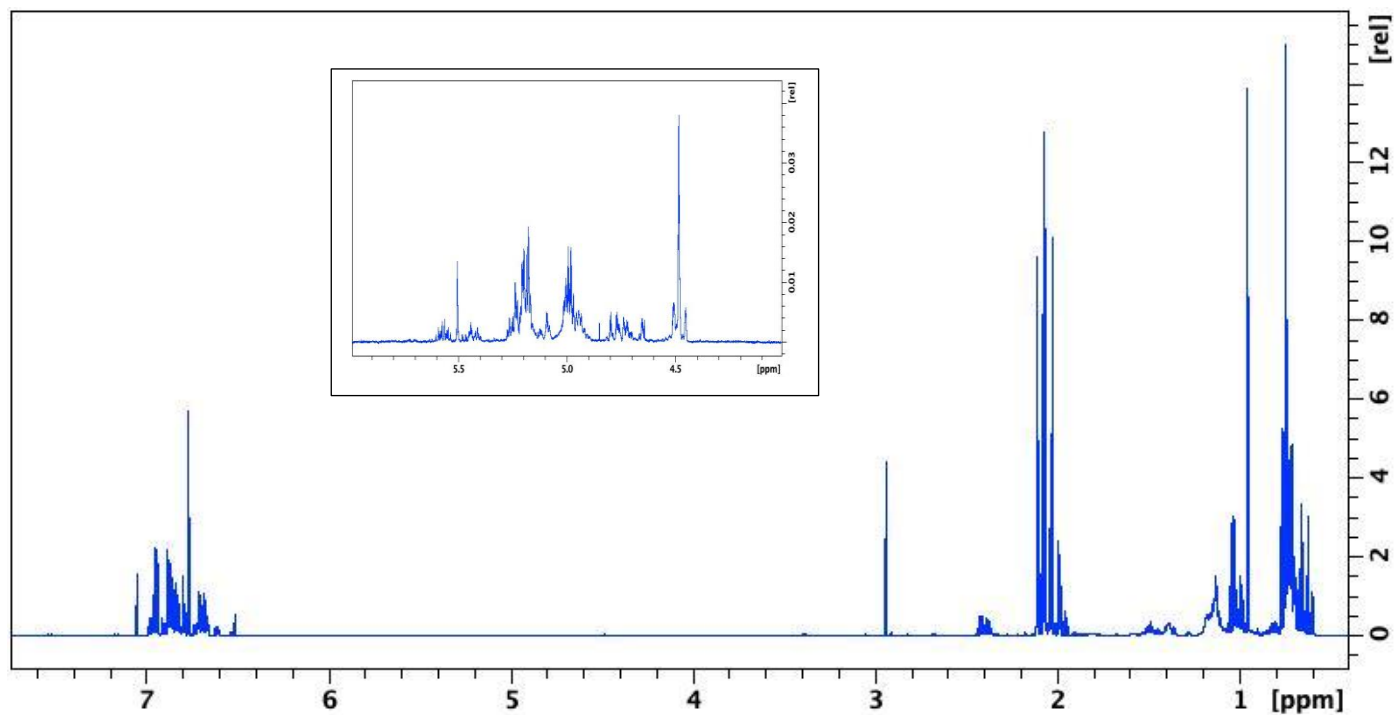
(vi) Texaco



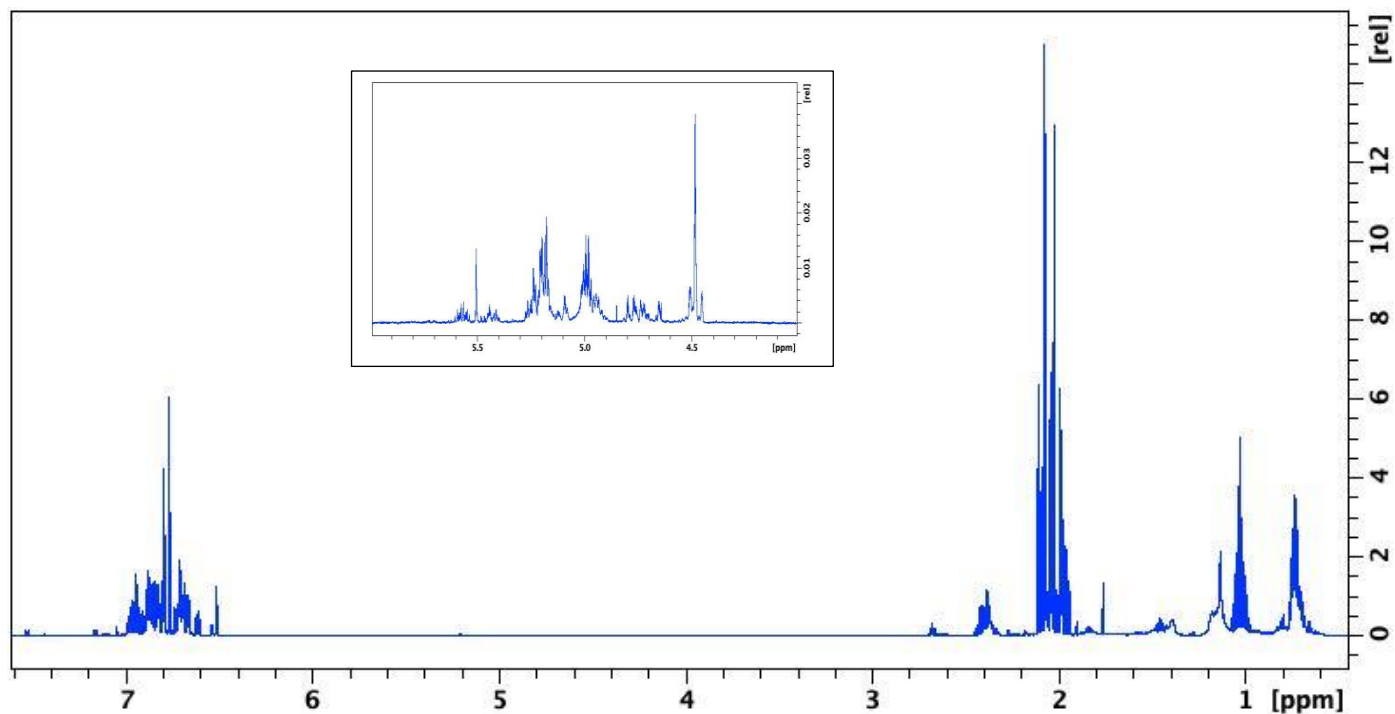
3.4 Identification of target olefin sets in burnt petrol samples on its own and on substrates.

Figure S5. The illustration of full and olefins area of ^1H NMR spectra of burnt on its own petrol in (i) Jet (ii) Esso (iii) Shell and (iv) Texaco (v) BP M

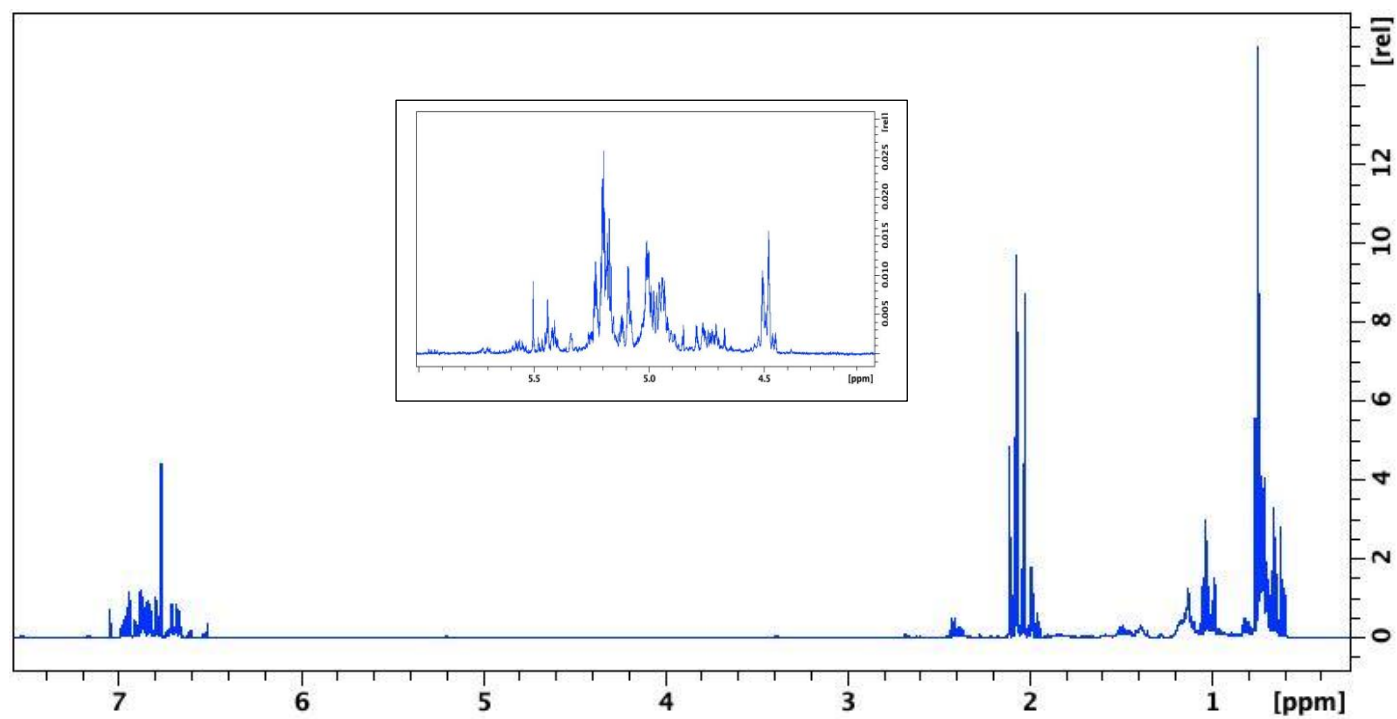
(i) Jet



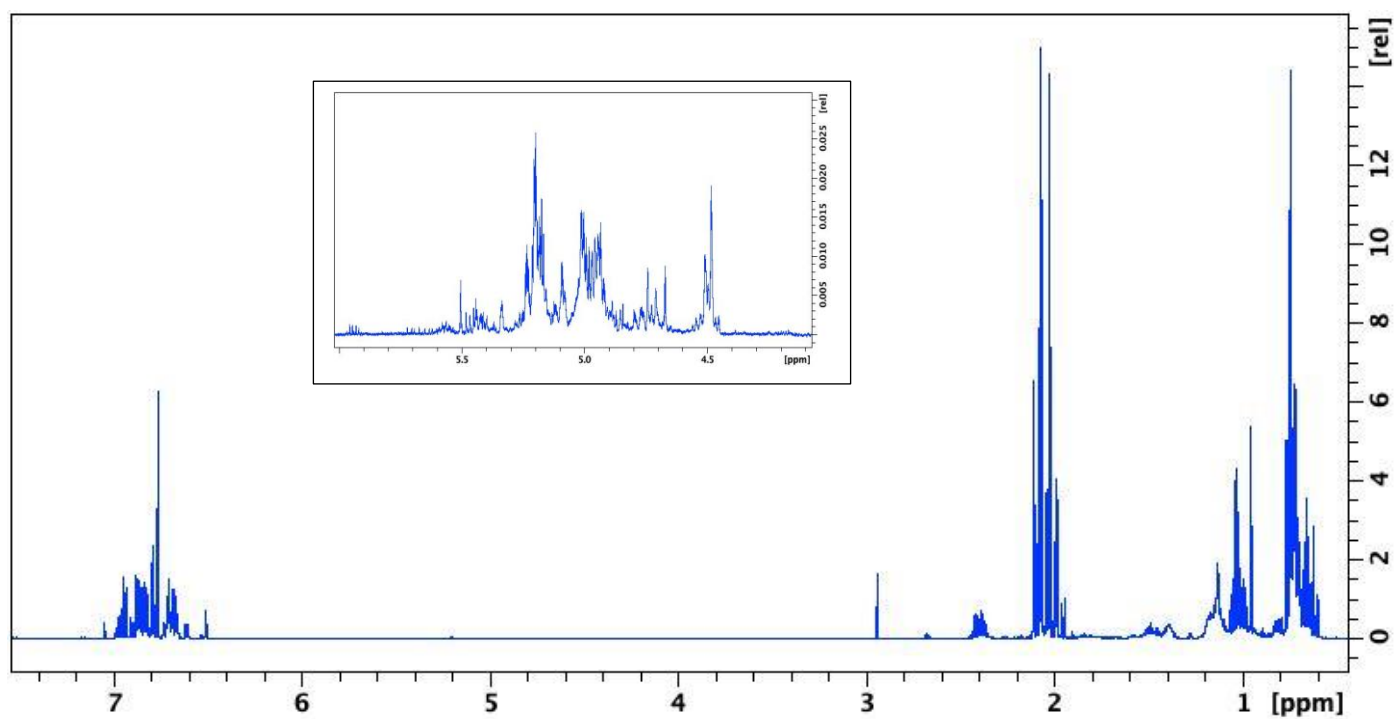
(ii) Esso



(iii) Shell



(iv) Texaco



(v) BP M

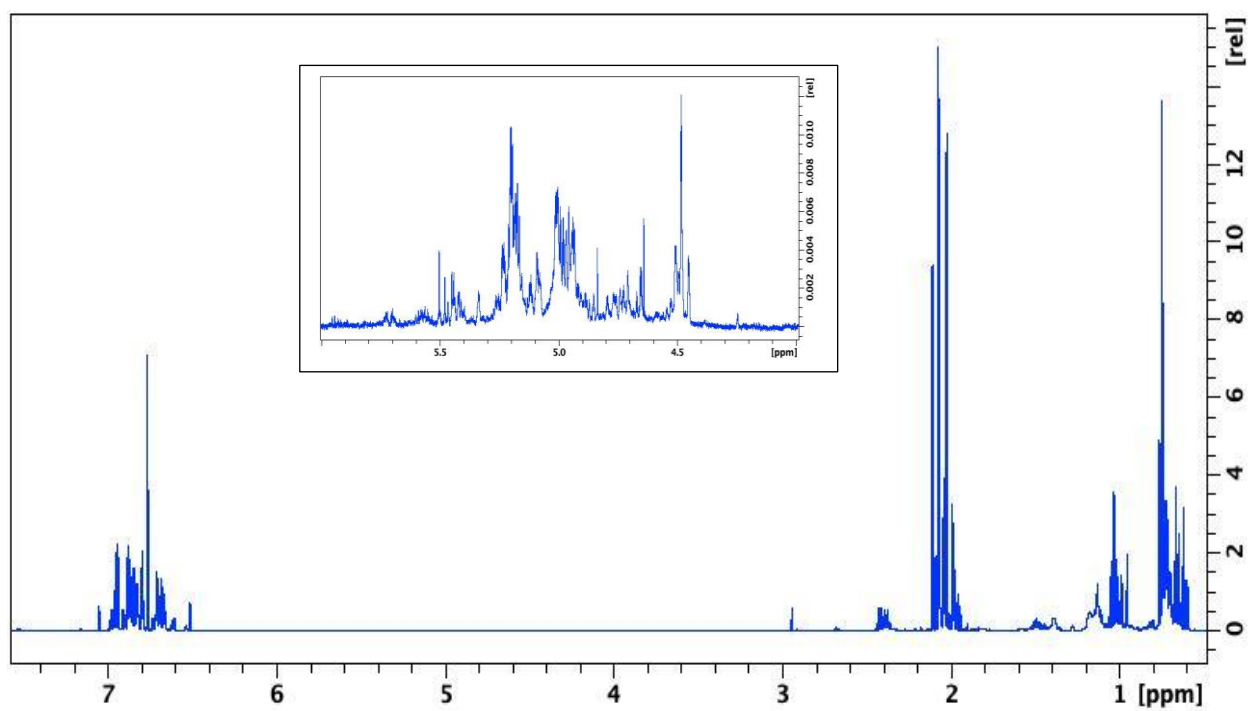
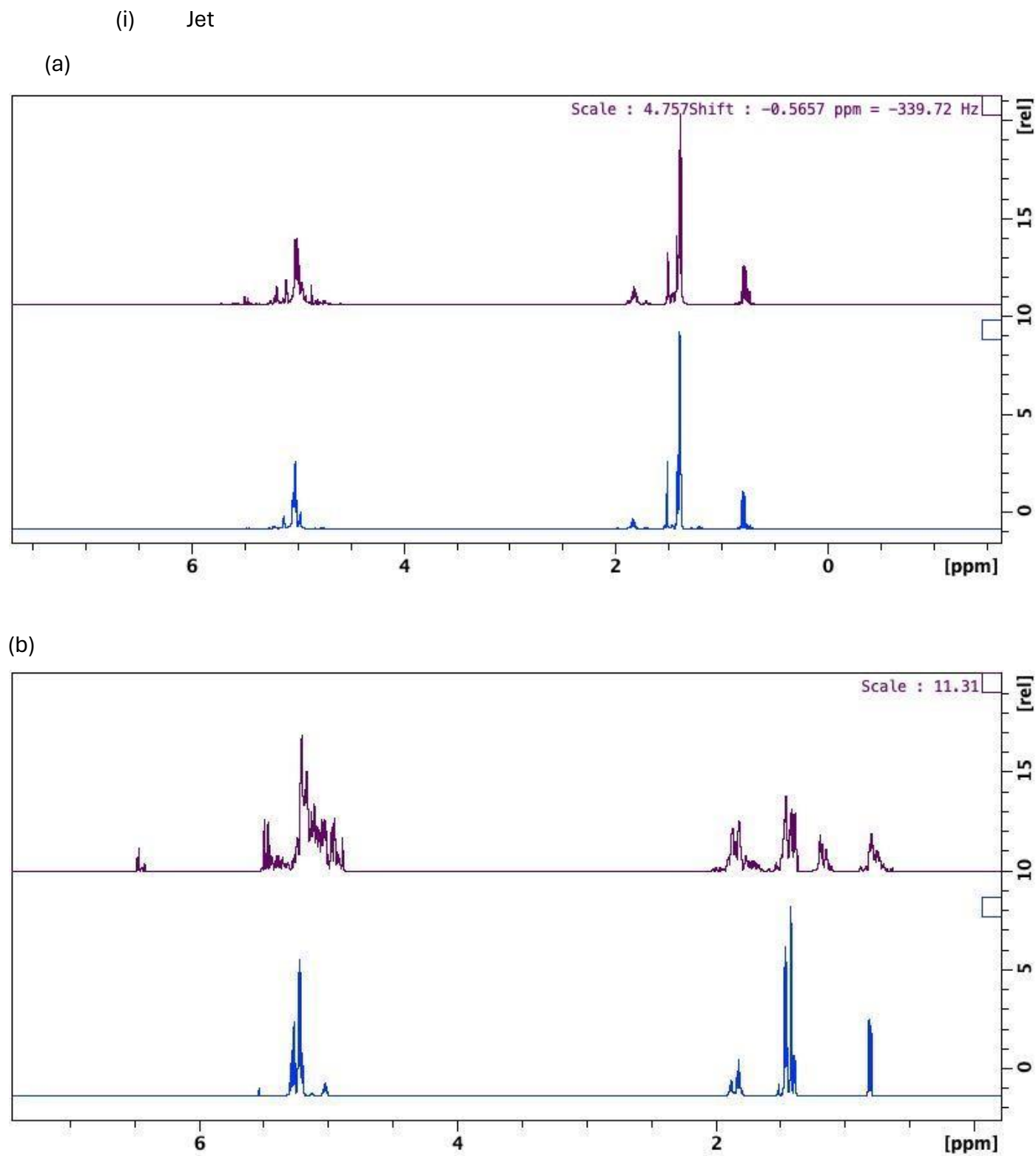
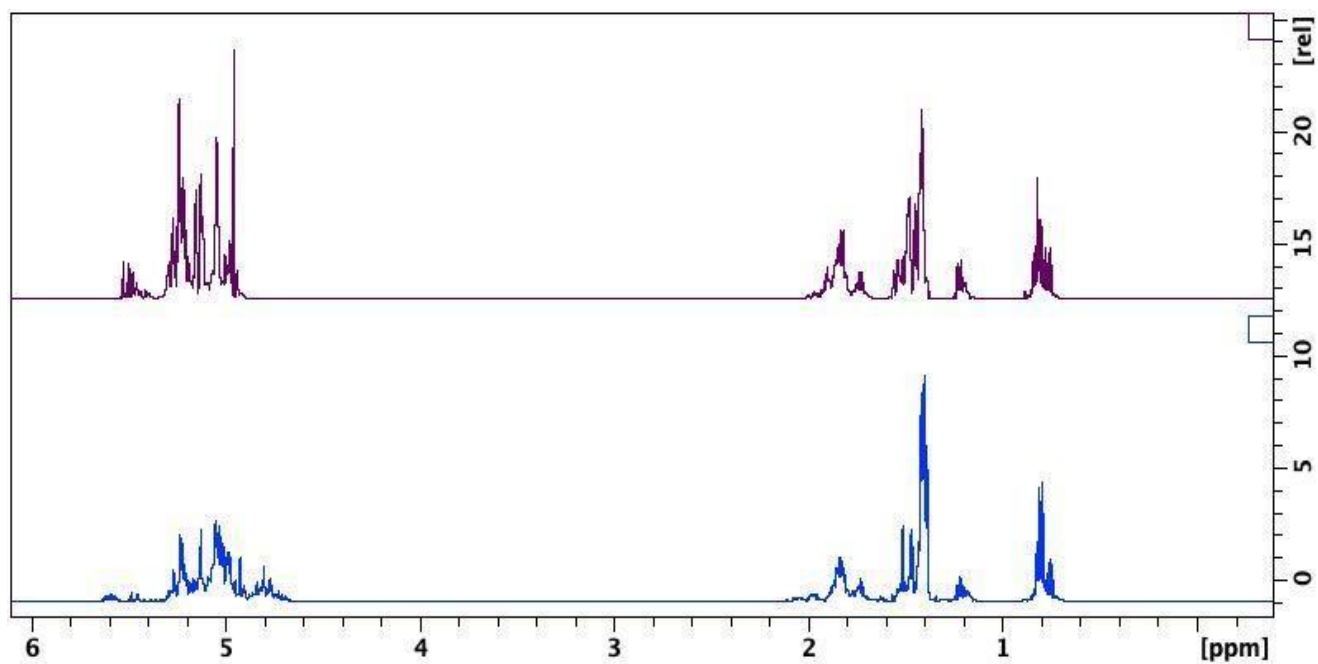


Figure S6. A representation comparison of ^1H selTOCSY spectra of a) 2-methyl-2-butene and b) the mixture of cis and trans-2-pentene in neat petrol vs in burnt cardboard in (i) Jet, (ii) Esso, (iii) Shell and (iv) Texaco, (v) BP M

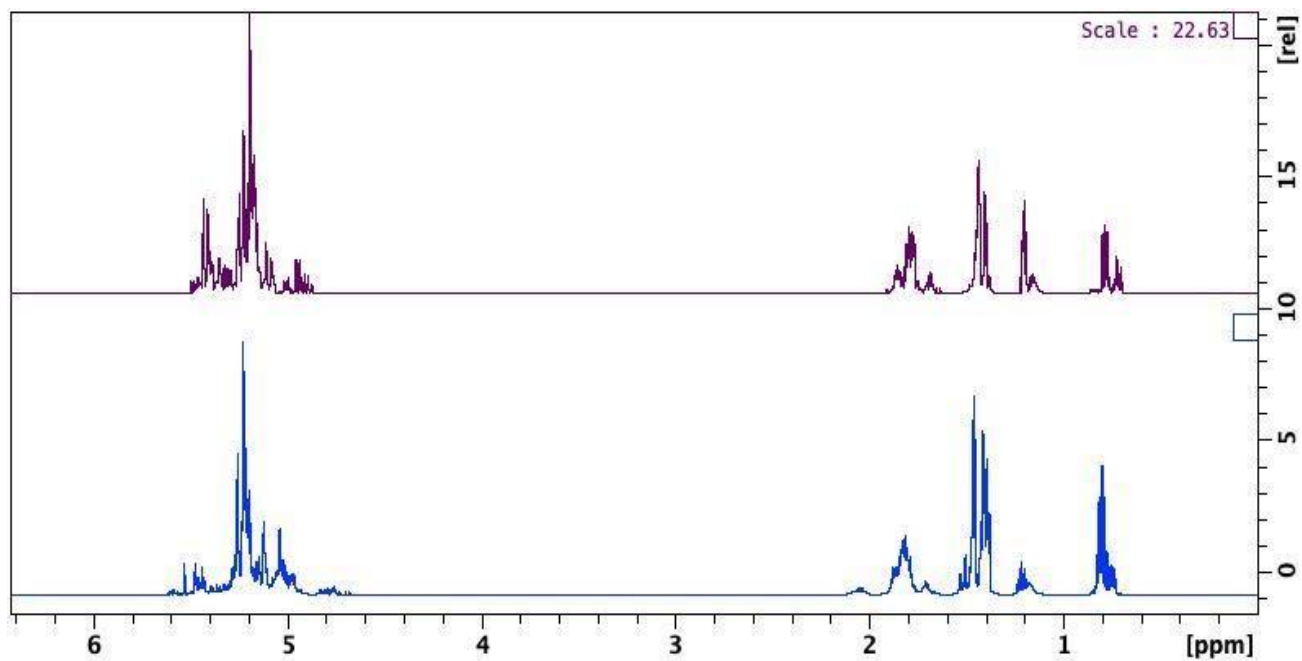


(ii) Esso

(a)

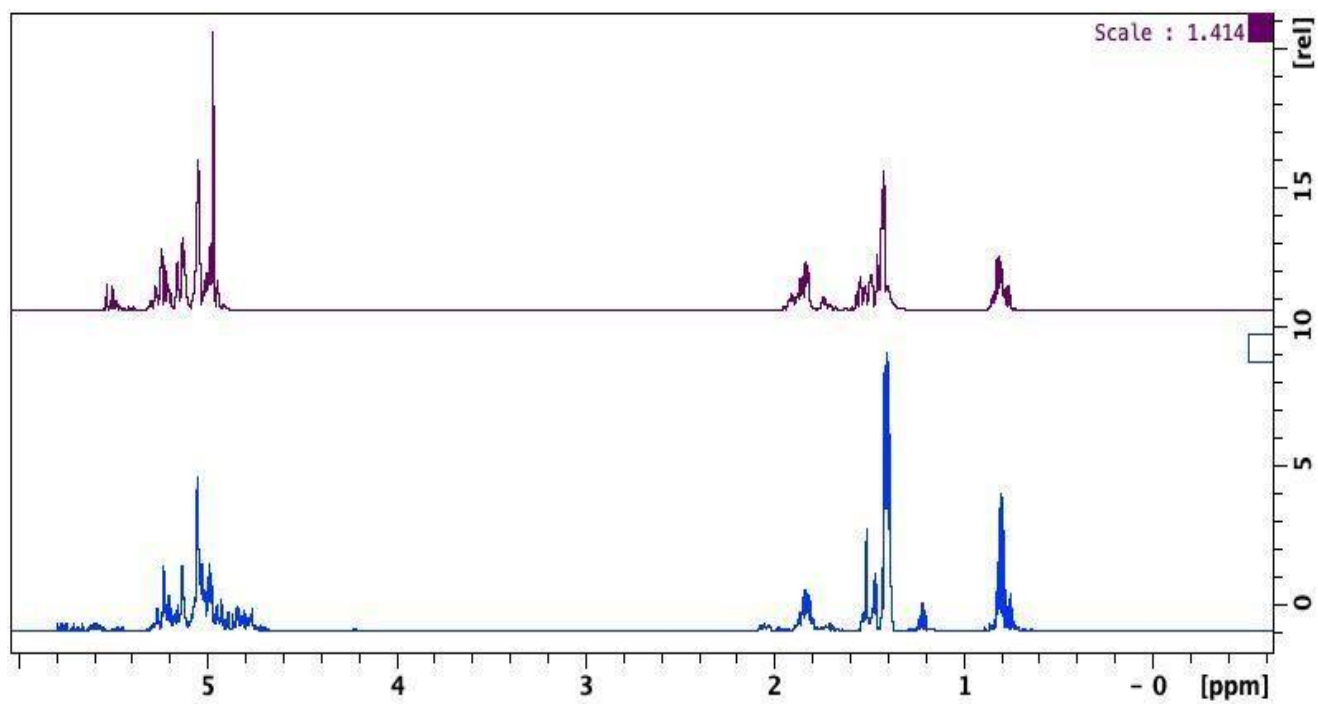


(b)

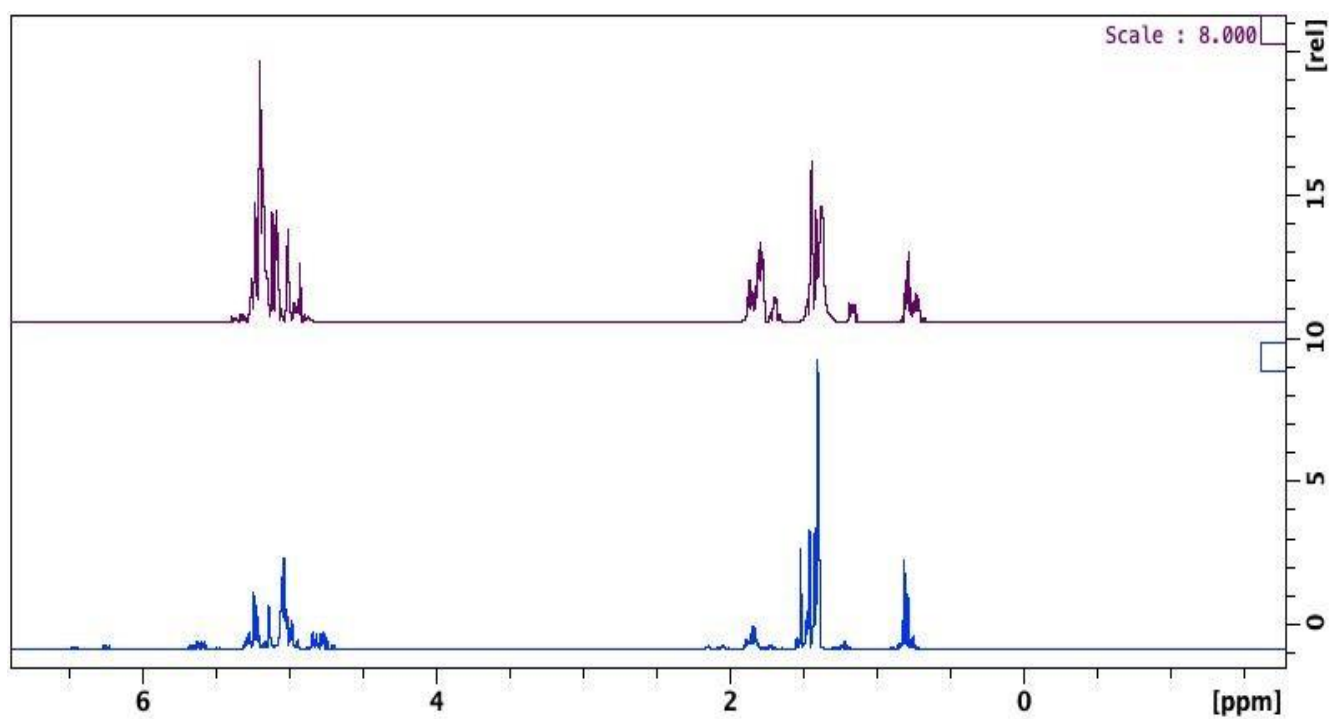


(iii) Shell

(a)

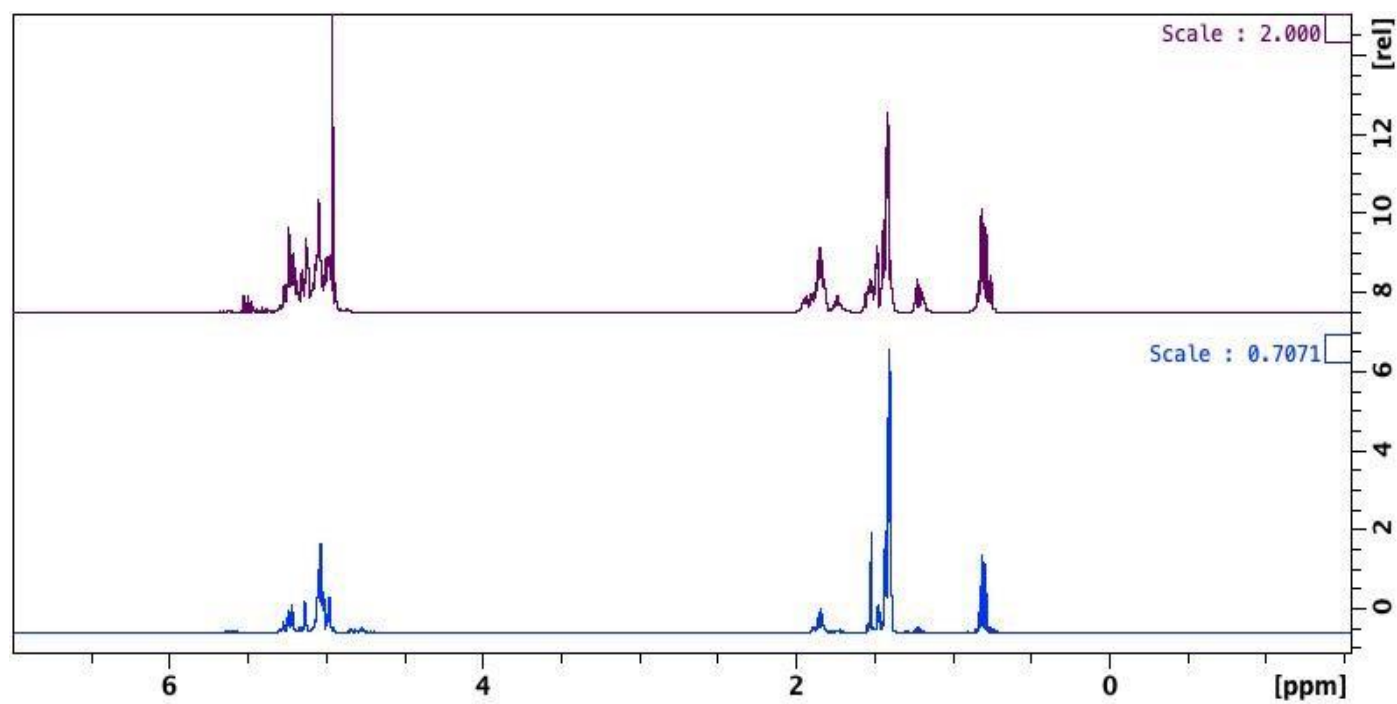


(b)

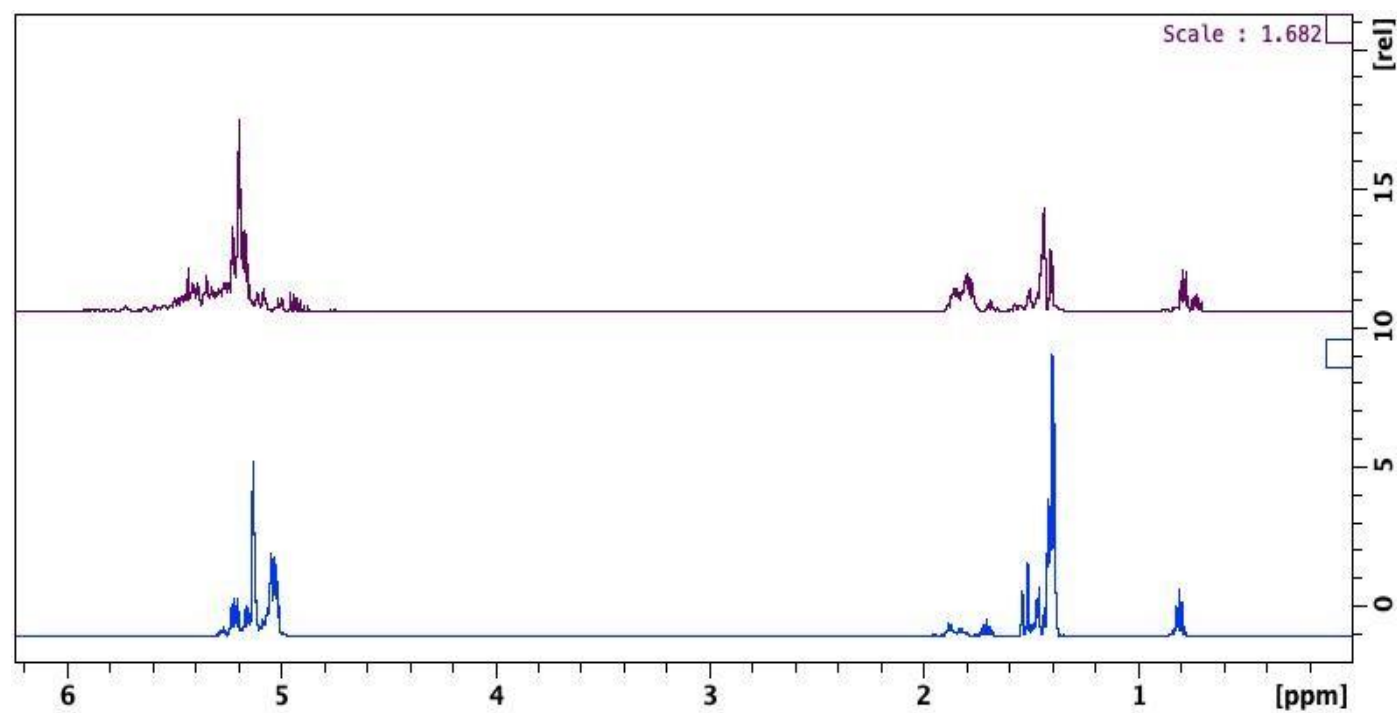


(iv) Texaco

(a)

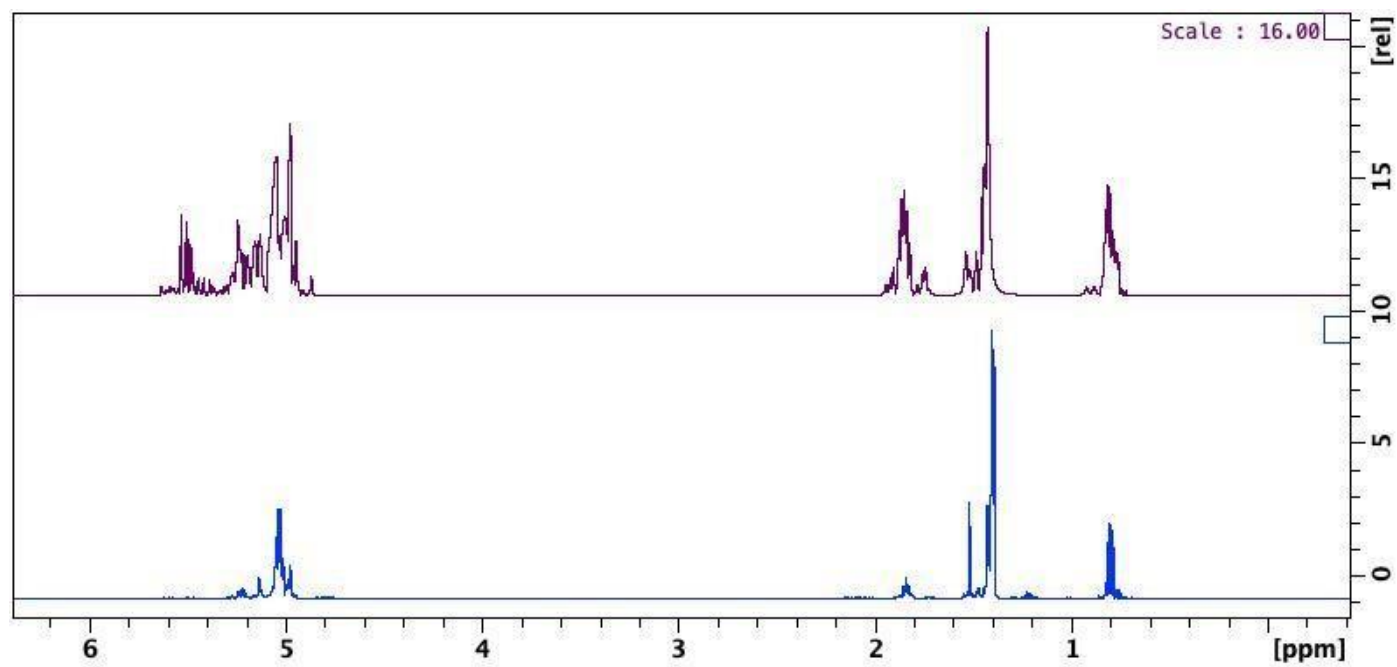


(b)



(v) BP M

(a)



(b)

